Thoughts re: Teaching about independent and dependent variables.
Molly Schauffler
1/31/15

Teaching students how to distinguish between independent and dependent variables seems to be high on teachers' list of anxieties when teaching about data. It seems to be one of the first decision points emphasized when students encounter a tabular data set: before graphing the data, decide which variable is dependent/independent -- which is of course what we do as scientists.

For students who haven't yet thought about the nature of the relationship between two variables, what if the idea that one variable "changes in response to the other, but not the other way around" arose instead later in the process of data analysis? Suppose it arose out of their reasoning about the data in a graph? Once students graph the data (regardless of which axis they put the variables on) they can see first if the variables *are* correlated. If there is correlation, students might then be prompted to consider if one variable might be changing in response to the other, but not the other way around.

Until students have opportunity to "own" this distinction through describing, interpreting, and reasoning about their graph and their data, they may not latch onto words that describe it.

Here's an example of prompts that might deepen students' thinking and help them realize the distinction between independent and dependent variables. Data are the amount of current (amps) produced by a solar panel under different amounts of sunlight (lumens). The student (against convention) put the predictor variable on the Y-axis, but the relationship is still evident. (Data from Chris Olson's daughter, accessed via Fathom Dynamic Data software data library).
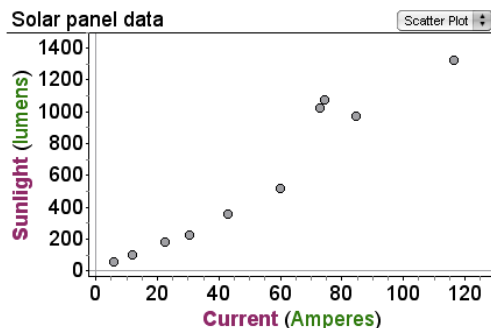

Prompts:

- *What most likely makes the amount of sunlight change?*
- *What most likely makes amps change?*
- *Which variable most likely affects the other variable?*

- *What will happen to sunlight if Watts changes?*

- *What will happen to Watts if sunlight changes?*

- *Can one variable be used to predict how the other will change? Which one? What is your reasoning?*

- *What made you decide to put sunlight on the Y-axis?*

By prompting students to ponder such questions while they are looking at their graphed data, they may begin to grasp the one-sided relationship that exists between the two variables. When they do, need arises for words that make that distinction (especially if they are going to write their Reasoning section). They may even adopt their own words for it. Once the predictor-responder / independent-dependent relationship is understood and practiced, the convention of putting the predictor on the X-axis is a simple thing to remember. Axes can easily be switched if needed. That there *is* a correlation is evident in the graph regardless of which variable is on which axis – that seems like the first cognitive step to take.

One pitfall we've seen to putting heavy initial emphasis on identifying independent and dependent variables: sometimes the relationship between two variables is neither independent nor dependent, such as a data set for arm-span and heights of students in a class.

Person	Arm span (in)	Height (in)
1	70	70.5
2	64	66
3	66.25	65.75
4	61.5	65
5	66	66.5
6	65	65
7	67	67.5
8	64	65
9	66	70
10	72	69
11	75	72
12	68	67
13	62	63

INT 188 Arm span data — Scatter Plot
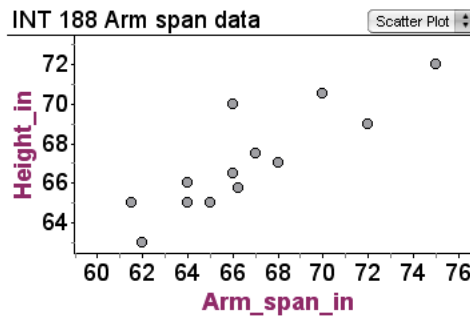
Arm span and height are strongly correlated, but neither one is dependent on the other, and yet (interestingly) either one can be used to predict the other!

In summary, I am wondering if putting less emphasis on independent/dependent variables at the front end of data analysis, and emphasizing it more on the reasoning end, once students are cognitively ready and in a position to contemplate the data while looking at a graph, will help alleviate the brain jam many students seem to have about this concept.