



The Food Insecurity Experience Scale

Development of a Global Standard for Monitoring Hunger Worldwide

The Food Insecurity Experience Scale

Development of a Global Standard for Monitoring Hunger Worldwide

Terri J. Ballard
Anne W. Kepple
Carlo Cafiero

Food and Agriculture Organization of the United Nations
Rome, 2013

Suggested Citation:

Ballard, T.J., Kepple, A.W. & Cafiero, C. 2013. *The food insecurity experience scale: developing a global standard for monitoring hunger worldwide.* Technical Paper. Rome, FAO. (available at http://www.fao.org/economic/ess/ess-fs/voices/en/).

Information on the authors:

Terri J. Ballard DrPH has worked on issues of food security and nutrition measurement for more than 10 years. She has been instrumental in validation of experience-based food insecurity scales and dietary diversity scores and has collaborated with the FANTA Project on developing the Household Hunger Scale, an impact indicator for the Feed the Future monitoring framework. She was organizer and scientific chair of the International Scientific Symposium on *Food and Nutrition Security information: from Valid Measurement to Effective Decision Making*, held in Rome in January, 2012.

Anne W. Kepple is a food and nutrition security specialist and qualitative researcher with a PhD in Nutrition from Cornell University, where she witnessed the seeds being planted for experiential food insecurity scales. Originally from the U.S., she has been a permanent resident of Brazil since 1993. She is a collaborating researcher at the State University of Campinas - UNICAMP, São Paulo, and has worked as a consultant for the Secretariat of Evaluation of the Brazilian Ministry of Social Development and Fight Against Hunger. Her research focuses on the challenges to intersectorial work aimed at guaranteeing the Human Right to Adequate Food and ways to promote the link between food security information and action.

Carlo Cafiero, agricultural economist, is a Senior Statistician in the FAO Statistics Division and Project Manager for the *Voices of the Hungry* project. He holds a PhD in Ag and Resource Economics from the University of California at Berkeley and a doctoral degree in Agricultural Policies from the University of Naples Federico II and has been teaching Statistics, Agricultural Economics and Agricultural Policy in Italy, the US, Syria and Bangladesh. His research interests include quantitative analysis of food security, with special interest to the dynamics of food price and to the measurement of food insecurity. In the last couple of years, he has been responsible for the revision of the FAO methodology to estimate the prevalence of undernourishment.

Table of Contents

List of Tables and Figures

Acknowledgements

The authors would like to dedicate this paper to Kathy Radimer, whose groundbreaking work, together with Cathy C. Campbell in the late 1980's, led the way towards developing a valid methodology for better understanding, defining and measuring food insecurity based on the perspective of people who experience it. Her work has been advanced through the efforts of many research organizations, including Cornell University, Tufts University, the Food and Nutrition Technical Assistance Programme, the Economic Research Service of the USDA, and numerous researchers in Latin America whose work led to a harmonized version of the Latin American and Caribbean Food Security Scale (Escala Latinoamericana y Caribeña de Seguridad Alimentaria - ELCSA). Much of the support for FAO's work with food insecurity experience scales has come from grants to the FAO Division of Nutrition since 2005 through different phases of a collaboration between the European Union and FAO on improving food security information for decision making (www.foodsec.org).

Inspiration for this technical paper and the development of the Food Insecurity Experience Scale was provided by the work of the ELCSA Scientific Committee composed of Ana Maria Segall-Corrêa (University of Campinas, Brazil); Martha Cecilia Álvarez-Uribe (University of Antioquia, Colombia); Hugo Melgar-Quiñonez (McGill University, Canada) and Rafael Pérez-Escamilla (Yale University, USA) and the Spanish language manual of the ELCSA which they wrote with support from FAO (FAO, 2012a).

We are grateful to Edward E. Frongillo of the University of South Carolina and Hugo Melgar Quiñonez of McGill University for their careful reviews of parts of earlier versions of this paper. We wish to express our heartfelt thanks to Mark Nord of the Economic Research Service, U.S. Department of Agriculture for his guidance in better understanding the conceptual underpinnings of this form of measurement and his technical assistance in the statistical and theoretical treatment of data to improve validity and comparability of the measurement.

Ellen Wielezynski was responsible for editing, formatting and final organization of the paper.

Funding for this technical paper on the Food Insecurity Experience Scale has been provided by the European Union through the European Union/FAO Improved Global Governance for Hunger Reduction Programme (outputs 2.1 and 2.4), with an additional contribution by the FAO Multi-Donor Fund for the *Voices of the Hungry* project (VOH) (http://www.fao.org/economic/ess/ess-fs/voices/es/).

Foreword

FAO has been a partner in the development, validation and use of food (in)security scales since 2006 and has had an important role in furthering the research on the Latin American and Caribbean Food Security Scale (Escala Latinoamericana y Caribeña de Seguridad Alimentaria - ELCSA) through financial support for regional conferences on food security measurement and capacity-building in developing countries regarding validation and use of these tools (Melgar-Quinonez, 2010; FAO, 2012a). Because no single instrument measures food (in)security in all its dimensions, there has been substantial research devoted to developing, refining and validating different approaches for measuring the state of food insecurity. The development of measures of whether people are experiencing food insecurity because of limited access to food, and if so at what level of severity, constitutes an important addition to the suite of commonly used food security measures.

Building on the experience of the Latin American scale, the FAO *Voices of the Hungry* project (VOH) has developed an experience-based food insecurity scale module called the Food Insecurity Experience scale (FIES), which is based on a short form of the ELCSA. The FIES will be used as a common metric for measuring food insecurity at several levels of severity, across different geographic areas and cultures.

Many efforts are going into the search for global indicators capable of measuring food insecurity in a comparable manner across different parts of the world, as seen by current discussions on indicators for the post-2015 development agenda (http://www.un.org/en/ecosoc/about/mdg.shtml). The FIES has a potentially important role for monitoring food security within this process. It is particularly well-suited to monitoring systems that meet principles recently identified by the Committee on World Food Security in the 2012 Global Strategic Framework for Food Security and Nutrition (CFS, 2012a).

Acronyms and abbreviations:

CFS	Committee on World Food Security
ELCSA	Escala Latinoamericana y Caribeña de Seguridad Alimentaria
FAO	The Food and Agriculture Organization
FIES	Food Insecurity Experience Scale
GWP	Gallup® World Poll
IRT	Item Response Theory
MDG	Millennium Development Goal
USDA	United States Department of Agriculture
US HFSSM	United States Household Food Security Survey Module
VOH	*Voices of the Hungry* project

1. Introduction

One of the most basic needs shared by all human beings is the need for enough food. Access to enough nutritionally adequate food was declared a basic human right at the World Food Summit[1] in Rome in 1996, reflecting an evolution in the discourse over a period of several decades and growing international political commitment to end hunger.

Yet despite progress in our understanding of how to guarantee this basic human right and how to conceptualize and monitor it, hunger continues to affect hundreds of millions throughout the world. FAO has estimated that one in eight people in the world (870 million) suffered from chronic undernourishment in 2010-2012 (FAO, 2012b). The social inequalities at the root of the problem – from the household to the international level - have proven resistant to change, while new factors such as climate change, demand for biofuels and food price volatility have emerged to exacerbate the problem.

Information regarding the distribution and severity of hunger and food insecurity in the population and the characteristics, circumstances, and location of those most affected can contribute to building political will, designing effective policies, and targeting allocation of resources. While information alone is clearly insufficient, it can be a powerful tool.

Efforts to measure, monitor, and eradicate hunger and food insecurity have been underway for decades, often reflecting the perspectives of different sectors: economy, agriculture, health and nutrition. Combined scientific and political efforts have converged on a growing consensus regarding definitions, terminology, conceptual frameworks, and measures of hunger that reflect a more integrated, multi-sectorial perspective (CFS 2012b). Because no single indicator can account for the many dimensions of food and nutrition security, suites of indicators are being proposed to capture this complexity within the diversity of different contexts, a useful step towards promoting multisectorial approaches for improving food security (FAO, 2012c; Coates, 2013).

The Food Insecurity Experience scale (FIES) is expected to make an important contribution to any suite of food and nutrition security indicators. It has particular potential as a cross-disciplinary indicator capable of promoting the link between different sectorial perspectives, for example, the link between nutrition and agriculture. It is an experience-based metric of severity of food insecurity that relies on people's direct responses to a series of questions regarding their access to adequate food. Accumulated evidence over the past two decades has convinced FAO of the potential for using this method of measurement to provide valid and reliable population estimates of food insecurity in the different countries of the world.

[1] http://www.fao.org/wfs/index_en.htm

The following section of this paper provides an overview of current terminology and conceptions related to food insecurity and lays the groundwork for a discussion of how the FIES contributes to the measurement of food insecurity– what it measures and how it can be used to complement other indicators. The theoretical basis underlying the scale and a summary of its evolution are also discussed. The *Voices of the Hungry* project (VOH), the FAO initiative aimed at exploring the potential of the FIES as a new global standard for measuring food insecurity, is described in Section 3. Section 4 addresses the need for careful linguistic adaptation of the FIES questionnaire module in major national languages, as a first step towards obtaining valid estimates of food insecurity. The method for linguistic and cultural adaptation applied by the VOH in four pilot countries during 2013 is described. Section 5 presents a discussion of the validity of the food insecurity construct underlying the FIES and the analytic approaches to be applied by the VOH to ensure cross-cultural comparability of severity levels of food insecurity, based on the Item Response Theory. The final section focuses on the link between information and action, summarizing potential uses of the information derived from the FIES in the global, national and subnational policy contexts.

2. Direct Measurement of Food Insecurity Using Experience-Based Scales

It is widely acknowledged that an array of measurement instruments is needed to account for the complex nature of food and nutrition security and to monitor its multiple dimensions. This section summarizes some of the different measures commonly used, and describes the particular contribution of experience-based food insecurity scales in identifying populations at risk of hunger and food insecurity. A brief history of such scales is provided, with particular focus on a food insecurity scale developed for use in Latin America and the Caribbean, which gave origin to the Food Insecurity Experience Scale (FIES) presented in this paper.

2.1 The many terms associated with hunger

Various terms are used, often interchangeably, to talk about hunger and food insecurity. While the word "hunger" speaks the most powerfully and clearly to most people, it ranges in meaning from short-term physical discomfort to life-threatening lack of food, and often encompasses broader issues related to food access and socioeconomic deprivation (Habicht *et al.*, 2004). These physiological as well as socioeconomic dimensions of hunger pose a conceptual challenge for measurement.

The FAO Prevalence of Undernourishment is one of the indicators selected to monitor progress toward achievement of the Millennium Development Goal (MDG) to halve the proportion of people who suffer from hunger by 2015. It is an estimate of the number of people who are likely not ingesting enough food to meet dietary energy needs, based on national estimates of: 1) total food energy available for human consumption, and 2) distributions of energy requirements and consumption in the population, (FAO, 2012b - SOFI Annex 2). This measure has served well to monitor national and regional trends in undernutrition estimates through an analysis of food availability in relation to estimates of needs, at the

level of national populations, but does not identify who the food insecure are or where they live within countries.

The other indicator selected to monitor the MDG regarding hunger is weight-for-age of children under five years of age, which measures another condition closely associated with hunger: malnutrition. The term "malnutrition" refers to both undernutrition (nutritional deficiencies) and overnutrition (consumption of too much energy in relation to energy requirements). This term was, until recently, associated primarily with poor child growth and nutritional deficiencies, in particular stunting and wasting, conditions resulting from insufficient dietary energy and nutrient intake and/or poor utilization of nutrients due to infections or other illnesses. As the nutrition transition became increasingly evident in countries throughout the world, characterized by decreasing prevalence of undernutrition and increasing rates of overweight among the less privileged classes, previous assumptions regarding the association between poverty, hunger, and undernutrition were challenged (Ruel, Haddad and Garrett, 2001; Monteiro, Conde and Popkin , 2004).

It was in this context that experience-based food insecurity scales emerged. Other measures were needed to account for a broader concept of hunger and its association with both undernutrition and overnutrition. Ethnographic research carried out in the USA to understand the lived experience of hunger revealed it to be a process characterized initially by worry about having enough food, followed by dietary changes to make limited food resources last, and finally, decreased consumption of food in the household (Radimer, Olson and Campbell, 1990; Radimer *et al.,* 1992). Although the original ethnographic study was based on a small number of households in a wealthy country, a review conducted years later of studies derived from many countries in different regions of the world concluded that these dimensions of the experience of hunger appear to be common across cultures (Coates *et al.*, 2006a).

The approach to conceptualizing the experience of hunger developed by Radimer and colleagues coincided with an increasing focus on unequal access to food and socio-cultural aspects of the experience of hunger (Sen, 1981). In the 2009 Declaration of the World Summit on Food Security, food security was defined in the following way*: Food security exists 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* (FAO, 2009[2]). Access to food was recognized as one of the pillars of food security. It is this dimension of food security – food access – that experience-based food insecurity scales are designed to measure in populations, based on data collected at the household and individual level.

[2] This definition of the Committee on Food Security (http://www.fao.org/cfs/en/) is slightly different from the 1996 World Food Summit definition: "when all people at all times have access to sufficient, safe, nutritious food to maintain a healthy and active life" (http://www.fao.org/docrep/003/w3613e/w3613e00.htm). The words "physical, social and economic" were added to characterize "access". See also CFS (2012b) for a discussion of the evolution of terminology.

2.2 Complementary uses of different food and nutrition security measures

It is easy to see from the breadth of the definition of food security why an array of measurement instruments is needed to account for its complex nature and to monitor its multiple dimensions. While each method generates a different type of indicator with different challenges for application, analysis and interpretation, the consensus is that the various indicators complement one another. For excellent comprehensive reviews and conceptual discussions of the state of the art of food security measurement, see Coates (2013) and Jones *et al.* (2013).

The FAO Prevalence of Undernourishment is an indicator of the relative adequacy of food availability across the population and can serve as a useful proxy for food security at the national level. However, it is statistically complex to calculate, is largely dependent on national capacity to generate reliable data on a regular basis and is not a direct indicator of access to food at the household or individual level. The same is true of composite measures of hunger like the Global Hunger Index, composed of indicators of undernourishment, child underweight and child mortality (Grebmer *et al.*, 2012). The time lag between data collection and availability of the results for these two indicators tends to be significant, which compromises the ability to take timely action.

Anthropometric measures, such as child weight-for-age (measure of underweight) and height-for-age (measure of stunting), provide invaluable information regarding the nutritional status of individuals, but are costly and require a relatively sophisticated level of expertise to collect and analyze the data. Given the complex relationship between food security and underweight *as well as overweight*, the utility of anthropometric measures as proxy indicators of household food security is questionable (Kac *et al.*, 2012; Finney *et al.*, 2010; Adams, Grummer-Strawn and Chavez, 2003; Alaimo, Olsen and Frongillo, 2001; Frongillo, 2003). One advantage of a direct measure of the experience of food insecurity, like the FIES, is that it can be used to complement anthropometric data and potentially identify vulnerable populations before malnutrition becomes manifest.

Other time-tested methods for assessing food security, which come closer to being direct measures of food access, include indicators of the quantity and quality of food consumed, such as food consumption scores, dietary diversity indicators, and food acquisition data obtained from household expenditure surveys (FAO, 2002a; Jones *et al.*, 2013). While food consumption and dietary diversity are a more direct measure of the adequacy of the diet, methods based on household expenditure surveys employ indirect methods to estimate food intake. Acquisition of food products during a reference period is not equivalent to actual consumption of the purchased products, and there is no guarantee that the food was consumed during the reference period. Results also do not reflect food purchased outside the home. As a general rule, similar to the FAO undernourishment measure and anthropometric measures, methods for measuring food insecurity based on expenditures and food consumption require significant technological expertise, time and resources to collect and analyze the data.

Experience-based food insecurity scales like the FIES represent a simple, timely and less costly method for measuring the access dimension of food insecurity based on data collected at the household or individual level. They do not provide specific information on actual food consumption, diet quality and food expenditures like household expenditure surveys and individual food intake surveys might do, but rather focus more broadly on reported food-related behaviors associated with the experience of food insecurity due to limited access to food. They should therefore not be seen as substitutes for but rather as complements to these other important measures.

Unlike aggregate measures, such as the FAO Prevalence of Undernourishment or the Global Hunger Index, the FIES measure of the severity of food insecurity can be used in surveys that allow disaggregation at sub-national levels and across different population groups, making it possible to identify more specifically who the food insecure are and their geographic distribution. Finally, the ease of application, analysis, and interpretation facilitates better communication of results to decision makers, leaders of civil society, and the general public.

