The Revised 30-Day Scale

The 30-day scale developed by Hamilton et al. (1997b) included 17 items—9 items indicating occurrence of conditions at any time during the previous 30 days and 8 items indicating recurrence of selected conditions on 5 or more days. In the CPS-FSS, affirmative responses to the 30-day questions are followed up with a question that asks, "In the last 30 days, how many days did…..[the experience or behavior occur]?" Among the 30-day questions, only the question about losing weight did not include such a follow-up. (Beginning in 1998 the follow-up to the question about children not eating for a whole day, was also dropped; see table 1.) Hamilton et al. dichotomized responses to the "how many days" follow-up questions, categorizing responses of 5 or more days as indicating recurrence of the conditions. These items were then included in the scale along with the 9 base items that indicate occurrence at any time during the previous 30 days. Goodness-of-fit statistics for all items were within an acceptable range, so all of the items, including the frequency-of-occurrence follow-ups, were retained in the scale.

Two issues are explored with regard to construction of the 30-day scale. The first is whether to include all of the items that Hamilton et al. included. The second is whether the "5+ days" cutoff is optimal, in terms of consistency with the 12-month measure, for identifying multiple recurrence of conditions. Based on resolution of these issues, a specific 30-day scale is proposed, with a metric that is consistent with that of the 12-month measure.

Which Items Should Be Included in the 30-Day Scale?

The Rasch model assumes that all items are mutually independent except for their relationship with the common underlying phenomenon being measured (Fischer and Molenaar, 1995; Baker, 1992; Hambleton et al., 1991; Wright, 1983). This assumption is clearly violated by the pairs of items consisting of a base item and its frequency-of-occurrence follow-up. The pervasive item dependencies created by the large

number of such pairs of items in the original 30-day scale may, therefore, be problematic. However, previous analysis has found that the item dependencies in the 12-month scale—three sets of dependent items similar to those in the 30-day scale—have negligible effects on item severity scores, although the dependencies bias fit statistics downward for the dependent pair and slightly upward for all other items (Nord and Fogarty, 2000). That analysis consisted of calculating three separate scales. The first included all 18 items. The second omitted the three frequency follow-up items. The third included the frequency follow-up items but omitted their base items. The three scales were then placed on the same metric based on the scores of the 12 items that do not have frequency follow-ups. Item scores and fit statistics for the sets of dependent items were then compared across the three scaling scenarios.

The dependent items in the 30-day scale were tested using a similar methodology and here, too, the effect on item severity estimates was found to be negligible (analysis not shown). Further, in the case of the 30-day scale as proposed by Hamilton et al. (1997b), because almost every item has a dependent partner, fit statistics are also not much affected by the item dependencies. Or, more precisely, the fit of all items is artifactually improved by the item dependencies, and this shows up as improved discrimination of the model overall. Since almost all items are similarly affected by item dependencies, their fit statistics, which compare fit of the item to overall discrimination of all the items in the model, are all near unity.

Thus, on statistical grounds, there is no compelling reason not to include all of the items originally suggested by Hamilton et al. However, the wholesale inclusion of frequency-of-occurrence follow-ups may not be justified on theoretical grounds. The three frequency-of-occurrence follow-ups that are included in the 12-month scale all have theoretical links to adult or child hunger. But more frequent occurrence does not, for all items, necessarily correspond to greater severity.[5] For this reason, and to

---

[5] On the other hand, the empirical results from the present analysis provide fairly strong evidence for the correspondence between greater frequency of occurrence and greater severity within a specific range. For all items with 30-day frequency-of-occurrence follow-ups the difference between very occasional occurrence and repeated

facilitate description and discussion of the 30-day scale *vis-à-vis* the 12-month scale in future reports and research, the revised scale proposed here includes only items corresponding to those in the 12-month scale.[6]

Data from the 1998 and 1999 CPS-FSS were fit to the Rasch model for both the original and revised item sets for the 30-day scale using joint maximum likelihood methods (Fischer and Molenaar, 1995; Wright, 1983). The items are presented along with their severity scores in table 2. The items that are excluded from the revised scale (those shaded in table 1) are mostly in the severe range of the scale. Additional items generally increase the precision of a Rasch-based scale. However, comparison of the measurement error of household scores on the two 30-day scales (not shown) indicates that the gain in precision achieved by including the additional 4 frequency-of-occurrence items is modest (about 10 percent) near the hunger threshold. The gain in precision is greater in the higher-severity range (15 to 20 percent), but added precision in this range does not seem to be important enough for research purposes to justify including the additional items.

