

Food and Agriculture Organization of the United Nations





Using the FIES App

A simple tool for the analysis of Food Insecurity Experience Scale data

Contents

| The FIES App: a User-Friendly approach to FIES data analysis | p. 1 |
|--|--|
| Section 1: Upload the data | |
| Find the app online Prepare your database Upload your data | p. 2 p. 2 p. 3 |
| Section 2: Analyze the data | |
| 4. Sample size 5. Item table 6. Respondent/Raw score parameters 7. Statistical validation 8. Select items for analysis 9. Equating 10. Differential Item Functioning (DIF) 11. Prevalence rates | p. 4 p. 4 p. 7 p. 9 p.14 p.14 p.19 p.19 |
| 12. Additional information for analysis | p.20 |

The FIES App: a user-friendly approach to FIES data analysis.

This manual will guide you through the analysis of Food Insecurity Experience Scale (FIES) data using an online open access application designed to facilitate this step. Additional information is included to help you interpret the output provided by the tool and to inform several decisions you will make when analyzing your FIES data. The app's features are divided into several tabs, with each containing a different step in the analysis.

The final result will be two indicators (prevalence rates) of food insecurity at different levels of severity.

FI_{mod-sev} The proportion of the population experiencing moderate to severe food insecurity

FI_{sev} The proportion of the population experiencing severe food insecurity

FI_{mod+sev} is SDG indicator 2.1.2: the "prevalence of moderate or severe food insecurity in the population, based on the Food Insecurity Experience Scale (FIES)".

For more on the uses of the FIES, how to collect data and how to present results, please refer to the e-learning course: "Using the Food Insecurity Experience Scale."

This document is divided in 2 macro sections indicated in the upper part of the page with a label:



It is designed to guide you step by step through the analysis. Tips are provided to help you investigate potential problems with the data, while boxes are there to clarify key concepts.

The Voices of the Hungry website <u>(link)</u> contains additional resources for using the FIES and reporting results.

For assistance using the app or for questions about the FIES, please contact fies-help@fao.org.

Now... let's get started.

1. Find the app online

Go to https://fies.shinyapps.io/ExtendedApp/

For output to appear, you will need to upload data.

2. Prepare your database

For the app to work, the database **must** be in **CSV format**, with the eight FIES items in the first eight columns, **in the order shown in Table 1**, with "Yes" responses coded as 1, "No" responses coded as 0, and "Don't Know" or "Refused" responses coded as "NA".

Household sampling weights ("wt") appear in the ninth column, and individual sampling weights ("wt. person") in the tenth column. Urban/rural area of residence appears in the eleventh column.

Region will appear in the twelfth column. Sampling weights (household or individual), as well as variables for area of residence or region will appear as "NA" for all cases if not available.

Urban/rural area of residence and Region can be coded as numbers, like in the example, or text indicating the area (for example "Urban" or "Rural").

Note that, for analytic purposes, cases with at least one missing response to the FIES items are automatically excluded from the analysis by the app.

| Worried | Healthy | Fewfood | Skipped | Ateless | Runout | Hungry | Whlday | WT | wt.person | Urban | Region |
|---------|---------|---------|---------|---------|--------|--------|--------|-------------|-----------|-------|--------|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.993458037 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.403475591 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2.297742844 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.124525988 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.681710492 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.425003597 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.798622261 | NA | 1 | 1 |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1.285738211 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.439688828 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 3.334660187 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2.137505395 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.197933391 | NA | 1 | 1 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.798622261 | NA | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0.740023222 | NA | 1 | 1 |

Table 1



What do you do with missing responses?

What are the characteristics of the cases with missing responses? If cases with missing responses are concentrated in **one sub-population** (e.g. language, gender, or culture), extra attention should be paid to translation and linguistic adaption or gender matching of respondent and interviewer in future rounds of data collection.

Did the missing responses come primarily from a particular item? If an item has **more than 10 percent** missing responses, this may indicate that it was difficult to understand, leading to a decision to **drop it from the scale**. This information can be found in the column "N_miss__if_ any_valid" in the Item Table described in part five below.

