PSYC 2101: Howell Chapters 1 & 2[©]

I. What is Measurement?

A. **Strict definition:** "any method by which a unique and reciprocal correspondence is established between all or some of the magnitudes of a kind and all or some of the numbers..."

1. **Magnitude:** a particular amount of an attribute
2. **Attribute:** a measurable property

B. Example:

```
              length
                        |---Y---Y--Y----Y---------------->
             numbers
<-----------------------|---+---+--+----+---------------->
                        0   2   4  5½   8
```

C. **Loose definition:** "the assignment of numerals to objects or events according to rules" (presumably any rules). A **numeral** is any symbol other than a word.

II. Scales of Measurement

A. **Nominal Scale of Measurement**

1. The function of nominal scales is to classify - <u>numerals</u> (symbols, such as 0, 1, A, Q #, ω) are arbitrarily assigned to name objects/events or qualitative classes of objects/events.
2. Does not meet the criteria of the strict definition of measurement.
3. Given two measurements, I can determine whether they are the same or not, but I may not be able to tell whether the one object/event has more or less of the measured attribute than does the other.
4. For example, I ask each member of the class to take all of e's (his or her) paper money, write e's Banner ID number on each bill, and then put it all in a paper bag I pass around class. I shake the bag well and pull out two bills. From the Banner ID numbers on them I can tell whether they belong to the same person or not.

B. **Ordinal Scale of Measurement**

1. The data here have the characteristics of a nominal scale and more. When the objects/events are arranged in serial order with respect to the property measured, that order is identical to the order of the measurements.
2. Given ordinal data on two objects, we can state whether they have equal amounts of the attribute measured or, if not, which has the greater amount. Given two pairs of scores we may not, however, be able to make statements about whether or not the difference between the one pair of scores (a-b) differs from the difference between the other pair of scores (c-d).
3. The most common type of ordinal scale is that of rank order.

4. Example: I throw those bills out the window. My associate, outside, my associate, outside, ranks the students in order of how long it took them to retrieve their money, where rank 1 belongs to the student who most quickly retrieved her money.

```
Error-Free Stopwatch      | 1   3 4     7         12     Minutes
                TIME      |-Y---Y-Y-----Y---------Y--------->
                RANK      | 1   2 3     4         5
```

For the ordinal data (the ranks), we can say that 3 represents more time than 1, but we cannot honestly say that the difference between rank 3 and rank 1 (3 minutes) is the same as the difference between rank 5 and rank 3 (8 minutes). Two rank points between ranks 5 and 3 represent more time than two rank points between ranks 3 and 1.

C. **Interval Scale of Measurement**

1. This is an ordinal scale plus. It has equal units – a change of 1 unit on the measurement scale corresponds to the same change in the measured attribute, no matter where on the measurement scale the change is.
2. We can make statements about (a-b) versus (c-d).
2. Example: Consider the measurements from the error-free stopwatch. The difference (3 minutes) between 1 and 4 minutes is the same as the difference between 4 and 7 minutes (interval data), but the difference between rank 1 and rank 2 (2 minutes) is not the same as between rank 2 and rank 3 (1 minute).

D. **Ratio Scale of Measurement**

1. An interval scale with a true zero point – a score of 0 means that the object/event has absolutely none of measured attribute.
2. With nonratio data we cannot make meaningful statements about the ratio of two measurements, but with ratio data we can.
3. Example: we cannot say that 20 degrees Celsius (interval scale) is twice as hot as 10 degrees Celsius, but we can say that 300 degrees Kelvin (ratio scale) is twice as hot as 150 degrees Kelvin. Please see my graphical explanation of this at http://core.ecu.edu/psyc/wuenschk/docs01/Ratio-Interval.pdf .

III. Variables

A. A variable is a quantity that can take on different values. A constant is a quantity that is always of the same value.

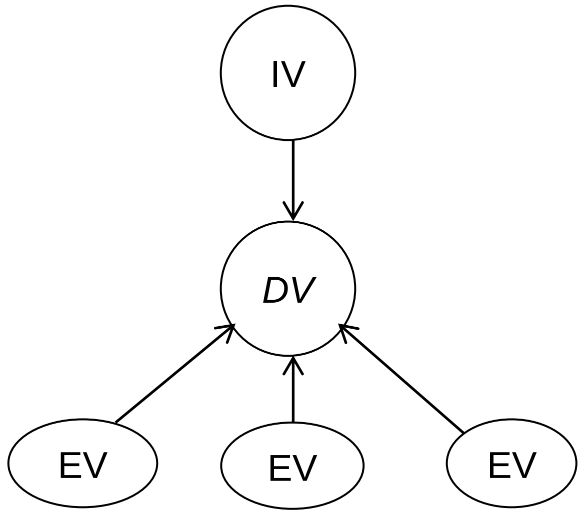
B. **Discrete variable** - one for which there is a finite number of potential values which the variable can assume between any two points on the scale. Examples: Number of light bulbs in a warehouse; number of items correctly answered on a true-false quiz.