2.3 Evolution of experience-based measures of household food insecurity: direct measures of the access component

The ethnographic research mentioned above, conducted by Radimer, Olsen and Campbell (1990) and Radimer *et al.* (1992) in the USA revealed that the experience of food insecurity is characterized initially by uncertainty and anxiety regarding food access, followed by changes in the quality of the diet as the situation worsens, such as a less balanced, more monotonous diet. With increasing severity, the quantity of food consumed decreases as portion sizes are reduced or meals are skipped (Radimer, Olson and Campbell, 1990; Radimer *et al.*, 1992; Radimer, 2002). Additional ethnographic studies aimed at understanding the experience of hunger from the perspective of the elderly in the U.S. and low-income families in Quebec, Canada, revealed similar patterns (Wolfe *et al.*, 1998; Wolfe, Frongillo and Valois, 2003; Hamelin, Beaudry, and Habicht, 2002).

This underlying theoretical construct of food insecurity formed the basis for the U.S. Household Food Security Survey Module (US HFSSM), which has been applied annually in the United States since 1995 to monitor the food security situation (Hamilton *et al.*, 1997). The US HFSSM, in turn, has served as a model for many other experience-based food insecurity scales in diverse countries around the world, including the FIES. It represented a significant change in approach to food insecurity measurement compared to traditional ways of assessing it indirectly, either through the determinants (such as food availability) or the consequences (such as anthropometric failures and other signs of malnutrition).

The items that compose the US HFSSM, as well as the FIES module, ask people directly about having to compromise the quality and quantity of the food they eat due to limited money or other resources to obtain food. Each item refers to a different situation and is associated with a level of severity according to the theoretical construct of food insecurity underlying the scale.

By asking the series of related questions that compose the FIES, it is possible to classify respondents at different levels of severity: "food secure" (those who answer "no" to all the questions about food insecurity-related experiences) or "food insecure" along a continuum of food insecurity severity, as shown in Figure 1.

Figure 1: Food insecurity severity along a continuous scale

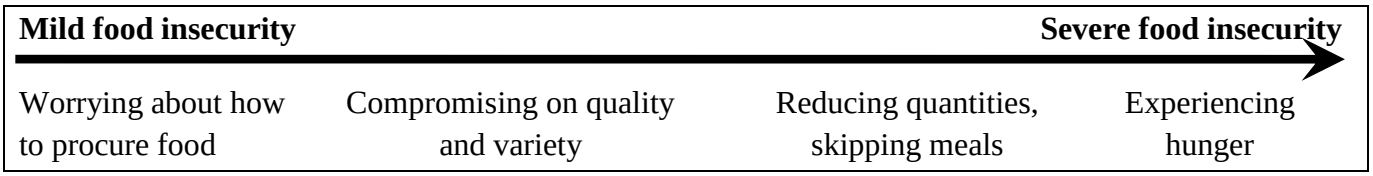
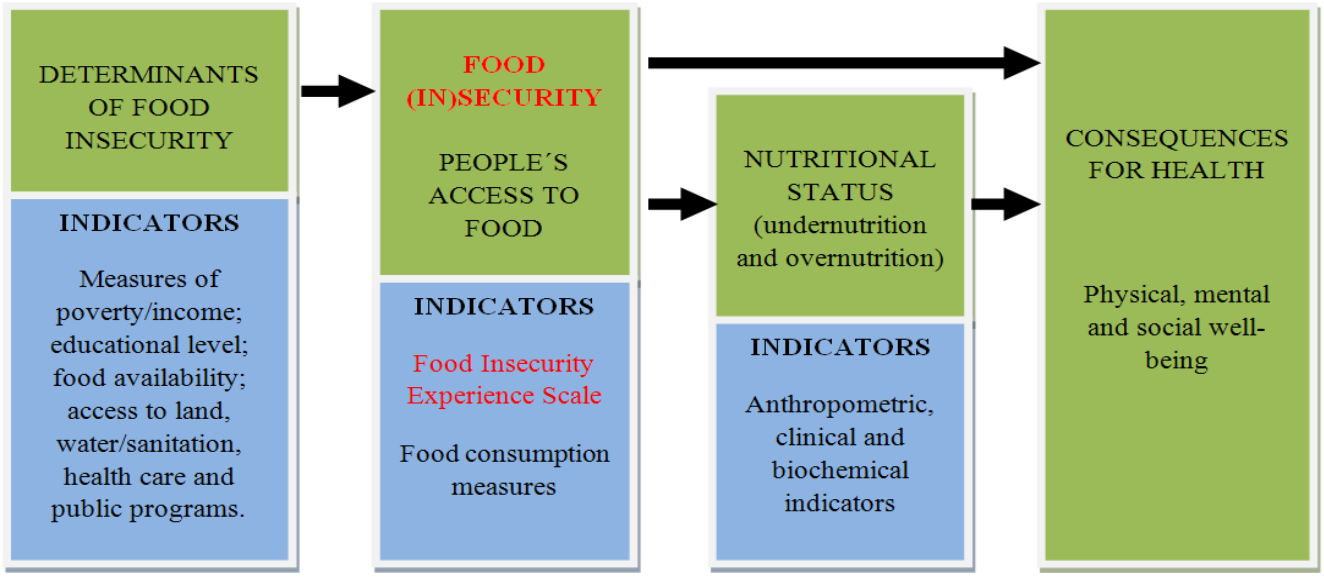


Figure 1 points to one of the unique contributions of experienced-based food insecurity scales: in addition to contemplating aspects related to deprivations in diet quality and quantity, they also capture an aspect of the experience of hunger and food insecurity that others do not, i.e. psychosocial aspects associated with anxiety or uncertainty regarding the ability to procure enough food. Food insecurity can affect health and well-being in many ways, with potentially negative consequences for mental and social in addition to physical well-being, even *in the absence of measurable negative effects on nutritional status* (Figure 2). Innumerable studies using experience-based food insecurity scales have documented negative psychosocial effects of food insecurity in women and children, as recently reviewed by Pérez-Escamilla and Vianna (2012).

Figure 2: Determinants and consequences of food insecurity at the individual level[3]



[3] Adapted from Campbell (1990).

With respect to the indicators cited in Figure 2, it should be kept in mind that determinants of household and individual food insecurity, as measured by experience-based food insecurity scales, and the consequences for individual health and well-being, are determined by local, regional, national and international factors.

While there is considerable evidence pointing to the validity and reliability of experienced-based food insecurity scales in diverse contexts, research to refine and improve them is ongoing. Most of the scales used in research and monitoring to date measure food insecurity at the household level, such as the previously described US HFSSM, the Latin American and Caribbean Food Security Scale (described in the next section), the FANTA Household Food Insecurity Access Scale (Coates, Swindale and Bilinsky, 2007) and the FANTA Household Hunger Scale (Ballard *et al.*, 2011). However, their applicability for measuring food insecurity of individual adults and children has also been investigated (Wolfe *et al.*, 1996; Nord and Hopwood, 2007; Fram *et al*., 2011; Bernal *et al.*, 2012).

Experts participating in the International Scientific Symposium on the Measurement and Assessment of Food Deprivation and Undernutrition, held at FAO Headquarters in 2002, emphasized the usefulness of experience-based measures of household food security for decision makers at various levels, as expressed by Eileen Kennedy in her keynote speech:

> *At the policy level, they are relatively uncomplicated to interpret and understand, which can be critical in policy analysis and presenting resonant messages to policy makers. At the programme or project level, they can be effective in targeting interventions (specifically for identifying populations or geographic areas, but not for identifying households or individuals) and in monitoring changes in food insecurity and hunger* (FAO, 2002b).

While the lack of internationally validated instruments that capture the experience of food insecurity was cited as a shortcoming in food security measurement at the 2002 Symposium, participants in the follow-up Symposium ten years later acknowledged that substantial progress has been made in the use and refinement of experience-based scales (FAO, 2012c).

In 2006, the Journal of Nutrition released a supplement that brought together many of the foremost researchers with experience on developing and applying food insecurity scales in widely different parts of the world (http://www.fantaproject.org/publications/hfias.shtml). This seminal publication, which set the stage for a new way of measuring food insecurity internationally, cited three main conceptual developments in food security measurement: "*1*) a shift from using measures of food availability and utilization to measuring "inadequate access"; *2*) a shift from a focus on objective to subjective measures; and *3*) a growing emphasis on fundamental measurement as opposed to reliance on distal, proxy measures" (Webb *et al.*, 2006). The Journal of Nutrition supplement for the first time defined food insecurity scales as "experiential" or "experience-based" measures that directly ask people about their experiences or behaviors in relation to food when money or other means to obtain food are limited,

reflecting "not only increased severity in food stresses but also the actual experience of becoming hungry" (Webb *et al.*, 2006).

The development of food insecurity scales over the past two decades has also included an evolution regarding how this type of instrument is labeled. In the proceedings of the 2002 FAO symposium, they were referred to as "qualitative measures". Webb *et al.* (2006) described a dichotomy between "objective/quantitative" compared to "subjective/qualitative" schools of thought. Food insecurity scales may be considered subjective measures in the sense that they are based on self-reporting, as is the case for much health, social and economic data obtained through survey work. A commonly cited misconception is that the scales measure the perception of food security; however, measurement is based on reported behaviours and experiences associated with compromised food access due to limited resources. They are sometimes referred to as qualitative because the experience of food security is not directly quantifiable, in the sense of directly observing and counting something. However, food insecurity scales are quantitative measurement tools that produce numerical scores with statistical properties, as will be discussed further in Section 5.

The evolution of food insecurity scales and efforts to validate this type of measurement tool in the United States, Canada and Latin America have been instrumental in demonstrating their validity and reliability as measures of what is essentially an unobservable phenomenon through the application of statistical models based on Item Response Theory (Frongillo, 1999; NAS, 2006; Nord, 2012; Faye *et al.*, 2011), described further in Section 5.

2.4 The Latin American and Caribbean Food Security Scale: A regional experience paves the way for a global measure

The FIES is the global version of an experience-based food insecurity scale that originated from a regional initiative in Latin America and the Caribbean. During the 2000s, several countries in Latin America independently began to adapt experience-based food insecurity scales for their own contexts. Recognizing the potential for a regional food insecurity measure, researchers combined their experiences to create a Latin American and Caribbean Food Security Scale (*Escala Latinoamericana y Caribeña de Seguridad Alimentaria*, or ELCSA) (Pérez-Escamilla *et al.*, 2007; FAO, 2012a), with roots in the US HFSSM, the Brazilian Food Insecurity Scale, a similar scale adapted for Colombia, as well as the Household Food Insecurity Access Scale (HFIAS) developed by the U.S. Agency for International Development (Coates, Swindale and Bilinsky, 2007). While the Brazilian scale is an adapted version of the US HFSSM (Pérez-Escamilla *et al.*, 2004), the instrument adapted for use in Colombia was derived from a slightly different scale that was a precursor to the US HFSSM (Wehler, Scott and Anderson, 1992; Lorenzana and Sanjur, 1999; Álvarez *et al.*, 2006; Hackett, Melgar-Quinonez and Álvarez, 2008). The experiences in Brazil and Colombia were the first national efforts to validate such scales in Latin America.

The ELCSA was thus conceived out of the combined experiences with food insecurity scales in various countries as well as the growing demand for tools to diagnose and monitor hunger and food insecurity in the region. A formal, interactive process of consultation was initiated in 2007 to promote the development of a single instrument capable of measuring household food insecurity in diverse national and sub-national contexts (Pérez-Escamilla *et al.*, 2007).

A 2010 report commissioned to the Ohio State University by FAO, with funding from the European Commission, gave a detailed historical review of the development and use of the different experience-based food insecurity scales in Latin American and the Caribbean that fall under the ELCSA umbrella. The report described the need for developing a harmonized scale that would enable cross-country comparison (Melgar-Quinonez, 2010). Based on findings of this report, FAO sponsored a workshop in Cuernavaca, Mexico, in 2010 with the participation of representatives from Mexico, Guatemala, Nicaragua, Honduras, and El Salvador, which resulted in a final, harmonized version of the ELCSA[4]. The harmonized ELCSA has since been applied in national population surveys in Guatemala and Mexico and in pilot studies in several other countries in the Latin American region, including Bolivia and Peru, and a manual has been produced in Spanish (FAO, 2012a). A linguistically adapted version of the ELCSA has also been tested in Albania (Albanian Centre for Economic Research, 2012).

Various experience-based food insecurity scales have been proposed, tested, and written about in the literature; however care must be taken when comparing results of studies using different versions of food insecurity scales. The ELCSA, as well as the Brazilian Food Insecurity Scale, and consequently, the FIES, are all closely based on the US HFSSM, and despite differences in their application[5,6], they measure food insecurity in essentially the same way.

Much of the success of the regional ELCSA experience can be attributed to the inclusive and intersectorial process that characterized its development, validation and dissemination. The ELCSA has been successfully applied from the local to the national level in government surveys, academic studies and public opinion polls. Application of the instrument has contributed significantly to a better understanding of the distribution, causes and consequences of food insecurity in Latin American and Caribbean. In countries where the ELCSA has been applied, the positive response from national-level decision makers and accumulated scientific evidence that experience-based food insecurity scales can measure food insecurity accurately and precisely have set the stage for proposing the use of a similar tool at the global level.

[4] http://www.foodsec.org/fileadmin/user_upload/eufao-fsi4dm/docs/ELCSA-exec-summary-english.pdf. http://www.foodsec.org/web/publications/pubshome/pubsdetail/ar/c/80691/

[5] While the items of the US HSFSSM remain the same, the manner of scoring responses and names of the food insecurity categories were changed several years ago, so comparisons must take that into consideration. See: http://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/definitions-of-food-security.aspx

[6] It should be noted that the reference period differs for each scale: 12 months for the US HFSSM, 3 months for the ELCSA, and 30 days for the Brazilian Food Insecurity Scale. Response categories vary somewhat as well; the US HFSSM has three additional affirmative response categories to characterize the frequency of occurrence, whereas the response categories for the ELCSA and the Brazilian scale are dichotomous (yes/no).

The US HFSSM has withstood intense scrutiny since it was first applied nationally in the USA in 1995, and has proven capable of reliably measuring food insecurity among various sub-populations. Extensive testing of the ELCSA in various countries in the Latin American and Caribbean region as well as sporadic use in other continents over the past ten years strongly supports the potential validity and reliability of this type of measurement instrument in diverse socio-cultural contexts.

It is on this solid basis of evidence that FAO is undertaking the initiative to develop, test and refine the FIES for application globally through the *Voices of the Hungry* project (VOH).

3. Voices of the Hungry project (VOH)

Accurate global measurement of food insecurity on an annual basis will contribute to monitoring of the development agenda objectives and identifying global, regional and national trends, as well as informing country-level decision making. The *Voices of Hungry* project (VOH) aims to explore the feasibility of using the FIES to provide timely and valid information regarding the degree of severity of food insecurity in populations world-wide. FAO has identified the FIES as the tool with the greatest potential for becoming a global standard capable of providing comparable information on food insecurity experience across countries and population groups to track progress on reducing food insecurity and hunger.

The version of the FIES currently being adapted and piloted in the VOH is presented in Table 1.

Table 1: Food Insecurity Experience Scale

FOOD INSECURITY EXPERIENCE SCALE Included in the 2013 pilot study delivered through the Gallup World Poll in Angola, Ethiopia, Malawi and Niger.		
Now I would like to ask you some questions about your food consumption in the last 12 months. During the last 12 MONTHS, was there a time when:		
Q1. You were worried you would run out of food because of a lack of money or other resources?	0 No 1 Yes	98 DK 99 Refused
Q2. You were unable to eat healthy and nutritious food because of a lack of money or other resources?	0 No 1 Yes	98 DK 99 Refused
Q3. You ate only a few kinds of foods because of a lack of money or other resources?	0 No 1 Yes	98 DK 99 Refused
Q4. You had to skip a meal because there was not enough money or other resources to get food?	0 No 1 Yes	98 DK 99 Refused
Q5. You ate less than you thought you should because of a lack of money or other resources?	0 No 1 Yes	98 DK 99 Refused
Q6. Your household ran out of food because of a lack of money or other resources?	0 No 1 Yes	98 DK 99 Refused

Q7. You were hungry but did not eat because there was not enough money or other resources for food?	0 No 1 Yes	98 DK 99 Refused
Q8. You went without eating for a whole day because of a lack of money or other resources?	0 No 1 Yes	98 DK 99 Refused

FAO will be leveraging on the Gallup® World Poll (GWP)[7], a branch of Gallup Inc. that has conducted nationally representative surveys in more than 140 countries annually since 2005, as a vehicle for data collection. Starting in 2014, the FIES will be incorporated into the World Poll questionnaire and the data will be used to derive estimates of the prevalence of food insecurity at different levels of severity. Through this measurement approach and the data collection service provided by the Gallup® World Poll, FAO will obtain cross-culturally comparable, real time information at a relatively limited cost, gathered from a nationally representative sample of adults in a large number of countries. The focus on food insecurity of individuals rather than households will enable the study of gender disparities in food access in the population. Data from the scale responses as well as a detailed food security analysis will be available to Governments and the public through an open-access platform.