Remember that doing so will affect the criteria for what is considered a "complete" case; respondents who answered all of the seven remaining items will then be "complete."

3. Upload your data

Once your database has been prepared in the correct format, click on "Browse" to select the file and upload it to the app.

| Upload FIES | data |
|-------------|-----------------|
| Browse | Trial_Data.csv |
| | Upload complete |

Now that your data file has been uploaded, we can examine the output that should appear automatically...

4. Sample size

Start with the **"Item and Raw Score Stat"** tab. The first section contains information on your sample. In the example below, although we started with 1000 cases, only 423 are complete (meaning they had no missing responses) and have a non-extreme **raw score** (meaning they had between one and seven "yes" answers to the FIES questions). The last row shows you how many were excluded due to missing responses.

| Number of cases (N) | |
|------------------------|------|
| Total N | 1000 |
| N complete non-extreme | 423 |
| N any missing | 7 |

Please remember that...

• Cases with missing responses are those that do not have either a YES or NO response to all eight of the FIES questions. Cases with any missing responses are removed from the analytical sample and **not used for the calculation of prevalence estimates**.

• The **raw score** is the sum of YES responses to the FIES questions. Cases with **extreme** raw scores are those who have answered NO to all eight questions (and have a raw score of zero) or those who have answered YES to all eight questions (and have a raw score of eight). These **will not be used for the validation** of FIES data described in the following section but they **will be included in the calculation of food insecurity prevalence estimates later on**.

5. Item table

Next you will find information on each of the FIES questions (i.e. items). The **item parameters** indicate the relative severity of each item in this specific application of the FIES, along with their standard errors.

Item parameters are estimated based on the overall **pattern of responses** given by all respondents. A question representing a **less severe experience will have a smaller item parameter** value, whereas a question representing a **more severe experience will have a larger item parameter value**.

> "Item" refers to question. "Respondent" refers to household or individual.

| Item statistics | | | | | | | |
|-----------------|----------|-------|-------|-----------|--------|-------------|----------------|
| | Severity | S.E. | Infit | S.E.Infit | Outfit | N_Yes_Compl | N_if_any_valid |
| Worried | -0.492 | 0.131 | 1.110 | 0.071 | 1.168 | 302.000 | 0 |
| Healthy | -0.365 | 0.129 | 1.164 | 0.069 | 1.356 | 292.000 | 0 |
| Fewfood | -0.997 | 0.141 | 0.837 | 0.079 | 0.621 | 335.000 | 0 |
| Skipped | 0.223 | 0.121 | 0.966 | 0.061 | 0.883 | 248.000 | 0 |
| AteLess | -0.702 | 0.135 | 0.943 | 0.074 | 0.972 | 316.000 | 0 |
| RunOut | 0.124 | 0.122 | 0.900 | 0.062 | 0.808 | 254.000 | 0 |
| Hungry | 0.537 | 0.118 | 1.030 | 0.057 | 1.050 | 229.000 | 0 |
| WholeDay | 1.673 | 0.120 | 1.041 | 0.050 | 1.139 | 119.000 | 0 |
| | | |) | | | | |

Table 2

The relative severity of the items is determined based on the understanding that **the more severe an item is, the less likely respondents are to report it**. In the example (as shown in Table 2), WholeDay has the highest severity, meaning that, of all the items, it received the fewest number of "yes" responses.

Now let's look at two more pieces of information about the items (see Table 3). N_Yes_Compl refers to the number of affirmative responses to this item among "complete" cases (those who have answered either YES or NO to each of the eight FIES items). The column N_miss_if_any_valid shows the number of cases where this item is missing, among cases that have valid answers to at least one of the FIES items.

