DOCUMENT RESUME

ED 038 695 CG 005 251

AUTHOR Tolsma, Robert J.; And Others
TITLE Measurement Characteristics of the HSCI.
INSTITUTION American Educational Research Association, Washington, D.C.; Iowa State Univ. of Science and Technology, Ames.
PUB DATE 2 Mar 70
NOTE 12p.; Paper presented at American Educational Research Association Convention, Minneapolis, Minnesota, March 2-6, 1970

EDRS PRICE EDRS Price MF-$0.25 HC Not Available from EDRS.
DESCRIPTORS Classroom Environment, Environment, *Environmental Research, *Evaluation, *High School Students, *Measurement, Measurement Instruments, Measurement Techniques, *Perception, Pupil Personnel Services

ABSTRACT

The report determines the characteristics of the High School Characteristics Index (HSCI) both as a group and person measuring instrument. Four objectives are listed: (1) determine the item characteristics of the index as a group measurement instrument by using the ratio between means variance to within means variance as an index of item discrimination; (2) contrast the above analysis with the inappropriate procedure of item scale correlation and the correlation of the item means for the total group with the mean scale scores as indices of item discrimination; (3) estimate the reliability of the index both as a group and person measuring instrument; and (4) check the hypothesized factor structure using multiple group factor analysis. The index was administered to 3365 junior and senior students in 16 Iowa high schools and tested as to item discrimination, reliability and factor analysis. Findings indicate that the HSCI should not be used as a measurement device to assess similarities or differences between groups of individuals. The index can be used as a measure of individual rather than group difference. [Not available in hard copy due to marginal legibility of original document.] (Author/MC)

MEASUREMENT CHARACTERISTICS OF THE HSCI

Robert J. Tolsma, M.S.
John W. Menne, Ph.D.
Gordon C. Hopper, Ed.D.
Iowa State University of Science and Technology

INTRODUCTION

The High School Characteristics Index is designed to assess student perceptions of their high school environment. The instrument consists of 300 true-false items. The items refer to teaching and classroom activities, to rules and regulations, to school organizations, and to features of the buildings and grounds, etc. The high school environment is described in terms of 30 scale scores which are yielded by the 300 dichotomous items. Definitions of the scales appear in the Appendix.

PURPOSE

The HSCI has, in the past, been used in several different research situations. For example, it has been used to assess an individual's perceptions of the environment for comparison with his assessed needs. In several instances the HSCI has been used not as an individual measuring instrument but to assess and compare the perceptions of groups of students from different classes or high schools.

The purpose of the research which led to the presentation of this paper was to determine the characteristics of the HSCI both as a person and as a group measuring instrument. It seemed unlikely that a measurement instrument such as this could yield meaningful data when applied to both situations. In order to

assess differences in perception among individuals it is necessary to construct items which discriminate one individual's perception from those of another. Thus, if the items comprising an individual measurement instrument are discriminating, considerable variance would be expected within the group of individuals being measured. If, on the other hand, the instrument is employed to assess differences in perceptions of groups of people the items must discriminate between the perceptions of members of one group from those of another group of individuals. Therefore, one would expect low variance within group members but a relatively larger variance between groups. The two uses made of the HSCI, individual and group measurement, are at cross-purposes. The current study was undertaken to determine the most appropriate employment of the HSCI.

In brief, the objectives of this study were to:

1. Determine the item characteristics of the HSCI as a group measurement instrument by using the ratio of between means variance to within means variance as an index of item discrimination.

2. Contrast the above analysis with the inappropriate procedure (for group instruments) of item-scale correlation and the correlation of the item means for the total group with the mean scale scores as indices of item discrimination.

3. Estimate the reliability of the HSCI both as a group and as a person measuring instrument. The former assessment to be done via analysis of variance, the latter by the coefficient

alpha estimate.

4. Check the hypothesized factor structure using multiple group factor analysis. Contrast the results with those obtained utilizing principal axes with Varimax rotation.

PROCEDURE

The High School Characteristics Index was administered to 3365 junior and senior students in 16 Iowa high schools.