C. **Continuous variable** - one which theoretically can assume an infinite number of values between any two points on the scale. Example: Weight of an object in pounds. Can you find any two weights between which there is no intermediate value possible?

D. **Categorical variable** – similar to discrete variable, but usually there are only a relatively small number of possible values. Also called a grouping variable or classification variable. Such a variable is created when we categorize objects/events into groups on the basis of some measurement. The categories may be nominal (female, male; red, green, blue, other) or they may be ordinal (final grade of A, B, C, D, or F).

IV. Bivariate Experimental Research

Consider the following sketch of a basic bivariate (focus on two variables) research paradigm.



"IV" stands for "**independent variable**" (also called the "**treatment**"), "DV" for "**dependent variable**," and "EV" for "**extraneous variable**." In **experimental research** we manipulate the IV and observe any resulting change in the DV. Because we are manipulating it experimentally, the IV will probably assume only a very few values, maybe as few as two. The DV may be categorical or may be continuous. The EVs are variables other than the IV which may affect the DV. To be able to detect the effect of the IV upon the DV, we must be able to control the EVs.

Consider the following experiment. I go to each of 100 classrooms on campus. At each, I flip a coin to determine whether I will assign the classroom to Group 1 (level 1 of the IV) or to Group 2. The classrooms are my "experimental units" or "subjects." In psychology, when our subjects are humans, we prefer to refer to them as "participants," or "respondents," but in statistics, the use of the word "subjects" is quite common, and I shall use it as a generic term for "experimental units." For subjects assigned to Group 1, I turn the room's light switch off. For Group 2 I turn it on. My DV is the brightness of the room, as measured by a photographic light meter. EVs would include factors such as time of day, season of the year, weather outside, condition of the light bulbs in the room, etc.

Think of the effect of the IV on the DV as a signal you wish to detect. EVs can make it difficult to detect the effect of the IV by contributing "**noise**" to the DV – that is, by producing variation in the DV that is not due to the IV. Consider the following experiment. A junior high school science student is conducting research on the effect of the size of a coin (dime versus silver dollar) on the height of the wave produced when the coin is tossed into a pool of water. She goes to a public pool, installs a wave measuring device, and starts tossing coins. In the pool at the time are a dozen rowdy youngsters, jumping in and out and splashing, etc. These youngsters' activities are EVs, and the noise they produce would make it pretty hard to detect the effect of the size of the coin.

Sometimes an EV is "**confounded**" with the IV. That is, it is entangled with the IV in such a way that you cannot separate the effect of the IV from that of the DV. Consider the pool example again. Suppose that the youngsters notice what the student is doing and conspire to confound her research. Every time she throws the silver dollar in, they stay still. But when she throws the dime in, they all cannonball in at the same time. The student reports back remarkable results: Dimes produce waves much higher than silver dollars.

Here is another example of a confound. When I was a graduate student at ECU, one of my professors was conducting research on a new method of instruction. He assigned one of his learning classes to be taught with method A. This class met at 0800. His other class was taught with method

B. This class met at 1000. On examinations, the class taught with method B was superior. Does that mean that method B is better than method A? Perhaps not. Perhaps the difference between the two classes was due to the time the class was taught rather than the method of instruction. Maybe most students just learn better at 10 than at 8 – they certainly attend better at 8 than at 10. Maybe the two groups of students were not equivalent prior to being taught differently. Most students tend to avoid classes at 8. Upperclassmen get to register before underclassmen. Some people who hate classes at 8 are bright enough to learn how to avoid them, others not. Campbell and Stanley (*Experimental and quasi-experimental designs for research*, 1963, Chicago: Rand McNally) wrote about the importance of "achieving pre-experimental equation of groups through randomization." Note that the students in the research described here were not randomly assigned to the treatments, and thus any post-treatment differences might have been contaminated by pre-treatment differences.

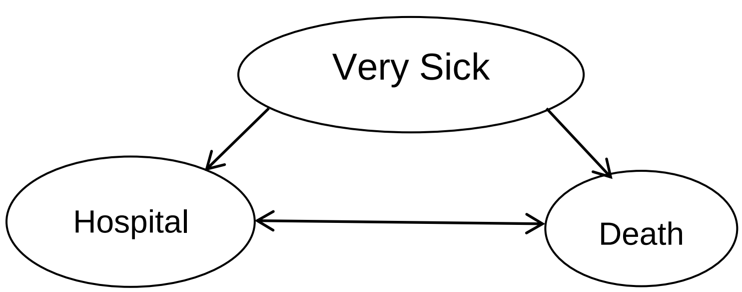
V. Nonexperimental Research

Much research in the behavioral sciences is not experimental (no variable is manipulated), but rather "**observational**". Some use the term "correlational" to describe such a design, but that nomenclature leads to confusion, so I suggest you avoid it. It is true that correlation/regression analysis (which we shall study later) is often employed with nonexperimental research.