The center columns (Infit, S.E.Infit, and Outfit) are item fit statistics you will learn more about in section seven.

| em statistics | | | | | | | |
|---------------|----------|-------|-------|-----------|--------|-------------|----------------|
| | Severity | S.E. | Infit | S.E.Infit | Outfit | N_Yes_Compl | N_if_any_valid |
| Worried | -0.492 | 0.131 | 1.110 | 0.071 | 1.168 | 302.000 | 0 |
| Healthy | -0.365 | 0.129 | 1.164 | 0.069 | 1.356 | 292.000 | 0 |
| Fewfood | -0.997 | 0.141 | 0.837 | 0.079 | 0.621 | 335.000 | 0 |
| Skipped | 0.223 | 0.121 | 0.966 | 0.061 | 0.883 | 248.000 | 0 |
| AteLess | -0.702 | 0.135 | 0.943 | 0.074 | 0.972 | 316.000 | 0 |
| RunOut | 0.124 | 0.122 | 0.900 | 0.062 | 0.808 | 254.000 | 0 |
| Hungry | 0.537 | 0.118 | 1.030 | 0.057 | 1.050 | 229.000 | 0 |
| WholeDay | 1.673 | 0.120 | 1.041 | 0.050 | 1.139 | 119.000 | 0 |

Table 3



Is the order of items always the same?

A fundamental feature of the FIES is that the order of the items cannot be considered fixed across countries or even sub-populations within countries.

There are two main reasons for this:

• Food insecurity may be experienced and/or managed differently in different cultures or livelihood systems.

• Nuances in translation may lead to slightly different interpretations of the questions.

In summary, the order of severity is not imposed a priori, but is determined by analyzing the specific data collected.

6. Respondent/Raw score parameters

Still within the "Item and Raw Score Stat" tab, scroll down to look at the table of "Raw score statistics" (see Table 4). Recall that **each respondent has a raw score**, the sum of "yes" responses he or she has have given to the eight FIES questions. All nine possible raw scores are shown in the first column.

| law sco | ere (RS) statistics | | |
|---------|---------------------|-------|---------|
| RS | Severity | Error | Wtd_N |
| 0 | -2.934 | 1.477 | 167.076 |
| 1 | -2.147 | 1.092 | 37.798 |
| 2 | -1.248 | 0.849 | 27.072 |
| 3 | -0.604 | 0.770 | 40.561 |
| 4 | -0.029 | 0.754 | 42.565 |
| 5 | 0.556 | 0.783 | 56.010 |
| 6 | 1.232 | 0.874 | 80.717 |
| 7 | 2.187 | 1.123 | 128.608 |
| 8 | 3.572 | 1.477 | 413.645 |

The second column shows the respondent/raw score severity parameter, or severity of food insecurity associated with each raw score, as well as the errors around them in the next column. **The weighted number of cases** with each raw score are shown in the final column, "Wtd_N."

Keep in mind that...

...the item and severity parameters are all located on the same underlying scale of food insecurity severity.

Difference between the raw score and the respondent parameter.

Looking only at raw scores in the image of the scale shown in Figure 1, you can see that a respondent with a raw score of six is more food secure than someone with score of seven and more food insecure than someone with score of five.



Using respondent parameters, however, you see that difference in severity is smaller between raw scores five and six than between six and seven (i.e. five and six are closer together than are six and seven on the scale).

Respondent parameters allow us to more precisely evaluate the relative differences in food insecurity severity along the scale, between the respondents with each raw score.

Statistical validation, which we are getting to next, allows us to see **if the data are consistent with assumptions** of the Rasch model. If they are, we can say that:

- the **respondent** parameter is an **interval-level measure**¹ of severity
- the **raw score** is an **ordinal measure**² of the severity of food insecurity.



What do you need to remember about item and respondent parameters?

There are two key concepts underlying the approach to FIES measurement: People answer YES or NO to the FIES questions depending on the severity of food insecurity **they experience**.

The items (questions) and the respondents (individuals or households) are located on the **same underlying continuum** of severity of food insecurity.

Interval Scale: In an interval measure, the order of and distance between the values are known and meaningful.
 Ordinal Scale: In an ordinal measure, the order of values is meaningful, such as low, medium and high, but the exact distance between these values is unknown.

