

**OPINION PIECE:**

## **Face validity: Is there a place for this in measurement?**

Dennis M. Roberts (Penn State University Dept. of Educational Psychology)

Recently, on the LTEST-L listserv, there was a debate about the concept of "face validity" (FV) and whether this was a notion whose time has passed, or, something that still has usefulness within the field of measurement. I have come down on the side that FV still can be useful when we are faced with decisions about what instruments to use in various assessment situations. With caution of course!

In short, the concept might be boiled down to a simple definition:

***"FV is making a decision about the appropriateness of use of some particular measuring instrument in a given assessment situation through the process of simple inspection OF that instrument."***

At the core of FV is the idea that ONLY simple viewing of the instrument is done in order to render this judgment. However, this leaves open the question as to WHO does the viewing: is this person a novice, with no real knowledge about the content field or, does this person have some expertise in the content field? Perhaps the typical view (according to some texts) is that it is a non-expert doing the viewing and making the judgment. A typical version of this might be Davis (1964), who described face validity as a "... casual inspection, especially by examinees or laypersons" (p. 34).

However, we do NOT have to define FV exactly that way and, I for one, do not. The main issue that differentiates this kind of validation from other types (predictive, etc.) is the fact that the judgment about the appropriateness of the instrument is made by inspection only, with little or no reference to any other kinds of information. To me, the unnecessary part of the usual FV definition is the additional constraint that the one doing the examining is a novice with respect to either the content or knowledge about measurement. My view is that the concept of FV is useful IF we stick to the original definition that I offered at the beginning. Certainly if the person who is in the position of rendering the judgment about the instrument is a real novice, the usefulness of FV judgments will be minimal at best (though at times still helpful as I will suggest in a moment). And of course, at worst such judgments could totally misrepresent the utility of using an instrument for some assessment purpose.

One situation where even a novice (say an examinee) can provide a useful perspective might be the following. On most tests, we are hoping that the examinee will give it their best shot. But,

what if a particular examinee, who is taking a test, were to perceive that this test contained items that seemed NOT to be that appropriate for the assessment being undertaken (even if the items were in a psychometric sense). Would it not follow that such an examinee might NOT assert him/herself to the maximum given the perception that the measurement will not really tap important information about his/her ability? In a case like this, the score would likely be lowered and, the "validity" of the score on the test will have been compromised. Thus, even a cursory inspection by real or potential examinees could provide useful information to someone making a decision about either using or not using some specific instrument.

Therefore, if we can accept the notion that FV can be defined as simple inspection of an instrument and NOT further assume that the examiner is a novice (though they might be), then what are a few possible benefits of doing a FV inspection? Here are three possible benefits of a good FV:

1. Given that there are dozens and dozens of instruments that might be used in some particular situation, a good FV can narrow down the domain to a subset, which would then have a more likely potential for being useful.

2. Given the intended purposes of some particular assessment need, a FV can provide information about caveats that need to be adhered to, and/or ways in which the use of an instrument could be augmented so as to make the overall assessment more meaningful to decision makers, students, teachers, and others.

3. And finally, a good FV investigation can provide (in the case where a decision has to be made quickly and without time to do more thorough documentation), some "evidence" as to whether the instrument will fulfill the intended purpose or not.

Thus, I would like to take the side that argues FV has a legitimate place in assessment, and making decisions about the potential use (or improvement) or NON-use of instruments for particular purposes. In fact, I would go out more on a limb and say that FV is an essential part of this entire process generally ASSUMING of course that the person doing the FV work has some content knowledge and decent understanding of the principles of measurement (but that might not always be the case). Therefore, if we can shed the baggage that the definition of FV seems to have attached to it being done by novices or laypersons, then the concept of FV deserves respect within the overall assessment process. I am NOT claiming that FV done by just anyone can be useful but, the converse that NO FV can be helpful seems to be equally extreme.

## References

Davis, F. B. (1964). *Educational measurements and their interpretation*. Wadsworth Publishing Co.: Belmont, California.

**HTML:** [http://www.jalt.org/test/rob\\_1.htm](http://www.jalt.org/test/rob_1.htm) **PDF:** <http://www.jalt.org/test/PDF/Roberts1.pdf>