Measuring food security in the U.S. for 25 years:  
History, methods, findings, and opportunities

Irma Arteaga and Parke Wilde

Introduction

For 25 years, the U.S. Department of Agriculture (USDA) has used Census Bureau data to scientifically estimate the annual prevalence of household food insecurity, a measure of food-related hardship (Coleman-Jensen et al., 2020). This anniversary provides an opportunity to reflect on the achievements of the U.S. food security measurement, contemplate lessons learned, and consider potential changes going forward, to better serve the purposes of federal investment in this policy area.

One purpose of federal household food security measurement is to monitor national progress toward reducing food insecurity and hunger (Wilde, 2004b). For example, in the fourth year of reporting, USDA found that 11.8% of U.S. households were food insecure in 1998. That same year, as part of a global initiative at the World Food Summit to set goals for hunger reduction, the United States adopted a target of halving household food insecurity to 6% by 2015 (Andrews and Nord, 2001). Unfortunately, in the most recent USDA report—showcasing the twenty-fifth year of U.S. food security statistics—10.5% of U.S. households still were food insecure in 2019.

A second purpose of federal household food security measurement is to help in evaluating federal nutrition assistance programs that may be expected to improve food security, such as the Supplemental Nutrition Assistance Program (SNAP). These evaluation efforts are challenging, because people who participate in these programs may differ systematically from those who do not. For example, in 2019, among households with income below 130% of the federal poverty line (a gross income threshold used in SNAP eligibility determination), 49.7% of SNAP participants were food insecure and 22.6% of SNAP non-participants were food insecure (Coleman-Jensen et al., 2020). Cross-sectional comparison shows that SNAP participants have comparatively high levels of food insecurity, while research that controls for self-selection into the program finds that SNAP strongly improves household food security (Bartfeld et al., 2015).

As part of a new small grants project and conference, titled “25 Years of Food Security Measurement: Answered Questions and Further Research,” with support from USDA’s Economic Research Service (ERS), this paper sets the stage by providing a brief introduction to relevant literature. Further information and a request for proposals are available at the project website (Tufts University, 2020). The submission due date is Feb 19, 2021. This paper provides some history of federal food security measurement, briefly summarizes selected findings, and reviews a few selected special topics in measurement methods. We aim to include high-profile sources and to mention some more idiosyncratic examples of recent research suggesting potential directions for further development. This project and conference will consider adaptations or modifications that have both research merit and practical suitability, keeping in mind the constraints and purposes of the federal food security measurement program.

For readers who consider applying for a grant through this program, this paper may help save effort in literature review and reduce the fraction of the proposal that must be devoted to background, allowing more space to focus on the proposal’s own content and approach. By
showing what literature we already have been reading, we recognize that this paper may reveal gaps or limitations in our own coverage. That is all for the best. Through this project and the related conference, we look forward to learning about new topics beyond those covered here.

**History of the food security measure**

The USDA food security measure was created after Congress passed the National Nutrition Monitoring and Related Research Act of 1990. USDA’s ten-year comprehensive plan developed under the act recommended a standardized instrument to measure food insecurity in the U.S. and at the state and local levels. Good sources for the history of the USDA household food security measure include Eisinger (1998), National Research Council (2006, chapter 2), Coleman-Jensen (2015), USDA Economic Research Service (2020), and others.

Previously, food-related measures of hardship were based on anthropometrics (wasting or stunting), household income (poverty), growth of emergency food services (food banks and food pantries), or a simple food insufficiency survey question, none of which met the need for a representative validated survey-based measure of food security. The USDA food sufficiency question predates the official household food security measure. It has remained as a screener question on federal surveys and has received increased attention in 2020 during the COVID-19 pandemic due to its inclusion in the Census Bureau’s high-frequency Household Pulse Survey, discussed in the next section (Methods).

In January 1994, the Food and Nutrition Service (FNS) of USDA and the National Center for Health Statistics of the U.S. Department of Health and Human Services (DHHS) sponsored the National Conference on Food Security Measurement and Research. Participants at that conference were academic experts that led nutrition and hunger projects, private researchers from Abt Associates, Mathematica Policy Research and similar organizations, as well as staff from federal agencies that studied and collected data related to food insecurity and insufficiency. The participants conceptually defined food security and recommended an approach to scientifically measure food insecurity using a national survey.