7. Statistical validation

Now it is time for statistical validation to **assess the quality** of your FIES data. Stay on the "Item and Raw Score Stat" tab, but move back up to the item statistics (see Table 5). Validation of FIES data involves applying the Rasch model to the FIES response data and assessing whether the data conform to the model's assumptions. **If they do**, you can conclude that the **data can be used** to calculate a valid measure of food insecurity.

Keep in mind that...

...applying the Rasch model will result in several pieces of information (output) that you can use to determine whether the FIES data is of acceptable quality. This information will help you to identify problems in the data and their possible causes.

Table 5

| erity S.E. 92 0.131 65 0.129 97 0.141 | 1.164 | S.E.Infit 0.071 0.069 0.079 | Outfit 1.168 1.356 0.621 | N_Yes_Compl 302.000 292.000 335.000 | N_if_any_valid 0 0 0 0 |
|---|--------------------|--------------------------------------|---|---|---|
| 65 0.129 97 0.141 | 1.164 | 0.069 | 1.356 | 292.000 | 0 |
| 97 0.141 | | | | | |
| | 0.837 | 0.079 | 0.621 | 335.000 | 0 |
| | | | | | |
| 3 0.121 | 0.966 | 0.061 | 0.883 | 248.000 | 0 |
| 02 0.135 | 0.943 | 0.074 | 0.972 | 316.000 | 0 |
| 4 0.122 | 0.900 | 0.062 | 0.808 | 254.000 | 0 |
| 7 0.118 | 3 1.030 | 0.057 | 1.050 | 229.000 | 0 |
| 3 0.120 | 1.041 | 0.050 | 1.139 | 119.000 | 0 |
| | 4 0.122 7 0.118 | 4 0.122 0.900 7 0.118 1.030 | 4 0.122 0.900 0.062 7 0.118 1.030 0.057 | 4 0.122 0.900 0.062 0.808 7 0.118 1.030 0.057 1.050 | 4 0.122 0.900 0.062 0.808 254.000 7 0.118 1.030 0.057 1.050 229.000 |

7.1: INFIT

Interpreting infit statistics.

Look at the "infit" column. Infit statistics are particularly useful **to identify items that did not perform well in a given population**. This may mean that the question was not understood well by respondents. The box below provides guidance on interpretation of the infit.

LOW INFIT (Below 0.7)

An item with a particularly low infit **may be somewhat redundant with other items**. This usually does not substantially worsen the measure of food insecurity, but means that the item does not add important information to the measure. Particular care should be taken in the formulation of these items in future surveys to more clearly differentiate them.

Acceptable fit: 0.7-1.3

An adequate fit to the Rasch model is indicated by an infit statistic of 0.7-1.3. For values between 1.2 and 1.3, the data can be retained for measurement but effort should be made to improve the question wording.

HIGH INFIT (Above 1.3)

These are items with unexpected response patterns. An item **with a high infit** is considered to be **performing poorly**, and should be examined to decide whether to **drop it from the scale for the current analysis**. This also suggests that work is needed to improve this item for future surveys.

As an additional investigation, you may want to also analyze the data separately for different subgroups of interest. You can then examine any differences in their infits to identify where questions may have been understood differently. If an infit is **high only for some sub-groups**, then attention should be given to improving translation that affects those groups **in future surveys**, but no action needs to be taken in the current analysis of the FIES data.

Should you drop items from the scale due to high infit?

The second reason to consider dropping an item from the scale (in addition to missing responses, discussed earlier) would be a high infit. In the example in the image in the previous page, all items have infits within the acceptable range, but **if an item's infit is high**, the decision may be made to drop the problematic item from the scale. This should **only be done after careful consideration** of:

The **standard error** around the infit statistic. A wide standard error around a HIGH infit indicates a less trustworthy result, and therefore provides weaker evidence for dropping the item.

The **number of affirmative responses** to the problematic item. You can find this information in the N_Yes_Compl column. This information is relevant, for example, in a very food secure country where, if you find a high infit for one of the most severe items in the scale, it has probably been answered "yes" by only a small number of respondents. In this case the high infit can essentially be ignored and the item retained in the scale, as this will not significantly alter the estimates of food insecurity prevalence rates.

