*Neuendorf*

# Scale Construction

Definition of a Scale (sometimes called an "index"; see Babbie for the distinction): Any combination of two or more indicators intended to measure one general concept or construct. What this usually means operationally is some addition of two or more individual questionnaire items, with the composite becoming a new variable in the data set.

**Some important sources**:

Carmines, E. G., & Zeller, R. A. (1979). Reliability and validity assessment. Newbury Park, CA: Sage Publications.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.). (1991). Measures of personality and social psychological attitudes. San Diego, CA: Academic Press.

Rubin, R. B., Palmgreen, P., & Sypher, H. E. (Eds.) (1994). Communication research measures: A sourcebook. New York: The Guilford Press.

DeVellis, R. F. (1991). Scale development: Theory and applications. Newbury Park, CA: Sage Publications.

**Decisions to make on how many items, and how to combine them**:

A. Theory, or ad hoc researcher decision

1. Straight addition*[1]
2. Standardize and add*
   - Vital when items are on different measurement scales–e.g., SES as constructed from income in dollars, education in years, and occupational prestige in 0-100 units
   - Also important to consider any time variances across items are different; if standardization is not done, the items with greater variance will be represented more strongly in the scale.
   - Remember that standardizing makes each item's mean zero, and its sd 1. Thus, the scores on any such scale will look strange–3.3, -0.7, -5.2, etc. The overall mean of the scale should be zero.
3. For previously developed scales, you should typically use the construction recommended by the authors of the scale. (See, for example, Rubin et al. or Robinson et al.)

B. Factor analysis based (see Appendix of Carmines & Zeller)–must make a decision on how many factors to extract (either forced number, or eigenvalue cutoff such as the "latent root criterion" (eigenvalue at

[1]*--Don't forget to reverse code negatively worded items, or else subtract them rather than add (these two options have the same result).

least 1.0)).

1. Straight addition*
   - Examine factor loadings for a set of variables. Add (or subtract, for negative loadings) the variables that "go together."
2. Standardize and add*
   - Again, examine factor loadings, standardize individual items, and add (or subtract).
3. Use factor score coefficients to weight the items after standardization*
   - In SPSS, the Scores/Save as Variables option does this automatically. The new variable will be added to your SPSS data set. Note that if you use this option, you do <u>not</u> have to reverse code, since the factor score coefficients are appropriately positive or negative.
   - Be aware that <u>all</u> items in the set are used when this process is employed, regardless of how many factors are discovered.

**<u>Reliability</u>**

- Important to assess for the scale no matter how the items are combined, although the decision about whether the reliability of a set of items is acceptable is a different ball game when factor analysis score coefficients have been used.

- When running reliability on SPSS (use Scale/Reliability Analysis), make sure all items are coded in the same direction (i.e., reverse code negatively worded items).

- Just a thought--Can the reliability of the items for an index be TOO high? Think about a possible antagonistic relationship between reliability and content validity. . .

**<u>Scale Development</u>**

For a review of how commonly used, standard scales such as the Marlowe-Crowne Social Desirability Scale, the Beck Depression Inventory, and the Personal Report of Communication Apprehension got their start, see Robinson et al. and Rubin et al. To see how you may work on developing and testing your own scale, see DeVellis' eight-step plan:

1. Define latent variable/construct

2. Generate item pool
   a. Redundancy (basis for internal consistency); this initial item pool can tolerate more than the final scale
   b. Content validity
   c. Positive and negative wording

3. Determine format for measurement (need to keep it consistent)
   a. Thurstone & Guttman scales: both made up of graded/ordered items; the focus is on a single affirmative response for Thurstone, the point of transition from affirmative to negative for Guttma

(e.g., Bogardus social distance scale)

b. Likert-type format:
e.g.,
SD D DK A SA

c. Semantic differential format:
e.g.s,

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| EASY | -3 | -2 | -1 | 0 | 1 | 2 | 3 DIFFICULT |
| ENJOYABLE | | * | * | * | * | * | UNPLEASANT |

4. Have initial item pool reviewed by experts

5. Consider inclusion of validation items
e.g., Social Desirability scale

6. Administer items to a sample from the population of concern

7. Evaluate the items via the data
   a. Variance?
   b. Covariance? –> Correlation? –> Alpha?
   c. Factor analysis –> Unidimensional?

8. Optimize scale length–drop "bad' items

(9. Final scale construction)

**Standardization**

Just a quick review of how a variable is standardized; the general form for standardizing a variable is:

ZX=(X-mean)/sd

in SPSS syntax-speak, it is:

COMPUTE ZQ1=(Q1-3.41)/1.2

Where Q1 is the original variable, ZQ1 is the standardized (z-score) variable, 3.41 is the mean of Q1, and 1.2 is the sd of Q1.

This creates a variable that has a mean of zero and a sd of 1. So a score on ZQ1 of 2.1, for example, would indicate a score that is 2.1 standard deviations above the mean. Standardization puts variables on the same "footing" for possible combinations. Standardization does not fix problems with distributions, such as skewness.

SPSS will create standardized versions of variables automatically with the Descriptives procedure (if you check a box asking for standardized versions of the variables to be saved).