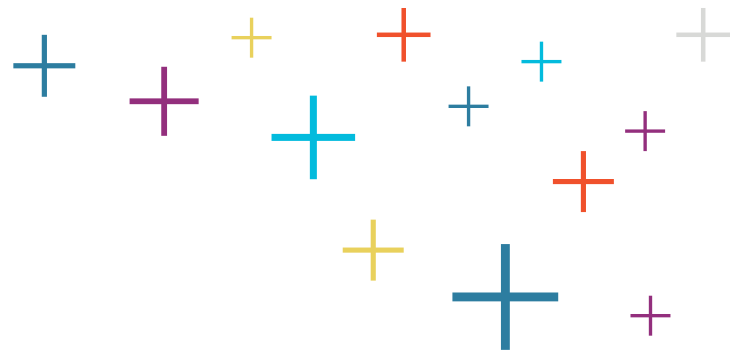
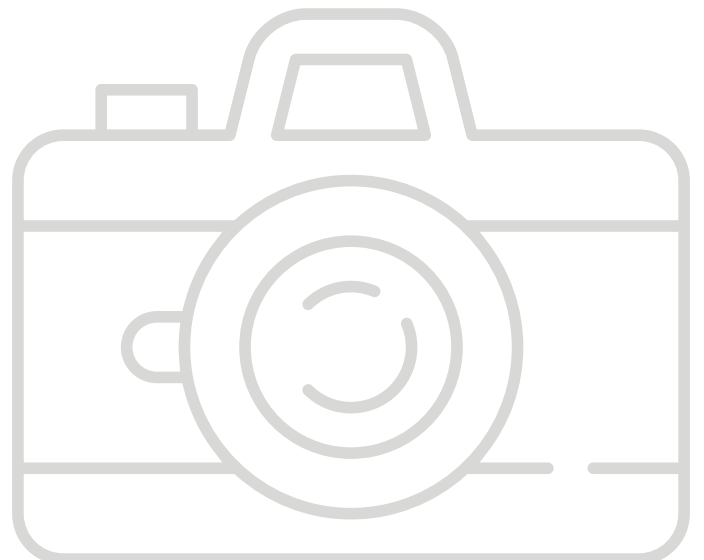
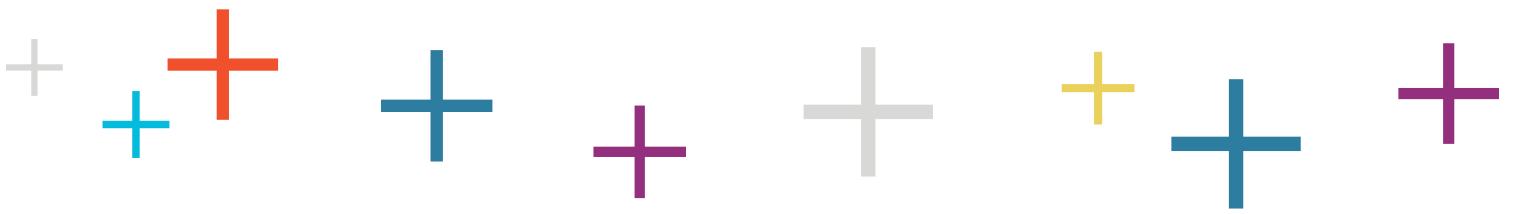