A. Item Discrimination - The items were first analyzed to determine whether or not they were useful in discriminating the responses of one high school group from those of other groups. Analysis of variance procedure was used as an index of item discrimination. The variance within each of the 16 groups was computed, summed, and averaged to obtain the average within means variance. The total variance, i.e. the variance over all groups, was obtained. A ratio of average within group variance to total variance was yielded by dividing the former into the latter. This is a more conservative criterion of selection than a statistical significant F ratio of within to between variance. This latter ratio approaches a critical level when there are as few as one hundred people in a sample, e.g. twenty groups of five persons each. Although it would, in most instances, be unlikely that such small groups would exist in the measurement situation this example nevertheless does show that statistical significance is not a rigorous enough criteria for item selection. For this reason a total variance/within variance ratio

of 1.2 was selected as the point below which items would be identified as "bad," i.e. non-discriminating items, above which the item was considered a "good" item, i.e. an item which was useful in helping to discriminate among groups. While we would like to have the ratio as high as possible, a ratio of 1.2 was considered minimal since of the total variance 17 percent is between group variance and 83 percent within group variance. Out of 300 items on the HSCI there were 22 items which met this criteria of item discrimination. Fifty of the items yielded ratios equal to 1.1 or above. Ratios for the remaining items were below 1.1. The evidence is clear that the HSCI cannot be employed to detect differences between groups of students in analyzing their perceptions of various school environments.

Items were also analyzed in regard to discriminating between the responses of individuals. The point-biserial correlations between items and the respective scales were computed. Prior to viewing these correlations the reliabilities of five HSCI scales were calculated in order to gain an idea of the size of correlations to expect. They tended to be low, consequently it was apparent the correlations would also tend to be somewhat lower than desirable, especially in view of the fact that each scale only consists of ten items. The average item-scale correlations and reliabilities of the first five scales are listed in Table 1.

The range for all item-scale correlations was from a -.027 to .551. A summary of all correlations is presented in Table 2. It should be noted that the correlations presented here have

Table 1. Average and range of item-scale point biserial correlations (uncorrected)

SCALE	AVG. CORR.	RANGE	RELIABILITY
1	.408	.260-.519	.702
2	.341	.261-.423	.561
3	.293	.092-.372	.611
4	.404	.274-.494	.679
5	.411	.318-.491	.749

not been corrected for the biasness which is present due to the fact that the item is part of the respective scale score with which it is correlated. Thus they are somewhat higher than what may be the actual situation. The corrected correlations which are highest represent the "good" items. Corrected correlations of above .20 are generally sought.

Table 2. Summary of item-scale correlations (uncorrected)

CORRELATIONS	FREQUENCY
.50 through .59	8
.40 " .49	74
.30 " .39	130
.20 " .29	70
.10 " .19	15
.00 " .09	2
Neg. Corr.	1

Another approach which was used as an index of item discrimination was that of correlating average item scores with average scale scores for the groups. This procedure was carried out on only two scales since it is not as appropriate for groups as is the analysis of variance method.

Table 3. Summary of avg. item - avg. scale score correlations

CORRELATION	SCALE 1	SCALE 2
.90+	1	
.80 through .89		
.70 " .79	3	
.60 " .69	2	
.50 " .59		
.40 " .49		
.30 " .39	1	
.20 " .29	2	3
.10 " .19		2
.00 " .09	1	4
-.10 " -.01		1

B. Reliability - Scale reliabilities of the HSCI when used as a group measuring instrument were not calculated due to the fact there were so few items which were discriminating. No scale contained more than three of the original ten items which could be used to discriminate between groups. Several scales

contained no discriminating items. The magnitude of the average reliability for the scales can be estimated by obtaining the reliability coefficient for the total test (300 items) and estimating the reliability for a test 1/30 as long. This computation results in an expected average scale reliability of .308. It can be concluded that the reliability coefficients, when the HSCI is used for group measurement, are not sufficiently high, for most if not all scales, to warrant their use in basic research and/or applied situations.

Reliabilities of the first five scales, when the HSCI was used as a person measuring instrument, were estimated using the coefficient α estimate. These reliabilities are presented in Table 1. These reliabilities are relatively low for the scales to be used for other than basic research.

Employing analysis of variance procedures the reliability of the HSCI as a group measuring instrument was estimated. This reliability coefficient has no practical utility since the entire test has no score. Scale scores are interpreted but there is no overall test score. The reliability was computed to serve as an example of the appropriateness of computing reliability using analysis of variance procedures. Scale score variance within each of the 16 groups was first computed. These variances were summed and averaged to obtain the within group variance over all scales. Each within group variance (mean square) was then multiplied by the respective degree of freedom for that group to obtain sums of squares for each group. The

S.S. for each group were summed to obtain the within sums of squares for all groups. Total sums of squares were obtained by getting the average scale variance for the entire sample and then multiplying by the respective degrees of freedom. The between group sums of squares were obtained by subtraction of within S.S. from total S.S. The between and within groups mean squares were obtained in the usual manner. The reliability was calculated by summing the within and between group variance and dividing the sum into the between group variance to obtain an over all test reliability of .933. The values obtained are presented in the ANOV table in Table 4.