Consider the following research. I recruit participants in downtown Greenville one evening. Each participant is asked whether or not e has been drinking alcohol that evening. I test each participant on a reaction time task. I find that those who report that they have been drinking have longer (slower) reaction times than those who were not drinking. In observational research like this, a nonmanipulated categorical variable, is often referred to as the independent variable, but this can lead to confusion. Better practice is to reserve the word "independent variable" for manipulated variables in experimental research, referring to the grouping variables as "**groups**" and the continuous variable as the "**criterion variable**" or "**test variable**."

It is important that you recognize that drinking and reaction time research described above is observational, not experimental. With observational research like this, the results may suggest a causal relationship, but there are always alternative explanations. For example, there may be a "**third variable**" involved here. Maybe some people are, for whatever reason, mentally dull, while other people are bright. Maybe mental dullness tends to cause people to consume alcohol, and, independently of such consumption, to have slow reaction times. If that were the case, the observed relationship between drinking status and reaction time would be explained by the relationship between the third variable and the other variables, without any direct casual relationship between drinking alcohol and reaction time.

As another example of a third variable, consider the relationship between going to the hospital and dying. They are positively related. Does this mean that going to the hospital <u>causes</u> death? I hope not. Hopefully the relationship between going to the hospital and dying is due to the third variable, being very sick. Very sick people are more likely to go to the hospital, and very sick people are more likely to die (especially if they do not go to the hospital).



The demonstration of a correlation between variables X and Y is necessary, but not sufficient, to establish a causal relationship between X and Y. To establish the causal relationship, you have to rule out alternative explanations for the observed correlation. That is, to establish that X causes Y you must show the following:

- X precedes Y.
- X and Y are correlated.
- Noncausal explanations of the correlation are ruled out.

VI. Other Definitions

A. **Data** - numbers or measurements collected as a result of observations. This word is plural. The single is "datum."

B. **Population** – in applied statistics, a population is a complete and well defined collection of measurements, aka scores. A population is often considered to be composed of an infinite number of cases. An example of a finite population is the amounts spent on lunch today by all students at ECU.

C. **Sample** - a subset of population. For example, the amounts spent on lunch today by members of my classes is a sample from that the population defined above.

D. **Parameter** - a characteristic of a population. These are typically symbolized with Greek symbols, such as μ_Y for the population mean of variable Y.

E. **Statistic** - A characteristic of a sample. These are typically symbolized with Roman letters. For example, the sample mean of variable Y can be symbolized as M_Y or $\overline{Y}$, Also called an **estimator**, as these are typically used to estimate the value of population parameters.

F. **Descriptive Statistics** - procedures for organizing, summarizing, and displaying data, such as finding the average weight of students in your statistics class or preparing a graph showing the weights.

G. **Inferential Statistics** - methods by which inferences are made to a larger group (population) on the basis of observations made on a smaller group (sample). For example, estimating the average weight at ECU by finding the average weight of a randomly selected sample of ECU students.

H. **Real Limits** - a range around any measure taken of a continuous variable equal to that measure plus or minus one half the (smallest) unit of measurement. For example, 5.5 seconds means between 5.45 and 5.55 seconds.

I. **Upper case Greek sigma** means sum. $\Sigma Y = Y_1 + Y_2 + ... + Y_n$

VII. Rounding

A. If the first (leftmost) digit to be dropped is less than 5, simply drop all unwanted digits.

B. If the first digit to be dropped is greater than 5, increase the preceding digit by one.

C. If the first digit to be dropped is 5

1. and any non-zero digits appear anywhere to the right of the 5, increase the preceding digit by one.
2. and no non-zero digit appears anywhere to the right of the 5
   a. if the preceding digit is even, simply drop all unwanted digits.
   b. if the preceding digit is odd, increase it by one.
   c. The reason for always rounding to even in this case is so that, over many instances, approximately half will be rounded up, half down, thus not introducing any systematic bias.

D. Examples, rounding to three decimal points
   1. 25.2764 rounds to 25.276
   2. 25.2768 rounds to 25.277
   3. 25.27650001 rounds to 25.277
   4. 25.27650000 rounds to 25.276
   5. 25.27750000 rounds to 25.278

- More on Independent and Dependent Variables