



Catalyzing Leaders for Health and Impact

Indexes of a Psychometrically Sound instrument:

The LeaderSource 5C Checkpoint instrument is a psychometrically sound instrument. This means that it conforms to the industry standard of being both *valid* and *reliable*.

An instrument is considered *valid* when there is a due process developed and followed to ensure the constructs purported to be measured are actually being measured. There are a number of methods for substantiating validity of an instrument.

The method used on this instrument was to give a team of fifteen to twenty "subject experts" the individual items and allow them to assign them to one of the various constructs being assessed in the instrument. When the results were collected from the subject experts, calculations were performed to determine whether a high percentage (close to or over 90%) of the respondents assigned the item to the intended construct. If the item did not meet such a high standard, then it was eliminated or rewritten, and another iteration was performed following the same procedure. The instrument is only considered valid when all of the items on the instrument meet this high standard.

An instrument is considered *reliable* if a respondent can take the instrument today, and then again after a short interval, delivering the same results each time. Another measure of reliability is that there is high consistency within the instrument itself; the first is called *external reliability* and the second is called *internal reliability*. As with validity, there are various methods for determining this.

The method selected for this instrument uses the statistical index called Cronbach Alpha (CA). CA is a measure of the internal reliability of a scale. Basically, it measures how a certain scale of the instrument is internally consistent. This would mean that if there are five items on one scale, all measuring the same thing (validity), then if one item is rated high, so should the other four items. The same thing would apply if rated low or somewhere in the middle. CA measures how consistent these items are to each other. For affective domains, CA is acceptable above .60 and for cognitive domains, CA is acceptable above .80.

The only deviation from the standard measures of validity and reliability we used for the LeaderSource instrument is that there are no internal scales. In other words, every item was considered a scale and these items were not summed up in any way to create a score.



Catalyzing Leaders for Health and Impact

Therefore, for validity purposes, we looked at how the subject experts assigned the items to one of the five Cs (Christ, Character, Community, Calling and Competencies).

However, when approaching reliability, we designed the process to include approximately 10 individuals who would respond to the same item as it relates to the same individual. These scores were then processed in the same way as standard reliability scores would be. In other words, was there consistency across the 10 respondents on the same item as it related to the same individual? If not, why not? In effect, this was the same as assigning 10 items to one scale and seeing if all of the items were consistent. We developed a rigorous methodology in advance of determining whether or not there was consistency.

Each of the items in the final instrument met or exceeded these standards.