# What is a Variable?

Think of your three closest friends, or three faculty with which you have taken courses. List their names below:

| Name |
|---|
| |
| |
| |

Now, please think of the following features of each person: 1)their sex; 2)their age; 3)their height, and 4)their haircolor (if any). Please place the information in the table below:

| Name | Sex | Age | Height | Hair color |
|---|---|---|---|---|
| | | | | |
| | | | | |
| | | | | |

A variable is something that varies. Age varies in that some people are 18, others are 22, others are 42, and so on. Different people have different heights; thus, height is a variable as well. In our society, about half of the people are male and half are female. Thus, sex varies across people.

In your list of three people above, some of the variables may not vary. Perhaps the three friends you listed all have red hair. In that case, the variable hair color *is* a variable, but in *your* list (your sample) it does not vary across individuals. In short, in *principle* hair color varies (in 2007) but in this particular case (this particular sample) it does not vary.

A Value of a Variable.

You may think of a variable as a question you can ask about a person. What is your sex? What is your age? What is your height? What is your hair color? The answers to those questions will differ for different people. In other words, each person will have a possibly different answer to the question, or, put another way, each person will have a particular *value* in the table above for each question. The full list of all possible answers (all possible values) makes up the logical range of the variable.

What is the logical range in the USA in 2007 for sex? What is the logical range in the USA in 2007 for hair color?

Some variables have very clear logical ranges, while others are murky and unclear. Research proceeds best when rules for dispelling the murkiness are clear, even when the phenomenon is murky. **What makes it best is not that the rules are right but, instead, that they can be identified and assessed**.

Precision of Measurement

We measure with different degrees of precision. There are four levels of precision in measurement:

1)Nominal

2)Ordinal

3)Interval

4)Ratio

*Nominal level measures.* A measure at the nominal level is one in which *names* are assigned to categories, but no relation between the categories can be inferred. For example, if we asked everyone in class to state what religion (if any) that they were taught when they were a child, we might obtain responses such as: 1)Presbyterian, 2)Judaism, 3)Episcopalian, 4)Baptist, 5)Muslim, 6)Catholic, 7)Buddhist, and so on. This variable--call it religious up-bringing--allows us to classify different people. But we cannot say something like "Baptists are closer to Catholics than Presbyterians". We cannot make such statements because the names simply categorize, but they do not summarize a relationship between the categories.

*Ordinal level measures.* Contrast the above example of nominal level measurement with ordinal level measurement. We can ask students in the class to answer the following question:

How much do you like sports?

1)Very much

2)Somewhat

3)Not very much

4)Not at all

The above question allows us to place people in *order* on the basis of how much they like sports. Ordinal level measures allow us to state that some respondents have more, did more, like more, etc. than other respondents. Hence, ordinal level measurement is *more precise* than nominal level measurement.

*Interval level measures.* Ordinal level measures have a weakness; ordinal level measures do not allow researchers to figure out the distance between categories. So, in the example above, the ordinal level measure does not allow us to know *how much more* students who like sports very much like sports compared to those students who like sports somewhat. Students who like sports very much may like sports 5 units more that students who like sports somewhat or only 1 unit more. We have no way of knowing.

Interval level measurement allows us to make such claims. In interval level measure not only does the order matter, but also the distance between categories. So, for example, if we asked students to report how much money they have in their pockets during class, we could state that Jenny, who has $4.57 in her pocket, has $1.23 more than Marcus, who has $3.34 in his pocket. This is possible because a penny is a penny is a penny, regardless of whether it is the first penny you obtain or the last penny obtained by a millionaire.[1] In short, the *units of measure* (in this case, pennies)

---

[1] Of course, people behave differently towards the first penny they get and the last. Part of economics is the study of how people behave per additional unit of money. Still, it is true that in terms of measurement, a

have the same meaning everywhere on the scale. Because interval level measures allow us to infer both the order of groups and the distance between groups, interval level measurement is *more precise* than both ordinal and nominal level measurement.

*Ratio level measurement.* For all intents and purposes ratio level measurement is the same as interval level measurement. Only one difference distinguishes ratio and interval level measures. Ratio level measures have an *absolute zero point*, interval level measures do not. So, in the example above, how much money a student has does not have an absolute zero point, because the student could owe thousands of dollars. In contrast, consider another measure that resembles an interval level measure--temperature. This is the classic example of a ratio level measure. Unlike in the case of money it is impossible for something to have a temperature below absolute zero on the Kelvin scale.[2]

<u>Why consider precision?</u>

Knowing the level of precision of the variable helps us decide what kind of statistics are appropriate. For example, if we asked all 100 students

---

penny is a penny is a penny.

[2]Absolute zero Kelvin equals -460° Fahrenheit ; in other words, -460° F is the absolute end of the scale of temperature. Further, 0 Kelvin is defined as the point at which all molecular motion stops. These facts mean that when the temperature rises from 40° F to 80° F the temperature did not double. In reality, moving from 40 to 80 Fahrenheit is only to move from 277.6 Kelvin to 299.8 Kelvin; however, moving from 40° K to 80° K would be to double the temperature.

in a class to report their annual earnings, and then we added all the earnings up and divided by 100, we would know the average amount of money a student in that class earns. This is a sensible strategy because every penny counts the same.

However, assume we take those same students and ask them to report their race as white (=0), black (=1), Asian (=2), and other (=3). We could not then add up the numbers, divide by 100, and learn the average race of the students in the class. If we found that the average was 1.3, that would *not* tell us that the average student is about 3/10ths of the way between black and Asian! In short, this statistical operation (adding up all the numbers and dividing by the total number of persons) is *not* defensible for ordinal or nominal level variables.

## Concluding Remarks

Variables are at the center of how we think about the social world. Quantitative analyses investigate the relationship between variables, but even qualitative analyses (such as Lareau's book on *Unequal Childhoods*) use variables (middle class versus poor) to identify the study population and to discover how people understand and navigate the world in which they live. The ubiquitous nature of variables makes the concept of a variable essential for understanding and conducting social research.