The FIES is being pilot-tested in Angola, Ethiopia, Malawi and Niger during the 2013 round of the GWP. The characteristics of the pilot study include the following:

1. Surveys are being conducted based on nationally representative samples of 1000 adult individuals, using a three-stage sampling framework, in each of the four pilot countries;
2. The linguistically and culturally adapted FIES questions are directed to adult individuals randomly selected at the 3rd stage who reside in sample households randomly selected in the 2nd stage from primary sampling units, which are in turn either randomly selected or selected based on probabilities proportional to population size (1st stage).
3. In households where children under 5 years of age reside, two additional questions are asked of the respondent about food insecurity experiences of these children (if they were not able to eat healthy or nutritious foods or if they were not given enough food to eat because of a lack of money or other resources). The information will be used to describe the context of food insecurity of children, but will not be included in analysis of the 8 adult items to determine food security status, including severity of experienced food insecurity.

As part of the 2013 VOH pilot study, FAO carried out linguistic adaptations in national languages of Angola, Ethiopia, Malawi and Niger [8] following a common methodology (presented in Section 4). The goal of these activities was to produce translations of the FIES that were linguistically and culturally appropriate and faithful to the intention of the FIES questions. The adapted versions have been pre-tested by Gallup, Inc. in preparation for their 2013 World Poll surveys in each country, which in several cases led to minor refinements to the wording of several items. The experiences of the 2013 pilot study

[7] http://www.gallup.com/strategicconsulting/en-us/worldpoll.aspx
[8] http://www.fao.org/economic/ess/ess-fs/voices/linguistic/es/

will provide invaluable information on linguistic challenges and inform subsequent approaches for accurate, albeit less intense adaptation of the scale for global use.

The FAO initiative, starting with the pilot study limited to the four sub-Saharan countries in 2013 and extended to global data collection in 2014, will provide a unique opportunity to explore the cross-cultural equivalency of the FIES. The project plans to carry out extensive validation studies with the country datasets from the pilot study and over the first two years of global operation in order to understand the potential biases in its performance across different countries and cultures. In collaboration with the Economic Research Service of the U.S. Department of Agriculture (USDA), FAO is developing innovative analytic methods to define the standard against which to score and classify cases (see Section 5). The goal of the validation work will be to identify a scale that is simple to use, yet measures the phenomenon of food insecurity experience across countries in a comparable way.

The VOH contributes to strengthening the ability of FAO to fulfill its mandate of monitoring food insecurity at the global level in a timely and consistent manner. Results of the application of the FIES through the World Poll will be available every year within a few months of data collection in more than 140 countries.

To ensure sustainability of the methodology for direct measurement of food insecurity, the VOH aims to promote use of the FIES in national surveys designed to allow disaggregation at sub national levels, thus enabling governments to produce their own statistics as part of national food security information systems. FAO will review the country-level FIES results with Governments and begin a parallel capacity development programme with national statistics agencies. The consistency of the method used by FAO globally and by national governments in more detailed, periodic population surveys will promote comparability of results.

The expectation is that these efforts will contribute to advances in food insecurity measurement on a global basis and play a significant role in the monitoring of the Zero Hunger Challenge set forth by the Secretary General of the United Nations in 2012 as well as any food security target that may be considered for the post 2015 Development Agenda.

4. Linguistic Adaptation of the Food Insecurity Experience Scale

An important pre-requisite towards developing a tool that measures the severity of food insecurity in a comparable way world-wide is careful linguistic adaptation. In this section, we discuss the methodology used by the VOH during the 2013 pilot study to adapt the FIES to different cultural and linguistic contexts.

4.1 Trade-offs between context-specific measures and a global measure of food insecurity

An important debate regarding experience-based food insecurity scales during the history of their development addressed the feasibility of creating an internationally valid instrument using a single scale for the many diverse cultural and socioeconomic contexts in the world. Originally grounded in ethnographic research in the USA, the applicability of the US HFSSM in other contexts, even after linguistic adaptation, was an open question. Many believed that a similar process of developing household food insecurity scales "from the ground up", based on ethnographic research, would be necessary in different contexts in order to adequately reflect the theoretical constructs of the food insecurity experience in each cultural context (Wolfe and Frongillo, 2001; Frongillo and Nanama, 2006). In a few countries, ethnographic research resulted in measures based on slightly different theoretical constructs (Coates *et al.*, 2006b; Frongillo and Nanama, 2006).

Aiming to address this question and motivated by the recognized value that a common instrument would provide for international comparisons, Coates *et al.* (2006a) compared 22 food insecurity scales and ethnographic studies derived from 15 countries in different regions of the world, ranging from "ground up" research to application of existing scales following linguistic adaptation. Their exploration of "cross-cultural commonalities" to the food insecurity experience confirmed that the theoretical constructs reflected in the FIES – "uncertainty and worry about food, inadequate food quality, and insufficient food quantity" – were common to all the cultures sampled, pointing to the feasibility of a universally applicable measure.

The VOH is based on the premise that the evidence of universal dimensions of the food insecurity experience (Coates, *et al*, 2006b; Swindale and Bilinsky, 2006), and accumulated research pointing to the cross-cultural validity and applicability of measures very similar to the FIES have paved the way for a standardized measure that enables international comparisons (Derrickson, Fisher and Anderson, 2000; Álvarez *et al*., 2008; Hromi-Fiedler *et al.*, 2009; Melgar-Quiñonez, 2010; Segall-Corrêa *et al.*, 2008; Segall-Corrêa, Marín-León and Pérez-Escamilla, 2010).

4.2 Linguistic adaptation of the FIES for the 2013 VOH pilot study

The FIES consists of a set of questions asked directly to respondents. Applying it on a global level requires linguistic adaptation in the many different contexts and languages to ensure that the questions are understood in the way they were meant to be understood and to modify the wording appropriately within the specific language and culture while maintaining their original meaning. The linguistic adaptation process undertaken for the VOH was aimed at producing culturally and linguistically appropriate versions of a standardized set of questions, not a "ground-up", ethnographic approach.

As part of the 2013 VOH pilot study, FAO carried out linguistic adaptations in national languages of Angola, Ethiopia, Malawi and Niger using a methodology similar to ones used successfully in a number of different settings (Derrickson and Anderson, 2000; Melgar-Quiñonez *et al.*, 2003; Segall-Corrêa *et*

al., 2008; Segall-Corrêa, Marín-León and Pérez-Escamilla, 2010). The process included consultations with country-level specialists and officials as well as focus group discussions with people representative of different population groups or regions in the pilot countries.

While the FIES questions used in the VOH are worded to be as universally relevant as possible (See Table 1, Section 3), cultural and language differences are likely to influence how the questions are understood and answered. Linguistic adaptation must go beyond literal translation of the questions, making sure that the original concepts captured by the questions are maintained in translation while using culturally appropriate terms and phrases. Careful linguistic adaptation reduces the risk of obtaining inconsistent information due to misunderstanding of the questions, improves ease of administration and helps guarantee comparability across different cultures as well as within countries.

The starting point for the linguistic adaptation work in Angola, Ethiopia, Malawi and Niger was to have a solid understanding on the intended meaning of each of the items, using the explanations presented in Table 2 as a guide. Based on these explanations, preliminary translations of the FIES questionnaire were first produced by independent translators or in collaboration with a group of knowledgeable professionals.

Table 2: Questions that compose FIES and explanations of the intended meanings

"*Now I would like to ask you some questions about your food consumption in the last 12 months.* *During the last 12 MONTHS, was there a time when:*	
Q1. You were worried you would run out of food because of a lack of money or other resources?	The question refers to a state of being **worried, anxious, apprehensive, afraid or concerned** that there might not be enough food or that the respondent would run out of food (because there was not enough money or other resources to get food.) The worry or anxiety is due to circumstances affecting their ability to procure food, such as: loss of employment or other source of income, or other reasons for not having enough money; insufficient food production for own consumption; disrupted social relationships; loss of customary benefits or food assistance; environmental or political crises. It is not necessary for the respondent or the household to have actually run out of food in order to answer affirmatively to this question, the rationale being that even just the concern and the consequent possible coping strategies are manifestations of food insecurity, even in cases when the actual food consumption is not compromised
Q2. You were unable to eat healthy and nutritious food because of a lack of money or other resources?	This question asks the respondent whether s/he was not able to get foods they considered **healthy** or those that make a **nutritious or balanced diet** (because there was not enough money or other resources to get food.) The answer depends on the respondent's own opinion of what *they* consider to be healthy and nutritious foods. This question refers to the quality of the diet and not the quantity of foods eaten.
Q3. You ate only a few kinds of foods because of a lack of money	The question asks if the respondent or any other adult in the household had to eat a

or other resources?	diet with a limited variety of foods or whether they had to eat the same foods or just a few kinds of foods every day because there was not enough money or other resources to get food. The implication is that the diversity of foods consumed would likely increase if the household had better access to food. This question refers to quality of the diet and not the quantity of foods eaten. It is important to stress the link to lack of money, to identify conditions of food insecurity, rather than customary habits to limit the variety of food for other circumstances (i.e., health or religion)
Q4. You had to skip a meal because there was not enough money or other resources to get food?	This question inquires about the experience of having to **miss** or **skip** a major meal (for example, breakfast, lunch or dinner depending on the norm for number and times of meals in the culture) that would normally have been eaten (because there was not enough money or other resources to get food.) This question refers to insufficient quantity of food.
Q5. You ate less than you thought you should because of a lack of money or other resources?	This question enquires about eating less than what the respondent considered they should, **even if they did not skip a meal** (because the household did not have money or other resources to get food). The answer depends on the respondent's own opinion or perception of how much *they* think they should be eating. This question refers to quantity of foods eaten and not the quality of the diet. This question does *not* refer to special diets to lose weight or for health or religious reasons.
Q6 Your household ran out of food because of a lack of money or other resources?	Referring to any experiences when there was actually no food in the household because they did not have money or other resources (for example, the household's own production or bartering) to get food.
Q7 You were hungry but did not eat because there was not enough money or other resources for food?	This question asks about the **physical experience of feeling hungry**, and specifically, feeling hungry and not being able to eat enough (because of a lack of money or resources to get enough food). It does not refer to dieting to lose weight or fasting for health or religious reasons
Q8. You went without eating for a whole day because of a lack of money or other resources?	This question asks about a specific behaviour—not eating anything all day (because of a lack of money and other resources to get food). It does not mean dieting to lose weight or fasting for health or religious reasons.

4.3 Expert panels

For the linguistic adaptation work in Angola, Malawi and Niger, panels of experts were formed composed of people from relevant government ministries, national statistics institutes, research institutions, or non-governmental organizations with experience in issues related to food insecurity, such as agriculture, social development, health and nutrition.

These panels discussed the preliminary version of the translated questions, suggested modifications to each question based on their experiences, and identified different ways to express key concepts which were later explored in the focus group discussions. Potential problems related to comprehension of the questions in the module, as well as its application in the field, were discussed. They also provided assistance in determining criteria for planning of focus group participants (e.g. whether to separate by age and/or gender) and identifying key socio-demographic differences, such as urban, peri-urban, rural, different ethnic groups and livelihood strategies. Since limited time and resources imposed limits on the number and geographical distribution of the focus groups conducted the four countries, the expert panels helped to identify priorities with respect to the key criteria.

The version of the translated questions reflecting modifications recommended by the panel of experts, including possible alternatives that needed to be explored further, served as the basis for the focus group discussions and key informant interviews.

4.4 Focus group discussions

To ensure that the translated FIES module was understood by respondents as intended, the questions were explored in semi-structured discussions with groups of individuals from the target population, using a qualitative research technique known as **focus group discussions**. Focus group discussions are designed to elicit the perceptions, opinions, beliefs, and attitudes of participants through a guided, interactive discussion. This step helped to identify potential sources of misunderstanding and enabled further refinement of the questions based on insights into how the questions were understood by potential survey respondents.

In all four countries of the pilot study, focus groups were planned so that each group was composed of people of similar socio-cultural backgrounds, generally from the same community. In general, participants were people who were at risk of hunger or food insecurity at the time or who had been in the past. The total number of focus groups in each country was planned to take into consideration the key socio-demographic differences likely to lead to different perspectives or understandings regarding the themes and questions in the FIES. When deemed culturally appropriate, focus groups were conducted separately for men and women, and younger and older adults, to ensure that everyone would feel comfortable speaking.

A summary of the number and composition of focus groups conducted in each country is presented in Table 3.

Table 3: Focus groups conducted in Angola, Malawi, Niger and Ethiopia

Type of community	Number of focus groups			Language	
ANGOLA	*Men*	*Women*			
Urban (Luanda)	1	1		Portuguese	
Peri-urban (Luanda)	2	2		Portuguese	
Rural farming community near Luanda	1	1		Portuguese	
Rural farming community – Kwanza Sul Province	1	1		Umbundu	
Rural fishing village - Kwanza Sul Province	1	1		Umbundu	
MALAWI	*Men*	*Women*	*Mixed youth*		*Key Informants*
Lilongwe (Rural)	1	1	1	Chichewa	1 (m)
Lilongwe (peri-urban)	1	1	1	Chichewa	
Rumphi (rural)	1	1	1	Chitumbuka	1(m)
Rumphi (peri-urban)	1	1	1	Chitumbuka	1(m)
Zomba (rural)	1	1	1	Chichewa	1 (f)
Zomba (peri-urban)	1	1	1	Chichewa	1(m)
NIGER					
Niamey – peri-urban	1 + 1	1 + 1		Haoussa and Djerma	
Dosso region (rural)	2 + 2	2 + 2		Haoussa and Djerma	
Maradi region region	2	2		Haoussa	
Tahoua region (peri-urban and rural)	2	2		Haoussa	
Tillabéri region (rural)	4	4		Djerma	
ETHIOPIA					
Urban Addis	1			Amharic	
Rural Addis	1	1		Amharic	
Rural Adama		2		Oromo	

The focus group discussions began with more general questions about obtaining food and eating habits followed by some open questions about the experience of not having enough food. Participants were asked to describe situations they had personally experienced or experienced by people they know, and to describe the distinguishing characteristics between households where people always have enough to eat and those that often do not.

Each question on the FIES questionnaire was then reviewed with the discussants to verify comprehension and identify potential discrepancies in relation to the intended meaning. Possible

alternative phrases suggested by the expert panels were explored. The objective was not to elicit answers to the questions that compose the FIES, but rather to explore how respondents understood the questions, to listen to how they talked about the experiences to which the questions refer and to note possible alternative phrasing.

De-briefing sessions were held as soon as possible following each focus group to discuss methodological issues and to identify emerging themes to be explored in greater depth in subsequent focus groups as well as phrases used by participants to express key concepts. In Malawi and Niger, Key Informant Interviews were also conducted with local leaders in the respective areas to further verify certain terms and phrases proposed by the team. After all the focus group discussions were completed, final de-briefing sessions were held with focus group moderators and assistants, supervisors and expert panel members to reach consensus on final versions of the scale based on focus group findings.

The final linguistically adapted versions of the FIES were pre-tested by Gallup in preparation for their World Poll surveys to be conducted in each country during 2013, leading to minor refinements in wording of a few items.

The experiences from the linguistic adaptation work in Angola, Ethiopia, Malawi and Niger highlighted some of the more difficult challenges inherent in adapting the FIES in diverse contexts, some of which are summarized in Table 4.

