



# ONE

## INTRODUCTION

### 1.1 INTRODUCTION TO THE CLT<sup>10</sup>

**T**HE CLASSIC LEARNING TEST (CLT) was introduced in 2015 as an alternative standardized test to the SAT and ACT. It largely aligns with a “Great Books” curriculum, one that emphasizes the humanities alongside STEM subjects and prefers to use original sources that anchor students with a passion for accurate learning in every field. This aids families, students, and schools who (in part or in full) favor this type of curriculum over Common Core, helping to foster greater educational choice.

The CLT10 was developed as the official preparatory exam for the CLT, comparable in role to the PSAT or to the ACT Aspire; it was first administered two years after the CLT entered use. The present norming study was conducted on the CLT10, to help ensure its validity for parents, students, educators, and lawmakers.

### 1.2 INTRODUCTION TO NORMING

**I**N EDUCATIONAL MEASUREMENT, SCORES CAN BE interpreted in either a norm-referenced or a criterion-referenced framework. For a norm-referenced interpretation, a student’s test score is compared with the normative or reference group, which is the population the test targets, and student performance is described in terms of the relative standing in the target population. For a criterion-referenced interpretation, scores are interpreted relative to content standards: this kind of interpretation tells what a student can do or cannot do in terms of the content standards. Normative information can be used to enhance the interpretability of test scores. When developing score scales, normative or content-related information is regularly built into the scale (Kolen & Brennan, 2004). The process of incorporating normative information into scale scores provides a means of aiding interpretation of the scores (Gardner, 1962).

College admission tests such as the Classic Learning Test (CLT), SAT, and ACT are used to select high-profile candidates. It is generally expected that an individual’s scores are compared with the distribution of scores

for one or more reference groups, to derive useful inferences about the person’s performance relative to other candidates. Test scores based on such comparison with a targeted national population are norm-referenced (*Standards for Educational and Psychological Testing*, American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014). Such norm-referenced information can be provided based on norming studies.

“The validity of norm-referenced interpretations depends in part on the appropriateness of the reference group to which test scores are compared” (*ibid.*, p. 97). The representativeness of the sample for norming studies is crucial to the interpretation of the results.

### 1.3 OVERVIEW OF CLT<sup>10</sup> NORMING SAMPLES

The CLT launched in December 2015 as an alternative to the SAT and ACT for college admission purposes. To measure student performance and track their progress for college-readiness, CLT introduced the CLT10 in 2017. The CLT10 is the official preparatory exam for the CLT, designed for 9th- and 10th-grade students. The CLT10 is developed to measure content similar to the CLT at an age-appropriate level and with fewer questions of the highest difficulty. Students may take the exam on their own laptop or tablet and receive their scores on the same day (<https://www.cltxam.com/products/clt10>). The CLT10 reports one total score and three subdomain scores: Verbal Reasoning, Grammar/Writing, and Quantitative Reasoning.

Users of CLT test scores also value norm-referenced interpretative information. Several states have requested a comparison between test scores on CLT exams, including the CLT10, and the corresponding national population of the non-district educated (NDE) programs (e.g., private schools, charter schools, and home schooling), to determine the academic achievement of CLT examinees and fulfil the requirements of state regulations. Thus, a norming study is needed to obtain information for users of the CLT suite of exams.

The purpose of the current norming study is to provide information about the performance of students in grades 9 and 10 on the CLT10, relative to the corresponding national population. More specifically, this consists of two sub-studies. The first one is a national norming study that provides information about CLT10 scores, referencing to the national NDE population. The second is a linking study, mapping the relationship between the CLT10 and the PSAT. This study provides a concordance Table that shows the correspondence between the scores on the two tests. After such mapping, the national normative information from PSAT can be used to further contextualize CLT10 scores.

### 1.4 PROCEDURES

In *Standards for Educational and Psychological Testing*, the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education clearly outline procedures for norming. “Norms, if used, should refer to clearly described populations. Reports of norming studies should include precise specification of the population that was sampled, sampling procedures and participation rates, any weighting of the sample, the dates of testing, and descriptive statistics. Technical documentation should indicate the precision of the norms themselves” (2014, p. 104).

The general procedures for the current study involve defining the reference population, selecting the appropriate methodology for data analysis, quality assurance of data, and quality control of data analysis. The results of these studies are summarized and discussed in this report. Each of the above facets is addressed and presented for both studies. At the end, a summary drawn from both studies is provided, and limitations in utilizing the study results are discussed.

Different techniques can be used to normalize scores, including linear or non-linear score transformation. As Kolen & Brennan (2004) indicate, percentile ranks for various groups of examinees are a non-linear transformation often used for national norming as an auxiliary score scale. In the transformation process, the distance between score points is compressed in the middle of the distribution and expanded at the upper and lower ends. They emphasize the importance of estimating score precision, reliability, and standard error of measurement, to support any interpretations.

The samples used in both the norming and linking studies were reviewed and multiple filters were applied to ensure that the sample of examinees were all from the NDE population, currently enrolled in grade 9 or grade 10, received a valid score on the CLT10, and took the online CLT10 (as distinct from the paper exam provided in exceptional cases).

For the national norming study, the target population was defined for the CLT10 sample of examinees using the 2016 national survey by the National Center for Education Statistics for private and home schools. The CLT10 sample was weighted based on several demographic variables, such as gender, race/ethnicity, and geographic location, to match the characteristics in the target population. Percentile ranks were developed from the normative sample.

For the linking study, a subsample of examinees who received a valid score on both the CLT10 and the PSAT10 were included. The statistical linkage was performed using the equipercentile approach, with post-smoothing by LEGS 2.0 (Center for Advanced Studies in Measurement and Assessment, University of Iowa). Various supporting statistics were provided from the analyses, such as effect size, standardized mean difference, linear conversion, and parallel conversion, as well as Standardized Root Mean Square Difference (RMSD) and Standardized Root Expected Mean Square Difference (REMSD) for the technical quality of the statistical linkage. The follow-up extrapolation was conducted through different approaches (e.g., linear, polynomial, and exponential). The results of those analyses were compared to settle on an appropriate method to extend scores in the lower and the higher ends along the scale. The process of sample selection and norm development was carefully reviewed for quality control.

“To be meaningful any test scores must be related to test content as well as to the scores of other examinees” (Ebel, 1962, p.18). Previous national efforts, such as the NAEP mapping study and scale anchoring, are good examples. The study of content correspondence between the CLT10 and the PSAT, included in this report, supports the mapping of CLT10 scores to PSAT scores presented in Chapter 3.