The  
**Enrollment  
Management**  
Association



Validity Evidence for  
**The Character  
Skills Snapshot**

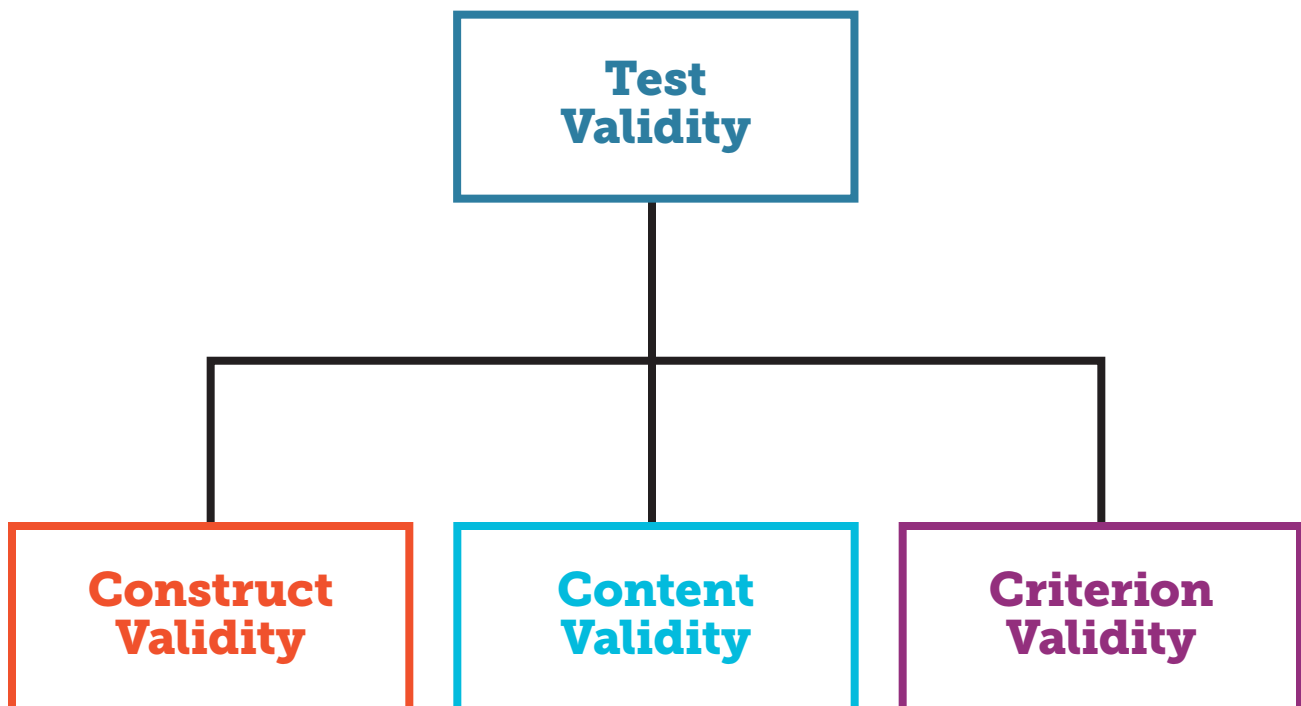


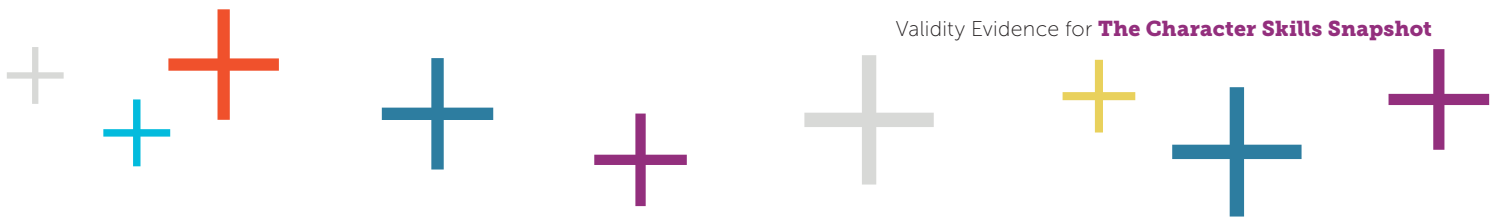


# Test Validity

Test validity refers to the degree to which evidence and theory exist to support the interpretation of test scores for particular purposes (AERA, APA and NCME, 2014). It is important to note that we validate a test score for a particular use (e.g., admission, placement), and that validity is not the property of a test in and of itself. This means that as opposed to talking about a test as simply valid or not valid, one should instead state, for example, “There is a great deal of validity evidence to support the use of The Character Skills Snapshot scores for independent school admission and/or placement decisions.” This also represents the notion that validity is a matter of degree and not absolute. It is therefore very important to gather validity evidence over time to either enhance, confirm or contradict previous findings.