Frequency-of-Occurrence Threshold To Indicate Recurring Conditions

The second issue is where to place the cutoff for repeated or recurring conditions for the three frequency-of-occurrence items that are retained in the 30-day scale. In the standard 12-month scale, the three "How often did this happen?" follow-up questions are scored as indicating recurring conditions or behaviors if the response indicated that it occurred in 3 or more months. The corresponding follow-up questions for the 30-day-referenced questions ask, "In the last 30 days, how many days did this happen?" The response is a number of days from 1 to 30. Hamilton et al. placed the cutoff for these items between 4 and 5 days.

---

occurrence of at least 3 or 4 times per month does, in fact, correspond to greater severity. That is, the Rasch fit-statistics of these "frequency" items are consistent with a good fit to the Rasch model.

However, this cutoff results in higher severity of the frequency-of-occurrence items compared with their corresponding items in the 12-month scale (figure 1). In this analysis, the metric of the 30-day scale was adjusted to that of the 12-month scale by a linear transformation that equated the mean and standard deviation of the 30-day base items only to the mean and standard deviation of the corresponding items in the 12-month scale. The calibrations of the base items are very nearly linear with respect to their corresponding items in the 12-month scale. However, each of the three frequency-of-occurrence items is about one-half logistic unit more severe than its corresponding 12-month item.

The scale was recalculated with the "recurrence" threshold set at 4+ days and at 3+ days (in two separate scale estimations; figure 2). The 3+ days cutoff was adopted for the revised scale because it results in relative severities of the frequency-of-occurrence items that are most nearly identical to those of their counterparts in the 12-month scale.

Finalizing the Revised 30-Day Scale

The final step in specifying the 30-day scale is to finalize its metric and calculate household scores for each raw score for households with complete responses. The objective of this process is to adjust the metric of the 30-day scale to that of the standard 12-month scale so that equal scores on the two scales represent (probabilistically) identical arrays of conditions and behaviors, except that they refer to different periods of time. That is, if a specific household score on the 12-month scale indicates a .75 probability that an adult in the household went a whole day without eating in the previous 12 months, the same score

---

[6] The follow-up to the most severe question, asking about children going a whole day without eating, has to be dropped from the scale in any case, as it is no longer asked in the CPS-FSS. The question was dropped because multi-day occurrence was reported by so few households that the data were not usable.

for another household on the 30-day scale will indicate a .75 probability that an adult in that household went a whole day without eating in the previous 30 days.[7]

This objective is accomplished by a linear transformation of the 30-day scale so that the mean and standard deviation of the scores of all items in the scale are equal to the mean and standard deviation of the scores of the corresponding items in the 12-month scale. Since the frequency-of-occurrence items have now been aligned with the base items in a way that is consistent with the 12-month scale, all items in the scale (rather than just the base items) are now used to adjust the metric to that of the 12-month scale (table 3). This adjustment makes it appropriate to use the same thresholds for food insecurity (Rasch scale score of 3.25) and for hunger (Rasch scale score of 6.38) as are used for the 12-month scale. Using these item scores and the discrimination parameter required to equate the standard deviation of the item scores to that of the corresponding items in the 12-month scale, household food security scores and status categories were calculated for each raw score for households with complete responses (table 4). Households that affirm 1 or 2 items are classified as food insecure without hunger, those that affirm 3 or more items are classified as food insecure with hunger.[8]

It is appropriate to consider households that affirmed 1 or 2 items in the 30-day scale to be food insecure without hunger. However, it is not appropriate to describe all households with raw scores of zero as food

---

[7] Note that this does *not* mean that two households, H12 with a score of 4.0 on the 12-month scale and H30 with a score of 4.0 on the 30-day scale, were equally food insecure during either period. H30 was at least as food insecure as H12 during the previous 12 months, but may have been more insecure. H30 experienced, during the previous 30 days, the same array of conditions and experiences that H12 experienced during their worst spell of food insecurity in the previous 12 months. But the last 30 days may not have included the most severe spell of food insecurity experienced by H30. By the same reasoning, during the previous 30 days, H12 could not have been more insecure than H30, but may have been less insecure—even fully food secure—because their worst spell of insecurity may have occurred several months earlier.