In April 1995, the first nationally representative food insecurity survey was administered by the U.S. Census Bureau, via a supplemental module to the CPS. Abt Associates analyzed the data from the 1995 survey; Mathematica Policy Research had a contract with FNS to determine the stability of the food security measure for the period 1995-1997. Thereafter, CPS annually collected food security data as part of a supplemental module. The 1995 technical report led by Abt Associates found that scaling techniques established three thresholds and four food security categories: food secure, food insecure with hunger not evident, food insecure with moderate hunger, and food insecure with severe hunger.

One group of questions in the first CPS Food Security Supplement addressed coping or resource augmentation, including actions that households might take to deal with scarce food resources. Examples of these questions are sending children to eat at a friend’s, putting off paying other bills in order to buy food, or obtaining meals from soup kitchens or food banks. The report concluded that these questions did not meet the statistical criteria for inclusion in the scale (Hamilton & Cook, 1997).
The wording of food security questions has not changed much since 1995. From the beginning, CPS had a battery of 18 questions to measure food insecurity for households with children and 10 questions to measure food insecurity for households without children (CPS-FSS). In 1998, ERS assumed responsibility for analyzing food insecurity measures, as well as sponsorship of the annual household food security survey module in the Current Population Survey. Since then, small changes have occurred, mainly related to the screening questions and ordering of items rather than the actual content of the questionnaire. USDA has conducted “split ballot” tests of variations in the wording of some questions, including child-referenced items and a “balanced” meals question. In addition, shorter 6-item and 2-item survey instruments were developed for particular uses, discussed in the Methods section.

In 2006, USDA consulted with the Committee on National Statistics of the National Academies on a high-profile review: Food Insecurity and Hunger in the United States: An Assessment of the Measure. Based on one recommendation from that report, the labels of all four food security status categories were changed to: (1) high food security, (2) marginal food security, (3) low food security and (4) very low food security. The first two categories indicate food security, while the last two indicate food insecurity. The term “hunger” was removed from the category labels. The expert committee explained that hunger might result from food insecurity, and measuring hunger required detailed and extensive information on physiological experiences of each household member.

Following another recommendation from the National Academies review, the ordering of items was changed for households with children, so that all adult-referenced items would precede all child-referenced items. Engelhard et al. (2018) estimate measures of statistical fit for food security items, with attention to the order in which questions were asked. The National Academies review made other recommendations about the statistical models used to assess household food security measures, as will be discussed further in the Methods section.

The USDA questions are used on most U.S. national surveys, including but not limited to CPS, Panel Study of Income Dynamics (PSID), National Longitudinal Survey of Youth (NLSY), Early Childhood Longitudinal Survey Birth Cohort (ECLS-B) and Kindergarten Cohort (ECLS-K), National Health Interview Survey (NHIS), National Health and Nutrition Examination Survey (NHANES), and the Medical Expenditure Panel Survey (MEPS) (USDA Economic Research Service, 2020).

The use of sound statistical methods to validate the USDA food security instrument and scale has motivated other countries and organizations to use the survey in its current form (e.g. Young Lives), or to adapt it (Ibok et al., 2014; Rafiei, Nord, Sadeghizadeh & Entezari, 2009; Vargas & Penny, 2010). Moreover, the USDA food security instrument has influenced the way we measure food security around the world. For example, in 2014 the United Nations, Food and Agriculture Organization’s (FAO) Voices of the Hungry project developed a measure of food insecurity, the Food Insecurity Experience Scale (FIES). It was modeled after the USDA food security survey module and consists of eight questions. FIES is a standardized measure of individuals’ direct experiences of food insecurity and is used to compare food insecurity around the globe (Cafiero, Viviani, and Nord, 2018). It is currently used by 146 countries.

Methods
The methods for federal household food security measurement are described in multiple sources, including USDA’s annual food security report series (Coleman-Jensen et al., 2020), ERS web pages on food security in the United States (USDA ERS, 2020), a guide to measuring household food security (Bickel et al., 2000), the National Academies report *Food Insecurity in the United States: An Assessment of the Measure* (National Research Council, 2006), and an ERS technical report in response to the National Academies report (Nord, 2012b). Reviewing these sources on measurement methods may suggest ideas for modification or improvement in the future. This section discusses conceptual definitions, the food security survey items and category thresholds that determine federal food insecurity prevalence statistics in the past 12 months, an alternate measure with 30-day in place of 12-month reference period, alternate measures for food insecurity in households with children, and item response theory (IRT) models used for several statistical purposes.

