A study concerning the development and validation of an instrument intended to measure Goal II of quality education is presented. This goal is that quality education should help every child acquire understanding and appreciation of persons belonging to social, cultural and ethnic groups different from his own. The rationale for measurement contains only a sample of possible definitions of this goal and the inventory measures only a sample of possible behaviors. Based on the findings from grade 5 and grade 11 assessment data, a decision was made to use the nine items common to the inventories of both levels as the nucleus of an item pool for a grade 7 instrument. An item example is: How would you feel about sitting in class next to a person whose skin color is different from your own? Item construction was subject to the following restraints: the situation described in each item should be both possible and plausible and related to an individual stimulus rather than a class of stimuli. In order to score the sociometric instrument, the race, IQ level, socioeconomic level, religion and physical handicaps of each of the students in the two 7th grade classes were obtained from school records. It is concluded that this study present a model for future validity studies. The attitude measure produced can serve as a reasonably reliable and valid measure of 7th grade pupils' attitudes toward others who are different from themselves on the dimensions of race, religion, IQ, socioeconomic status, and physical characteristics. (CK)
MODEL FOR USE OF SGCIONETRY TO VALIDATE ATTITUDE MEASURES

Thomas P. McGuiness
and
Peggy L. Stank

Bureau of Educational Research
Pennsylvania Department of Education

The Pennsylvania Department of Education has developed and is implementing a plan to assess the quality of education in the public schools of the Commonwealth (Campbell and Beers, 1970). These ten goals, broadly stated and loosely defined, were the product of the cooperative efforts of the State Board of Education, Educational Testing Service and a committee of citizens representing leadership in many areas of the state. Although they are not ranked in any order of emphasis or importance, these ten goals of quality education represent goals which include more than the usual academics as represented by the three R's (Beers, 1970).

Instrument packages to measure the ten goals of quality education have been prepared and validated for grades 5 and 11. The grade 7 package has been prepared and validation procedures are currently being completed. The development of a grade 3 package has just begun. Data were collected in the fall of 1969 from 20,026 grade 5 pupils and 17,415 grade 11 pupils to develop norms for each of the goals (Hertzog, Beers and Campbell, 1970). In the fall of 1971, data were collected from more than 2,700 pupils in grade 7.

The study reported here concerns the development and validation of an instrument intended to measure Goal II: Quality education should help every child acquire understanding and appreciation of persons belonging to social, cultural and ethnic groups different from his own. The rationale for measurement contains only a sample of possible definitions of this goal and the inventory measures only a sample of possible behaviors. It is left to school district personnel to further define the goals in light of their own program objectives.
Rationale

If Coal II of quality education is to be realized, schools must provide the experiences necessary for students to achieve an attitude of acceptance of others who differ from themselves in racial, religious, intellectual, socioeconomic and physical characteristics. An attitude can either be the determinant or the consequent of an individual's beliefs and behavioral intentions. The immediate antecedent of overt behavior is the behavioral intention of the individual (Fishbein, 1967).

Instrument Development

Based on the findings from the grade 5 and grade 11 assessment data, a decision was made to use the nine items common to the inventories of both levels as the nucleus of an item pool for a grade 7 instrument. Employing a five point Likert response scale, the item format was in the form of a question. Examples of items used are as follows:

1. How would you feel about sitting in class next to a person whose skin color is different from your own?
2. How would you feel if your family moved into a neighborhood where most of the families were richer than yours?

Item construction was subject to the following restraints: the situation described in each item should be both possible and plausible and related to an individual stimulus rather than a class of stimuli. Seventy additional items were generated to measure attitudes toward race, religion, intelligence, socioeconomic status and physical handicaps. These items were examined by a panel of 7th grade teachers who reviewed the instrument for readability, clarity and content validity. Upon the recommendations of the review panel, 16 items were rejected. The remaining 54 items were reviewed and approved by the guidance personnel from a small city junior high school.
The Goal II instrument was administered to 7th grade classes in three
local school districts: a rural school with all white students, three levels
of IQ and SES, and multiple religious backgrounds; a small city school with
all white students, three levels of IQ and SES, and multiple religious back-
grounds; and a small town school with all white students, three levels of IQ
and SES, and multiple religious backgrounds.

A Likert t was computed for each item using total data from all the
classes tested. Items were scored on a continuum from 1 to 5 for this analysis.
All of the items discriminated between top-scoring 27 per cent and low-scoring
27 per cent of the subjects. Additional analysis of the total data, utilizing
a frequency count of responses to items, showed that some items failed to
discriminate among schools and among students within schools. These items
and all of the items showing a Likert t of less than 4.417 were discarded.

A split-half reliability coefficient was also computed on the total
data, and an r = .732 was found.

Validation Procedures

Based on rationale that behavioral intention is both an antecedent of
overt behavior and either the determinant or consequent of attitude, a socio-
metric instrument was constructed for use as a criterion for the validity of
the Goal II instrument.

