Scoring Likert Type Items

This scale has been designed so that you can rate a patient on his abilities in certain mental health areas. Please respond to every item. In each case, draw a circle around the letter which represents what you think his abilities are as follows:

SA if you strongly agree with the statement
A if you agree but not as strongly
N if you are neutral
D if you disagree but not too strongly
SD if you strongly disagree

1. The patient hears voices.   SA   A   N   D   SD

2. The patient shaves himself.   SA   A   N   D   SD

3. The patient interacts appropriately. SA   A   N   D   SD

Face reliability scoring.

Decide which items are positive and which are negative in relation to the goals established. That is, if the goal is for the person to be competent, then the item, "The patient shaves himself" would be positive. On the other hand, "The patient hears voices" is negative. If you have an item which you cannot decide whether it is positive or negative, then you should discard it or change the wording to make it either positive or negative. If you wish to keep the item as it is, you can item analyze the items, and it will be determined. After you have decided whether the item is positive or negative, give values to the letters as follows:

Likert items with SA, A, N, D, and SD should have the following weights:

SA = 5
A = 4
N = 3
D = 2
SD = 1

How many weights should there be on the Likert scale. The true-false item has two weights. The five point scale above has 5 weights (usually 1, 2, 3, 4, and 5). There are some studies to indicate that reliability improves as the number of weights increase up to about 15. The improvement begins to wane at about 7 or 8. When people are making subjective judgments they tend to give fractional weights when the judgment is between two numbers. For example, when the judgment is either 1 or 2 and the person making the judgment is in between will indicate 1 and a half. This will happen more frequently when there is no middle weight (some people are truly undecided on in the middle--to force them one way or another causes unreliability). It happens in another way: when using a 10 point scale judges will sometimes report 7 and a half. So that this is half

way between the "half-way" point and the highest point of the scale. There is some evidence that people can make this "half-way" judgment three time. A nine point scale (0 1 2 3 4 5 6 7 8) allows such a possibility. Four is half-way between 0 and 8, 2 is half-way between 0 and 4, and finally 3 is half-way between 2 and 4. At any rate the 9 point scale (0 through 8) is recommended.

The weights can have different qualitative descriptors. For example, the above Likert scales are based on the strength of agreement (Strongly Agree to Strongly Disagree). The descriptors can be used for various items.

The following descriptors indicate the about of time spent performing an activity.

INSTRUCTIONS: For each item draw a circle around the number that you think best describes the setting according to the following scale.

none of the time		a little of the time		some of the time		a lot of the time		all of the time
0	1	2	3	4	5	6	7	8

OR

When people are in this setting they are:

1. 0   1   2   3   4   5   6   7   8   tense

2. 0   1   2   3   4   5   6   7   8   satisfied

A more detailed descripter of the amount of time:

never hardly ever once in a while little of the time some of the time a lot of the time fre- quent-ly most of the time all of the time

never	hardly ever	once in a while	little of the time	some of the time	a lot of the time	fre- quent- ly	most of the time	all of the time
0	1	2	3	4	5	6	7	8

OR

Degree of satisfaction

Completely Satisfied		Somewhat Satisfied		Neutral		Somewhat Dissatisfied		Completely Dissatisfied
0	1	2	3	4	5	6	7	8

Degree of Importance.

Not Important		NotVery Important		Somewhat Important		Important		Very Important
0	1	2	3	4	5	6	7	8

These scales have the following format:

INSTRUCTIONS: This scale is designed so that you can indicate how often you experience various emotions. For each item circle the number that best represents how frequently you experience the emotion according to the following scale:

never	rarely	infrequently	occasionally	sometimes	commonly	frequently	usually	always
0	1	2	3	4	5	6	7	8

1. Happiness   0 1 2 3 4 5 6 7 8
2. Sadness   0 1 2 3 4 5 6 7 8
3. Anger   0 1 2 3 4 5 6 7 8
4. Worthwhile   0 1 2 3 4 5 6 7 8

Then each item would have the score of the number circled. For a total score, these numbers would then be added together. This score would be an estimate of the positive emotions of the patients that had been rated.

There are a number of problems with this score. First, two of the items represent positive emotions while two items represents a positive emotions. If you wanted to add the items together to a total test score this issue would need to be resolved. Lets assume that the purpose of the test is to measure positive affect. The items "happy" and "worthwhile" could be added together but the items "sad" and "anger" are not positive emotions. Consequently, the weight on these items would contribute to negative emotions. A zero (0) on the negative emotion should be changed to an 8 on positive emotion, while a score of eight (8) on negative emotion should be scored zero (0) on positive emotion. This is sometimes referred to as "reversing the item." A two (2) becomes a 6, and a 6 becomes a 2.