**Conceptual definitions**

USDA defines household food security as having “access at all times to enough food for an active, healthy life for all household members.” Households that are not food secure may be called “food insecure.” These food-insecure households may be subclassified as having “low” or “very low” food security. USDA defines very low food security as “the more severe range of food insecurity where one or more household members experienced reduced food intake and disrupted eating patterns at times during the year because of limited money and other resources for obtaining food” (Coleman-Jensen et al., 2020).

In operational practice, the classifications used for estimating food security prevalence using survey data are related to, but not precisely the same as, the conceptual definitions (Wilde, 2004b). USDA notes, “This operational measure does not specifically address whether the household members’ food intake was sufficient for active, healthy lives—the conceptual definition of food security” (Coleman-Jensen et al., 2020).

**Food security items and category thresholds**

The empirical categories are based on the number of affirmative responses to survey questions, so it is useful to become familiar with the specific survey items and their response frequencies. The food security classification system used in the high-profile statistics for monitoring national progress, cited previously in the introduction, is based on 18 survey items (Coleman-Jensen et al., 2020, p. 5). These 18 items are referred to as the U.S. Household Food Security Survey Module.

Ten of the items use survey questions asked of all households. The first 3 most frequently affirmed of these 10 items are “household items” (percent affirmative among all households in 2019 from the report’s statistical supplement in parentheses):

- Worried food would run out before (I/we) got money to buy more (13.9%),
- Food bought didn’t last and (I/we) didn’t have money to get more (11.4%),
- Couldn’t afford to eat balanced meals (11.8%).
The remaining 7 of the 10 items asked of all households are “adult items,” describing the experience of the respondent or other adults in the household:

- Adult(s) cut size of meals or skipped meals (6.4%)
- Respondent ate less than felt he/she should (6.7%)
- Adult(s) cut size or skipped meals in 3 or more months (4.9%)
- Respondent hungry but didn't eat because couldn't afford (3.5%)
- Respondent lost weight (2.2%)
- Adult(s) did not eat for whole day (1.5%)
- Adult(s) did not eat for whole day in 3 or more months (1.2%)

The final 8 of the 18 items are “child items,” based on questions asked only of households with children (percent affirmative among all households with children in 2019 in parentheses):

- Relied on few kinds of low-cost food to feed child(ren) (11.3%)
- Couldn't feed child(ren) balanced meals (6.8%)
- Child(ren) were not eating enough (2.9%)
- Cut size of child(ren)'s meals (1.2%)
- Child(ren) were hungry (0.8%)
- Child(ren) skipped meals (0.5%)
- Child(ren) skipped meals in 3 or more months (0.4%)
- Child(ren) did not eat for whole day (0.2%; Coleman-Jensen et al., 2020b)

For most of these items, the survey asks whether the household experienced the hardship never, sometimes, or often. The current USDA approach makes a binary distinction between whether the item is affirmed (sometimes or always experienced) or not affirmed (never experienced). At times, as discussed later in this section, researchers have explored using statistical models that treat the survey questions as “polytomous,” with 3 separate classifications for never, sometimes, or often.

Food security modules in federal surveys are administered with internal screeners: any household that does not affirm the initial screeners or one of the first three items is not administered any of the remaining (household) items and those items are counted as non-affirmations. Similarly, for any household administered the fourth through the eighth item, any household that does not affirm at least one of those is not administered the final two. Similar screeners are present for households with children. Initial screeners are also used, including an income screener. This mode of administration operationalizes an important assumption of the underlying model:
namely, that households that affirm items indicating severe hardship should also have affirmed some subset of less-severe items.

