

## **INTERPRETING TEST SCORES**

### **Interpreting Test Scores**

All types of research data, test result data, survey data, etc is called raw data and collected using four basic scales. Nominal, ordinal, interval and ratio are four basic scales for data collection. Ratio is more sophisticated than interval, interval is more sophisticated than ordinal, and ordinal is more sophisticated than nominal. A variable measured on a "nominal" scale is a variable that does not really have any evaluative distinction. One value is really not any greater than another. A good example of a nominal variable is gender. With nominal variables, there is a qualitative difference between values, not a quantitative one. Something measured on an "ordinal" scale does have an evaluative connotation. One value is greater or larger or better than the other. With ordinal scales, we only know that one value is better than other or 10 is better than 9. A variable measured on interval or ration scale has maximum evaluative distinction.

### **Interpreting Test Scores by Percentiles**

The students' scores in terms of criterion-referenced scores are most easy to understand and interpret because they are straightforward and usually represented in percentages or raw scores while norm-referenced scores are often converted to derive standard scores or converted in to percentiles. Derived standard scores are usually based on the normal curve having an arbitrary mean to compare respondents who took the same test. The conversion of students' score into student's percentile score on a test indicates what percentage of other students are fell below that student's score who took the same test. Percentiles are most often used for determining the relative standing position of any student in a population. Percentile ranks are an easy way to convey a student's standing at test relative to other same test takers.

### **Interpreting Test Scores by Percentages**

The number of questions a student gets right on a test is the student's raw score (assuming each question is worth one point). By itself, a raw score has little or no meaning. For example if teacher says that Fatima has scored 8 marks. This information (8 marks) regarding Fatima's result does not convey any meaning. The meaning depends on how many questions are on the test and how hard or easy the questions are. For example, if Umair got 10 right on both a math test and a science test, it would not be reasonable to conclude that his level of achievement in the two areas is the same. This illustrates, why raw scores are usually converted to other types of scores for interpretation purposes. The conversion of raw score into percentage convey students' achievements in understanding and meaningful way. For example if Sadia got 8 questions right out of ten questions then we can say that Sadia is able to solve  $100 \times \frac{8}{10} = 80\%$  questions. If each question carries equal marks then we can say that Sadia has scored 80% marks. If different questions carry different marks then first count marks obtained and total marks the test. Use the following formula to compute % of marks.

Total Marks /Marks Obtained  $\times$  100= % marks

### **Interpreting Test Scores by ordering and ranking**

Organizing and reporting of students' scores start with placing the scores in ascending or descending order. Teacher can find the smallest, largest, rang, and some other facts like variability of scores associated with scores from ranked scores. Teacher may use ranked scoes to see the relative position of each student within the class but ranked scores does not yield any significant numerical value for result interpretation or reporting.

### **Measurement Scales**

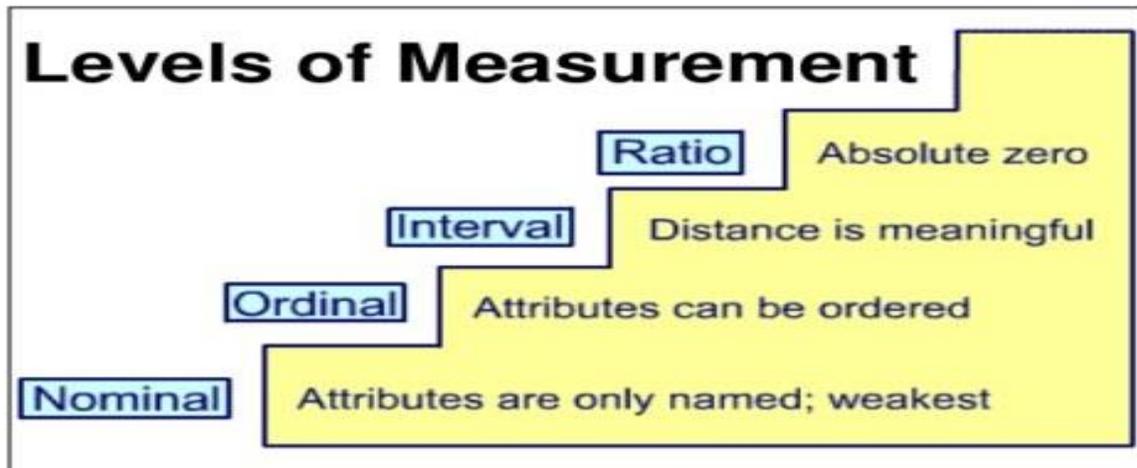
Measurement is the assignment of numbers to objects or events in a systematic fashion. Measurement scales are critical because they relate to the types of statistics you can use to analyze your data. An easy way to have a paper rejected is to have used either an incorrect scale/statistic combination or to have used a low powered statistic on a high powered set of data. Following four levels of measurement scales are commonly distinguished so that the proper analysis can be used on the data a number can be used merely to label or categorize a response.