Table 4: Linguistic challenges addressed during the 2013 pilot study

PHRASE	**EXAMPLES FROM THE FOCUS GROUPS: FINDING THE RIGHT PHRASE AND EXAMPLES**
Past [year /12 months]	In Malawi, many people thought that the past 12 months referred to the calendar year 2012, so the phrase way modified to reflect the period between 12 months ago and present (e.g. if interviewed in August, the period would be "from last September until present...".
Lack of money and other resources	In Angola, "lack of means" was understood better than "lack of money and other resources". In Niger, in both Djerma and Haoussa languages, lack of money and other resources was understood as the lack of any mean to cope with a certain situation, to find a solution to a problem. Participants to the focus group discussions stressed the need of having money to procure food, especially during the lean season, as well as the importance of relying on other resources, such as livestock or other informal social safety net systems (e.g. barter, called *troc*, or regular collection of money to be distributed to the vulnerable families called *tontine*) as a way to escape extreme poverty.
Household	In Malawi, the communities in central region were traditional Chichewa is spoken, referred to household as "*pakhomo"* meaning people living together and sharing food and other resources. The communities referred to "*pabanja"*, another term that was

	suggested, meaning someone's blood line or clan, and therefore not appropriate. However, this was different in the southern region; there was no distinguishing between the two words. The team decided that both words be used when administering the questions.
Food	In Malawi, **food** was generally referred to as the staple maize in all the regions. Some respondents indicated that even if they ate rice but not maize that day they had not eaten (Rumphi and Lilongwe), while in the Southern region (Zomba) they acknowledged that rice and cassava were food. In the Central region, food was also defined as the main staple (maize) and was tied to manual labor (farming) and that nsima (maize meal) is the food that provides energy to work in agricultural production. Similarly as in Malawi, in Niger food was generally associated with the main staples (those which provide the majority of required energy) which are millet, sorghum and the residual bran of these cereals. In some communities of herders milk is also considered as a basic, important food. As an example, picking and eating wild leaves is also a source of food supply but it is considered more as a feeding strategy which compensates for the scarcity of food during the lean season.
Healthy and nutritious food	In Angola, "healthy food" to some people referred to food that is hygienic and safe, while "nutritious food" is associated with having a varied diet. In all districts in Malawi, question 2 (not eating healthy and nutrition foods), and question 3 (having to eat only a few kinds of food) yielded a considerable amount of debate. There seemed to be a thin line between "different kinds of foods" and "healthy and nutritious foods". In all the 3 districts the phrase "healthy and nutritious foods" was interpreted as "food that gives energy" and even after probing the respondents indicated that if food is healthy and nutritious it should be the one that will "give you energy to do your farming activities". Most respondents indicated that different kinds of foods are part of healthy and nutritious diets. In Niger healthy and nutritious food was associated with 'food that is not harmful to the health of the person (healthy) and builds the body (nutritious)". It contains everything the body needs and it helps to have strong, healthy and shiny skin. Healthy and nutritious diets are diets including different kinds of food.
Eating only a few kinds of foods	During focus group discussions in Niger, respondents had no problem understanding and giving examples of what it means for them to eat different kinds of food. Eating few kinds of foods means having very limited variety of foods during the same meal. On the contrary, to have a diversified diet means being able to eat all the food (quantity) that the person wants (quality) such as rice, meat, fish, salad, cabbage, mangoes.
Skipping a meal	In some languages, such as Djerma in Niger and Chichewa in Malawi, there is no single term for meal or way to express skipping a meal. In both languages, the question was modified to ask if food was skipped in the morning, afternoon or evening, depending on the interval that those communities take their meals.
Ran out of food	In Niger, participants associated this expression with the worry, the doubt, the fear of not

	having enough food left in their stocks and the difficulty of buying more due to lack of resources. However, discussions revealed that men and women might not have the same perception of this concept. For women, "being out of food" is synonymous with having a monotonous diet while among men, it means being unable to feed their family, not having food in their stocks.
Eating less than you thought you should	In Malawi, some respondents indicated that they maintained the same frequency of meals but cut down on the portions, hence eating less than expected. Some respondents indicated that they had to skip meals, hence eating less than what they thought they should have eaten. For the two above phrases, the Chichewa and Chitumbuka versions were adjusted to the different contexts correctly without changing the meaning. In Niger, virtually all respondents in all the visited villages reported that the number of meals is strongly associated with the cropping season. In the different regions visited, participants of both groups (men and women) indicated that during "normal" periods, when there is good food availability, following the harvest season, they take three meals a day. During the lean season, the frequency decreases and varies between 1-2 meals per day. However, when crops fail the number of meals can be drastically reduced to 1.
Hungry but did not eat	The communities in the northern region of Malawi referred to the initial translation as long-term hunger, famine, or implied drought. The "*mukaziya*" depicts feeling hungry.
Go a whole day without eating	In Niger, this question was adapted with the expression of "not being possible to eat from sunrise to sunset" which corresponds to "not taking any of the three meals of the day: breakfast, lunch and dinner".
Enough food	In Djerma, respondents understood the expression eating enough food as eating until "they have a full stomach" (quantity) while in Haoussa it was possible to almost make a literal translation with the expression "eating enough food".

The linguistic adaptation experiences in Angola, Ethiopia, Malawi and Niger provided invaluable information on the challenges involved, corroborating experiences in other countries regarding phrases and concepts which require more careful adaptation. While it will not be logistically possible to conduct such in-depth adaption work in every country and language, these experiences will be used to inform approaches for accurate, albeit less intense adaptation and translation of the scale for global use.

5. Towards a Valid Global Standard for Food Insecurity Measurement

In science and statistics, validity is the extent to which a concept, conclusion or measurement is well founded and corresponds accurately to the real world. […]

> *The validity of a measurement tool […] is considered to be the degree to which the tool measures what it claims to measure.*[9]

This section is concerned with establishment of the validity of the FIES as a food-insecurity measurement tool for global monitoring.

Validity is a quite general term, related to adherence of a concept, the result of a scientific inquiry or a measure to reality. In the physical sciences there is usually no significant ambiguity regarding what the reality is, and therefore validation of a measurement tool reduces to comparison of the measures obtained from the tool with some available objective reference or benchmark, sometimes referred to as the "gold standard". In recognition of the fact that measures are always possibly affected by errors, the quality of a measurement tool is assessed with reference to two different aspects, referred to as "reliability" and "accuracy" respectively, which could also be described as the attributes of "being right" and "being precise". By comparing the results of repeated applications of the measurement tool to the benchmark object, the measure of which is known, one can assess both reliability and accuracy of the measures, thus establishing validity in its broader meaning. Obviously, a valid measurement tool produces measures that are both right (that is, it produces, on average over repeated applications, the correct measure) and precise (each of the produced measures is quite close to the true magnitude of the "thing" that one aims at measuring).[10]

When applied to many concepts in social sciences, however, this process is not so simple. Assessing the quality of a measurement tool poses the immediate problem that the object of interest may be an *immaterial*, *unobservable thing* (commonly referred to as a *construct* or a *latent trait*) for which no *objective* benchmark or reference exists. Lack of an objective benchmark has the consequence that it may become particularly difficult to disentangle the two components that contribute to validity: that of the underlying conceptualization of the construct (i.e., "reliability") and that of the measure (i.e., "precision"), which can only be assessed if the construct has been conceptualized correctly. In fact, in debates over food security measurement, the two processes of defining the object to be measured and evaluating the various proposed tools to measure it have been so inextricably linked that judgments on the appropriateness of an indicator have sometimes been confused with value judgments on the underlying concept.

In addressing the question of the validity of experience-based food insecurity scales in general, and of the FIES in particular, we consider the two issues in turn. We discuss first the *propriety of the tool as a food security monitoring instrument*, which concerns the validity of the food insecurity construct that the

[9] (http://en.wikipedia.org/wiki/Validity_(statistics), accessed on August 4, 2013)

[10] Judgment of the overall validity of a tool must thus balance the two properties. While it is desirable that measures be as precise as possible, excessive focus on precision may sometimes lead to the risk of preferring instruments that are … "precisely wrong" over those that are "approximately right," a risk that we propose should be adamantly avoided.

tool is intended to measure. We then tackle the question of the *accuracy of the produced measures*, that is, *how well* the proposed tool is able to measure the specific concept of food insecurity.

The above distinction is important, as in some cases, the term "validity" is used to simply imply precision without also considering the adequacy of the underlying concept of food insecurity. Ignoring this important aspect of the validation process may rightfully lead to some skepticism. Statistical tests and modeling may confirm the "validity" of experience-based food insecurity scales because the response patterns conform to certain criteria. Without an understanding of the underlying construct of the food insecurity scale that the tool is intended to measure, some erroneous conclusions may be drawn. Experiential food insecurity scales are sometimes perceived as being unrelated to nutrition, for example. However, as discussed in Section 3, nutritional aspects associated with changes in food intake resulting from restricted access to food are indeed an important part of the underlying construct of food insecurity. On the other hand, food insecurity scales are not intended to quantify changes in food or nutrient intake.

For the first part of this discussion, we refer back to the summary in Section 2 of the history of the debate regarding the use of the food security definition for practical purposes such as monitoring progress towards the MDGs or evaluating projects and programmes. For the second part of the discussion, as no specific evidence on the application of the FIES in its current form is available yet, we provide a review of the significant amount of research, undertaken by numerous investigators in different settings aimed at establishing the validity of other experience-based scales. We critically review those efforts through the lens of the proposed two-layer analytic framework highlighting if, and to what extent, the results obtained for the US HFSSM, the HFIAS or the ELCSA also apply to the FIES and its potential applications.

We conclude that the collective evidence from existing validation work is sufficient to suggest that the FIES is indeed founded on a valid concept of food insecurity. It covers domains that are common across cultures and socio-economic conditions and thus has *the potential* to form the basis for a valid measure worldwide. It is also clear, however, that a conclusion regarding whether or not such a measure can be established *in practice*, in each and every country in the world is premature if based only on the existing evidence so far. Successful practical implementation of a global measurement standard requires additional research, to verify, for example, whether or not the FIES in its current form is able to measure the prevalence of different levels of severity of food insecurity comparably across populations that differ in their linguistic, cultural and socio economic aspects. It is safe to say, however, that this is indeed a worthy endeavor, and that the development of such a global standard will generate enormous benefits for the entire community of researchers, analysts and policy makers interested in eradicating food insecurity and hunger. This is precisely what the projected work to be carried out by the FAO VOH focuses on. The objective is to promote the FIES as a global standard through extensive validation of the results obtained in the initial phases, accompanied by the development of the necessary analytic

procedures to compute severity of food insecurity indicators based on the responses to properly adapted questionnaires.

5.1 What is meant to be measured? Defining the "severity of food insecurity"

The FIES is being proposed as a metric for the **severity of food insecurity** (access dimension) in population groups, based on data collected at the household or individual level. Establishing its validity implies finding agreement on a definition of this food insecurity construct that can be measured along a scale of severity. In other words, it requires being able to speak legitimately of subjects not only in terms of being food insecure or not, but also as being *more* or *less* food insecure than others.[11] Food insecurity experiences may occur at either the household (economic unit) level or the individual level. While the experiences of reduced food intake, disrupted eating patterns, poor diet, and hunger are essentially individual, food provisioning and the adequacy of food stores in terms of quantity and quality inherently apply to the household-level.

To investigate the validity of the FIES as a measure of food **in**security, we recall the definition of food security in the 2009 Declaration of the World Summit on Food Security presented in Section 2: *Food security exists 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.*

Arguably, this has been a useful definition for advocacy and to inform discussions at the higher political level, but it is also leaves, as put very recently, "sizeable gaps between the food security definition and its measurement in practice" (Coates, 2013, p. 190). One of the major effects of the World Food Summit definition is that such an encompassing definition of the concept of food security suggests multiple possible definitions of the inverse concept, that of "food insecurity", which is what we really want to measure at its different levels of severity. The construct of food insecurity that the FIES aims to measure is a more restricted concept within the broader, more encompassing definition. The advantage of the FIES is that it enables measurement of the food insecurity which can then be analyzed together with indicators of its determinants and consequences to contribute to a more comprehensive understanding and inform more effective policies and interventions.

Measuring food insecurity using experienced-based scales like the FIES puts peoples' experiences and behavioral responses at the core of the definition of what food insecurity *means*. According to a number of authoritative commentators, this was a long due and welcome change in perspective that would fill the gap in existing methods (Barrett, 2002; FAO, 2002b). This line of thinking has been described, by Coates *et al.* (2006a), as one way to promote a shift towards "third generational indicators" [the first two generations being those characterized by attention to determinants and to outcomes of food insecurity,

[11] This is not obvious, and the problem is particularly relevant if we consider that the two official indicators used to monitor the hunger target within the Millennium Development Goal (MDG) framework only provide measures of prevalence in the population, leaving the question of the *severity* of food insecurity largely unaddressed.

respectively proposed by Barrett (2002)]. The substantial innovation is represented by the possibility of capturing food insecurity directly and analyzing it from a behavioral perspective.

The essential arguments in favor of directly measuring food insecurity are that (a) food insecurity is a condition that may affect overall welfare in different ways, depending on the type and strength of the constraints faced by people affected; (b) it is characterized by behavioral responses, and (c) both the responses and the welfare impact can be located along a scale of severity. As described in Section 2, Radimer's work identified a consistent pattern: the lived experience of food insecurity was characterized initially by anxiety, associated with worry about being able to get enough food; then, as conditions worsened, it resulted in decreased amount of stored food in the home, followed by worsening quality and diversity of the diet, decreased quantity of food eaten per meal, and, finally, in being forced to skip meals and feel hungry for an extended period (Radimer *et al.,* 1992). We are thus presented with a construct that involves three domains - uncertainty and worry about food, inadequate food quality, and insufficient food quantity and a hypothesis on the relative position of the domains on an underlying scale of severity as shown in Figure 1 (Section 2).

The meta-analysis of food insecurity scales and ethnographic studies in different regions of the world described in Section 4 revealed that the same basic dimensions of the theoretical construct identified by Radimer were common to all the cultures sampled (Coates *et al.*, 2006a). While the analysis does not confirm that a universally applicable questionnaire may be developed, it does suggest that these common domains are robust across cultures in their broad definition and ranking of severity (Swindale and Bilinsky, 2006). These are precisely the domains covered by the items included in the FIES, as shown in Table 5, where each item of the FIES is presented with the domains of the underlying construct and level of severity of food insecurity to which it is assumed to correspond.

Table 5: The 8 FIES items by domain of the theoretical construct of food insecurity and assumed level of severity

FIES order of items	Scale items	Domains of the food insecurity construct	Assumed severity of food insecurity
1	Felt anxiety about having enough food at any time during the previous 12 months (this time period applies to all 8 scale items)	uncertainty and worry about food	**Mild**
2	Not able to eat healthy and nutritious food because of lack of money or other resources to get food	inadequate food quality	**Mild**
3	Consumed a diet based on only few kinds of foods because of lack of money or other resources to get food	inadequate food quality	**Mild**

4	Did not eat breakfast, lunch or dinner [or skipped a meal] because there was not enough money or other resources to get food	insufficient food quantity	**Moderate**
5	Ate less than they thought they should because of lack of money or other resources to get food	insufficient food quantity	**Moderate**
6	Household ran out of food because of lack of money or other resources to get food	insufficient food quantity	**Moderate**
7	Felt hungry but didn't eat because there was not enough money or other resources for food	insufficient food quantity	**Severe** (hunger)
8	Went without eating for a whole day	insufficient food quantity	**Severe** (hunger)

5.2 Lacking a proper reference, how can we determine accuracy of the measures?

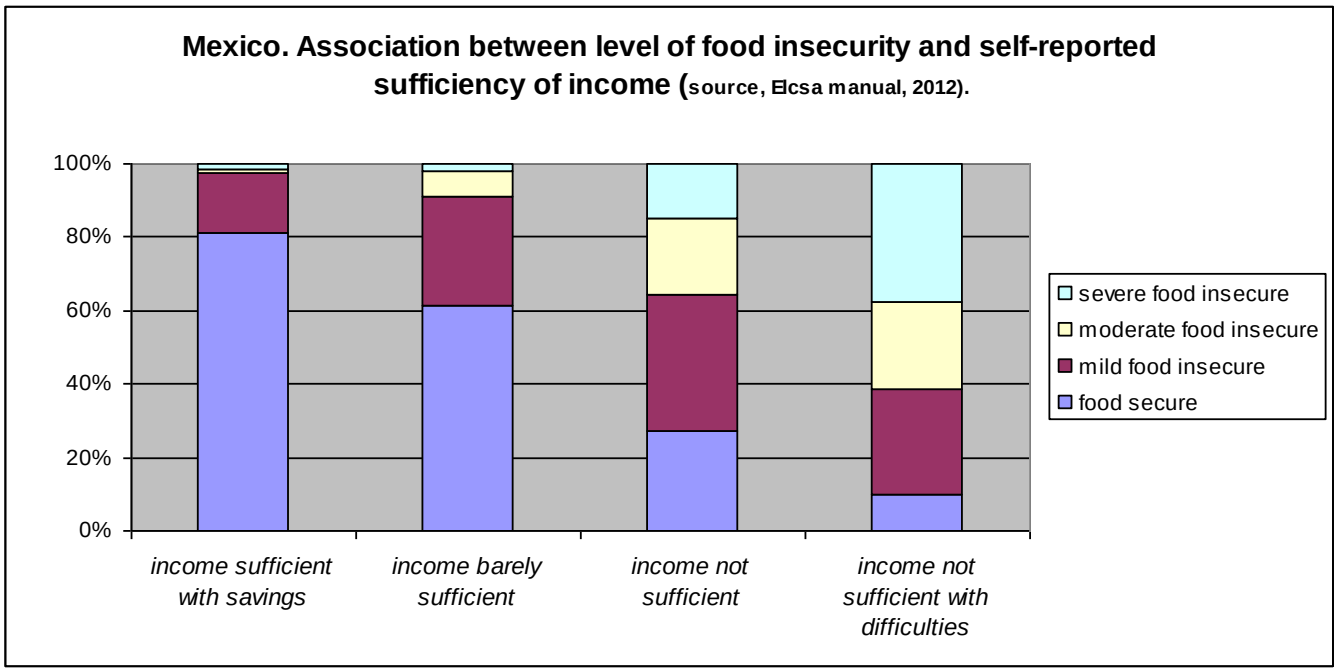
Experience-based scales were developed to fill a recognized gap in the ability to measure food insecurity, but this creates a conundrum when one tries to empirically validate the accuracy of the measures that they produce. According to Frongillo, "Accuracy [of a food insecurity measure] is assessed by in-depth analysis, and by relating the measure to a criterion measure, which may be a more definitive measure, determinant or consequence" (Frongillo, 1999, p. 508S). The ideal approach would be to assess the accuracy of a scale by using a definitive measure of the underlying theoretical construct against which the measurement tool can be compared (validated). However, there is no definitive measure. Food insecurity is a social as well as biological, nutritional, and economic phenomenon (Frongillo, 1999), and experience-based food insecurity scales capture social aspects that other measures do not. These types of scales are direct measures of the access component of food security: the experience of not having enough money or adequate resources to get food in sufficient quantity and/or quality, for which no alternative direct measure exists.