The empirical food security categories use a “raw score,” the number of affirmative responses to these survey items. In some cases, the threshold for a particular category differs for households with and without children. As noted in the annual report:

Households without children are classified as food insecure if they report 3 or more indications of food insecurity in response to the first 10 questions; they are classified as having very low food security if they report 6 or more food-insecure conditions out of the first 10 questions. Households with children are classified as food insecure if they report 3 or more indications of food insecurity in response to the entire set of 18 questions; they are classified as having very low food security if they report 8 or more food-insecure conditions in response to the entire set of 18 questions (Coleman-Jensen et al., 2020).

The intent of USDA’s current approach is to have the categories represent approximately equal levels of severity in households with and without children, so that a single food security status classification can be used for all households.

*Food security status for households with and without children*

USDA researchers have explored separate classifications for households with and without children, as an alternate experimental approach (Coleman-Jensen, Rabbitt, and Gregory, 2017; Nord and Coleman-Jensen, 2014). Adult food security status may be determined by responses to the 10 household- and adult-referenced items, while child food security status may be determined by response to the 8 child-referenced items.

Under this alternate approach, a question arises about how to classify the overall food security status of the household. In households with children, the household could be considered food insecure if either adults or children or both were classified as food insecure. Or, one could simply classify household food security status based on the 10 household- and adult-referenced items in all households. USDA research into experimental food-security classification methods discusses indications of internal and external validity, with some strengths and weaknesses for each option (Coleman-Jensen, Rabbitt, and Gregory, 2017; Nord and Coleman-Jensen, 2014).

*12-month and 30-day measures*

The food security measures described above reference the previous 12 months. Because participants enter and exit federal nutrition assistance programs from month to month, an alternative measure with a shorter reference time period is more useful for program evaluation purposes. The statistical supplement to USDA’s annual report provides estimates of food security during the 30 days prior to interview, in households with and without participation in federal nutrition assistance programs (Coleman-Jensen et al., 2020). This measure has been validated. The hardship described in a survey item is more likely to have occurred in the past 12 months than in the past 30 days.

*Shorter instruments*

Shorter instruments for food security measurement, including a 2-item and a 6-item survey, have been used to assess household food insecurity (Hager, 2010). Both versions have been validated.
as screeners (Blumberg, 1999). The advantages of these short instruments are the decreased respondent burden and minimal bias compared to longer versions. The disadvantages are that they do not measure the most severe levels of food insecurity, they are less precise than the 18-item instrument, and they do not ask questions regarding children in the household. Shorter 2-item instruments may also be more feasible in clinical healthcare settings.

For the 2-item survey, the Hunger Vital Signs, the following two questions have been used:

- “We worried whether our food would run out before we got money to buy more.” Was that often true, sometimes true, or never true for your household in the last 12 months?
- “The food we bought just didn't last and we didn't have money to get more.” Was that often true, sometimes true, or never true for your household in the last 12 months?”

At least one alternative question would be needed for a 2-item survey that aimed at assessing hunger.

The 6-item short form of the food security survey module has had minor technical improvements to wording since the 1995 module. Blumberg et. al. (1999) found that the 6-item survey was useful not only in a clinical setting but had applications for national survey use as well.

The Household Pulse Survey administered during the COVID-19 pandemic is a 19-item questionnaire assessing employment status, food sufficiency, housing security, education disruptions, and physical and mental wellbeing (Ziliak, 2020). Food-related hardship is assessed using several questions, including a version of the long-standing USDA food sufficiency question:

- Getting enough food can also be a problem for some people. Which of these statements best describes the food eaten in your household?
  - Enough of the kinds of food (I/we) wanted to eat
  - Enough, but not always the kinds of food (I/we) wanted to eat
  - Sometimes not enough to eat
  - Often not enough to eat

The current Household Pulse Survey provides a unique example of a shorter household food sufficiency measurement instrument, with a short 7-day reference period, incorporated into a larger more general survey, with relatively specific questions in the context of the pandemic and the need for recent data in order to feel the “pulse” of the nation during a national emergency (Ziliak, 2020).

