

THREE KEY ELEMENTS OF A STRONG CONCORDANCE

September 2023



CLT

CLASSIC LEARNING TEST

CLTEXAM.COM

Introduction

In April 2023, Classic Learning Test published a new concordance table between the CLT and the SAT. This brief intends to summarize the key features of a strong concordance study and to help stakeholders locate these features in our report. The key features of a concordance study are:

1. Construct similarity between the concorded tests;
2. High and similar test reliability across the concorded tests;
3. A sufficiently large and representative sample.

1. The concorded tests measure similar constructs.

“The strength of the relationship between the two measures determines whether we can successfully estimate the score on one test using the score on another” (Dorans & Walker, 2007, p. 198).

For a concordance relationship to allow meaningful comparisons between two tests, the tests that are concorded must measure similar constructs (e.g., skills, abilities, etc.). For example, a concordance table between math scores and shoe size would be meaningless because math ability and shoe size are (presumably) unrelated traits. Construct similarity is established by evaluating the alignment between the tests’ content and by examining the correlations between the scores. Dorans (2004) noted that “High correlations, which are helped by high reliability, are a necessary and sufficient condition for useful concordances...” (p. 244). Correlations range between 0 and 1, with 1 indicating a perfect relationship. Dorans and Walker (2007) suggested that a correlation of 0.866 is needed to estimate a concordance relationship between two tests.

- ❖ Our study begins with a section on content alignment and shows that **the CLT and the SAT measure very similar skills**, including reading comprehension, grammar, algebra, and mathematical problem solving. The fact that CLT demands a deeper literacy and numeracy does not mean that it measures skills different from reading, writing, and mathematics.
- ❖ On the following page, Table 1 shows the correlations between the CLT and the SAT for each section, **demonstrating that the two tests measure highly similar constructs.**

Table 1. The Correlations Between CLT and SAT Scores

Section	Correlation
CLT Total – SAT Total	0.89
CLT VR + GW – SAT EBRW	0.90
CLT QR – SAT Math	0.87

2. The concorded tests have high and similar levels of reliability.

“Poor measures of anything are even poorer measures of each other” (Dorans, 2004, p. 236).

The reliability of a test determines the amount of measurement error in the test scores. More reliable tests produce more consistent and precise scores. A concordance table between unreliable tests would be unreliable as well. Like correlations, the most commonly used reliability indexes take values between 0 and 1. Suggestions of minimum values range between 0.70 and 0.95 in the psychometric literature (Tavakol & Dennick, 2011). That said, reliability is affected by test length, and a maximum reliability of 0.90 has been recommended since higher values may indicate that the test has redundant items (Tavakol & Dennick, 2011).

Table 2 shows the reliability of each CLT section for the most recent CLT base form along with the range of reliabilities reported in the SAT Technical Manual (College Board, 2017). Although the concordance study combined the Verbal Reasoning (VR) and Grammar & Writing (GW) scores to be linked to SAT’s Evidence-Based Reading and Writing (ERW) scores, reliabilities are reported for each section separately because the reliability indexes used by the two organizations are most comparable at this level.

- ❖ Both the CLT and the SAT have reliabilities in the 0.85-0.92 range, which means that **both tests are highly reliable** (CLT, 2023; College Board, 2017).
- ❖ CLT uses Item Response Theory and automated test assembly procedures in test development to **ensure that different test forms measure the same content** at a consistent level of difficulty (CLT, 2023).

Table 2. Reliabilities of CLT and SAT Sections and Tests

Section / Test	CLT	SAT
CLT VR - SAT R	0.86	0.88-0.89
CLT GW - SAT WL	0.85	0.84-0.89
CLT QR - SAT MTS	0.86	0.90-0.92

3. The sample used in the concordance study represents the population of students who will use the concordance table and is large enough to produce precise results.

“... every attempt should be made to collect a concordance sample that is as similar as possible to the population to whom the results will be applied” (Pommerich, 2007, p. 204).

Sample size and sample representativeness serve two important but distinct aims. A large sample minimizes the random error introduced by the sampling process and increases the precision of the concorded scores. That is, larger samples will lead to a smaller margin of error in the concorded scores. Kolen and Brennan (1995) suggest that a sample size of 1,500 provides sufficient precision for the equipercentile linking method used in our study.

On the other hand, a representative sample minimizes systematic error – also known as bias – and ensures that the concorded scores do not systematically underestimate or overestimate the true concordance relationship that holds in the target population. For the concordance results to be unbiased, the study sample should represent the population of students who will actually use the concordance table (Pommerich, 2007).

- ❖ To minimize random error, our study used a sample of 4,375 students for the total scores and 1,551 for the section scores, **exceeding the sample size recommended** by Kolen and Brennan (1995).
- ❖ To minimize bias, our study incorporated data from the two groups of students who are the **most likely to represent the future users of our concordance table**: a) students who have actually taken the CLT in the past; b) public and charter school students who will likely form a larger proportion of our test takers in the future. To include the latter group in the study, we organized a special CLT administration in March 2023 which provided 435 official CLT and SAT scores from public school students.

Conclusion

While concordance studies vary in their sample size, target population, and statistical techniques, there are clear, agreed-upon principles that establish the validity of any effort to link two tests using concordance. This brief was meant to outline those principles, and to show that each of them was attended to in the CLT-SAT Concordance Study conducted this spring.

Specifically, we discussed three key elements of a strong concordance: construct similarity between the concorded tests, high test reliability, and a large, representative sample:

1. Construct similarity was demonstrated by high correlations between the two tests;
2. High test reliability was established by the technical reports published by the two organizations;
3. The study sample included both past CLT test takers and public school students to represent the students who are likely to use the concordance table, and the sample size exceeded the recommendations in the psychometric literature.

In short, these findings verify that the concordance between the CLT and the SAT meets industry standards for validity and reliability.

References

- College Board. (2017). *SAT Suite Assessments Technical Manual*.
<https://satsuite.collegeboard.org/media/pdf/sat-suite-assessments-technical-manual.pdf>
- Dorans, N. J. (2004). Equating, Concordance, and Expectation. *Applied Psychological Measurement*, 28(4), 227–246. <https://doi.org/10.1177/0146621604265031>
- Dorans, N.J., & Walker, M.E. (2007). Sizing Up Linkages. In N.J. Dorans, M. Pommerich, & P.W. Holland (Eds.), *Linking and Aligning Scores and Scales* (Statistics for Social and Behavioral Sciences). Springer, New York, NY.
- Kolen, M. J., & Brennan, R. L. (1995). *Test equating methods and practices*. Springer.
- Pommerich, M. (2007). Concordance: The Good, the Bad, and the Ugly. In N.J. Dorans, M. Pommerich, & P.W. Holland (Eds.), *Linking and Aligning Scores and Scales* (Statistics for Social and Behavioral Sciences). Springer, New York, NY.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>