Hot and Cold: Whose ruler is it?

In educational measurement, we don't yet know if we are measuring heat *or* temperature.
A man with one watch knows what time it is; a man with two watches is never quite sure. Lee Segal

Meaning comes from experience and experience comes from ignorance. You learn what hot means by touching the stove; you learn what cold is by not wearing your mittens. Nothing here answers the question of where cold ends and hot begins; that point is subjective, arbitrary, and personal; hopefully not capricious.

Temperature is one of the first lessons we learn: what things are too hot to touch? When is the weather too cool to not wear a jacket? When is it warm enough to go barefoot? How much fever warrants staying home from school? These concepts may define meaningful temperature levels, but "*because mom says so"* is not very objective, and definitely not measurement.

Eventually, some very clever scientists realized you could do a better job of monitoring temperature by measuring the length of a liquid in a glass tube than you could with the venerable, and more direct, hand-on-the-forehead method. It took a little longer for the science to explain why this less direct approach worked; i.e., to provide a theoretical validation by describing the underlying mechanism, moving beyond (but not replacing) empirical validation like when the liquid is longer, I feel hotter.

If you have the only liquid-filled tube in the world and want to use it to monitor the ambient temperature of your house, make some marks along the tube, count the number of marks below the level of the liquid every day, and you're done. Eventually you'll come to understand that if the level is below, say, eight marks, you want to add another log to the fire or if it is above the 12^{th}, to open a window. You have a perfectly valid thermometer for the purpose.

If you want to do a little more, perhaps answer the question is it warmer (or colder) in Minneapolis or Novosibirsk, you need a liquid-filled tube in each location and a process to *equate* them. The equating always requires a *link* of some sort between the two instruments; obvious candidates for the linking are the *common-place* and *common-law* methods.

With the *common-place* approach, the tubes are placed side-by-side. The tubes could be different sizes and contain different liquids, within limits. To equate, the lengths of the liquids are marked, observed over a range of temperatures. We can label the marks in some nice, orderly way and have two serviceable and equated thermometers that we could carry about the countryside. We still wouldn't know what any of the marks actually mean but we would know which mark on the first corresponds to which mark on the second. We know this because we "calibrated" the tubes at same time in the same place, ensuring identical conditions.

And there is an alternative.

With the *common-law* approach, the tubes could be in different places, say Minneapolis and Novosibirsk or anywhere else the laws of physics apply. We just need to agree on some standard conditions and how to create those conditions independently. One might, for example, define one as the coldest you can make salt water and have it stay liquid and, perhaps, another condition as "normal" human body temperature. For convenience, we might call these two points zero and 100. This is more or less what Fahrenheit did, but the final scale got tweaked a little between then and now; he may not have been feeling well when he did this. Or we might do what Celsius did, which was to use the point where water changed from liquid to solid and the point where water changed from liquid to gas, and also called them zero and 100, and again it got tweaked a little[1]. Since anyone can replicate either set of conditions, anyone anywhere could create and calibrate a thermometer and confidently compare scale scores across time and distance.

Either the *common-place* or *common-law* strategy could work, even if the tubes were different sizes or contained different liquids, or if methods other than the liquid-filled tube were used if we agree on the location or control the conditions. But there are still some issues. Depending on the liquid, there are upper and lower bounds for the temperatures that can be monitored without exploding the glass. Temperatures outside those limits require different, perhaps very different methods. However, we can equate any two valid instruments regardless of the approach to measurement taken, if the ranges overlap somewhere and, by induction, we could equate any number of valid instruments in an overlapping sequence.

To *equate* here means to connect the scales so that the scores are interchangeable; measurements from one instrument can be compared to, i.e., subtracted from measurements from another. *Equated* does not imply that the instruments are created equal or that they are always appropriate for the same objects. One might try to control the manufacturing of instruments so closely that they are physically identical so they require no equating. This is standard operating procedure in the manufacture of ordinary thermometers today but it is not how Fahrenheit or Celsius started and is problematic, unnecessary, and maybe undesirable, in the manufacture of alternate forms of cognitive assessments.

Whether you are in Minneapolis or Novosibirsk, you may want to wear a sweater today.

---

[1]Zero on the scale that bears the Celsius name is now the *triple-point* of water, which is slightly different than the freezing point but no one but a physical chemist knows or much cares.