**Item Response Theory and scaling methods**

Some decisions in federal measurement methods, including decisions about the raw score thresholds for food security status classification, draw on scaling tools from item response theory (IRT), a field of statistics with long-standing applications in educational testing and psychometrics. The National Academies report mentioned above describes several potential IRT methods, including the Rasch Model, which was originally used to develop the federal food security measurement approach, and several alternatives (National Research Council, 2006).
The Rasch model is akin to a logistic regression model showing the log-odds of affirmative response to the survey items as a function of the “severity” of the item and a food security scale score for the household. With the Rasch model, the raw score (the number of affirmative responses, an integer variable) is a “sufficient statistic” for the underlying food security scale score (a continuous variable). This means one can use the raw score to determine the food security scale score without needing other information about which questions were answered affirmatively. The Rasch model is unidimensional, which means that the survey items are taken to provide information about a single underlying latent variable. USDA chose to develop a unidimensional measure of food security that captures the central dimension of food security, food sufficiency. As a result, respondents to the household food security survey model are assumed to formulate their responses to the food-security questions based on the food available to them relative to their subjective needs (Bickel et al., 2000).

Because the raw score is a sufficient statistic for the scaled Rasch score, the official food security status classifications are based on raw score thresholds -- which are straightforward to interpret -- rather than direct estimates from an IRT model. In this sense, the IRT models do not directly determine household food security prevalence estimates. The IRT models do help determine what thresholds should count as equivalent for households that are asked different questions, most notably households with and without children (Wilde, 2004b), and these models are heavily used in validation research.

Previous research has found that Rasch model estimates for the severity of survey items differ in households with and without children, which violates one of the model’s technical assumptions and complicates efforts to use the Rasch model for setting equivalent raw score thresholds among these groups (Wilde, 2004a; National Research Council, 2006; Nord and Coleman-Jensen, 2014; Coleman-Jensen, Rabbitt, and Gregory, 2017). Formally, the Rasch model rules out differential item functioning (DIF), which is differences across populations in the frequency of affirming items when the underlying food security score is held constant. Noting other evidence of this type of differential item functioning with respect to several household characteristics, the National Academies report investigated more general statistical models in the IRT family (National Research Council, 2006).

However, these alternate IRT models each present additional difficulties in estimation and ease of explanation to policy-makers. For example, the alternatives do not permit one to estimate an underlying scale score directly from the raw score as the Rasch model does. Nord (2012b) reviewed multiple technical suggestions from the National Academies, finding in most cases that they would provide only minor statistical benefits and would not greatly influence food security prevalence estimates. For example, the most complex model he explored would make three changes: (a) two coefficients would be used to describe the severity of each survey item (instead of one coefficient per item in the Rasch model), (b) the survey questions would be treated as polytomous with 3 categories for never, sometimes, or often experiencing hardship (instead of a binary outcome in the current approach), and (c) household food security scores would be assigned probabilistically (instead of determined based on the raw score). After estimating food security models with and without these changes, Nord concluded: “The extent of improvement in precision that might be realized is not likely to justify the loss of transparency, simplicity, and communicability that would result from use of the more complex model” (Nord, 2012b). The alternatives were not adopted for official use.
Findings

Since 1995, food security data have been consistently and scientifically collected in the US every year. USDA’s annual household food security report presents many findings about the prevalence and severity of household food insecurity by year and the correlates of food insecurity, such as income, race/ethnicity, region of the country, and household size (Coleman-Jensen, Rabbitt, Gregory & Singh, 2020).

Beyond descriptive studies, food security research has addressed important policy questions. For example studies on the effects of the American Recovery and Reinvestment Act on food security (Nord and Prell, 2011) and specifically on the changes of SNAP benefit sizes on food security and other health outcomes (Kim, Rabbitt & Tuttle, 2019; Kim, 2016) have guided the federal response to the COVID-19 pandemic that included over $2 billion per month in emergency SNAP supplements to take all participating households up to the maximum amount. Moreover, with the majority of schools closed in 2020 due to the pandemic, SNAP has also been used to distribute money to families to replace meals that children would have received through the National School Lunch Program (NSLP). Such studies build on decades of research into the effects of food assistance programs on food security and health outcomes by academic researchers, health experts and physicians (Frank, Bruce & Ochoa, 2020). Prior literature has studied the effects of safety net programs on food insecurity (Schmidt, Shore-Sheppard & Watson, 2016) including the effects of SNAP (Yen et al., 2008; Depolt et al., 2009; Mykerezi and Mills, 2010; Gundersen and Krieder, 2008; Gundersen, Kreider, Pepper, 2017; Kreider et al., 2012; Bitler, 2016; Gregory, Rabbitt & Ribar, 2015; Gundersen, Kreider & Pepper, 2018; Nord, 2012; Ratcliffe, McKernan & Zhang, 2011), NSLP (Huang & Barnidge, 2016; Kang & Moffitt, 2019; Nord & Romig, 2006), and the Special Supplemental Nutrition Program for Women, Infants and Children -WIC (Arteaga, Hefflin & Gable, 2016; Kreider, Pepper and Roy, 2016;). SNAP participation is found to mitigate the negative association between food insecurity and health outcomes (Pak & Kim, 2020). In addition, research suggests that states that require schools to offer school breakfast programs have reduced food security for young children (Fletcher & Frisvold, 2017). This body of research points to the importance of policy action such as increase take-up rates for WIC and SNAP, and to making changes to the recertification process in terms of administration and length (Hefflin, Hodges and Ojinnaka, 2020).