Although a scale with fewer than eight items may be used for the analysis, a minimum of six items must be retained to produce an acceptable measure of food insecurity severity.

7.2: OUTFIT

Interpreting outfit statistics

While the main statistic to focus on is infit, interpretation of the **outfit** statistic is also worthwhile. Outfit statistics are sensitive to the presence of even just a few cases with highly unexpected response patterns, and are therefore useful to flag the presence of outliers.

An outfit of >2 is considered "high"

The interpretation of a high outfit entails investigation to determine whether the high outfit results from:

A few **highly unusual respondent response patterns**: when less than 0.25 percent of cases, e.g. one case out of a sample of 400, it can be ignored or the decision may be made to drop the unusual cases and refit the model, checking that the item outfit for the remaining cases is acceptable.

A more **pervasive set of somewhat unusual patterns**: If there is a larger number of moderately to highly discrepant responses to the item, the decision may be made to drop the unusual cases and refit the model, checking that the item outfit for the remaining cases is acceptable.

What is an expected pattern of responses?

An expected pattern of responses would be "Yes" answers to the less severe items followed by "No" to the more severe ones (see Figures 2 and 3). A respondent who reports not having faced the less severe experiences yet affirms a more severe item is considered to have an unusual pattern of responses.

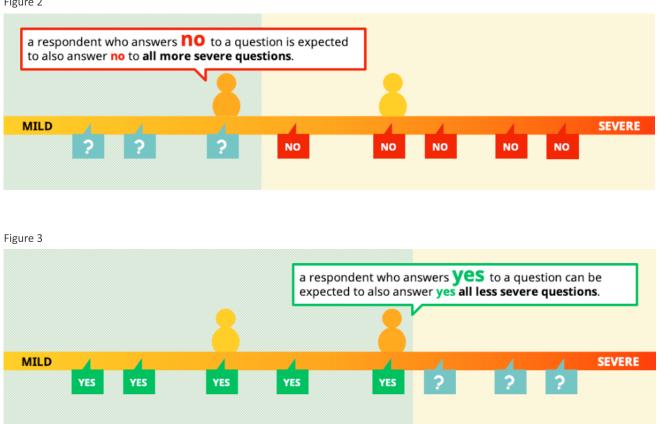


Figure 2



You can download the item parameters and statistics, as well as other output, directly from the app!

Lownload Item Statistics

7.3: RELIABILITY

Interpreting Rasch reliability.

Go to the "Reliability and residual correlation tab."

The Rasch reliability statistic provides information about the discriminatory power of the overall scale, measuring **the proportion of variability in the data that is explained by the Rasch model**.

For an **eight-item** FIES scale, a Rasch reliability value **of at least 0.7** is considered acceptable. For a **seven-item** scale, a Rasch reliability value **of at least 0.6** is considered acceptable.

Rasch reliability

0.70

7.4: RESIDUAL CORRELATION

Interpreting residual correlation matrix.

A residual correlation between a pair of items is considered high if it is >|0.4| (absolute term)

Under the Rasch model assumptions, all correlations among items should result from their common association with the latent trait (in this case food insecurity).

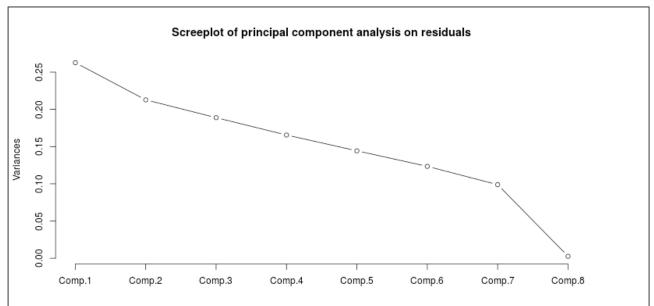
This assumption is assessed by comparing residual correlations among items (see Table 6 and 7), after having eliminated the correlation that exists because the items contribute to measure the same latent trait. One **possible cause of high residual correlation** between one pair of items is **when two items basically overlap in meaning**, which may be a reflection of inaccurate translation or wording. Each item is meant to capture a different aspect of food insecurity, so having **redundant questions** weakens the ability to measure food insecurity accurately.