Frongillo *et al.* (1997) and Frongillo (1999) used an approach for developing and applying a criterion measure that involved consensus among independent researchers who classified households using a very rich data set, composed of food security status as well as demographics, factors contributing to food insecurity, coping strategies, fruit and vegetable consumption, disordered eating behaviors, dietary recall, and an inventory of household food stores. A similar approach to developing a criterion measure was also used in other settings (Wolfe *et al.*, 1998, 2003; Hamelin, Beaudry and Habicht, 2002; Frongillo and Nanama, 2006).

However, developing criterion measures for validating the FIES in new settings would require a lengthy and thorough process to obtain detailed databases. A more feasible approach involves relating measures produced by the scale to existing measures of determinants and/or consequences of food insecurity. It has been standard practice to test accuracy of the scale by studying associations between the scale and variables that are theoretically part of the same construct as food insecurity, and/or that would vary in an expected way across different levels of food insecurity, when these variables are measured concurrently in the same subjects.

Showing that the FIES or similar scales correlate with other variables in the expected direction (such as observing an increase of food insecurity severity across quintiles of increasing poverty) supports the conclusion that whatever is being captured by the scale is consistent with the intended underlying construct of food insecurity (Cook and Beckman, 2006). Consistent relationships have been found across countries, as is illustrated in the following graph of data from a study in Mexico that looked at food security levels calculated from the ELCSA across levels of self-reported sufficiency of income (Figure 3). There is a clear trend of a worsening food security situation as income becomes insufficient.

Figure 3: Association between level of food insecurity and self-reported sufficiency of income



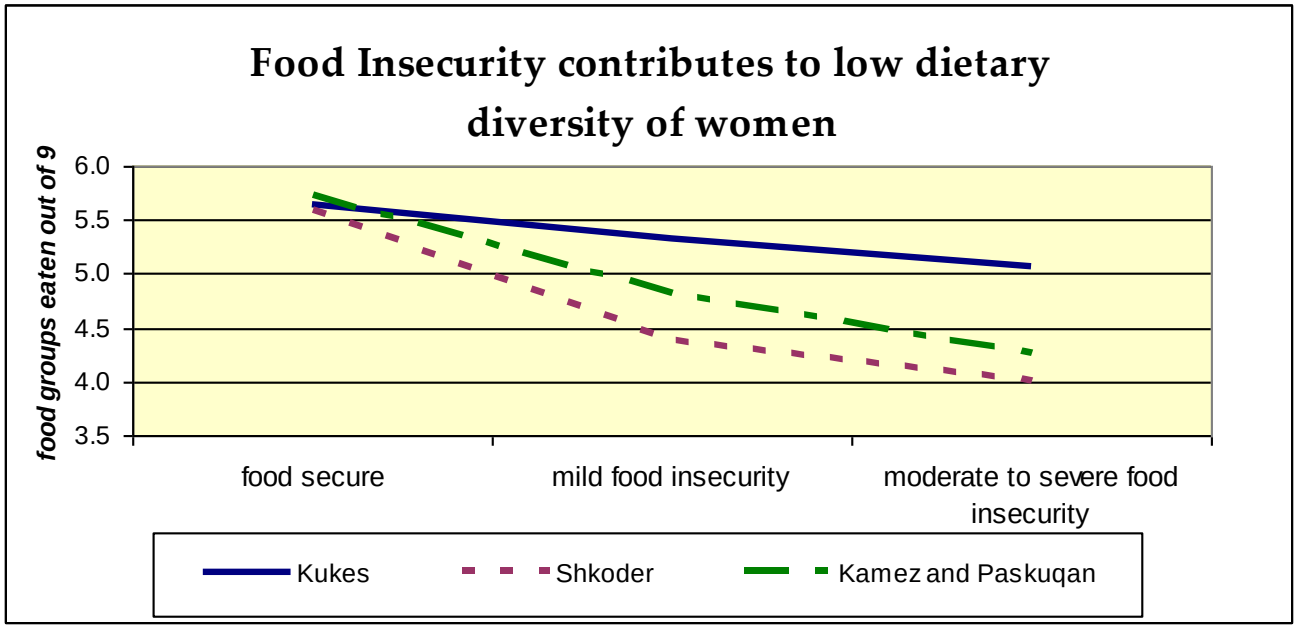
Source: Adapted from Figure 4, ELCSA manual (pg. 55), FAO 2012a.

Examining associations between food insecurity and indicators of dietary intake is also useful to verify the accuracy of experience-based food insecurity scales, since changes to dietary quantity and quality are part of the underlying construct of food insecurity. National studies conducted in Brazil (Segall-Corrêa *et al.*, 2008) and Mexico (Pérez-Escamilla, Parás and Vianna, 2012) showed an inverse relationship between severity of household food insecurity and consumption of high nutritional quality

foods such as fruits, vegetables, meat, and dairy products. Another study in Brazil using a food frequency questionnaire found that 85% of food secure households had consumed meat and 75% had consumed fruits, while among the most food insecure households, consumption of these foods was 12% and 8%, respectively (Panigassi *et al.*, 2008).

In Albania, the diversity of women's diets was associated with food security status. A study showed that dietary diversity (number of food groups consumed in a 24 hour period) in three vulnerable areas declined with increasing severity of household food insecurity, consistent with the underlying theoretical construct of food insecurity. (see Figure 4 below).

Figure 4: Food insecurity and dietary diversity of women - Albania



Source: United Nations Albania. Baseline nutrition and food security survey. 2012. Tirana. Chapter 4, figure 4.12, page 59. http://www.unicef.org/albania/Baseline_Nutrition_and_Food_Security_Survey.pdf

It is often of interest to policy makers to know how food insecurity relates to child malnutrition. While this relationship can be explored using the FIES and similar scales, anthropometric measures and other biological indicators of malnutrition would be inappropriate for validation studies of experience-based food insecurity scales for a number of reasons. Results of such comparisons are likely to be misleading and difficult to interpret, particularly in the case of the FIES, which provides prevalence estimates of adult food insecurity. Stunting and wasting, especially in the case of young children, are influenced by many other factors besides access to food, including sanitation, maternal care and access to health and social services. Equally important are temporal aspects, as stunting, in particular, reflects nutritional and environmental conditions in the past, thus complicating further the analysis of the relationship between stunting and severity of food insecurity (which in the case of the FIES, has a 12-month reference

period). Thus, while it is valid to explore the relationship between biological indicators of nutritional status and experience-based measures of food insecurity, it would not be an appropriate to use the former to explore the validity of the latter. It should also be kept in mind that the experience of food insecurity, even in the absence of observable negative effects on nutritional status, is a serious problem in itself, indicating a violation of the Human Right to Adequate Food.

In conclusion, we can say that there are strong empirical bases to claim that the fundamental construct, best described as *resource-constraint food insecurity* (Ohls, Radbill and Schirm, 2001, p.6), is indeed a meaningful and useful concept. The findings by Coates *et al*. (2006a) and the scrutiny to which experience-based food insecurity scales have been submitted suggest that the items included in the FIES questionnaire can serve as a valid starting point to create a globally valid measurement tool. However, most of the discussion thus far has proceeded under the assumption that the food insecurity measures produced by the scales are indeed reliable measures of the underlying concept. By being fully aware of the potential risk of circularity of the arguments, we must recognize that meaningfulness of the concept is not synonymous with meaningfulness of the measures. In other words, establishing the validity of the concept is not sufficient to claim validity of the measures. This is a crucial aspect in the analysis of validation, to which we now turn.

5.3 From a concept to a measure: application of the Item Response Theory (IRT).

The next step in the development and validation of an actual measurement tool based on the concept of food insecurity experiences is the identification of methods to obtain and process the information that will lead to food insecurity *scores*, along the continuum defined by the arrow in Figure 1 (Section 2, page 6). In bridging the gap between defining the concept and developing a measure, the lead was taken by a USDA-led team of collaborating researchers while developing the US HFSSM.[12] Based on the results of the qualitative research done at Cornell and elsewhere in the U.S., it was evident that probably the best way to obtain the necessary evidence on peoples' experience was through *self-reported data* on perceptions, behaviors and experiences associated with different levels of food insecurity. It was then proposed that the analytics for processing the data and computing the scores could be successfully borrowed from the toolkit of *Item Response Theory* (IRT) models, commonly used in the educational and psychological testing fields.

The foundation of modern IRT is the assumption that a quantitative measure of an underlying, unobservable construct (i.e., a *latent trait*) can be inferred from a set of dichotomous variables (1/0, for positive/negative, correct/incorrect, affirmed/denied) obtained as the result of a test.[13]

[12] See Bickel *et al*., 2000 for a description of how this has been accomplished.

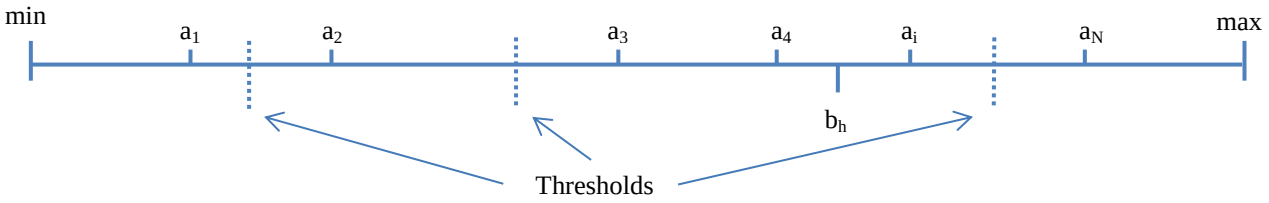
[13] In educational testing, the underlying construct is the "proficiency" or "ability" level of the subjects, and a test is composed of a collection of items clearly linked to the underlying construct and characterized by different levels of difficulty. In psychological testing, the underlying construct is an "attitude" or a "character", and test items are represented by "symptoms" or "manifestations" typically associated with that attitude or trait. In experience based food insecurity measurement, the underlying trait is the severity of resource-constrained food insecurity, and the items are typical experiences associated with increasingly binding constraints.

The fundamental assumption of this theory of measurement is that the probability of testing positively, responding correctly, or affirming an item associated with a certain level of difficulty or severity depends on the unobservable true ability, attitude or condition of the subject. Formally, by using a_i for the item parameter and b_h for the respondent's, the dependency can be represented by a function, *F(·)*, termed the *item characteristic curve* (ICC) or *item response function* (IRF):

$$\text{Prob}\left(x_{h,i} = 1 \middle| b_h, a_i \right) = F(b_h, a_i) \qquad (1)$$

Once the function *F(·)* is specified, the likelihood of obtaining a certain set of responses to different items - each characterized by a different difficulty level - from a subject with a certain degree of proficiency can be ascertained. This provides the statistical basis for estimating both the items' difficulty parameters and the respondents' ability scores. The estimated parameters will correspond to locations along the underlying scale of abilities. When necessary, thresholds can also be defined on the same scale to identify groups of classes of severity. (See Figure 5)

Figure 5: Thresholds



By comparing Figures 1 and 5, the parallel with the problem of measuring the severity of food insecurity should be evident: the model in (1) can be interpreted as describing the probability that a respondent whose unobservable food insecurity level is b_h would report having gone through the experience or behavior corresponding to the severity a_i. When a set of data is available on the responses from a number, say *T*, of respondents to the same set of, say *N*, questions, both the *N* different *item parameters*, a_i, and the *T* different *respondents' parameters*, b_h, can be estimated, most commonly through maximum likelihood procedures.

The idea is rather simple, yet powerful: by looking at the way in which the many respondents report on one of the experiences, one can establish a measure of the severity associated with that experience, i.e., where that experience is located on the scale (intuitively, experiences reported by a larger number of subjects are deemed less severe, and vice versa). Then, once the severity level characteristic of each question has been established, each respondent can be located on the same scale by looking at the answers provided to the whole set of questions, that is, how many and which experiences have been reported. Based on their position along the continuum of the food insecurity severity scale represented in Figure 1, they can be classified in different classes of severity levels.

The Rasch model

Modern IRT provides a battery of models that can be used to address various issues common in any attempt at measuring latent traits, which explains why IRT is becoming the dominant framework for measurement in many social sciences, in particular in psychology and educational testing.[14] Among them, one of the simplest formulations is the **One Parameter Logistic Model**, also known as the classical **Rasch model** (Rasch, 1960).

In a Rasch model, the probability that a subject with ability b_h responds correctly to a test item characterized by difficulty level a_i is modeled as a logistic function of the distance between b_h and a_i:[15]

$$\text{Prob}\left(x_{h,i} = 1 \middle| b_h, a_i \right) = F(b_h - a_i) = \frac{e^{b_h - a_i}}{1 + e^{b_h - a_i}} \qquad (2)$$

Application of the model to the measurement of food insecurity severity entails interpretation of the *a* parameters as reflecting the severity associated with the experience captured by the different questions, and of the *b* parameters as the measure of the level of food insecurity experienced by the respondents. Assuming that the probability of affirming any item by each respondent is (locally stochastically) independent from the probability of affirming any other item by the same respondent, or from the probability of affirming the same item by any other respondent, the likelihood function of any sample of 0/1 responses, $\mathbf{X} = \{x_{h,i}\}$ (where $x_{h,i} = 1$ if the *h*-th respondent reported having experienced item *i*, and 0 otherwise) can be determined by the product of the probabilities of each $x_{h,i}$.[16] The likelihood function can then be maximized with respect to the vectors of parameters, ***a*** and ***b,*** yielding the required measures.[17]

The major attractiveness of the simplest formulation of the traditional Rasch model as in (2) resides in the fact that the measures of individual severity (i.e., the estimated *b* parameters) are linked monotonically (albeit not linearly) to the *raw score*, that is the number of affirmed items. This means that raw scores are a sufficient basis to represent the severity of food security of the respondents on an

[14] Also in economics, since the pioneering work of McFadden (1974), methods for latent variable measurement are becoming standard with application of Random Utility Models in the analysis of both stated and observed choices (Louviere, Hensher and Swait, 2000). The analytics of discrete choice models parallels very closely that inspiring IRT.

[15] Notice that, because the model is specified in terms of the difference between *a* and *b*, only the distance between two parameters is meaningful. This means that in practice one is free to choose any reference point as the "zero" of the measurement scale and to express the "units" as it is most convenient. This is certainly not a problem for any stand-alone assessment, i.e. one conducted with one dataset only and when results are interpreted with reference to that particular sample of subjects. When different assessments made with the same or similar set of items are to be meaningfully compared, however, some conventions need to be followed to set both the zero of the scale and the units. When applied to food insecurity measurement, the *b* parameter is interpreted as the food insecurity severity to which a respondent is exposed and corresponds to the severity associated with the condition that the respondent has a 50% chance of having experienced over the reference period.