The next problem with these weights is that the actual weights are unknown. For example, is a score of 5 on Item #2 worth the same as a score of 5 on Item #3? The above scoring method assumes that it does.

Third, it assumes that different people will agree when they rate the same patient on the same item. For example, reliability assumes that two or more people will rate the same patient the same on item #3.

A fourth problem is that it is assumed that a certain score of 3 would have some kind of meaning. The only way to know this is to compare a certain score with other scores.

These can be solved or at least the error made can be estimated by testing reliability, standardized the best, and weighing the items. However, that is time consuming and you may want to evaluate the program and risk unreliability. And furthermore, you may be able to show validity by accounting for variance later in the program. On the other hand, you may want to check reliability particularly if you have been through the program and your methodology did not account for much of the variance. To check reliability and further standardize the test. If this is not your first time through the system and you accounted for much of the criterion variable, but suspect its validity, then check validity.

none of the time	a little of the time	some of the time	a lot of the time	all of the time
0 1	2 3	4 5	6 7	8

never	hardly ever	once in a while	little of the time	some of the time	a lot of the time	fre- quent- ly	most of the time	all of the time
0	1	2	3	4	5	6	7	8

Procedure for scoring Likert tests using SPSS

The following is an example of a Likert test to be scored.

INSTRUCTIONS: This scale is designed so that you can indicate how often you experience var emotions. For each item circle the number that best represents how frequently you experienc emotion according to the following scale:

never	rarely	infrequently	occasionally	sometimes	commonly	frequently	usually	alwa
0	1	2	3	4	5	6	7	8

1. Happiness   0 1 2 3 4 5 6 7 8
2. Sadness   0 1 2 3 4 5 6 7 8
3. Anger   0 1 2 3 4 5 6 7 8
4. Worthwhile   0 1 2 3 4 5 6 7 8

Assume that the purpose of the test is to assess positive emotions of the respondent. The items cannot be simply added together for a total score because two of the items represent positive emotions while the other two items represents negative emotions. The items "happy" and "worthwhile" could be added together but the items "sad" and "anger" are not positive emotions. Consequently, the weight on these items would contribute to negative emotions. A zero (0) on

negative emotion chould be changed to an 8 on positive emotion, while a score of eight (8) on negative emotion chould be scored zero (0) on positive emotion. This is sometimes referred to as "reversing the item." A two (2) becomes a 6, and a 6 becomes a 2.

Before any calculations are made the items that respondents left blank must be resolved. These items left blank are referred to as missing values. The following jobstream sets missing values and scores the test of positive emotions.

```
get file = "\rdda\samp1.sav".
missing values happy sad anger worth (9).
compute sadr=8-sad.
compute angerr=8-anger.
compute tot=mean(happy, sadr, angerr, worth).
EXECUTE.
```

[Now you really are going to need syntax files. Computing the mean using the "click procedure" just doesn't work very well.]

Click to review the procedure for creating syntax files. Use the "back arrow" to return to here.

Statements 2 through 5 set the missing values to 9 and "reverse" the negative emotion items.

The original data was as follows.

HAPPY	SAD	ANGER	WORTH
2	5	9	9
3	4	9	4
9	9	9	9
9	6	9	4

The above jobstream generated the following additional variables: SADR, ANGERR, and T

HAPPY	SAD	ANGER	WORTH	SADR	ANGERR	TOT
2	5	9	9	3	.	2.5
3	4	9	4	4	.	3.67
9	9	9	9	.	.	.
9	6	9	4	2	.	3

Reversing Items

The method is slightly different when the scale is a 1 to 5 scale rather than a 0 to 8 scale. The example above is for a 0 to 8 scale.

When the scale is SA A N D SD they are usually numbered SA = 5, A = 4, N=3, D=2 and SD = 1. To reverse these items the 5 needs to become a 1, a 4 becomes a 2 and etc. The compute statement for such a reversal would be as follows:

Assume the original variable name was DEPRES the compute statement would be:

COMPUTE DEPRESR = 6 - DEPRES.
EXECUTE.

After all of this folderol the actual scoring of a Likert questionnaire is quit simple. You simply take the mean of the items of the subtest or test. The SPSS computer statement is:

compute total=mean(happy, sadr, angryr, worth).
total contains the subtest or total test score of each respondent.