One high residual correlation requires no action, however **multiple high correlations** suggests you should try eliminating one of the problematic items that is recurrently correlated with the others and re-running the analysis.

Table 6

| Residual cor | rrelation | | | | | | |
|--------------|-----------|---------|---------|---------|--------|--------|----------|
| | Healthy | Fewfood | Skipped | AteLess | RunOut | Hungry | WholeDay |
| Worried | -0.02 | 0.14 | -0.13 | 0.04 | -0.08 | -0.16 | -0.08 |
| Healthy | | 0.12 | -0.12 | -0.08 | -0.11 | -0.15 | -0.15 |
| Fewfood | | | 0.08 | 0.23 | 0.05 | -0.02 | -0.02 |
| Skipped | | | | 0.06 | 0.13 | -0.04 | 0.13 |
| AteLess | | | | | -0.05 | 0.04 | -0.09 |
| RunOut | | | | | | 0.19 | 0.13 |
| Hungry | | | | | | | -0.01 |





8. Select items for analysis

Figure 4



The validation tests may suggest that you drop one or more items from the scale for the analysis if they have not performed well. Go to the tab "Select items for analysis" and **un-check any items you have chosen to omit,** based on the four pieces of information from application of the Rasch model (infit, outfit, residual correlation and reliability).

Notice that when you do omit an item, the analysis is then performed with fewer items, so your output (the item and respondent severity parameters and the fit statistics we have just reviewed) will, in fact, change! It's a good idea to take a look to understand just how much they may have changed.

9. Equating

Equating is a procedure that must be done every time we wish to **compare measures** across **two different applications** of the FIES,³ such as when comparing results obtained in two different countries. This is because the relative position of the items, in terms of severity, **depends on the specific data collected**.

Application of the Rasch model on a single dataset produces a scale that is, to some extent, arbitrary and idiosyncratic to that country/population. Before you can make valid comparisons, it is necessary to **calibrate the two scales on a common metric** by equating the mean and the standard deviation of **the set of items that are common between the two scales**.

At the top of the "Equating" tab, you will now encounter another set of tick boxes that will allow you to select which items you will consider "common," meaning that (after adjustment to the same metric) **their severities are adequately similar**.

It is important to note that this app is designed to produce globally comparable estimates of the prevalence of food insecurity. Therefore, equating is automatically done to the FIES global standard scale. **For equating to another scale**, please refer to the R software available on the Voices of the Hungry webspace.

<u>3.</u> Equating should also be done when comparing the results obtained with applications of the same scale in two different points in time, or when using different scales for different population groups (for example, in households with and without children).

SFYI

Figure 5

If you want to compare FIES data between two countries, equating is necessary.

Simply comparing the numbers obtained from the analysis of FIES data in two different countries would be like comparing temperatures measured in Fahrenheit and Celsius. It has no real meaning.

It is important not to get this step confused with the selection of items for analysis. The choice to drop an item that performed poorly is a decision you based on the results of statistical validation.

The choice of items to use in equating is an important decision, as it affects the position of the thresholds used to classify people into categories of food insecurity, and thus affects the final prevalence rates that will be produced.

Your decision will be informed by the three pieces of information provided on the "Equating" tab. Start with all items selected (see Figure 5).

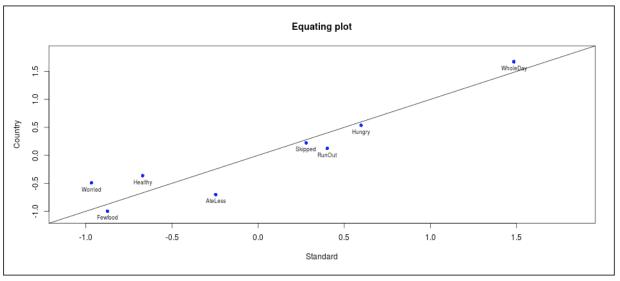
| Choose common items | |
|---|----------|
| ✓ Worried ✓ Healthy food ✓ Few kinds of food ✓ Skipped a meal | Ate less |
| Run out of food V Hungry V Whole day without eating | |

All points should now **appear blue on the plot** of the two scale's item severities. The **correlation among the common items** is 93.4 percent (not shown here). By both looking at Table 9 and also looking at Table 8 (both on page 16), showing the absolute differences in severity between the items on the global and country scales, you will notice that WORRIED is the most discrepant, or different in severity between the two scales.