[16] Local stochastic independence means that the probability of responding yes to one of the questions, conditional on having a certain level of food insecurity, is independent from having responded yes to any other of the questions, or from the fact that others have responded yes to the same question. In other words, it means that the food insecurity level of the respondent is the only variable that establishes the observed pattern of responses.

[17] For a detailed discussion of the assumptions needed to obtain consistent estimates of the unknown parameters, see any of the excellent references available on the Rasch model, e.g., Fischer and Moleenar (1995).

ordinal scale. Classes of food insecurity severity can be defined simply in terms of raw score, irrespective of the actual pattern of responses, with a degree of precision that is comparable to the one produced by more sophisticated models, but with obvious returns in terms of communication and simplicity of use (Nord, 2012).

This possibility of using the raw score as a measure of respondent's severity is the result of essentially two restrictions that are imposed on the model: one is the *conditional independence of the responses*; the other is the *equal discriminatory power of the questions* included in the questionnaire. To evaluate the validity of measures obtained via applications of a Rasch model, the question is whether or not these restrictions are rejected empirically: if the evidence supports the hypotheses, the measures produced are defensible also on empirical grounds if not, one has to look for more flexible models that relax some of the assumptions embedded in the simple Rasch formulation.

The empirical evidence so far seems to support application of the Rasch model to food insecurity measurement in various countries and conditions, which is one of the reasons why we believe that the OPLM may well be well-suited to constitute the basis for measuring the severity of food insecurity. The conviction is reinforced by the additional consideration that some of the problems highlighted in the literature on educational and psychological testing that have suggested the need for less restrictive, more sophisticated models will not arise in the case of food insecurity measurement.

The rapid evolution of IRT and the proliferation of existing models (including 2 parameter models, 3 parameter models, partial credit, mixture models, etc.), in fact, have been developed mainly in the educational testing field to address issues that are rather common in that area of measurement, such as for example the fact that an answer to a test may be only partially correct or the possibility of "guessing" the right answer. Analogous issues are much less relevant in food insecurity measurements. Here respondents are not expected to provide the "correct" answer, but only to report on whether or not they have experienced a well-described situation. Even though the possibility exists that somebody could make an honest mistake in self-reporting an experience, this may be more likely due to misunderstanding of the question, rather than to the different status of the respondent in terms of food insecurity (which would invalidate the measure). Significant empirical departures from the assumptions underlying the Rasch model are therefore much more likely due to insufficient care in formulating and asking the questions, than to problems with the theory of measurement that underpins it. [18] All of this

[18] The rapid evolution of IRT and the proliferation of existing models (including 2 parameter models, 3 parameter models, partial credit, mixture models, etc.) might well lead a casual reader to develop the wrong impression that the Rasch model is insufficiently sophisticated to capture the essential elements of the needed measurement in practice. To see things in the proper perspective, it should be considered that IRT models have been developed mainly in the educational testing field, where more sophisticated models than the one originally proposed by Rasch may be needed to address issues that are rather common in that area of measurement (such as the possibility of "guessing" the answer on a test). Analogous issues are much less relevant in food insecurity measurements, when respondents are not expected to provide the "correct" answer to a quiz, but only to report on whether or not they have experienced a well-described situation. Even though a probability exists that somebody could make a mistake in self-reporting an experience, this could be linked more to misunderstanding the question, rather than to a differential ability of the respondent.

points to the importance of careful linguistic and cultural adaptation of the questionnaire and to proper implementation of the survey.

Testing the assumptions required for proper measurement

One of the benefits of developing a measurement tool strongly rooted in a probabilistic theory is that the restrictions imposed on the model can be tested against the data. In this section we shall first review the meaning of the fundamental restrictions imposed by the OPLM formulation, and then outline the procedures for empirical testing.

Conditional independence is the assumption that allows forming the likelihood function used to estimate the model's parameters as the product of the conditional probabilities of affirming each item. It can be stated by saying that the probability of affirming one experience by a respondent, conditional on his or her food insecurity level, does not depend on whether or not he or she has affirmed other experiences, or on whether or not other respondents have affirmed the same experience. In other words, it amounts to saying that the respondents' food insecurity severity levels are sufficient to explain the structure of correlation that exists in the matrix of responses.

In educational testing, failures of conditional independence may arise because of problems in the actual test design that induce patterns of correlation in the responses which are not explained solely by the ability levels of the respondents (for example, because some test takers have "cheated" by peeking at the responses given by others, and therefore their probability of responding correctly depends also on the ability of neighboring test takers). In the context of food insecurity experience scales, conditional independence may fail to be reflected in the data because of, for example, inadvertently prompted responses by the interviewer, or of confusion between questions that are perceived by the respondent as being equivalent (so that the answer given to one of them conditions the answer given to the second). While failures of conditional independence can be detected in the data (see below), there is little scope for treating them once the data has been collected. Given that they most likely derive from problems in design and implementation of the test, such failures, if detected, will point to the importance that due attention is paid to preparation and implementation of the survey. In the context of the FIES questionnaire, these include careful linguistic and cultural adaptation of the questions and administration by properly trained enumerators, to avoid problems due to anchoring of the responses, inducing bias, etc.

Uni-dimensionality requires that the items included in the scale used for measurement capture one clearly identified dimension: the one that corresponds to the latent trait of interest, and not others. In the context of food insecurity measurement, this dimension is the *severity* of the food insecurity situation as related to the resource constraints that prevents people from achieving satisfactory levels of food consumption. Defining a scale to measure this dimension does not imply neglecting the existence or the importance of other dimensions, but only to stress that efficient measurement requires focusing on the object of interest, avoiding confounding elements.

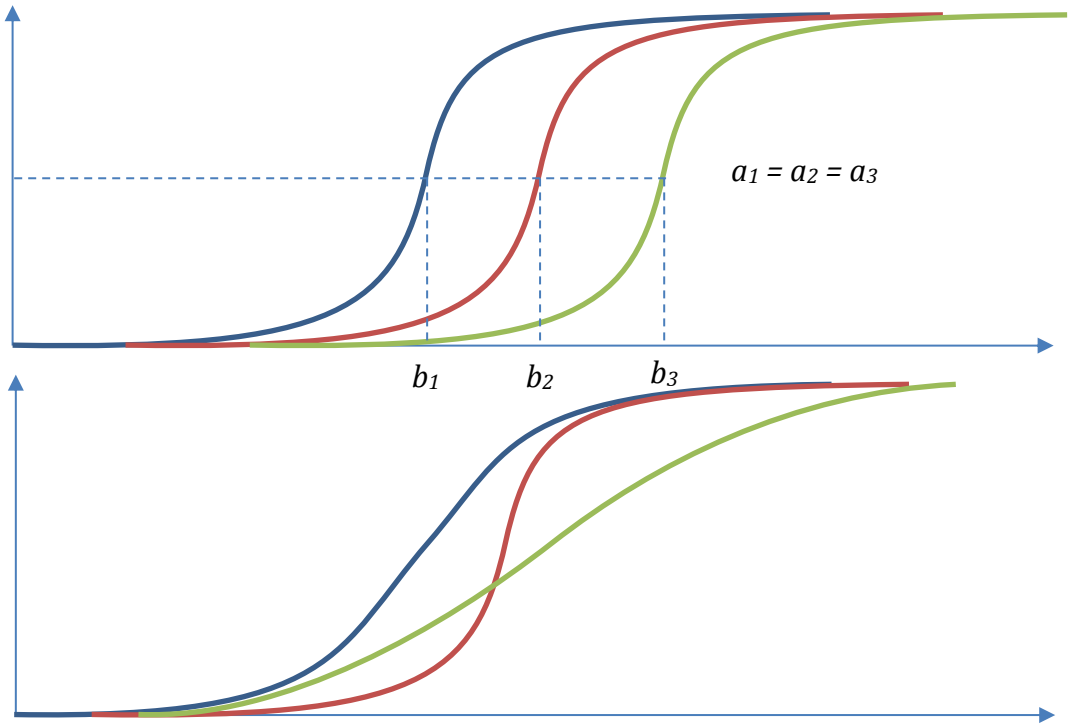
We have already discussed at some length the fact that the severity of resource-constrained food insecurity is a valid construct and that it is amenable to measurement, clarifying also that it should not be taken as an all-encompassing synthetic index of food and nutrition security. If two or more dimensions of a problem or of a latent trait are deemed relevant, effort should be made to define *different uni-dimensional* scales to be used for classification along a multi-dimensional lattice, rather than forcing the two dimensions on a single scale.

The discussion is relevant here because some of the existing food insecurity experience modules (such as the US HFSSM or the HFIAS) intended to capture the severity of food insecurity also contain follow-up questions on the related (but conceptually distinct) dimension of the *frequency* of the experienced food insecurity condition. While, in general, asking questions on both severity and frequency should not create problems, validity and efficiency of the food insecurity measure depend on the choice of the appropriate scoring and classification methods. Rather than trying to combine considerations of severity and frequency of food insecurity on a single scale[19], a more appropriate methodology seems to be identifying which items provide information on each of the dimension, and to use only those to score that particular dimension, as is currently done by the Economic Research Service of the USDA in their analysis of Household Food Insecurity in the United States (Nord, 2012).

Equal discrimination of the items means that the items in the scale have similar power in discriminating among respondents, which is seen, by proponents of the Rasch model, as a necessary prerequisite for a scale to reflect construct validity. The best way to describe this assumption is by reference to a diagram with Item Characteristic Curves of different items. The discriminatory power is reflected in the "steepness" of the ICC at the midpoint (reflected in the "*a*" parameters). In the top panel, the ICCs of three items with different severity (as reflected in the "*b*" parameters) are simple horizontal displacements of the same graph (corresponding to a logistic cumulative distribution function) which implies that the three items will be ranked in the same order of severity by any respondent independently from the location of the respondent on the food insecurity scale: the probability of affirming the most severe item is always lower than the probability of affirming a least severe one (the order in the vertical position of the three ICC, or the ranking in terms of severity of the three items, is always the same).

Figure 6: Item Characteristic Curves for three items with equal discriminatory power (panel (a)) or with different discriminatory power (panel (b))

[19] This is suggested in the proposed methods to determine food insecurity scores from the data collected through the HFIAS (HFIAS manual), or in Recommendation 5-1 in National Research Council, 2006 to improve the US assessment. Confusion arises from the suggestion to use polytomous IRT models based on data collected on the *frequency* of experiences to derive a scale of *severity*, essentially treating frequency as a sub-dimension of severity. The problem with this approach is that it would be imposing an unwarranted *quantitative* comparison between severity and frequency (as if being exposed twice to a certain experience is equivalent to being exposed once to an experience that is twice as severe).



Violation of equal discrimination (as revealed in the data by estimating significantly different slopes of the ICCs, and that would appear as in the second panel of Figure 6) implies the possibility that ranking of the items in terms of severity may change depending on the position of the respondent. According to the prevailing interpretation within the Rasch measurement school of thought, this would seriously question construct validity. In fact, violation of equal discrimination implies that, when confronted with the same set of experiences, people differing in their condition will rank them in different order of severity, which makes the "severity" construct somehow ambiguous. This is why proponents of the Rasch model insist that equal discrimination is something that should not be easily given up, and attention should be focused instead on using a set of items that discriminate equally, even if that means sacrificing some of the data by not using it for scoring some of the items whose discrimination parameter is mostly different from average. Low discriminating items may result, for example, from difficult interpretation and possible misunderstanding of the meaning of the experience being presented in a question, stressing once again the need for proper design and careful adaptation of the questionnaire to local languages and cultures.[20] Even in cases where some of the items reveal degrees of

[20] For example, in the FIES, if insufficient emphasis is given to the qualifier "for lack of money or other resources" to the experience of skipping meals, the item can be affirmed with high probability also by relatively food secure people, who skip meals for religious, health related or other reasons. The result will be a flatter ICC, located at very low levels of food insecurity, thus completely missing the point.

discrimination different from average, this does not mean that Rasch theory has to be abandoned, as there is always the possibility to identify a subset of questions that could form a proper basis for scoring in a way that is theoretically consistent with Rasch theory. This will entail a loss of efficiency (as some of the question asked will not contribute to the measurement) but can still allow for meaningful assessments, especially if there is convincing evidence that problems with the items that are discarded may be the result of the way in which the questions have been asked, and identifying them can provide useful information on how to possibly improve on the set of items in subsequent applications.

The conclusion of this brief discussion on the restrictions imposed by the OPLM (the Rasch model) is that several potentially problematic features of the data could be the result of faulty implementation of the measurement tool, rather than signs of the fact that the model is inadequate as a basis for providing valid measures of food insecurity severity. Nevertheless, whether or not the restrictions imposed by the OPLM formulation are relevant remains an empirical matter that can be addressed once the data has been collected. If the evidence is such that the restrictions must be rejected, proper consideration must be given to implementation issues. If no problems are identified in the survey implementation, it may be necessary to resort to more sophisticated versions of IRT models that relax some of the assumptions.[21]

How to test, in practice

All indications thus far point to the fact that the OPLM is indeed an appropriate formalization for meaningful measurement of the severity of food insecurity. In practice, adequacy is assessed through three kinds of tests to verify the suitability of the model: two to assess whether the data adequately fit the Rasch assumptions or not, and the third one to assess the overall reliability or "performance" of the measure.

All these tests are based on the analysis of the "residuals", that is the differences between the observed responses and those that would be expected under the truth of the estimated model. Given estimated values for item ($\hat{a}$) and person ($\hat{b}$) parameters, the probability of affirming the *i*-th item by the *h*-th respondent can be predicted as:

$$E_{h,i} = \frac{e^{\hat{b}_h - \hat{a}_i}}{1 + e^{\hat{b}_h - \hat{a}_i}}$$

A residual $r_{h,i}$ can then be calculated as the difference between the actual response, $x_{h,i}$, and the expected value $E_{h,i}$ for each item/person combination, and a matrix of residuals $R = \{r_{h.i}\}$ can be formed. As the responses are coded as zero/one and the predicted probabilities are strictly included within the $\{0\ ,1\}$ interval, the residuals will always be values between -1 and +1, and will be positive when the item has been affirmed, and negative when it has been denied. Moreover, a residual will be larger, in absolute value, either when an individual with a low estimated parameter (i.e., found to be relatively food

[21] The slightly more sophisticated 2 parameter logistic model (2PLM), for example, relaxes the assumption of equal discrimination of the items and could be estimated with relatively minimal added complexity, even though it will require a different approach to scoring cases.

'secure') affirms severe items, or when individuals with high parameters (i.e., relatively more food 'insecure') deny the less severe parameters. High values of fit statistics are thus evidence of response patterns that are difficult to reconcile with the estimated severity of the items and characteristics of the respondent.

As with the matrix of actual responses, also the matrix of residuals (organized with each row representing one respondent and each column representing one item) can be used for a differential analysis of the items and of the respondents. First, by looking at how the residuals are distributed across the sample of respondents, one can derive indications on the relative "performance" of each of the items; then, analyzing the way in which residuals are distributed across items, one can identify potential problems with one or a group of respondents.

In practice, the row and column averages of the squared residuals can be taken as measures of the extent of misfit of respondents and of items, respectively.[22] Under the truth of the model, these statistics have an expected value of one. Ideally, a scale for meaningful measurement of a uni-dimensional underlying trait ought to be formed by items whose fit statistics are between 0.8 and 1.2. In practice, once the data has been collected, values in the range of 0.5 to 1.5 are considered acceptable, identifying items that are "productive for measurement" (Wright and Linacre, 1994). Detected differences between the in-fit and out-fit values for a given item are a sign that responses to that item may be idiosyncratic for a small group of respondents, perhaps revealing problems with the way the question has been differently understood by that group of people.

The available empirical evidence on application of the Rasch modeling to food insecurity measurement, based on both published and unpublished results of the analysis of various scales derived from the US HFSSM (both for the US and for other countries) points clearly towards the possibility to identify a limited set of items that ensures equal discrimination. A cursory review of the results from those applications in various countries and in different settings reveals in-fit values between 0.7 and 1.3 for most of the items (Nord, personal communication).[23]

The analysis of fit statistics provide a sufficiently solid basis to assess the appropriateness of the model and of the data for measurement purposes, especially if measurement is aimed at obtaining estimates of the populations' prevalence of a limited number of classes of food insecurity. If more reliable measures of *individual* food insecurity measures are necessary, then more extensive tests are available that can be useful in deciding whether more sophisticated models may be needed.