Both the Goal Ia instrument and the sociometric measure were adminis-
tered to the two 7th grade classes on the same day at the same time. To control
for the effect of the order of administration of the two measures, in one class
the Goal II instrument was administered first, in the other class the socio-
metric measure was administered first.
A high correlation between responses on the Goal II instrument items and actual choices on the sociometric measurement for each of the pertinent factors (race, religion, intelligence, socioeconomic status and physical handicaps) was to be considered evidence of criterion-related validity for the items on the Goal II instrument.

Scoring the Instruments and Analysis of Data

In order to score the sociometric instrument, the race, IQ level, socioeconomic level, religion and physical handicaps of each of the students in the two 7th grade classes were obtained from school records.

The hypergeometric distribution was used to compute, for 5-, 4-, 3- and 2-choice items, the probability that various combinations of choices could occur by chance. These probability figures were computed for all factors for each 7th grade class. Any choice combination with a probability greater than .05 was accepted as evidence of the presence of bias relevant to the specific factor being considered.

Both the sociometric and Goal II measurement of each factor were scored so that a total factor score of 0 indicated absence of bias and a 1 indicated presence of bias.

Table 1 shows the scoring procedures used for the sociometric instrument.

Table 2 shows the scoring procedures for the Goal II instrument.

The absence of physical handicaps within the classes tested made it impossible to compute a sociometric score for this area. However, it was possible to compute a Pearson Product-Moment Correlation for the factors of race, religion, socioeconomic status and intelligence. These correlations and the levels of significance are shown in Table 3.
Table 1
Scoring the Sociometric Measure

<table>
<thead>
<tr>
<th>Probability of Choice Combination</th>
<th>Item Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. $p &lt; .05$ and choices of &quot;others who are the same&quot; predominate</td>
<td>-</td>
</tr>
<tr>
<td>2. $p &lt; .05$ and choices of &quot;others who are different&quot; predominate</td>
<td>+</td>
</tr>
<tr>
<td>3. $p &gt; .05$ that a choice combination happens by chance selection</td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency of + Scores for Each Factor</th>
<th>Factor Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_t &gt; 50%$</td>
<td>0 (absence of bias)</td>
</tr>
<tr>
<td>$f_t &lt; 50%$</td>
<td>1 (presence of bias)</td>
</tr>
</tbody>
</table>

Table 2
Scoring the Goal II Measure

<table>
<thead>
<tr>
<th>Item Response</th>
<th>Item Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I cannot say</td>
<td>0</td>
</tr>
<tr>
<td>2. I would dislike it</td>
<td>-2</td>
</tr>
<tr>
<td>3. I would rather not</td>
<td>-1</td>
</tr>
<tr>
<td>4. I wouldn't mind it</td>
<td>+1</td>
</tr>
<tr>
<td>5. I would like it</td>
<td>+2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mean for Each Factor</th>
<th>Factor Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{x}_F &gt; .50$</td>
<td>0 (absence of bias)</td>
</tr>
<tr>
<td>$1...5$</td>
<td></td>
</tr>
<tr>
<td>$\bar{x}_F &lt; .50$</td>
<td>.1 (presence of bias)</td>
</tr>
<tr>
<td>$-1...5$</td>
<td></td>
</tr>
</tbody>
</table>
Table 3

Factor Scores Correlated with Sociometric Scores

<table>
<thead>
<tr>
<th>Factor</th>
<th>r</th>
<th>Level of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>.4454</td>
<td>.01</td>
</tr>
<tr>
<td>IQ</td>
<td>.7344</td>
<td>.01</td>
</tr>
<tr>
<td>SES</td>
<td>.2851</td>
<td>NS</td>
</tr>
<tr>
<td>Religion</td>
<td>.3563</td>
<td>.05</td>
</tr>
</tbody>
</table>

All of the factors with the exception of socioeconomic status showed significant correlation between the Goal II instrument factor scores and sociometric factor scores. Socioeconomic status did approach the .05 level of significance ($r_{.05} = .288$, $r_{SES} = .285$).

This validity study resulted from data collected during September, 1971. A slightly refined version of the instrument was readministered to the same group of pupils in April of 1972. These administrations constitute, in essence, a test-retest situation with an intervening period of seven months. The Pearson Product-Moment Correlation between the scores on the two administrations was .58 ($p < .01$).

Since the original validity study, additional information concerning the final form of the instrument has become available. In October, 1972, approximately 2,600 pupils from 90 schools in the Commonwealth were randomly selected and the final Goal II instrument was administered, and the school mean was selected as the unit of analysis. These data revealed the following technical characteristics:

- Number of Pupils = 2,609
- Number of Items = 50
Mean = 166.059
Variance (unbiased estimate) = 48.665
Standard Deviation (estimated) = 6.976
Reliability (coefficient alpha) = 0.952
Standard Error Measurement = 1.524
Estimated Average Inter-Item Correlation = 0.283
Range = 33.990

When the range of the 2,609 grade 7 pupils were correlated with scores received on the Crown-Marlowe Social Desirability Scale, a correlation coefficient of -0.072 indicated that the Goal II instrument is not related to social desirability.

This study presents a model for future validity studies. The attitude measure produced can serve as a reasonably reliable and valid measure of 7th grade pupils' attitudes toward others who are different from themselves on the dimensions of race, religion, IQ, socioeconomic status and physical characteristics.