Operationalizations

A problem I have noticed in student operationalizations in the past is that they tend to be content merely to define a concept by listing characteristics it ought to have in order to be what it is rather than something else. So, for example, a student operationalizes "democratic nation" by saying that a democratic nation would have "free and open elections" and "guaranteed freedom of speech" or even "majority rule."

All that is fine, except that such ideas simply procrastinate. How, for example, will that student measure how "free and open" elections are? Her so-called operationalizations turn out to require operationalization themselves!

Let's start over, by trying some simpler cases, keeping in mind that:

> **Operationalizing a concept means specifying how you intend to detect its presence and measure how much of it is there.**

For example:

1. Operationalize "length."

WRONG: I would use "inches" or feet."

RIGHT: To determine the "length" of an object, I will place one end of this tape at one end of the object, stretch it to the other end of the object, and read the number on the tape which that other end reaches.

2. Operationalize "weight."

WRONG: I will say that a person's "weight" is how heavy he or she is.

SATISFACTORY: I will ask people to tell me how much they weigh, in pounds, and write down the answer.

BETTER: I will have people step on this bathroom scale, and report the number it registers as their weight.

3. Operationalize "honesty."

WRONG: "Honest" people do not lie. I will check that first of all!

RIGHT: I will drop a $50 bill on the street, wait until someone picks it up, and then ask that person if he or she found my "lost $50 bill." People who say "yes," I will score as "honest," while people who deny having found it I will score as "dishonest." And then punch them.

Take a look at this last one. The correct operationalization does not pretend to <u>define</u> what "honesty" really is. Instead, it merely asserts that people who have that quality will generally be observed to have certain other <u>distinct and measurable</u> characteristics: they will return found money to its rightful owner. We will use <u>these</u> qualities as our operationalization. The trick, of course, is to discover what they are.

Now one can certainly argue about whether the proposed operationalizations are good ones—whether they really do measure what we think they do ("validity")—and certainly social and natural scientists have these arguments all the time. Try it: it's fun. But you cannot deny that one <u>can</u> actually <u>measure</u> how many times people admit that a found $50 bill is not really theirs! And that's what we're after: definite, measurable surrogate characteristics.

The natural sciences, quite frankly, have it pretty easy: how difficult is it to operationalize "length," after all, or even "temperature"? Even some of their trickier concepts, such as "lethality" or "electromotive force" are child's play compared to some of the most elementary social science concepts. A physicist once rather breathlessly told me how devilishly difficult it was to understand just how much entropy there was in a system. I merely smiled and wondered exactly how much justice was contained in the Universal Code of Military Justice. Sort of reminds me of the comedian Steven Wright talking about the furious argument he once had at the roulette table, over what <u>he</u> considered to be an "odd" number. . . .

Here's an interesting question: how do I operationalize "operationalization"? That is, when I ask my students to write operationalizations, how do I measure their success? What characteristics do I look for that tell me, "Aha! They've got it! <u>This</u> is a good operationalization!"? Well, in my mind, a good operationalization should have eight detectable qualities:

1. <u>Specificity</u>. The characteristics used should be definite, unambiguous, and specific. Do not measure "just compensation" by referring to "fairness" or something else just as difficult to pin down as the original concept.

2. <u>Objectivity</u>. The measurable characteristics should not depend upon someone's opinion about whether they exist, <u>unless it is discovering opinion that is your goal</u>. If you want to know how much people like President Obama, it is perfectly acceptable to ask them, but if you want to know whether a particular Obama policy is effective, asking people what they think might not be the best approach.

3. <u>Measurability</u>. You should explain the actual physical means of detecting and measuring the above characteristics. Don't just say you want to note whether "freedom of the press" is there: say how you will actually detect it! (e.g., I will classify all opinion articles in the major newspapers according to whether they criticize the government. If at least 20% of them do, I will say that freedom of the press exists).

4. <u>Relevance</u>. There should be a clear explanation of how the presence of those detectable characteristics relate to the operationalized concept. Explain why you think there is a rational link between the number of televisions owned per household and the "standard of living."

5. <u>Proportionality</u>. There should be a clear understanding of how the <u>amount</u> of those characteristics you measure relate to the concept. Does more of the measured phenomenon indicate more of the concept, or less? Does a greater number of newspaper articles about Jews indicate <u>more</u> anti-Semitism in a culture, or less? Is the relationship likely to be linear, or not?

6. <u>Relatedness</u>. If you propose multiple measures or a scale of measures, explain how they interact. Are they scalar? Are you using one to serve as a validity check on another?

7. <u>Validity</u>. "Validity" means the extent to which your measures measure <u>what you want them to measure and nothing else</u>. Acknowledge any obvious weaknesses in this area, and tell how you might try to avoid them. What might contaminate a measure of "criminal activity" which relies on FBI statistics, for example?

8. <u>Robustness</u>. "Robustness" is a measure of how sensitive your operationalization is to the particular circumstances of your investigation. The more sensitive, the less robust. The less robust, the less useful for others.

As I say, these are desiderata for perfect operationalizations and, almost by definition, no operationalizations are perfect. What we are

trying to do it get a good as possible. What? You thought science was easy?

In my mind, I think what I do is sort of score an operationalization on each of the eight qualities above, giving it a -1 if it totally lacks that quality, +1 if it clearly has it, and a 0 otherwise. An acceptable operationalization seems to me to require a score of at least 4, and a really good one should rate a 6 or above.

Try scoring these operationalizations, using the above method. It will help you become clear on the concepts.

1. "For the purposes of this study, we measure 'civil unrest' by counting the number of times per month a front-page article of *The New York Times*, *The Washington Post* or *The Los Angeles Times* reports on a protest gathering. Each such protest or gathering may be counted only once, however."

2. "For the purposes of this study, "long hair" will be considered any hairstyle which touches the collar in back <u>and</u> covers at least half of the ear."

3. From the Boy Scouts of America Lifeguard exam (Requirement 5 of 25): "Starting in the water, swim 20 yards using a front crawl or breaststroke, surface dive 7 to 10 feet, retrieve a 10-pound object, surface, swim with the object 20 yards back to the starting point, and exit the water, all within 1 minute, 40 seconds."