[22] Two types of statistics are traditionally reported in IRT analyses and labeled as *in-fit* and *out-fit* respectively, the difference being that, in the first case, the sums of row (respectively, column) squared residuals are weighted by the variance of responses in the corresponding columns (rows) of the original data matrix. Both are measures of the contribution of each item/respondent to overall misfit, but in-fit statistics are more robust to the presence of a few "outliers," a reason why they are preferred to assess the performance of the items.

[23] The results of these analyses have, for the most part, not been included in the published reports because they did not fit the type of publication, focusing more on the results than in the technical aspects of the measure.

Conditional independence across items, for example, is typically assessed by examining the pattern of correlations of residuals across items, based on the implication that, under the truth of the Rasch model, residuals should not show any discernible pattern of correlations. Once again, if some residual correlation is detected in the residuals, this may be used as a guide to select only a subset of items used for scoring, and to guide on the possible improvement of the items.

In the context of experience based food insecurity measurement, all the testing and resulting refinement of the questions will be greatly favored by the availability of datasets coming from different countries and contexts, one of the outputs expected from the Voices of the Hungry project.

5.4 Equalization of the scale and methods for classification

When proposing the FIES as the basis for a global standard to monitor food insecurity, establishing the appropriateness of the measurement theory is not enough. Another challenge is to use the information provided by the answers to the FIES questionnaires to classify cases into food insecurity classes in a way that is meaningful and comparable over time and across countries and socioeconomic contexts. Doing so requires: (a) establishing the metric equivalence of the scale, and (b) classifying cases into different food security levels taking into account possible differences in the severity of some items in some of the countries.

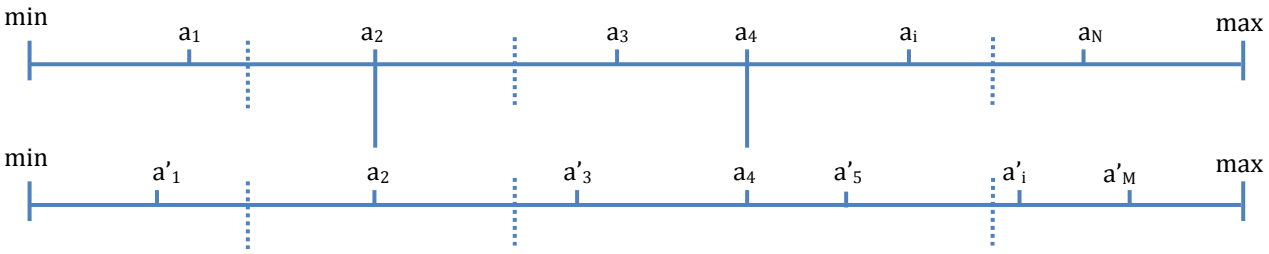
The need to ensure the *metric equivalence of the scale* derives from a fundamental indeterminacy of the IRT models (and of the Rasch model in particular): as the probability of affirming an item is a function only of the *difference* between the person and the item parameter, one is free to arbitrarily set the "zero" of the absolute scale. Also, as the model is defined only in terms of *relative* probabilities, only *N*-1 item parameters can be estimated. This means that, in a sense, each dataset of responses from a food insecurity experience set of questions can be used to measure the *relative* position of respondents on a scale that is strictly valid for *that* dataset in *that* instance, and this relative position would not change if the scale is transformed through an affine function. When comparing estimates from different applications (for example when analyzing data from different countries) metric equivalence must be ensured. In the contexts of unidimensional IRT models such as the Rasch model, this requires appropriate *anchoring* and *scaling* of the person parameters estimates obtained in different contexts to make sure that they are all expressed on a scale with the same zero and the same unit of measure. This is most commonly achieved by rescaling severity estimates of a group of *anchoring* items so that the group of items will have the same mean and standard deviation of severity. While only items that show the same severity in different countries can be used as anchoring, it is not necessary to require that *all* items in the scale are anchoring items in order to obtain proper comparisons of the severity measures. This is important as one of the reservations raised against the possibility of defining a globally valid food insecurity scale, when analyzing the experience with the HFIAS, has been based on the fact that only three items proved robust to cross country comparisons (Deitchler *et al.*, 2010). That analysis, however, is conducted on the presumption that there are no alternatives to discrete assignment of cases to food security classes based on raw scores, which may be an inappropriate limitation. In reality,

differences in item severity across countries do not necessarily imply rejection of the underlying concept of an experience-based scale or of the capacity to compare prevalence rates based on such measures across countries.

This will require, unfortunately, abandoning the simplest classification based on discrete assignment of food security status using raw scores, as it may produce biased results. Discrete classification by raw score is only one of several possible methods for comparing measures on the same latent trait across countries and language groups. Two more flexible methods are available that retain the simplicity of the Rasch model. These require probabilistic assignment of food security status based on the severity parameter and measurement error for each raw score, which may differ from country to country.

A third method abandons the simplest Rasch model in favor of a two-parameter IRT model, relaxing the assumption of equal item discrimination and may provide even more precise comparability of the actual individual level measure, and could be explored for research purposes, when the objective of the analysis goes beyond estimation of prevalence rates.

Figure 7: Two test forms yielding comparable measure on a scale



Full commonality of the items actually included in a measurement test is not a mandatory requirement. Modern theory of measurement in educational testing recognizes this, for example, since in this case, there is often the opposite need to ensure that no two instances of the same test are identical, but that the various instances produce fully comparable measures. This is accomplished by developing different test forms, and by considering the specific set of items included in one form as being selected from a larger set of available items, all of which relate to the same underlying latent trait. Results are then processed in a way that allows equalization of the measures obtained. The figure represents two forms of an equivalent test, one made of N items, and the second made of M items. The presence of (at least) two common items (anchoring) in the two forms allows equalization of the two scales.

As evidenced by this discussion of the statistical methods that may be required for classification of severity of food insecurity, further research is needed to confirm the validity of a global experience-based food insecurity scale and ensure cross-cultural comparability of results. The VOH aims to

undertake the challenge of exploring the issues raised here based on global data made available for the first time to enable such analyses. It is the next step in the already impressive evolution of experience-based food insecurity scales.

6. Uses of the FIES: Linking Information and Action

The above discussion presents the evidence that the FIES is based on a well-grounded theoretical construct, and that reliable and accurate measures of the experience of food insecurity are possible when based on careful application of the principles of Item Response Theory. However, information is but one element that contributes to effective decision making. The real challenge is to link the information to action.

The FIES can be used to provide information for purposes ranging from advocacy and policy formulation to basic research. Some of the potential uses of the FIES, and possible modifications for different purposes, are discussed below.

6.1 Estimation of food insecurity prevalence

The VOH aims to use the FIES to estimate the prevalence of food insecurity of individuals at different levels of severity across countries and regions of the world. However, it can also be used to estimate prevalences at the sub-national level, for example regions of a country, states or provinces, cities and even neighborhoods. When included in a survey that allows disaggregation beyond the national level, the prevalence of food insecurity can also be estimated for sub-groups of these populations, such as rural or urban residence, age groups, and different racial or ethnic groups.

Prevalence is typically reported as a percentage of respondents surveyed who have the characteristic being studied. The different levels of severity of food insecurity (i.e. mild, moderate and severe) are sometimes combined for analysis and reporting, often to increase the power of the statistical inferences. However, they should be reported separately whenever possible to provide more complete information, as the differences between them are meaningful, theoretically as well as empirically.

6.2 Targeting and defining priorities for programmes and resources

Results from surveys that include the FIES can also be used to inform decisions regarding priorities for targeting programmes and resources. While it is not appropriate to use the FIES to identify individual beneficiaries for programmes, the information provided by population surveys that include the FIES can serve to identify vulnerable sub-populations or geographic areas that are more affected by food insecurity.

National survey data in Brazil in 2004, for example, clearly illustrated marked differences in household food insecurity levels among the five geographical regions of the country (BRASIL-IBGE, 2006; Segall-Corrêa, *et al*., 2007). Evidence of these stark regional inequalities convinced the Brazilian government to direct resources and public policies toward the more vulnerable regions.

6.3 Monitoring trends in food insecurity

With repeated application of the FIES in the same population, it is possible to monitor changes in food insecurity levels over a period of time. Several countries have already incorporated experience-based food insecurity scales into their national monitoring systems, including the U.S.A., Canada, Mexico, and Brazil.

When using trend data, it is important to report the different levels of food insecurity as mentioned above because this reveals subtle but meaningful changes in people's food insecurity experience from one time period to the next. For example, in Brazil, a decrease in prevalence of severe food insecurity was accompanied by an increase in moderate food insecurity, suggesting that households had moved out of a more serious food access problem into a slightly improved situation. In the same way, families previously classified as moderately food insecure appeared to have moved into the mild food insecurity category (FAO, 2012a). Such changes have real significance for the lives of those at risk of food insecurity. Simply reporting "food secure" and "food insecure" can mask the shift in the magnitude of food insecurity over time, which may be a positive outcome of programmes and services.

Looking at food insecurity experience across a range of severity, and how this changes over time, provides valuable information for accurately reporting the wellbeing of citizens and for adjusting programmes and services to better reduce vulnerability and food insecurity. Likewise, timely monitoring of food insecurity, with special attention to the different levels of severity, could possibly serve as an early warning indicator. A sudden surge in mild and moderate food insecurity without a simultaneous drop in severe food insecurity may be an alert that levels of hunger may rise, as well, calling for timely action.

6.4 Identifying risk factors and consequences of food insecurity

The underlying causes of food insecurity and hunger are social injustice, inequalities and the lack of guarantees of the economic, social, cultural, and environmental rights of the population, including the human right to adequate food. This context gives rise to more immediate risk factors such as poverty and lack of access to education and good jobs with living wages.

It is thus important to keep in mind that the phenomenon of food insecurity encompasses much more than what the FIES captures; it includes aspects ranging from social, economic and agriculture policies at the international and national levels to livelihood strategies, basic sanitation, food habits and nutritional status at the household level. Indicators associated with these additional aspects should be

considered in broad-based studies to complement the information derived from the FIES, aiming to build a better understanding of the relationships among various factors, including the risk factors and the consequences of food insecurity.

Consequences may range from negative psychosocial effects or going without other essential items in order to get food, to nutritional problems related to consuming a diet of poor quality, such as micronutrient deficiencies or obesity, to frank undernutrition, such as underweight, stunting and wasting, and in extreme cases to prolonged hunger or starvation.

The advantage of having a direct experiential measure of food insecurity like the FIES is that it can be used in combination with other indicators to build a better understanding of this complex phenomenon.

6.5 Modifications to the FIES in other survey contexts

There are several modifications that can be made to the FIES questionnaire to meet the specific needs of the users, including altering the reference period and using households rather than individuals as the unit of data collection.

The FIES being used in the VOH is a measure of food insecurity collected at the individual level, which provides population prevalence estimates of people's experiences with food insecurity. This approach makes it possible to frame the experience of hunger from the perspective of individuals whose human right to adequate food may be violated. However, many national population surveys collect data at the level of the household, such as Household Income and Expenditure Surveys, Household Budget Surveys and Living Standard Measurement Studies. When used in household-level surveys, the FIES items can easily be adapted to capture experiences at that level (by changing the wording to the questions slightly as follows: Was there was a time when "you or any other member of your household"….).

The choice of the reference period depends on the objectives for collecting the data. Where seasonal changes are of interest, a three-month period might be the best choice to investigate how food insecurity varies in different seasons. However, this would limit comparability if data are collected across different areas that do not have the same climatic or agricultural calendar. In the case of national surveys with the objective of estimating overall prevalence of food insecurity in different parts of the country, or multi-national surveys that include countries with different environmental and climatic zones, a 12-month period is recommended to avoid the influence of seasonal effects and to improve comparability of the measure across different areas. The FIES makes use of a 12 month reference period for this reason. However, if the survey is being carried out in settings undergoing a humanitarian crisis, the investigators may be interested in capturing the most recent time period, in which case a one-month recall would be appropriate.

The information resulting from the uses of the FIES described above could be relevant for a variety of audiences at many levels, including advocates, community leaders and activists, researchers, programme

managers and government officials at all levels. FAO is committed to making this link between information and action more effectively. Prevailing models of policy-making concede a potentially influential role for survey-based information. However, there is also consensus that policy-making is a messy process characterized by negotiation among competing interests and diffuse decision making processes (Jones, 2009; Weiss, 1983). Those responsible for producing the information are therefore called upon to strive to improve their political savvy and reflect on the uses and appropriations of the resulting information.

7. Conclusion

While it cannot be said that the FIES gives voice to the hungry per se, it does represent a step closer to framing the experience of hunger from the perspective of those whose human right to adequate food is being violated. It is a measurement tool with great potential for contributing to monitoring systems that meet the five principles recently identified by the Committee on World Food Security in the 2012 Global Strategic Framework for Food Security and Nutrition (CFS, 2012a):

> a) Food Security and Monitoring Systems should be human-rights based, with particular reference to the progressive realization of the right to adequate food;
>
> b) They should make it possible for decision-makers to be accountable;
>
> c) They should be participatory and include assessments that involve all stakeholders and beneficiaries, including the most vulnerable;
>
> d) They should be simple, yet comprehensive, accurate, timely and understandable to all, with indicators disaggregated by sex, age, region, etc., that capture impact, process and expected outcomes;
>
> e) They should not duplicate existing systems, but rather build upon and strengthen national statistical and analytical capacities.

Research being undertaken by the VOH over the next several years will contribute more evidence regarding whether or not the FIES represents the best basis for a valid experience-based measure of food insecurity *worldwide*. However, despite the particular challenges of using it as a global measure, experience in many countries leaves little doubt regarding its potential to be adopted by national governments and used successfully to guide policy at the national and sub-national levels. Its ease of application and analysis makes it accessible to people at many levels and from diverse fields. Local governments, non-governmental organizations and advocacy groups can also appropriate this relatively simple instrument to monitor food insecurity locally or regionally, engaging diverse stakeholders in the process, and building bridges between people of different backgrounds. This may in fact be where their greatest potential lies to effect change and contribute to guaranteeing the human right to adequate food.

References

Adams, E.J., Grummer-Strawn, L. & Chavez, G. 2003. Food insecurity is associated with increased risk of obesity in California women. *Journal of Nutrition*,133: 1070-4.

Albanian Centre for Economic Research. 2012. *Linguistic adaptation of the Albanian food and nutrition security scale.* United Nations of Albania. Tirana.

Alaimo, K., Olson, C.M. & Frongillo, E.A. 2001. Low family income and food insufficiency in relation to overweight in U.S. children: Is there a paradox? *Arch Pediatr Adolesc Med*., 155(10):1161-7.

Álvarez, M.C., Estrada, A., Montoya, E.C. & Melgar-Quiñonez, H. 2006. Validación de escala de seguridad alimentaria doméstica en Antioquia, Colombia. *Rev Salud Pública de México*, 48:474-481.

Álvarez, M.C., Melgar-Quiñonez, H., Estrada A. & Vélez, O.L. 2008. Validación interna y externa de la escala Latinoamericana y Caribeña para la medición de la Seguridad Alimentaria en el Hogar. Editorial Divergráficas. Medellín.

Ballard, T.J., Coates, J., Swindale, A. & Deitchler, M. 2011. *Household hunger scale: indicator definition and measurement guide.* Washington, DC: FANTA-2 Bridge, FHI 360.

Barrett, C. 2002. Food security and food assistance programs. *In* B. Gardner and G. Rausser, eds. *Handbook of Agricultural Economics*, Volume 2. Elsevier Science B.V.

Bernal, J., Frongillo, E.A., Salgado, T., Candela, Y., Herrera, H. & Rivera, J.A. 2012. *Design and validation of experienced-based tools that measure food insecurity and hunger in vulnerable populations of children and seniors in a developing country*. Presentation at the FAO International Scientific Symposium, Rome, 18 January. (available at http://www.foodsec.org/fileadmin/user_upload/eufao-fsi4dm/docs/German_1400_JBernal.pdf).

Bickel, G., Nord, M., Price, C., Hamilton, W. & Cook, J. 2000. *Guide to measuring household food security*. Revised March, 2000. Alexandria, VA., U.S. Department of Agriculture (USDA), Food and Nutrition Service.

Brasil/IBGE. 2006. Pesquisa Nacional por Amostra de Domicílios-2004, Suplemento de Segurança Alimentar; Rio de Janeiro. (available at http://biblioteca.ibge.gov.br/visualizacao/monografias/GEBIS%20-%20RJ/segalimentar/suguranca_alimentar2004.pdf).