Table 4. Reliability of the HSCI as a Group Measuring Instrument

ANOV

Source	S.S.	d.f.	M.S.	E.M.S.
Total	21,260.48	3,364		
Within	19,838.99	3,349	6.32	$\sigma_E^2 + \sigma_w^2$
Between	1,421.49	15	94.77	$\sigma_E^2 + \sigma_w^2 + \sigma_n^2$

$$\text{Reliability} = \frac{\sigma_b^2}{\sigma_w^2 + \sigma_b^2} = \frac{88.45}{6.32 + 88.45} = .933$$

C. Factor Analysis - Another psychometric approach undertaken was factor analysis of the scales comprising the HSCI. The purpose of this analysis was to determine if the scales actually measure different aspects of student perception or do some of them measure essentially the same thing. Thirty factors

were sought from the 30 scale scores using the principal axis method of analysis. The majority of scales loaded heavily on the first factor. This factor alone accounted for 35.39 percent of the variance. This means that the scales tend to be measuring the same thing. The first four factors account for over sixty percent of the variance. It took twenty-two factors to account for 26 percent of the remaining variance. Had the scales been measuring separate aspects of the perceived environment one would expect to obtain thirty factors which would more equally account for the variance. Instead for measurement purposes the thirty scales yeild only four factors.

The factors were not rotated as originally planned since the majority of items comprising the scales are poor. With poor items the scales are questionable. There were not enough good items to check hypotheses as to the constructs being measured.

CONCLUSION

The final conclusion to be drawn is that the HSCI should not be used as a measurement device to assess similarities or differences between groups of individuals. Any study in which it is explicitly or implicitly indicated that groups of high school students differ or do not differ in their responses on the HSCI as a function of their group membership has likely generated erroneous conclusions.

The scales of the HSCI can more appropriately be used as measures of individual than of group differences. Some items

need to be replaced with "good" items. The reliability coefficients of many of the scales are sufficiently large to warrant use in basic research but are of questionable magnitude for use in applied situations.

Certain steps need to be taken to develop a more useful instrument for measurement of groups of high school students. For example multipoint items need to be written to replace the dichotomous items. This would aid in improved item discrimination and scale reliability. Items which do not discriminate or discriminate poorly need to be discarded. Keep in mind that a test of around 150 items is approximately the length desired. A test of this length can be administered in one class period. The factor structure of the revised instrument needs to be determined with the goal of arriving at four or five scales.

APPENDIX

Need-Press Scale Definitions[1]

1. aba Abasement--ass Assurance: self-depreciation versus self-confidence.
2. ach Achievement: striving for success through personal effort.
3. ada Adaptability--dfs Defensiveness: acceptance of criticism versus resistance to suggestion.
4. aff Affiliation--rej Rejection: friendliness versus unfriendliness.
5. agg Aggression--bla Blame Avoidance: hostility versus its inhibition.
6. cha Change--sam Sameness: flexibility versus routine.
7. cnj Conjunctivity--dsj Disjunctivity: planfulness versus disorganization.
8. ctr Counteraction--inf Inferiority Avoidance: restriving after failure versus withdrawal.
9. dfr Deference--rst Restiveness: respect for authority versus rebelliousness.
10. dom Dominance--tol Tolerance: ascendancy versus forbearance.
11. e/a Ego Achievement: striving for power through social action.
12. emo Emotionality--pic Placidity: expressiveness versus restraint.
13. eny Energy--pas Passivity: effort versus inertia.
14. exh Exhibitionism--inf Inferiority Avoidance: attention-seeking versus shyness.
15. f/a Fantasied Achievement: daydreams of extraordinary public recognition.
16. har Harm Avoidance--rsk Risktaking: fearfulness versus thrill-seeking.
17. hum Humanities, Social Science: interests in the humanities and the Social Sciences.
18. imp Impulsiveness--del Deliberation: impetuousness versus reflection.
19. nar Narcissism: vanity.
20. nur Nurturance--rej Rejection: helping others versus indifference.
21. obj Objectivity--pro Projectivity: detachment versus superstition (Al) or suspicion (El).

---

[1]Stern, George C. Scoring instructions and college norms. Activities index and college characteristics index. Psychological Research Center, Syracuse, New York. 1963. Pp. 2-3.

22. ord Order--dso Disorder: compulsive organization of details versus carelessness.
23. ply Play--wrk Work: pleasure-seeking versus purposefulness.
24. pra Practicalness--ipr Impracticalness: interest in practical activities versus indifference.
25. ref Reflectiveness: introspective contemplation.
26. sci Science: interests in the Natural Sciences.
27. sen Sensuality--pur Puritanism: interest in sensory and esthetic experiences.
28. sex Sexuality--pru Prudishness: heterosexual interests versus their inhibition.
29. sup Supplication--aut Autonomy: dependency versus self-reliance.
30. und Understanding: intellectuality.

Presented at AERA 1970