What is a Top-box score (TBS)?

A top-box score (TBS) refers to the percentage of respondents who selected *the most* positive response to a survey question (e.g., the 'Always' response from the choices Never, Sometimes, Usually, Always). A percent positive score (PPS) refers to the percentage of respondents who selected *any* positive response to a survey question. Both TBS and PPS are group-level measures calculated from a subset of respondents who provided a valid response to an item (excluding "don't know" and "no answer" responses). When the scores are calculated using weighted responses, TBS and PPS are predicted percentages in the target population, not the actual percentages in the sample.

Is Top-box scoring applicable to individual people?

Top-box scores and percent positive scores are group-level measures. These scores are designed to describe responses from a group of individuals and should not be used to interpret or compare responses from individual people.

When to use Top-box scores

- When comparing questions with different response categories.
- When interested in the percentage of respondents with the best possible score.
- When interested in the percentage of respondents with the worst possible score (bottom-box).
- When summarizing results of many questions in limited space.

Advantages of Top-box scores

- Simple to understand.
- Allow for comparisons to be made across questions with different response categories.

Disadvantages of Top-box scores

- Loss of information.
- Disregard and mask variability in responses.
- Difficult to track improvement or change.

Alternatives to Top-box scores

- Compute measures of central tendency such as the mean (average) or median (middle score or 50th percentile).
- Compute measures of variability such as the standard deviation (a standardized measure of the average squared difference from the mean) or inter-quartile range (IQR; the difference between the 25th and 75th percentiles).
- Tally all response categories, showing counts, percentages, and valid percentages.
- Visualize full response distributions with histograms or bar charts.

Can I report on a Single Item (Question)?

PREMs items are often designed to be used on their own or as part of a dimension (group of questions that address a similar topic). One may report results only at the item level, only at the dimension level, or both, to answer their questions.

This is in contrast to some PROMs instruments (e.g., VR-12), where the focus is often on the measurement of overall composite scores (e.g., overall health status) or composite scores based on particular dimensions or domains (e.g., physical, mental and social health), and less so on individual items. When using standardized PROMs instruments, one should adhere to the test publisher prescribed procedures, such as scoring algorithms, interpretation of scores, and whether item level responses can be reported on their own.

Top-box Scoring for Items (Questions)

The percentage of respondents who selected the most positive response to a survey item is computed by summing the number of times that the most positive response for that item was selected (i.e., the total frequency of the most positive response for the single item; numerator) and dividing this value by the total number of valid responses to the item (denominator). This resulting ratio is then multiplied by 100 to obtain a percentage.

$$\frac{\sum \text{most positive response}}{\sum \text{valid response}} \times 100$$

Top-box Scoring for Dimensions

The percentage of respondents who selected the most positive response for *all survey items that make up a dimension* is computed by summing the number of times that the most positive response for all items within a dimension was selected (i.e., the total frequency of the most positive response for all items within a dimension; numerator) and dividing this value by the total number of valid responses to these items (denominator). The resulting ratio is then multiplied by 100 to obtain a percentage. Note that this describes the "grand sum" method of calculation; however, one could also use the "avg. of sum" method wherein the average of all TBS for items within a dimension is computed.

Avg of Sum	Grand Sum
$\dfrac{\sum TBS}{N_{TBS}}$	$\dfrac{\sum \text{Most positive response for all Qs}}{\sum \text{valid response for all Qs}} \times 100$

Percent Positive Scoring for Items (Questions)

The percentage of respondents who selected any positive response to a survey item is computed by summing the number of times that any positive response for that item was selected (i.e., the total frequency of any positive response for the single item; numerator) and dividing this value by the total number of valid responses to the item (denominator).

$$\frac{\sum \text{any positive response}}{\sum \text{valid response}} \times 100$$

Percent Positive Scoring for Dimensions

The percentage of respondents who selected any positive response for *all survey items that make up a dimension* is computed by summing the number of times that any positive responses for all items within a dimension was selected (i.e., the total frequency of any positive responses for all items within a dimension; numerator) and dividing this value by the total number of valid responses to these items (denominator). Note that this describes the "grand sum" method of calculation; however, one could also use the "avg. of sum" method wherein the average of all PPS for items within a dimension is computed.

Avg of Sum	Grand Sum
$\dfrac{\sum PPS}{N_{PPS}}$	$\dfrac{\sum \text{Most positive response for all Qs}}{\sum \text{valid response for all Qs}} \times 100$

Small Example

Respondent	Q16	Q17	Total
Kerry	Never	Always	
Mary	Always	Usually	
Barry	Always	Never	
Terry	Always	Sometimes	
Jerry	Usually	DK	
Top-box	3	1	4
Valid	5	4	9
	3/5 = 60%	1/4 = 25%	
Top-box Score	(60% + 25%) / 2 = 42.5% Avg of Sum		4/9 = 44.4% Grand Sum

The "avg. of sum" and "grand sum" methods can be considered two different ways of computationally capturing the same thing. In practice, the two methods produce similar dimension scores unless the number of valid responses varies significantly by question.