Campbell, C.C. 1990. Food insecurity: a nutritional outcome or a predictor variable? *Journal of Nutrition*, 121:408–415.

Coates, J. 2013. Build it back better: Deconstructing food security for improved measurement and action. *Global Food Security*, Vol 2 Issue 3, Sept. p 188-194. (available at http://dx.doi.org/10.1016/j.gfs.2013.05.002).

Coates, J., Swindale, A., & Bilinsky, P. 2007. *Household Food Insecurity Access Scale (HFIAS) for measurement of household food access: indicator guide* (v. 3). Food and Nutrition Technical Assistance Project, Academy

Coates, J., Wilde, P.E., Webb, B., Rogers, L. & Houser, R.F. 2006a. Comparison of a qualitative and a quantitative approach to developing a household food insecurity scale for Bangladesh. *Journal of Nutrition*, 136: 1420S-1430S.

Coates, J., Frongillo, E.A., Rogers, B.L., Webb, P., Wilde, P.E. & Houser, R. 2006b. Commonalities in the experience of household food insecurity across cultures: what are measures missing? *Journal of Nutrition*, 136: 1438S-1448S.

Committee on World Food Security (CFS). 2012a. Global Strategic Framework for Food Security and Nutrition, First Version. Committee on World Food Security, Thirty-ninth Session, Rome, Italy, 15-20 October.

Committee on World Food Security (CFS). 2012b. Coming to terms with terminology. Committee on World Food Security, Thirty-ninth Session, Rome, Italy, 15-20 October, Item V.a. (available at http://www.fao.org/docrep/meeting/026/MD776E.pdf).

Cook, D.A. & Beckman ,T.J. 2006. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am Journal Med.* 119:166.e7 – 166.e16.

Deitchler, M., Ballard, T.J., Swindale, A. & Coates, J. 2010. *Validation of a measure of household hunger for cross-cultural use.* Technical Report. Washington, DC: FANTA-2. Academy for Educational Development. Washington, DC.

Derrickson, J.P., Fisher, A.G. & Anderson J.E.L. 2000. The core food security module scale measure is valid and reliable when used with Asians and Pacific Islanders. *Journal of Nutrition*. 130: 2666–2674.

Derrickson, J. & Anderson, J. 2000. Face validity of the core food security module with Asians and Pacific Islanders. *Journal of Nutrition Education*, 32: 21-30.

FAO. 2002a. Summary of Proceedings: Measurement and Assessment of Food Deprivation and Undernutrition. International Scientific Symposium, 26-28 June 2002, Rome. (available at http://www.fao.org/docrep/005/Y4249E/y4249e00.htm)

FAO. 2002b. Keynote speech by Eileen Kennedy, International Scientific Symposium, 26-28 June, Rome. (available at http://www.fao.org/docrep/005/Y4249E/y4249e0c.htm#bm12).

FAO. 2009. Declaration of the World Summit on Food Security, World Summit on Food Security, Rome, 16-18 November. (available at ftp://ftp.fao.org/docrep/fao/Meeting/018/k6050e.pdf).

FAO. 2012a. Escala Latinoamericana y Caribeña de seguridad alimentaria (ELCSA): Manual de uso y aplicaciones. FAO Regional Office, Santiago. (available at http://www.rlc.fao.org/es/publicaciones/elcsa/).

FAO. 2012b. State of Food Insecurity. Rome. (available at http://www.fao.org/docrep/016/i3027e/i3027e.pdf).

FAO. 2012c. Proceedings of the International Scientific Symposium on Food and Nutrition Security information: from Valid Measurement to Effective Decision Making. 17-19 January 2012. Rome. (available at http://www.fao.org/docrep/017/i3244e/i3244e.pdf).

Faye, O., Baschieri, A., Falkingham, J. & Muindi, K. 2011. Hunger and food in security in Nairobi's slums: an assessment using IRT models. *Journal of Urban Health*, 88 (Suppl 2): s235 – s255.

Finney Rutten, L.J., Yaroch, A.L., Colón-Ramos, U., Johnson-Askew, W. & Story, M. 2010. Poverty, food insecurity, and obesity: a conceptual framework for research, practice, and policy. *Journal of Hunger & Environmental Nutrition,* 5:4, 403-415. (available at http://dx.doi.org/10.1080/19320248.2010.527275).

Fischer, G.H. & Molenaar, I.W. (eds). 1995. *Rasch Models. Foundations, Recent Developments, and Applications.* Springer-Verlag: New York, NY.

Fram, M.S., Frongillo, E.A., Jones, S.J., Williams, R.C., Burke, M.P., DeLoach, K.P & Blake, C.E. 2011. Children are aware of food insecurity and take responsibility for managing food resources. *Journal of Nutrition*, 141(6): 1114-9.

Frongillo, E.A. 1999. Validation of measures of food insecurity and hunger**.** *Journal of Nutrition,* 129(Suppl 2S)**:** 506S-509S.

Frongillo, E.A. 2003. Understanding obesity and program participation in the context of poverty and food insecurity. *Journal of Nutrition*,133: 2225-2231.

Frongillo, E.A. & Nanama, S. 2006. Development and validation of an experience-based measure of household food insecurity within and across seasons in northern Burkina Faso. *J. Nutrition*. 136: 1409S–1419S.

Frongillo, E. A., Rauschenbach, B. S., Olson, C. M., Kendall, A. & Colmenares, A.G. 1997 Questionnaire-based measures are valid for the identification of rural households with hunger and food insecurity. *Journal of Nutrition*, 127:699-705.

Grebmer, K., Ringler, C., Rosegrant, M.W., Olofinbiyi, T., Wiesmann, D., Fritschel, H., Badiane, O., Torero, M., Yohannes, M., Thompson, M., Oppeln, C. & Rahall, J. 2012. *Global Hunger Index: The Challenge of Hunger: ensuring sustainable food security under land, water and energy stresses.* International Food Policy Research Institute (IFPRI), Welthungerhilfe and Green Scenary, Concern Worldwide; Bonn/ Washington, DC/Dublin . (available at: http://www.ifpri.org/sites/default/files/publications/ghi12.pdf).

Habicht, J.P., Pelto, G., Frongillo, E. & Rose, D. 2004. *Conceptualization and instrumentation of food insecurity.* Paper prepared for National Academy of Sciences workshop, 15-16 July 2004. (available at http://www.nap.edu/openbook.php?record_id=11578&page=41).

Hackett, M., Melgar-Quiñonez, H. & Álvarez, M.C. 2008. Internal validity of a household food security scale is consistent among diverse populations participating in a food supplement program in Colombia. *BMC Public Health*, 23(8):175.

Hamelin, A.M., Beaudry, M., Habicht, J.P. 2002. Characterization of household food insecurity in Québec: food and feelings. *Soc Sci Med.*, 54:119–32.

Hamilton, W.L., Cook, J.T., Thompson, W.W., Buron, L.F., Frongillo, E.A., Olson, C.M. & Wehler, C.A. 1997. *Household food security in the United States in 1995*. Summary Report of the Food Security Measurement Project. USDA, Washington, DC. (available at: http://www.fns.usda.gov/oane/MENU/Published/FoodSecurity/SUMRPT.PDF).

Hromi-Fiedler, A, Bermúdez-Millán A, Segura-Pérez S, Damio G & Pérez-Escamilla R. 2009**.** Adaptation of the U.S. Food Security Survey Module for Low-Income Pregnant Latinas: Qualitative Phase. *Hunger Environ Nutrition*, 4(1):62-80.

Jones H. 2009. *Policy-making as discourse: a review of recent knowledge-to-policy literature*. ODI-IKM Working Paper No. 5, London, ODI.

Jones, A.D., Ngure, F.M., Pelto, G. & Young, S.L. 2013**.** What are we assessing when we measure food security? A compendium and review of current metrics. *Adv. Nutr.*, 4: 481–505.

Kac, G., Melendez, G., Schlussel, M.M., Segall-Correa, A.M., Silva, A.A.M. & Pérez-Escamilla, R. 2012. Severe food insecurity is associated with obesity among Brazilian adolescent females. Public Health Nutrition. 15(10):1854-60.

Lorenzana, P.A. & Sanjur, D. 1999. Abbreviated measures of food sufficiency validly estimate the food security level of poor households: measuring household food security. *Journal of Nutrition*, 129(3): 687-92.

Louviere, J.J., Hensher, D.A. & Swait, J.D. 2000. Stated Choice Methods. Analysis and Application. Cambridge University Press: Cambridge, UK

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior, in P. Zarembka (ed.), *Frontiers in Econometrics*, 105-142, Academic Press: New York, 1974

Melgar-Quiñonez, H. 2010. *Informe sobre la aplicación pasada y presente de la Escala Latinoamericana y Caribeña de Seguridad Alimentaria (ELCSA) y otras herramientas similares en América del Norte, Centroamérica y El Caribe.* FAO/Ohio State University. (available at http://www.foodsec.org/fileadmin/user_upload/eufao-fsi4dm/docs/elcsa_aplicacion.pdf).

Melgar-Quiñonez H., Kaiser L.L., Martin A.C., Metz, D. & Olivares, A. 2003. Inseguridad alimentaria en Latinos de California: observaciones de grupos focales. Salud Publica Mex. 45: 198-205.

Monteiro, C.A., Conde, W.L. & Popkin, B.M. 2004. The burden of disease from undernutrition and overnutrition in countries undergoing the nutrition transition: a view from Brazil. *American Journal of Public Health*, (94):433-4.

National Research Council. 2006. *Food insecurity and hunger in the United States: an assessment of the measure.* Panel to Review the U.S. Department of Agriculture's Measurement of Food Insecurity and Hunger, Gooloo S. Wunderlich and Janet L. Norwood, Editors, Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC, The National Academies Press.

Nord, M. 2012. *Assessing potential technical enhancements to the U.S. Household Food Security Measures*. Technical Bulletin 1936, December. Washington DC, USDA/ERS.

Nord, M. & Hopwood, H. 2007. Recent advances provide improved tools for measuring children's food security. *Journal of Nutrition* 137:533-536.

Ohls, J., Radbill, L. & Schirm, A. 2001 *Household Food Security in the United States, 1995-1997: Technical Issues and Statistical Report.* Washington, DC, USDA/ERS.

Panigassi, G., Segall-Corrêa, A.M. & Marin-León, L., Pérez-Escamilla, R., Maria de Fátima Archanjo Mde Sampaio, & Maranha, LK. 2008. Food insecurity as an indicator of inequity: analysis of a population survey. *Cadernos de Saude Publica* 24(10):2376.

Pérez-Escamilla, R. &Vianna, R.P.T. 2012. Food insecurity and the behavioral and intellectual development of children: a review of the evidence. *Journal of Applied Research on Children: Informing Policy for Children at Risk*: (3): Iss. 1, Article 9. (available at: http://digitalcommons.library.tmc.edu/childrenatrisk/vol3/iss1/9).

Pérez-Escamilla, R., Parás, P. & Vianna, R. 2012. *Food security measurement through public opinion polls: The case of ELCSA-Mexico.* Presented at the International Scientific Symposium on "Food and Nutrition Security Information: From valid measurement to effective decision-making," 17-19 January, Rome, FAO.

Pérez-Escamilla, R., Melgar-Quiñonez, H., Nord, M., Alvarez Uribe, M.C. & Segall-Correa, A.M. 2007. Escala Latinoamericana y Caribeña de Seguridad Alimentaria (ELCSA) [Latinamerican and Caribbean Food Security Scale]. *Perspectivas en Nutrición Humana*, Colombia. 2007 (supplement):117-134.

Pérez-Escamilla, R., Segall-Corrêa, A.M., Kurdian Maranha, L., Sampaio Md Mde, F., Marín-León, L. & Panigassi, G. 2004. An adapted version of the U.S. Department of Agriculture Food Insecurity module is a valid tool for assessing household food insecurity in Campinas, Brazil. *Journal of Nutrition*, 134:1923-8.

Radimer, K.L. 2002. Measurement of household food security in the USA and other industrialized countries. *Public Health Nutrition*, 5(6A) :859-64. (available at http://www.ingentaconnect.com/content/cabi/phn).

Radimer, K.L., Olson, C.M. & Campbell, C.C. 1990 Development of indicators to assess hunger. *Journal of Nutrition,* 120 Suppl (11):1544-8.

Radimer, K.L., Olson, C.M., Greene, J.C., Campbell, C.C. & Habicht, J.P. 1992**.** Understanding hunger and developing indicators to assess it in women and children. *Journal of Nutrition Education* (24) Suppl:36-45.

Rasch, G. 1960. *Probabilistic models for some intelligence and attainment tests.* Danish Institute for Educational Research, Copenhagen.

Ruel, M.T., Haddad, L. & Garrett, J.L. 2001**.** *Rapid urbanization and the challenges of obtaining food and nutrition security*. In: Semba RD, Bloem MW, eds. Nutrition and Health in developing countries. Totowa, NJ, Humana Press, 2001:465-82.

Segall-Corrêa, A.M., Marín-León, L. & Pérez-Escamilla, R. 2010. Escala Brasileira de Medida da Insegurança Alimentar e Nutricional – EBIA: Validação, usos e importância para políticas públicas. In: *Fome Zero. Uma História Brasileira*. Ministério de Desenvolvimento Social e Combate à Fome – Brasil. Vol.3, p. 26-43.

Segall-Corrêa, A.M., Marín-León, L., Sampaio, M.F.A., Panigassi, G. & Pérez-Ecamilla, R. 2007. *Insegurança Alimentar no Brasil: do desenvolvimento do instrumento de medida aos primeiros resultados nacionais*. Ministério do Desenvolvimento Social e Combate à Fome. Avaliação de políticas e programas do MDS.1, v.01, p. 385-407.

Segall-Corrêa, A.M., Pérez-Escamilla, R., Marín-León, L., Yuyama, L., Vianna, R., Coitinho, **D., Sampaio, M., Panigassi, G., Maranha, L.K., Gubert, M., Leão, M.M., Guimarães, L.V., Bergamasco, S., Aier de Oliveira, J & Montanari Corrêa, M.L.** 2008. Evaluation of household insecurity in Brazil: validity assessment in diverse sociocultural settings. En: Memoria del Primer Concurso de Investigaciones REDSAN, Santiago, Chile.

Sen A. 1981. Poverty and Famines: An Essay on Entitlement and Deprivation. Oxford: Clarendon Press.

Swindale, A. & Bilinsky, P. 2006. Development of a universally applicable household food insecurity measurement tool: process, current status, and outstanding issues. *Journal of Nutrition.*136: 1449S-1452S.

United Nations Albania. 2012. *Better diets for health: monitoring dietary diversity in Albania to improve nutritional status.* Tirana.

Webb, P., Coates, J., Frongillo, E.A., Rogers, B.L., Swindale, A. & Bilinsky, P. 2006. Measuring household food insecurity: why it's so important and yet so difficult to do. *Journal of Nutrition* 136: 1404S-1408S.

Wehler, C.A., Scott, R.I. & Anderson, J.J. 1992. The community childhood hunger identification project: a model of domestic hunger - demonstration project in Seattle, Washington. *Journal of Nutrition Education*, 24 Suppl:29-35.

Weiss, C.H. 1983. Ideology, interest, and information: the basis of policy decisions In *Ethics, the Social Sciences, and Policy Analysis*, D. Callahan & B. Jennings (Eds.), pp. 213-245. New York: Plenum.

Wolfe, W.S. & Frongillo, E.A. 2001. Building household food security measurement tools from the ground up. *Food Nutrition Bulletin*, 22(1): 5-12.

Wolfe, W.S., Frongillo, E.A. & Valois, P. 2003. Understanding the experience of food insecurity by elders suggests ways to improve its measurement. *Journal of Nutrition*, 133: 2762-2769.

Wolfe, W.S., Olson, C.M., Kendall, A. & Frongillo, E.A. 1996. Understanding food insecurity in the elderly: a conceptual framework. *Journal of Nutrition Education*, 28(2): 92-100.

Wolfe, W. S., Olson, C.M., Kendal, A. & Frongillo, E.A. 1998. Hunger and food insecurity in the elderly: its nature and measurement. *Journal of Aging and Health* 10(3): 327-350

Wright, B.D. & Linacre, J.M. 1994. Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8:3 p.370.

Contact:
Statistics division (ESS)
The Food and Agriculture
Organization of the United Nations
Viale delle Terme di Caracalla
00153 Rome, Italy

http://www.fao.org/economic/ess/ess-fs/voices/en/

As583/1/10.13