Measurement: Reliability and Validity of Measures

A project is underway to develop and test a questionnaire to measure the health status of adolescents aged 11 to 17 years.

The questionnaire was designed in sections that cover 6 domains considered important areas of health and functioning. These domains include

- Activity
- Discomfort
- Satisfaction
- Disorder
- Achievement
- Resilience

For the most part, the individual items came from other questionnaires, although some were created especially for this measure development project.

When the full set of questions was compiled, a panel of experts in child/adolescent health was convened to do two tasks:

1. The experts were given all the questions in random order, and asked to sort them into domains and elements within domains. The researchers, to see how close the panel members' sortings came to the initial questionnaire, examined the results.
2. The experts reviewed the draft questionnaire to see whether any important items, elements, or domains were omitted.

After revising the questionnaire, several pretest trials were conducted with carefully selected groups of adolescents. Each group was composed of 10 to 12 teenagers with specific characteristics:

1. Healthy adolescents
2. Adolescents with chronic, non-disabling conditions
3. Adolescents with disabling conditions

Comparisons were made among the survey responses of each group to see whether there were differences. Responses across different domains and elements were also analyzed to see whether they intercorrelated as expected.

The fourth task was to review the medical records of the teenagers participating in the pretest. Information abstracted included diagnosis and service utilization over the previous 6 months. These data were compared with the questionnaire responses. The record abstracting was repeated a year later to see whether questionnaire responses were associated with subsequent use of services.

1. What was the purpose of the expert panel?
   a. Internal consistency
   b. Test-retest reliability
   c. Sampling validity
   d. Construct validity
   e. Convergent validity

f. Split-half reliability
g. Face validity
h. Content validity
i. Short-form/long-form reliability
j. Predictive validity
k. C, G, H
l. D, F, J
m. A, E, H

2. How were the small groups used to test reliability?
   a. Test-retest reliability
   b. Split-half reliability
   c. Short-form/long-form reliability
   d. Internal consistency
   e. A, C, D
   f. A, B, C
   g. B, D

3. What aspect(s) of validity did the small groups test?
   a. Sampling validity
   b. Construct validity
   c. Face validity
   d. Convergent validity
   e. Concurrent validity
   f. Discriminant validity
   g. Content validity
   h. Predictive validity
   i. D, F
   j. B, E
   k. G, H
   l. B, C

4. What role did the medical record abstract data play?
   a. Sampling validity
   b. Construct validity
   c. Face validity
   d. Convergent validity
   e. Concurrent validity
   f. Discriminant validity
   g. Content validity
   h. Predictive validity
   i. F, G
   j. D, H
   k. B, C

5. What aspect(s) of reliability or validity were NOT addressed?
   a. Internal consistency
   b. Test-retest reliability
   c. Sampling validity
   d. Construct validity
   e. Convergent validity
   f. Split-half reliability
   g. Face validity
   h. Content validity
   i. Short-form/long-form reliability
   j. Predictive validity
   k. C, H, I
   l. A, D, J
   m. B, F, I

ANWERS:

1. **K:** The panel was used to test content validity (were any important items omitted?), face validity (did the experts assign questions to the same domains and elements that the researchers did?), and sampling validity (were any key domains or elements omitted?)

2. **D:** The analysis compared the groups to see whether patterns of response were consistent with the adolescents' health status. Questions within domains tested to see whether they correlated the way the researchers expected.

3. **I:** Analysis of responses from the small groups looked at intercorrelations among items within elements and domains (convergent validity), and conversely, the extent to which items in one domain did not correlate with items in another (discriminant validity). Convergent and discriminant validity are aspects of construct validity.

4. **J:** Medical record data were taken as a "gold standard" against which to measure the performance of the questionnaire. Two questions were asked, a) did responses from teens about their health or use of services match what the records said (convergent validity) and b) how accurately did the responses predict accurately which teens had more illness or higher levels of service use over the following year (predictive validity).

5. **M:** There was no test-retest, split-half, or short-form/long-form reliability testing.