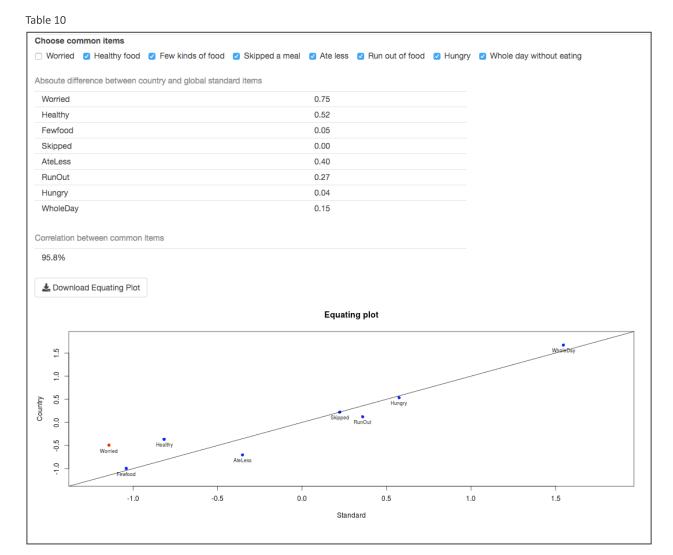
Table 8

| Absoute difference between o | country and global standard items |
|------------------------------|-----------------------------------|
| Worried | 0.60 |
| Healthy | 0.39 |
| Fewfood | 0.15 |
| Skipped | 0.07 |
| AteLess | 0.57 |
| RunOut | 0.35 |
| Hungry | 0.08 |
| WholeDay | 0.24 |

Table 9

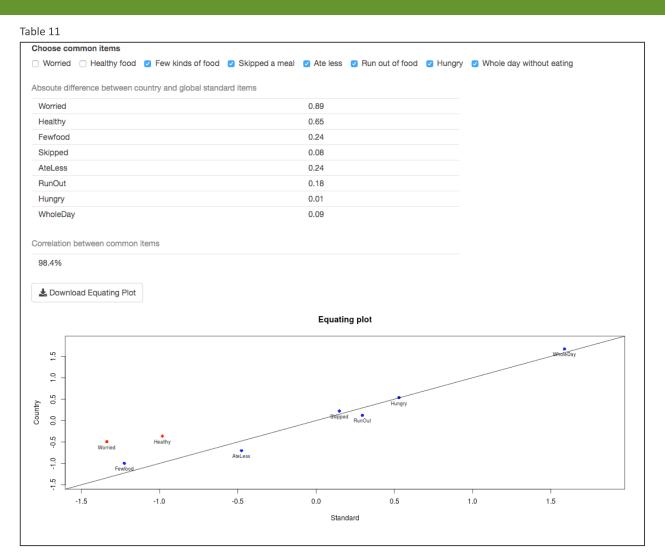


Try de-selecting this item (WORRIED) from the tick-box list and watch (Table 10 on page 17) as the plot, absolute differences, and correlation change! WORRIED is now highlighted as being "unique," while the correlation has increased from 93.4 percent to 95.8 percent.



You may then want to try de-selecting HEALTHY (see Table 11 on page 18) as it now has the highest absolute difference and appears farthest from the diagonal line in the plot. By doing so, you arrive at the following result: the remaining items have moved closer to the diagonal line in the plot, while the correlation has increased to 98.4 percent! You may then decide that this is the best equating scenario and make no further changes, or you might try selecting a different combination of items as common/unique, however **you must retain at least 5 common items for equating to WORK.** You should aim for the maximum number of items that result in the best alignment for equating.

