

- SUSIE:** Hello, Dr. Metric. I am looking forward to talking with you today about test validity.
- DR. METRIC:** Good, Susie. Test validity is a very important and interesting topic.
- SUSIE:** What is test validity, Dr. Metric, and why is it important?
- DR. METRIC:** There are many technical definitions of test validity, but a simple definition is-- the extent to which a test measures what it claims to measure. Test validity is important because if a test does not have validity, you cannot be confident that the test scores are accurate or useful for making decisions.
- SUSIE:** It certainly does seem to be important and simple.
- DR. METRIC:** It is really not that simple, Susie. When we examine what is really meant by test validity, we see a much more complex picture. We need to look beyond the simple statement, measuring what it claims to measure. There is much more to test validity.
- SUSIE:** What do you mean, Dr. Metric? What else does it mean other than ensuring that the test actually measures what it claims to measure?
- DR. METRIC:** Well, Susie, we are concerned with more than the test itself. We want to know how accurate the test scores are and whether the inferences or decisions made on the basis of the test scores are legitimate.
- SUSIE:** That sounds complicated, Dr. Metric. What do you mean by the inferences are decisions made based on the test scores?
- DR. METRIC:** It is not really that complicated, Susie. We often use tests to draw inferences or make a decision. For example, we often infer whether or not a student is proficient based on a test score. We might say, Sam received a score of 72 on a reading test and, from that, infer that he as proficient in reading. Or we might take Sam's score of 72 on that reading test and decide to place them in a remedial reading program.
- SUSIE:** Yes, Dr. Metric, but what does making an inference or decision have to do with validity?
- DR. METRIC:** That is a very good question, Susie. Think of it this way-- if we are going to draw inferences or make decisions based on the test, we really want to have evidence that proves that those

inferences or decisions are legitimate. This is what modern approaches to validity are all about.

SUSIE: I understand, Dr. Metric. But what kind of evidence would we need to support the decisions or inferences?

DR. METRIC: There are many kinds of evidence that we should have to support our inferences and decisions. The specific types of evidence depend on the situation. Every situation is different. We will leave a detailed discussion about types of evidence for another time. But there are some common types of evidence we can talk about.

SUSIE: Types of evidence?

DR. METRIC: Yes, there are several common types of evidence. The first type of evidence is often called content validity. You must first clearly define the knowledge and skills you believe your test measures. To gather this evidence, we would first clearly and specifically define what we are measuring on the test. We might then support that by having experts in the content included on the test systematically review it. It would also be a good idea to review the literature to find evidence from prior research, supporting what we are measuring.

SUSIE: Dr. Metric, what other types of evidence can be collected to support test validity?

DR. METRIC: Another type of evidence that is often talked about in testing is construct validity. Construct validity covers a wide range of types of information that show whether or not the test is performing in ways that would be expected, based on how we have defined the construct it is measuring.

SUSIE: What are some examples of construct validity evidence?

DR. METRIC: Well, Susie, we might want to do some research on how the scorers on the test we are validating compared to scorers on similar tests. For example, if we have a reading test, we would expect students taking our test to get scored similar to those on other reading tests. We might also want to see if scores on our test predict some future performance. For example, if we made a good reading test, then our test should predict future behaviors that require reading skill. Perhaps we would collect evidence that students who scored highly on our reading test did better when they go to college in later years.

SUSIE: Dr. Metric, what are other sources of evidence?

DR. METRIC:

We should probably leave the discussion for another day when we have more time. I am hoping that from today's discussion you understand the importance of test validity and understand that test validity is a process of collecting evidence from various sources to support the inferences and decisions you're making on the basis of your test.