Research has also examined food security for vulnerable populations such as elderly, immigrants, children, and more recently women. In the case of seniors, food security is associated with adverse health outcomes including health status, self-reported health, depression, anxiety, and disabilities (Pak & Kim, 2020; Steiner et al., 2018; Vilar-Compte, Gaitán-Rossi & Pérez-Escamilla, 2017; Hefflin, Altman, & Rodriguez, 2019; Coleman-Jensen and Nord, 2013; Coleman-Jensen, 2020). Research on food insecurity and immigrants have focused on Hispanic/Latino subjects (Rabbitt, Smith & Coleman-Jensen, 2016). Food insecurity is associated with weight increase with longer US residence for immigrant women (Ryan-Ibarra, et al., 2017). Relative to Hispanic natives, Hispanic immigrant households and children are more likely to be food insecure (Arteaga, Potochnick & Parsons, 2017; Chilton et al., 2009; Kalil & Chen, 2008). While SNAP have been found to be a protective factor for food insecurity, with the passage of the Personal Responsibility, Work, and Reconciliation Act, non-citizens became ineligible for food assistance programs and mixed families reduced their allotment of SNAP, which
contributed to the increase of food security (Van Hook & Balistreri, 2006). Recent high levels of food insecurity among immigrants can be explained by level of acculturation, and economic stress (Potochnick et al., 2019), maternal education, and household size (Kalil & Chen, 2008), as well as the great recession, language barriers to get access or keep access to food assistance programs, and increasing anti-immigrant sentiment (Perreira & Pedroza, 2019).

Just recently research in the US has focused on examining the relationship between gender and sexuality, food insecurity and health outcomes and found that women of color, elderly women and the LGBTQ+ population experience disproportional rates of food insecurity (Hernandez et al., 2017). Key variables that explain food insecurity for women are domestic violence (Chilton et al, 201), and the way they shop and make their meals (Foster et al., 2019). In addition, social capital and geographical location explain food security experiences for elderly women (Leddy et. al, 2020; Lane et. al, 2014). Transgender and gender nonconforming people experience high rates of food insecurity largely due to lack of steady employment opportunities and layered discrimination (Russomanno et. al, 2019).

Cognitive testing

Cognitive interviews are widely used to pretest survey instruments by identifying comprehension and response problems (Collins, 2003). This approach is used to create new survey questions and to understand and interpret existing survey instruments. For the original USDA food security module, researchers pretested survey items (Alaimo, 1997; Hamilton & Cook, 1997). A decade later, researchers also used cognitive testing when improving their children’s food security measure (Nord & Hopwood, 2007), and in 2019 USDA’s Economic Research Service contracted with Census to conduct cognitive testing of the entire CPS Food Security Supplement questionnaire (Office of Management and Budget, undated).

Several items in the food security module have been a special focus for cognitive testing. “Running out of food” and “balanced meals” might not be clearly understood or might have different meaning for different populations (Alaimo, Olson & Frongillo, 1999). Studies from Hawaii, India, and Indonesia show that the term “balanced meals” depend on cultural contexts. Suggested variations of that phrase based on cognitive testing are: “healthy and varied diet” or “nutritious meal” (Derrickson, Sakai, & Anderson, 2001; Sethi, Maitra, Avula, Unisa, & Bhalla, 2017; Studdert, Frongillo Jr, & Valois, 2001). Cognitive testing may suggest alternative answers when questions have multiple choices. Cognitive testing in the US with Latino immigrant mothers suggests that the option “hardly ever” should be used instead of “never” (Kuyper et al., 2006).