There must be a minimum of six items in the scale to produce an adequate measure of food insecurity. The requirement to have a minimum of five items for equating is different. When you omit an item during equating due to it being discrepant (having a too-different severity level between the two scales being equated), you are still using that item in the measure of food insecurity, and have not dropped it from the scale entirely.



BTIP

A **good exercise** would be to go back to the "Equating" tab and select all items as common.

Look at the prevalence rates, then **try different equating scenarios** (different choices of common items) and see how the prevalence rates change. It is important **not to do this with the simple objective of obtaining the lowest prevalence rates.** Remember, **you can choose which items to treat as "common"** either based on a tolerance level, where you define as "unique" any items with absolute differences in severity beyond a certain threshold or you can also consider **the correlation between common items or the plot of item severities** to inform your decision.

These choices may involve a level of informed judgement, but they are nevertheless guided by principle. **As long as** you clearly report the factors that went into your choice of common and unique items, you will provide a solid justification for your results.

10. Differential Item Functioning (DIF)

DIF analysis is needed to make sure that the scale works similarly in subsets of the data, in order to ensure measurement invariance (an important measurement property) and comparability of prevalence rates at sub-national level. To do so, item severity parameters of each subgroup are plotted against eachother, to provide a visual inspection of how well they align.

If no items show particularly different severity between the subgroups, it can be concluded that the scale worked in the same way independent of the context. Therefore, there is sampling invariance of the tool and prevalence rates of the subgroups can be considered comparable.

Fit statistics by subgroup are also shown to inspect whether some items are problematic only in a specific context.

11. Prevalence rates

Under the "Prevalence Rates" tab you will find your final results: two globally-comparable indicators of the prevalence of food insecurity at different levels of severity, along with their respective margins of error.

Severe (FI_{sev}) is the proportion of the population experiencing severe levels of food insecurity.

Moderate or severe (FI_{mod+sev}) is the SUM of the proportion of the population experiencing moderate or severe levels of food insecurity, combined. This is SDG indicator 2.1.2.

Margins of error.

<u>Under the "Prevalence Rates" tab</u>, there is a place for you to specify the design effect and confidence level for the margins of error (MoE). This allows you to obtain an estimate of the amount of uncertainty around the prevalence. The default value for the design effect is two, but if you have the relevant information for your survey, you can modify it by changing the value in the box. Similarly, you can change the width of the confidence interval and will notice that it changes the margins of error around the estimates.

Households or individuals

Depending on if you have household or individual weights available, you will be able to obtain prevalence rates referring to a percent of households and/or a percent of individuals in the population, the latter of which is most relevant for SDG monitoring.

Disaggregated results

If you have data on area of residence and/or region, you will be able to see prevalence rates for each of these categories.

12. Additional information for analysis

The "Additional Information" tab allows you to download a data file containing computed variables, including **the probability** of moderate-or-severe (middle column) or severe (right column) food insecurity for a respondent **with each of the nine possible raw scores** (see Table 12).

The individual-level probabilities and raw scores can be used in regression analyses **to study the determinants and consequences** of food insecurity.

| robability of f | ood insecurity by raw score | |
|-----------------|---|-------|
| 0 | 0.000 | 0.000 |
| 1 | 0.041 | 0.000 |
| 2 | 0.119 | 0.001 |
| 3 | 0.322 | 0.003 |
| 4 | 0.614 | 0.022 |
| 5 | 0.847 | 0.118 |
| 6 | 0.955 | 0.386 |
| 7 | 0.985 | 0.734 |
| 8 | 0.995 | 0.921 |
| djusted thres | holds of food insecurity on the latent trai | t |
| Moderate or | Severe | -0.25 |
| Severe | | 1.49 |

You will also be able to view the adjusted thresholds used to classify respondents into categories of food insecurity severity on this tab.

Now it's time to use the app with your FIES data.

For questions or technical assistance, please contact fies-help@fao.org

Voices of the Hungry: one metric for the world.

For more information about the Food Insecurity Experience Scale write to: fies-help@fao.org