Validity is a unified concept, yet there are multiple types of validity evidence that may contribute to the demonstration of validity. These types of evidence may include evidence based on test content—the relationship between the content of a test and the construct or characteristic it is intended to measure (Content Validity). Evidence based on internal structure refers to the degree to which test items and test components conform to the construct (e.g., score reliability, relationship of items and test dimensionality, etc.) (Construct Validity). Another type of validity evidence is based on relations to other variables, such as measures of criteria that the test is expected to predict (e.g., First Year GPA) or relationships to other tests measuring related constructs (Criterion Validity).





Ultimately, the validity of The Snapshot depends on how it is used by individual schools. As with the SSAT, schools are encouraged to conduct their own validity study whenever possible. The Enrollment Management Association provides a validity study service to all its members to assist in this endeavor.

Studies to evaluate The Snapshot validity are ongoing. Here is some evidence that has already been collected to support the validity of The Snapshot:

- + Content Validity Evidence:** The situational judgment items were written by teachers from independent schools in order to draw from real experiences. Teachers were given detailed instructions, numerous example items, and complete descriptions of the constructs in order to produce valid item content. Edited items were then administered to students as part of think-aloud sessions to ensure that students perceived the situations and associated responses as realistic and plausible. The final set of items went through a fairness review by Educational Testing Service as a final check on threats of bias. ETS edited content as needed to ensure cultural fairness.
- + Content Validity Evidence:** The statements used to develop the forced-choice items were created by research staff who had previous experience writing items for character assessments. A technical advisory group comprised of several experienced researchers reviewed the forced-choice items. As with the situational judgment items, the final set of forced-choice items went through the ETS fairness review process.
- + Criterion Validity Evidence:** Students with high intellectual curiosity scores are more likely to self-report that they want to obtain a master's or doctoral/professional degree. This is important because pursuit of graduate-level education suggests an interest in more in-depth and focused learning, a key characteristic of intellectual curiosity.
- + Criterion Validity Evidence:** Students with high initiative are more likely to self-report that they are never late for class. This is important because it suggests that students who are self-motivated won't miss opportunities to learn by being late for even a part of the classroom experience.
- + Criterion Validity Evidence:** Student scores on The Snapshot have weak relationships with SSAT scores. This is important because it suggests that a student's academic ability does not depend on the student's character (or vice versa). The Snapshot measures life skills, therefore it captures characteristics about the student beyond what she or he experiences or learns in school alone. This is also consistent with existing research looking at other noncognitive characteristics and the relationship with GPA and academic performance (Proapat, 2009; Crede, Tynan & Harms, 2016).

*Continued*



**+ Construct Validity Evidence:** When we look at the relationships between the constructs on The Snapshot, the relationships that we believe to be theoretically smaller are (teamwork and intellectual curiosity) and the relationships we believe to be larger are (e.g. responsibility and self-control). This is important because it suggests the relationships between the 8 character skills are consistent with theoretical frameworks of psychological research.

**+ Construct Validity Evidence:** Multiple-group analyses with The Snapshot data revealed that the constructs measured by the forced-choice and situational judgment components of The Snapshot are expressed in the same way for all groups compared. This includes comparisons of ethnic/racial groups (Hispanic/Latino, black, white, Asian, and other), gender groups (girl, boy), grade levels (middle grades, high grades), student type (enrollee, applicant), and pilot test (fall 2015, fall 2016). This is important because we would not want to compare student performance on an assessment if systematic (i.e., by-group) differences in student responses are causing a group to respond differently to certain items. This ultimately changes the meaning of the construct for that group, rendering comparisons with members from other groups invalid because we can no longer be confident that we are measuring the same thing for every group.

## References

- Crede, M., Tynan, M. & Harnes, P. (2016). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, 113(3):492-511
- Poropat, A. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135, 322–328.