There has been limited cognitive testing of the U.S. food security module in languages other than English. About 42 million people five and older speak Spanish at home in the U.S. Chinese languages (Mandarin, Cantonese, and others) are spoken by about 3 million people. While formal cognitive testing has not been conducted for the Spanish version of the CPS-FSS food security items, researchers at USDA-ERS have validated the Spanish version using Rasch analysis (Rabbitt & Coleman-Jensen, 2017). In the past, a lack of standardized Spanish translation led to the use of different Spanish variants of the CPS-FSS module For example, the “balanced meals” item was translated variously as “comida balanceada” and “comida nutritiva” (Harrison, Stormer, Herman, & Winham, 2003), while the new USDA standardized Spanish
module now uses “comida variada y nutritiva” (Rabbitt & Coleman-Jensen, 2017). However, for the Chinese version of the CPS-FSS module, no cognitive testing or Rasch analysis has been conducted to date. In Chinese no equivalent words are found for “balanced meals” or “eat less than you feel” (Kuyper et al., 2006; Kwan, Napoles, Chou, & Seligman, 2015).

Among areas for future research, formal cognitive testing of the Spanish and Chinese versions of the USDA food security instrument can use Hispanic and Chinese subjects from different parts of the country and who speak different dialects. There is limited research on whether a shorter version of items is more convenient for certain populations. Harrison (2003) and Connell et al. (2004) found that low-income Spanish speakers and children prefer shorter versions of the items because they can easily understand them. Cognitive testing can also be used to examine the appropriate age when a child can start reporting their own food security.

**Differences in household food security by different subgroup populations**

The prevalence of household food insecurity varies greatly across subgroup populations (Coleman-Jensen, Rabbitt, Gregory & Singh, 2020). One reason is that some populations experience higher poverty rates and greater symptoms of food-related hardship than others do. Another possible reason is that response frequencies to food security items might vary across subgroup populations even when food insecurity is held constant. This possibility has been the focus of a substantial body of validation research.

While early validation of the USDA Household Food Security Measure (HFSM) and scaling found similar patterns of responses by race/ethnicity, income and household composition (Kendall, Olson & Frongillo, 1995; Frongillo, 1999), similar analysis was not conducted by gender groups. A few years later, using CPS data from 1995, 1997 and 1999 and a generalized linear model, researchers tested the validity of the USDA food security measure. Although their overall results validated the HFSM, the study also found evidence that responses vary among subpopulation groups. Specifically, their findings suggested that minority respondents, Spanish speaking respondents, male respondents and those living in metropolitan areas respond to questions differently than other households (Opsomer, Jensen & Pan, 2003). Research by Mark Nord (2012b), described earlier in the Methods section, noted some evidence of differential item functioning (DIF), but found that it may not strongly affect food security prevalence estimates. Engelhard, Rabbitt and Engelhard (2018) tested the item fit and overall fit of the food security measure using household model-data fit. Findings suggests that higher levels of misfit are predicted by gender, educational attainment, Spanish as the only language for adults, and participation in SNAP in the last 12 months. The percentage of one-parent only male-headed householders, the percentage of Hispanic households, and the percentage of people living in metropolitan areas have increased in 4, 6, and 6 percentage points respectively from 1995 to 2019 (US Census Bureau, 1995, 1998, 2019), so this issue may grow in importance.

When response frequencies vary across populations, even holding constant the level of food insecurity, the validation literature sometimes has interpreted these findings as evidence that different populations understand the meaning of the questions differently (for example, they might have different understandings of “balanced meals”). However, another possibility is that different populations actually experience food-related hardships with different relative frequencies in the real world (for example, one population may be comparatively more likely to
have difficulty acquiring balanced meals and comparatively less worry about running out of food). One of the assumptions of the Rasch model is that response frequencies are the same across populations when the food security score is held constant, so this validation research indicates a potential difficulty for either the validity of the questions, or the Rasch model, or both.