#### **1. Nominal Scale**

Nominal scales are the lowest scales of measurement. A nominal scale, as the name implies, is simply some placing of data into categories, without any order or structure. You are only allowed to examine if a nominal scale datum is equal to some particular value or to count the number of occurrences of each value. For example, categorization of blood groups of classmates into A, B, AB, O etc. In The only mathematical operation we can perform with nominal data is to count. Variables assessed on a nominal scale are called categorical variables; Categorical data are measured on nominal scales which merely assign labels to distinguish categories. For example, gender is a nominal scale variable. Classifying people according to gender is a common application of a nominal scale.

#### **2. Ordinal Scale**

Something measured on an "ordinal" scale does have an evaluative connotation. You are also allowed to examine if an ordinal scale datum is less than or greater than another value. For example rating of job satisfaction on a scale from 1 to 10, with 10 representing complete satisfaction. With ordinal scales, we only know that 2 is better than 1 or 10 is better than 9; we do not know by how much. It may vary. Hence, you can 'rank' ordinal data, but you cannot 'quantify' differences between two ordinal values. Nominal scale properties are included in ordinal scale.



### 3. Interval Scale

An ordinal scale has quantifiable difference between values become interval scale. You are allowed to quantify the difference between two interval scale values but there is no natural zero. A variable measured on an interval scale gives information about more or better as ordinal scales do, but interval variables have an equal distance between each value. The distance between 1 and 2 is equal to the distance between 9 and 10. For example, temperature scales are interval data with 25C warmer than 20C and a 5C difference has some physical meaning. Note that 0C is arbitrary, so that it does not make sense to say that 20C is twice as hot as 10C but there is the exact same difference between 100C and 90C as there is between 42C and 32C. Students' achievement scores are measured on interval scale

### Ratio Scale

Something measured on a ratio scale has the same properties that an interval scale has except, with a ratio scaling, there is an absolute zero point. Temperature measured in Kelvin is an example. There is no value possible below 0 degrees Kelvin, it is absolute zero. Physical measurements of height, weight, length are typically ratio variables. Weight is another example, 0 lbs. is a meaningful absence of weight. This ratio hold true regardless of which scale the object is being measured in (e.g. meters or yards). This is because there is a natural zero.

### References

Suggested Readings Huff, D. (1954). How to lie with statistics. New York: Norton.

- Bertrand, A., & Cehula, J. P. (1980). *Tests, measurement, and evaluation: A developmental approach*. Reading, MA: Addison-Wesley. Chapter 7 provides an innovative presentation of most of the topics covered in the present chapter.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed.). Englewood Cliffs, NJ: Prentice-Hall. Chapters 7 through 12 are especially useful for helping teachers develop classroom achievement tests. Chapter 14 discusses observation and informal data collection techniques. Chapters 16 through 18 provide useful information on using standardized tests.
- Hills, J. R. (1986). *All of Hills' handy hints*. Columbus, OH: Merrill. This is a collection of articles originally published in *Educational measurement: Issues and practice*. The articles offer practical and interesting insights into fallacies in the interpretation of test scores. (Incidentally, the original journal provides theoretically sound guidelines that are easy to understand.)
- Kubiszyn, T., & Borich, G. (1987). *Educational tests and measurement: Classroom application and practice*. Glenview, IL: Scott, Foresman and Company. The chapter on data presentation provides useful and practical guidelines for communicating data effectively through graphs and diagrams.
- Lyman, H. B. (1986). *Test scores and what they mean* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall. This book provides a detailed discussion of the interpretation of test scores.
- Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphic Press. This book offers interesting examples of how to display information and discusses strategies for presenting data graphically.
- Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher*, 21, 14-23. This article presents strategies for employing and interpreting sophisticated yet understandable graphs to display quantitative data.