[8] The revised scale specifications were based on theoretical and statistical considerations. The overall effects of the revisions on the performance of the scale were small. Cross-tabulation of the revised scale with the original scale based on Hamilton et al. (1997b) found a high level of agreement between the two. The revised scale classified a total of 2.57 percent of households in the April 1995 CPS Food Security Supplement as food insecure with hunger during the previous 30 days compared with 2.40 percent by the original scale. Of all households in the Supplement, 2.29 percent were classified by both scales as food insecure with hunger; the revised scale classified an additional

secure. The lowest score that can be measured by this 30-day scale is well above the food insecurity threshold because the scale lacks the less severe items that are needed to reliably identify food insecurity. Therefore, a substantial proportion of households that were food insecure during the 30-day period are not identified as food insecure by this scale. Comparison of similar scores on the 12-month measure suggests that about 40 percent of all households that would be classified as food insecure (with or without hunger) by a full-range 30-day scale will be incorrectly classified as food secure by this truncated 30-day scale. This 30-day scale should, therefore, only be used to monitor the prevalence of hunger, not of food insecurity. The lower threshold (1 or more affirmatives) may be useful for analytic purposes, but if it is to be used for monitoring, appropriate language will need to be developed to describe the ranges of severity below and above that threshold so that the meaning of the threshold is not confused with that of the food-insecure threshold.

The 30-day food security scale is somewhat less reliable than the 12-month scale, but the difference is not great and affects only households with children. Near the hunger threshold, the measurement standard error of the 30-day scale is around 0.9 logistic units for households with children and 0.94 units for households without children (table 4).[9] The corresponding standard errors for the 12-month scale (not shown) are 0.77 units for households with children and 0.96 units for households without children. The lower reliability of the 30-day scale is almost completely due to the smaller number of items in the scale. The discrimination of the individual items is nearly identical in the 30-day and 12-month scales. The main limitation of the 30-day scale is that it does not cover the less severe range of the food insecurity continuum. In the range that it does cover, however, it is sufficiently reliable for routine monitoring and

---

0.28 percent as food insecure with hunger that were not so classified by the original scale; and the reverse was true for 0.11 percent of households.

[9] All statistics are subject to measurement error (distinct from sampling error) in surveys. An advantage of multiple-indicator measures such as the food security measure is that it is possible to estimate the size of the measurement error. The standard error of measurement of a Rasch-based measure is based on the assumed logistic relationship between the "true" food security of the household and the probability of affirming each item. It is the inverse of the square root of the sum of the information function ($p*(1-p)$) across items.

research applications. The measurable range of the scale extends almost 7 logistic units (from 4.90 to 11.77), and the measurement standard error is less than 1 unit across most of the measurable range.

During the period 1998 to 2000, prevalence rates of food insecurity with hunger during the 30 days prior to the CPS-FSS ranged from 2.0 to 2.7 percent (table 5). The 30-day prevalence rates were in the range of two-thirds to three-fourths of the prevalence rates during the 12 months prior to the surveys. For both measures, the prevalence rates of hunger in 1999 were lower than the corresponding averages of 1998 and 2000, and this pattern was substantially more pronounced for the 30-day scale than for the 12-month scale. This is consistent with the hypothesis advanced by Andrews et al. (2000) and Nord et al. (2002) of a seasonal effect on the measurement of food insecurity and hunger. The measured (12-month) prevalence rates of food insecurity and hunger in 1995, 1997, and 1999, when the food security surveys were conducted in April, were lower than would have been expected based on rates in the alternate years, when the surveys were conducted in August or September. Such a seasonal effect on the 12-month measure would result from measurement error—due to respondents remembering recent events more reliably than events that occurred nearly a year ago. It was expected that a 30-day measure would show a larger seasonal effect since it would be sensitive to the actual seasonal prevalence rates. The prevalence rates for 1998-2000 provide at least a preliminary confirmation of this hypothesis.