Interestingly, studies from other countries suggest different responses to food security questions for different subpopulations. A study from Canada suggests differences in response frequencies in similar married or cohabiting households, where women tend to report higher levels of food insecurity than men (Matheson & McIntyre, 2014). A more recent study that explored gender differences in food insecurity globally using a modified Oaxaca-Blinder decomposition gap approach found that a nontrivial share of the gender gap remained unexplained; the authors suggest that the willingness to report food security experiences might vary by gender (Broussard, 2019). Similar gender differences examining the United Nations food insecurity experience scale (FIES) with a multilevel explanatory Rasch model was found by Wang et al. (2020).

Prior studies have also explored differences by geographic areas and race and ethnicity. For example, Quandt et al. (2001) examined food insecurity for rural elders in North Carolina using quantitative and qualitative approaches and their findings suggest that self-sufficiency and pride might lead them to underreport their levels of food security.

There is a consensus among researchers who study food insecurity in indigenous populations that by using the USDA food security module, there is underreport of the levels of food insecurity for the indigenous population (Sowerwine et al., 2019). Indigenous populations require specific household food security measurement methods, which have increased in use in the United States and globally in the last two decades. Researchers have noted that such measures should include unique cultural and traditional aspects of food consumption, agricultural production, and community-developed tools (Lambden, Receveur, & Kuhnlein, 2007; Power, 2008; Sowerwine et al., 2019).

Conclusions

This review provides just an introduction to the lively and energetic body of research on household food security measurement in the United States, while also communicating a sense of how much remains to do. For this project on “Measuring Food Security in the U.S. for 25 Years,” the RFP describes a dozen areas of potential future work, but these should be taken as merely a rough indication of the types of proposals that are relevant.

1. Development and assessment of a short food security measure and/or a discussion of available short alternative food adequacy measures and their strengths and drawbacks.
2. Cognitive testing of existing and potential household food security items, with implications for expanding our understanding of food insecurity.
3. Practical approaches to scaling or item response theory (IRT) models, with opportunities for methodological improvements that prioritize ease of communication.
4. Improvements in methods for using food security measures in evaluating federal nutrition assistance programs, noting the distinction between 12-month and 30-day instruments.
5. Implications of using the existing Household Food Security Survey Module, either the 30-day or 12-month measure as the main outcome in interventions, especially when those interventions target specific individuals in households (e.g., interventions targeting children).

6. Strengths, limitations, and opportunities for using food security measures in national goal setting and progress monitoring.

7. An assessment of USDA’s food security data investments including identifying gaps and opportunities and implications for research.

8. Qualitative research on food security instruments that allows low-income Americans to express their experiences of food insecurity and hunger in their own voices, with a consideration of how these experiences relate to existing food security measures.

9. Economic models of household food spending and food consumption decision-making, with a consideration of implications for food security research and measurement.

10. Differences in household food security by region, geography, income strata, race, ethnicity, national origin, age, and/or gender, with implications for social justice, diversity, equity, and inclusion. Such work should include a consideration of how issues of diversity, equity and inclusion affect food security measurement, research and data collections.

11. Nutrition science and public health aspects of the association between food security and health outcomes, with implications for food security measurement methods.

12. Psychological connections between stress, depression, and household food security, with implications for understanding measurement effects on household food security.

A running theme is that we are interested in research with potential constructive lessons for the federal government’s food security measurement methods. This focus has a couple implications. First, beyond the literature reviewed here, there is an even larger body of work on programs and policies that reduce household food insecurity and hunger; our attention here is on measurement as a stepping stone. Second, there are many opportunities for increasingly complex instruments and analytical methods with considerable research merit; our attention here is on methods that one could explain concisely to a reasonably capable policy-maker. At this 25th anniversary of this area of work, we are enthusiastic about the next stage of conversation, imagination, and progress.

Acknowledgments

The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or US government determination or policy. We are grateful to Friedman School graduate students Sarah Wiemert and Ian Moore for literature abstraction, editing, and contributions to our discussion of short survey instruments, differences by gender, and differences for indigenous people.
References


Pak, T. Y., & Kim, G. (2020). Food stamps, food insecurity, and health outcomes among elderly Americans. Preventive Medicine, 130, 105871.


United States Census Bureau (2019). Table H1: Households by type, age and tenure of householder for selected characteristics 2019. Retrieved from:


