



CLT

2022 NORM REFERENCE REPORT

The Classic Learning Test

SEVENTH AND EIGHTH GRADE



2022 NORM REFERENCE REPORT
The Classic Learning Test

SEVENTH AND EIGHTH GRADE

CONTRIBUTORS

The following individuals helped to create this technical report. We greatly appreciate all their good and thoughtful work to provide this reference on the Classic Learning Test.

<i>Technical Manual Lead</i>	Tracy Gardner, Ph.D.	<i>Writers/Reviewers</i>	Gabriel Blanchard
<i>Psychometrics Lead</i>	Hong Jiao		Stephanie Beattie
<i>Chapter Leads</i>	Stephanie Beattie		Natasha P. Wilson, Ph.D.
	Tracy Gardner, Ph.D.	<i>Leadership Reviewers</i>	Tracy Gardner, Ph.D.
	Natasha P. Wilson, Ph.D.	<i>Graphic Designer</i>	Meg Prom

Letter from the President

I can't say I've always dreamed of starting a standardized testing company. As a boy, I hoped one day to become a football player, a cowboy, or a police officer. These professions seemed to offer adventure and fulfillment whereas my experience in public school, grades K-12, felt tiresome and stifling. Not once did a teacher or book present any of the questions that could have made school more meaningful.

By the time I entered school in the mid 1980s, any question that carried moral or ethical implications, or any question about the purpose of life, sacred responsibilities, or where to find human happiness, had been removed from the classroom. The education I experienced had been designed with purely utilitarian ends in mind. Any transcendent idea had been gutted from the curriculum and as a result, like most of my classmates, I was painfully bored. It wasn't until graduate school that I came to appreciate the holistic education previous Americans had received. The founding fathers of the United States revived my imagination. They were deeply interested in philosophy, human nature, political theory, and the pursuit of happiness. The education they received was aimed, most fundamentally, at making a person more fully human.

As I questioned how such a beautiful concept of education had been lost, I came to the conclusion that high-stakes testing, especially the SAT and ACT, were partially to blame. Not once since the launch of CLT has someone refuted the idea that high-stakes testing drives secondary curriculum. David Coleman, CEO of the College Board, has stated publicly that "teachers will teach towards the test. There is no force on this earth strong enough to prevent that." If teaching to the test is an inescapable reality, then shouldn't the most important test engage students with some of the most important ideas, texts, and subjects? CLT was born in response to this question. We hope that by offering a new standard that puts students in front of the thinkers and questions that have most meaningfully shaped our culture for the past two millennia, we can be a catalyst for renewal in education nationwide.

Jeremy Tate

Jeremy Tate,
CLT President



TABLE OF CONTENTS

Contributors	III
Executive Summary	VIII
1. Introduction	1
1.1 Norming	1
1.2 The CLT8	1
1.3 The CLT8 Norming Study	3
2. Norming.....	5
2.1 Overview of CLT8 national norms	5
2.2 National population targets	6
2.3 Development of CLT8 Normative Sample	7
2.4 CLT8 national percentile norm	10
2.5 Limitations and future considerations	11
2.6 Conclusions	12
3. Concordance between CLT8, CLT10, and PSAT	13
3.1 Overview	13
3.2 Content Alignment Study	14
3.3 Data Preparation	21
3.4 Concordance Relationship CLT8, CLT10 Scores	23
3.5 Limitations and future considerations	29
3.6 Conclusions	30
4. Summary and Discussions.....	33
4.1 Summary	33
4.2 Discussions	38
Screenshot for running the LEGS program	42
CLT8 score frequency for linking	43
The Mapped CLT10 Scores and the PSAT Scores	44
10th grade percentiles	45

Norming for the CLT8

EXECUTIVE SUMMARY

This norming study report provides normative information about interpreting the CLT8 scores against different target populations. Two groups of target populations are referenced: the CLT8 population for private schools and homeschools, and the general student population for the PSAT. Two sub-studies are carried out in this study: 1) a norming study where norms for the CLT8 scores are developed by referencing a nationally-matched sample of the CLT8 population against the Non-Public Education program, focusing on private schooled and homeschooled students (see Chapter 2); and 2) normative information about CLT8 scores and associated projected PSAT scores, as compared to PSAT nationally representative samples (see Chapter 3).

The norming study presented in Chapter 2 uses a representative sample obtained from the Spring 2021 CLT8 test administration, targeting the demographics of the CLT8 national population derived from the 2016 national survey conducted by the National Center for Education Statistics (NCES) for private schools and home schools and published in the Digest of Education Statistics. This norming study presents the percentile rank of a CLT8 score compared to the national normative sample created based on the Spring 2021 CLT8 test data.

Chapter 3 presents another source of normative information obtained from a linking study which establishes the concordance relationship between the CLT8 and CLT10 scores first, in order to produce the concordance relationship between the CLT8 and PSAT scores. The mapping between CLT8 and PSAT scores relates the percentile ranks of the CLT8 scores to two 2018 PSAT normative samples. The linking study develops a concordance relationship between CLT8 and CLT10 scores through the equipercentile linking method¹. The concordance table shows how each CLT8 score maps onto a CLT10 score. Based on the concordance relationship between CLT10 scores and PSAT scores, developed in 2019, each CLT8 score then maps onto a PSAT score. Then, the normative information constructed for the Spring 2018 PSAT is used to compare a CLT8 score with a mapped PSAT score. The percentile ranks of each CLT8 score are obtained relative to two normative samples: the PSAT nationally representative sample, and the PSAT/NMSQT and PSAT 10 test-takers, respectively, for the 2018 norm development of the PSAT.

Along the majority of the CLT8 score scale, the percentile ranks corresponding to each CLT8 score compared to the CLT8 nationally representative sample are lower than those relative to the PSAT nationally representative sample and the PSAT/NMSQT and PSAT 10 users group. Such differences are expected, as the norming results are sample dependent. The norming and concordance studies used different samples which differ in terms of score distributions and some key demographic variables. The score distributions for the norming sample and the concordance sample differ from each other, with the concordance sample generally representing a high-performing group of students. In addition, the norming sample resembles the target CLT8 population (i.e., Non-Public Education program focusing on private schooled and homeschooled students), while the concordance sample differed from the target CLT8 student population in terms of key demographic variables. Thus, it is expected that the percentile rank associated with the same CLT8 score is lower for the CLT8 nationally representative sample than those from the CLT8 concordance sample in reference to the PSAT nationally representative sample and

¹ Kolen, M. J., & Brennan, R. L. (1995). *Test equating*. New York: Springer.

the PSAT/NMSQT and PSAT 10 users group.

The studies presented in this report are initial steps toward creating national norms in order to interpret CLT8 scores relative to known target populations, represented by the Non-Public Education Program and the PSAT populations. In general, the norms developed from this study provide normative information about the interpretation of CLT8 scores compared to a CLT8 nationally representative sample of students which approximates the Non-Public Education Program, the PSAT nationally representative sample, and the PSAT/NMSQT and PSAT 10 test-takers. It should be noted that the samples used in the studies are convenience samples, which may limit generalization of the results to the intended population. Moreover, the disruptions in instruction during the COVID-19 pandemic may have impacted student performance on the CLT8 norming sample. Caution should be exercised in interpreting the norms reported in this current report. As the CLT8 test-taker population grows and evolves, CLT plans to replicate this study after collecting a larger, more representative sample of the test-taker population.



1. INTRODUCTION

1.1 NORMING

In large-scale high-stakes educational tests, scores are often interpreted in either a norm-referenced or criterion-referenced framework. When developing a test score scale, normative or criterion-related information is often built into the scale to facilitate score interpretation (Kolen & Brennan, 2004). Normative information can be used to enhance the interpretability of test scores. The process of incorporating normative information into scaled scores as a means of aiding interpretation of the scores is called norming (Gardner, 1962). For the norm-referenced interpretation, a student’s test score is compared with a norm or reference group, which is often considered as the targeted reference population of the test. A student’s performance can be described in terms of their relative standing in the target population. On the other hand, for the criterion-referenced interpretation, scores are interpreted relative to content standards. This interpretation provides data on student performance on content standards.

In standardized large-scale tests, norms are often developed to facilitate score interpretation and increase the utility of test scores for stakeholders. In norming, a test-taker’s score is compared to the distribution of scores for one or more target populations, to make meaningful inferences about the student’s performance as compared to other candidates. The interpretation of scores based on a comparison with a targeted national population is considered norm-referenced (Standards, 2014). Such norm-referenced information can be obtained from norming studies using a national representative sample.

It is worthy of note that “the validity of norm-referenced interpretations depends in part on the appropriateness of the reference group to which test scores are compared” (Standards, 2014, p. 97). The representativeness of the sample used in a norming study is critical to the validity of the score interpretation, as it impacts the types of claims or inferences that can be made from the scores.

1.2 THE CLT8

The Classic Learning Test (CLT), like SAT and ACT, is a college admission test to assess students’ college readiness. CLT scores are used to identify high-achieving high-school

candidates for undergraduate studies. The Classic Learning Test 8 (CLT8) is designed for 7th and 8th grade students. It is currently the lowest-level test in the CLT suite of exams. CLT8 has similarities to the PSAT 8/9, in that it may be used to determine the readiness of students for high school study and track student academic growth in early grades. Because much of the form design of the CLT8 matches that of CLT10 and CLT (adjusted for the ability level of the target population), it can also prepare students to take the CLT10 and CLT.

The CLT8 launched in May 2018 as an addition to CLT10 and CLT. Much like the CLT10 and CLT, the CLT8 was developed with the goal of reconnecting education with virtue. This is done by designing a test with two verbal-based sections focused on classic texts and a math section focused on problem solving and logic. The CLT8 is designed to test both a student's aptitude and achievement at the 7th and 8th grade level. While the CLT10 is a preparatory test for the CLT that targets 9th and 10th grade students, the CLT8 is an assessment tool to track the progress of 7th and 8th grade students as they prepare for high school and future college admissions tests.

The CLT8 was developed to measure similar content domains and subdomains as CLT10 and CLT at an age-appropriate level (<https://www.cltxam.com/test-content>), with fewer questions of the highest difficulty as summarized in Figure 1. Students may take the exam on their own computer and receive their scores within five business days. Like CLT10, CLT8 reports one total score and three subdomain scores, namely, Verbal Reasoning, Grammar/Writing, and Quantitative Reasoning.

Figure 1. Content coverages of CLT8, CLT10, and CLT.

WRITING 20 QUESTIONS		
Structure (8)	Style (8)	Word Choice (4)
GRAMMAR 20 TOTAL QUESTIONS		
Agreement (10)	Punctuation and Sentence Structure (10)	
PRE-ALGEBRA AND ALGEBRA 14 TOTAL QUESTIONS		
Arithmetic and Operations (7)	Algebraic Expressions and Equations (7)	
GEOMETRY 10 TOTAL QUESTIONS		
Plane Geometry (3)	Properties of Shapes (7)	
MATHEMATICAL REASONING 16 TOTAL QUESTIONS		
Logic (8)	Word Problems (8)	

1.3 THE CLT8 NORMING STUDY

Normative data provides information about a test-taker's performance relative to a target population. For college admission, such norm-referenced interpretation of test scores is important in making an admission decision. When CLT scores are used for college admission, the availability of normative data of the test scores increases the utility of CLT scores for high-stakes decision making. Different stakeholders of the CLT test series request information to allow them to compare test scores on the CLT exams, including CLT8 and CLT10, to their corresponding national population (the Non-Public Education program focusing on private schooling and homeschooling), in order to evaluate the academic achievement of the CLT test-takers and to meet state regulations. Thus, a norming study is needed for producing normative information for the users of the CLT tests. A study conducted in 2019 produced normative information for CLT10 scores, compared to the PSAT nationally representative sample and to PSAT/NMSQT and PSAT 10 users. This current report focuses on providing normative data for CLT8 scores within the norm-referenced framework, compared to two sets of national representative norming samples.

The Standards for the Educational and Psychological Testing stipulated by American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME) outline the 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. (2014, p. 104).

Two norming studies are presented in this report. The purpose of both studies is to provide CLT8 score users with information about 7th and 8th grade students' performance on the CLT8 compared to the target national population for the CLT8 and to two PSAT reference groups representing the general high school student population. More specifically, the first study is a norming study that provides test-takers with normative information about their performance on the CLT8 relative to the national population in Non-Public Education (NPE). The second study is a concordance study that investigates the relationship between the CLT8 scores and CLT10 scores, ultimately linking CLT8 scores to projected performance on the PSAT. The results of this study present a concordance table that shows the relationship between the scores on the two CLT tests. Essentially, each CLT8 test score from the concordance sample is mapped onto a CLT10 score through a linking method. Then, utilizing the concordance between the CLT10 scores and the PSAT scores, the correspondence between a CLT8 total score and a PSAT score is established. After mapping the CLT8 scores and the projected PSAT scores, the national normative information from PSAT can be utilized for interpreting the CLT8 test scores relative to the PSAT nationally representative sample and the PSAT/NMSQT and PSAT 10 users as a reference group respectively. It should be noted that the projected PSAT score derived from this study assumes that the student will take the PSAT as a 10th grader, not as an 8th grader.

The implementation of the planned norming studies involves identifying a reference population or group, selecting a technically defensible methodology for data analysis, and summarizing the results based on the empirical evidence collected. This current report summarizes two related studies. Each study is presented in its own chapter. In the end, a summary comparing the findings from both studies is provided. The limitations of the studies and the implications of the limitations on interpreting the norming results are summarized and discussed in this report.

Different perspectives can be used in producing normative scores, including linear or non-linear score transformation. Kolen and Brennan (2004) indicated that the percentile ranks for various groups of

examinees, often used for national norming as auxiliary scores, are based on a nonlinear transformation. In the transformation process, the distance between score points are compressed in the middle of the distribution and expanded at the upper and lower ends. They emphasize the importance of estimating the score precision, reliability, and conditional standard error of measurement to support the interpretations.

For the norming study presented in Chapter 2, the target national population was defined for the CLT8 sample of examinees using the 2016 national survey by the National Center for Education Statistics (NCES) for private and home schools. The CLT8 sample was weighted based on several demographic variables, including gender, school type, ethnicity, and geographic location, to match the distribution in the target population. Subsequently, the national percentile norm was established for CLT8 scores.

For the concordance study, the cleaned data based on the same exclusion rules was used to develop the concordance table between the CLT8 scores and CLT10 scores based on the single-group linking design. Equipercentile linking with post-smoothing was implemented in LEGS 2.0 developed by Center for Advanced Studies in Measurement and Assessment at the University of Iowa. Additional extrapolation was conducted through different approaches, including linear, polynomial, and exponential functions. The model fit was evaluated and compared across different models. The model with the best fit was selected to extrapolate the scores at the ends of the scale.

The interpretation of test scores is closely related to the test content as indicated by Ebel (1962). The CLT8 and CLT10 content coverage comparison is included in this report as well, to provide validity evidence to support the statistical linkage for the appropriate interpretation of the normative information.



2. *NORMING*

2.1 OVERVIEW OF CLT8 NATIONAL NORMS

This chapter presents the results of a norming study to develop initial estimates of CLT8 national norms. The national norm provides a reference for interpreting a student’s CLT8 test score relative to the target student population. For the CLT8 norm development, a national representative sample (normative sample) was derived that resembled the target student population of students in the Non-Public Education program, which focuses on private-schooled and homeschooled students. This sample forms the peer group with whom a student’s CLT8 score can be compared. The norm for the CLT8 scores is based on the performance distribution of the normative sample (e.g., private-schooled and homeschooled students). Percentile rank is used as the normative score for its straightforward interpretation and direct practical implications. National percentile rank of a test score is the percentage of scores that are at or below it. It indicates that the student performs at the same or above the level of a certain percentage of students in the population.

To develop norms for a large-scale test, the best practice would call for deliberate recruiting of a representative sample for the norming study. Due to significant practical constraints, however, sometimes convenience samples¹ are used. In such a case, stratification is needed to derive a normative sample from the convenience sample in the most unbiased way possible. For the CLT8, a convenience sample from the Spring 2021 administration is used. The convenience sample may not represent the target population. However, the development of norms based on the current available sample may still provide test stakeholders with some preliminary information about the normative information of the CLT8 scores relative to the targeted population. As the base of CLT8 test takers continues to grow, we plan to conduct additional studies in order to collect additional evidence to improve the interpretation of the normative data.

A series of steps were taken to derive the CLT8 normative sample from the Spring 2021 administration. These steps aim to maintain a sufficient sample size for norming while ensuring the representativeness of the target population. First, a set of data cleaning rules were applied to include only grades 7 and 8 CLT8 test-takers with valid demographic information. Sample stratification was conducted to match the national

¹ <https://research-methodology.net/sampling-in-primary-data-collection/convenience-sampling/>

population of students in home schools, private schools, and charter schools. Several key demographic variables such as gender, ethnicity, geographic region, locale, and school type were compared during stratification. The stratified sample was then used to create the CLT8 national percentile norms.

2.2 NATIONAL POPULATION TARGETS

The CLT8 national population target demographics were derived from the latest (2016) national survey by the National Center for Education Statistics (NCES) for private schools and homeschools. Specifically, Table 206.10 (retrieved from https://nces.ed.gov/programs/digest/d19/tables/dt19_206.10.asp), Table 206.20 (retrieved from https://nces.ed.gov/programs/digest/d19/tables/dt19_206.20.asp), and Table 206.30 (retrieved from https://nces.ed.gov/programs/digest/d19/tables/dt19_206.30.asp) published on the Digest of Education Statistics were used.

NCES Table 206.20 reports, in 2016, 3.31% and 8.80% of grade 6-8 students were homeschooled and in private schools, respectively. Table 206.30 indicates 5.08% grade 6-8 students were enrolled in charter school. Therefore, 17.19% (i.e., 3.31+8.80+5.08) of the total grade 6-8 students in the US in 2016 make up the CLT8’s target population. Thus, for the CLT8’s target population, 19.26% (or 3.31/17.19), 29.55% (or 5.08/17.19), and 51.19% (or 8.8/17.19) are in homeschools, charter, and private schools, respectively.

NCES Tables 206.10 and 206.30 provide student demographics for Homeschool and Charter/Private schools, respectively. These demographics percentages were then weighted based on the above calculated proportion of school types to get the target population demographic. Regional information is not available for homeschool students, thus the percentages for region were weighted based on charter and private school proportions. Table 2.1 presents the student demographic percentages obtained from NCES Table 206.10 and Table 206.30 and the weighted percentages used as CLT8’s target population demographics. For example, the percentage 50.1 for Male is calculated from $0.2955*51.4+0.5119*50.3+0.1926*47.8$.

Table 2.1 Demographics percentages by school type and population targets for CLT8 normative sample

VARIABLES	SCHOOL TYPE			WEIGHTED CLT8 POPULATION TARGET %
	CHARTER ² (29.55%)	PRIVATE ² (51.19%)	HOMESCHOOL ¹ (19.26%)	
Gender				
Female	48.6	49.7	52.5	49.9
Male	51.4	50.3	47.8	50.1
Ethnicity				
White	30.4	62.1	59.2	52.2
Hispanic	36.0	15.0	26.3	23.4
Black	26.2	12.1	7.8	15.4
Asian/Pacific Islander	4.7	6.3	2.6	5.1
Other	2.7	4.4	4.1	3.8

Region				
South	16.7	23.3		20.9
West	45.9	24.1		32.1
Midwest	20.6	26.5		24.3
Northeast	16.8	26.1		22.7
Locale				
Rural	4.0	11.8	21.8	11.4
Suburb	34.8	48.1	38.5	42.3
City	57.8	37.2	29.2	41.7
Town	3.4	3.0	10.5	4.6

¹Table 206.10 (https://nces.ed.gov/programs/digest/d19/tables/dt19_206.10.asp)

²Table 206.30 (https://nces.ed.gov/programs/digest/d19/tables/dt19_206.30.asp)

Note: Total percentage may not add up to 100% due to rounding.

2.3 DEVELOPMENT OF CLT8 NORMATIVE SAMPLE

2.3.1 Initial sample

CLT8 normative sample was drawn from the Spring 2021 administration test data (in file CLT8_1023_psychometric_data_with_zipcodes.csv). The raw sample includes a total of 3,541 valid CLT8 total scores. Several exclusions rules were first applied in data cleaning, including:

1. Remove students from grades other than 7 and 8 (510 test scores removed; remaining 3,031 students with gradYear 2025 and 2,026).
2. Remove students with missing gender (193 test scores removed; 2,838 remaining)
3. Remove students with missing ethnicity (330 test scores removed; 2,508 remaining)
4. Remove students not in Charter school, Private school, or Homeschool (61 test scores removed; 2,447 remaining)
5. Remove students with missing or non-US zip code (167 test scores removed; 2,280 remaining)

The initial sample includes the remaining 2,280 students. Table 2.2 presents the demographics for the initial sample with comparisons to the population targets. The initial sample is not representative in a few categories. Specifically, the sample is composed of a much higher percentage of Caucasian, students from South, and Rural areas. Meanwhile, Black and Hispanic students, as well as students from Northeast, and cities take much lower percentage than the population target.

Table 2.2 Percentages of students by school type and key demographics: initial sample vs. population

	INITIAL SAMPLE %(N=2,280)	POPULATION TARGET %	DIFFERENCE %
School Type			
Charter	8.2	29.6	-21.4
Private	60.9	51.2	9.7
Homeschool	30.9	19.3	11.6

Gender			
Female	50.4	49.9	0.5
Male	49.6	50.1	-0.5
Ethnicity			
White	75.4	52.2	23.2
Hispanic	10.2	23.4	-13.2
Black	3.0	15.4	-12.4
Asian/Pacific Islander	5.4	5.1	0.3
Other	6.0	3.8	2.2
Region			
South	46.4	20.9	25.5
West	27.2	32.1	-4.9
Midwest	22.2	24.3	-2.1
Northeast	4.2	22.7	-18.5
Locale			
Rural	52.5	11.4	41.1
Suburb	29.9	42.3	-12.4
City	17.5	41.7	-24.2
Town	0.2	4.6	-4.4

Note: Total percentage may not add up to 100% due to rounding.

2.3.2 Sample stratification

Both large overrepresentation and underrepresentation are observed in the initial sample. To prepare a representative national normative sample while maintaining sufficient sample size for norming, sample stratification steps included duplicating scores of students from under-represented categories (e.g., Black and Hispanic) and eliminating scores from over-represented categories (e.g., White, students from South, and students from Rural areas).

Duplication and elimination both have inherent limitations. One assumption in the duplication process is that students from the same demographics would perform similarly. Given the large score range in CLT8 (i.e., 0-120), this assumption might be likely to hold at the group level when sample size is large enough, but unlikely to be met at the individual score point being duplicated. In the meantime, elimination risks removing information from the score distribution.

The use of percentile norm, which is based on the group performance distribution, could mitigate some limitations introduced by the duplication process. Additional controls were also used to further limit the impact of the stratification process on the true score distribution and avoid introducing systematic bias. Specifically, duplication was strictly used only on critically under-represented demographics and was capped at 3, that is no single score could be counted more than 3 times. For elimination, when multiple scores were candidates for elimination, random selection was used.

During the stratification process 58 students were triplicated, 286 students were duplicated, and 1,444 students were eliminated. The final normative sample includes a total of 1,238 students. Table 2.3 shows descriptive statistics for the CLT8 total score from the initial sample and the normative sample. The score range and shape of the score distribution in the normative sample remain similar to the initial sample. A slight (2.1 points) drop in mean score in the normative sample from the initial sample is expected given

the demographic categories where stratification process adjusted. For example, the initial sample is over-represented by White students and students from Private and Homeschool settings. Typically, students with these demographics have higher performance. The stratification process lowered the percentage of students in these categories, resulting in a lower average score.

Table 2.3 Descriptive statistics for CLT8 total scores for initial and normative sample

SAMPLE	N	MEAN	STD	MIN	MAX
Initial	2,280	70.7	17.9	18	116
Normative	1,238	68.6	19.3	18	116

2.3.3 CLT8 normative sample

Table 2.4 shows the comparison between the national norm sample and the population targets on school type and key demographics. The national norm sample is within 5% variance on all but three categories. The largest discrepancies are in geographic categories, with Northeast and Rural off by 9.3% and 6.9%, respectively. However, both are much improved from the initial sample where Northeast is under-represented by 18.5% and rural is over by 41.1%. The difference between the CLT8 normative sample and the population targets are consistent with industry standard and practice for norming sample composition. The final sample size is sufficient for percentile norm development.

Table 2.4 Percentages of students by school type and key demographics: normative sample vs. population

NATIONAL NORM SAMPLE % (N=1,238)	POPULATION TARGET %	DIFFERENCE %	
School Type			
Private	51.7	51.2	0.5
Homeschool	20.7	19.3	1.4
Charter	27.6	29.6	-2.0
Gender			
Female	50.1	49.9	0.2
Male	49.9	50.1	-0.2
Ethnicity			
White	54.8	52.2	2.6
Hispanic	17.5	23.4	-5.9
Black	12.2	15.4	-3.2
Asian/Pacific Islander	7.2	5.1	2.1
Other	8.3	3.8	4.5
Region			
South	25.3	20.9	4.4
West	35.1	32.1	3.0
Midwest	26.2	24.3	1.9
Northeast	13.4	22.7	-9.3
Locale			
Rural	18.3	11.4	6.9

Suburb	40.9	42.3	-1.4
City	40.3	41.7	-1.4
Town	0.5	4.6	-4.1

Note: Total percentage may not add up to 100% due to rounding.

2.4 CLT8 NATIONAL PERCENTILE NORM

To develop the percentile norm for CLT8, the 1-99th percentile rank for the total score in the normative sample was first calculated using Python (`numpy.percentile` with `interpolation=lower`) (step 1). Then, each unique total score is mapped to a percentile rank. Total scores not directly associated with a percentile rank from step 1 were placed into the adjacent rank down. For example, total score 34 was associated with 3rd percentile and 36 was associated with 4th percentile, thus, total score 35 was placed with 34 into the 3rd percentile.

Table 2.5 presents the percentile rank correspondence for each obtainable CLT8 total score. For ease of use, total scores associated with the same percentile rank are grouped together in the table. With percentile norm, each performance level is mapped directly to the distribution of the CLT8 scores for the normative sample. This norm score has a straightforward interpretation. A norm score indicates the percentage of students in the normative sample a total score is at or above. For example, a student who scores a 91 on CLT8 is at 85th percentile, this student is at or above 85% of the normative sample. This straightforward interpretation of performance with reference to the peers makes it easy to identify students in need.

Table 2.5 CLT8 national percentile norm

CLT8 TOTAL SCORE	NATIONAL PERCENTILE RANK	CLT8 TOTAL SCORE	NATIONAL PERCENTILE RANK
0-30	1	71	56
31-33	2	72	58
34-35	3	73	61
36	4	74	62
37-39	5	75	64
40	7	76	66
41-42	8	77	67
43	9	78	69
44	10	79	70
45	11	80	72
46	12	81	73
47	13	82	75
48	14	83	76
49	16	84	77
50	17	85	78
51	19	86	80
52	21	87	81
53	23	88	82
54	24	89	83

55	26	90-91	85
56	28	92	86
57	30	93	87
58	31	94	88
59	33	95	89
60	35	96	90
61	36	97	91
62	39	98	92
63	41	99	93
64	43	100-101	94
65	45	102	95
66	48	103	96
67	50	104-105	97
68	51	106-108	98
69	53	109-120	99
70	54		

2.5 LIMITATIONS AND FUTURE CONSIDERATIONS

Three limitations warrant some caution in the use of the current CLT8 norms for this initial exploratory analysis. First, the CLT8 user base as of Spring 2021 is relatively small and overly representative of students in certain demographic areas (e.g., rural communities in the south). The initial sample, as shown in Table 2.2 has large under and over representation in a few demographic categories. Most of the large differences to population targets were balanced out through the stratification process, however, at the cost of large sample size reduction. The implementation of restrictions in the stratification ensured a sufficient normative sample size with close resemblance to the key national demographic targets. Nevertheless, a larger and more representative initial sample would be more desirable for future studies that replicate this initial exploratory analysis.

Second, the school year 2020-2021 introduced a unique limitation in using Spring 2021 test scores for norm development. Due to Covid-19 pandemic, the majority of the students who would normally have in-person instruction went through an entire school year with online instruction. The turbulence caused by the pandemic and the changes to instruction format could have resulted in a shift in end-of-grade performance, measured by tests like the CLT8. This is the first CLT8 national norm developed using the CLT8 test scores that might be impacted by the pandemic. It is worthwhile to investigate the direction and magnitude of the pandemic impact on student performance post-pandemic when school instruction returns to normal. Accordingly, the normative scores could be adjusted for short-term.

Lastly, as the CLT8 user base grows substantially in the future, performance shifts in both overall level and in score distribution are expected. This study will be replicated in the future to investigate the impacts of a larger user base, familiarity with the assessment, and as schools begin to normalize again after the pandemic.

Given the potential impact of the pandemic and the expected CLT8 user base growth, re-evaluation of the normative performance in the subsequent years is critically important to ensure the norm maintains meaningful reference to the current CLT8 users. Norm update should be conducted when the performance shift becomes statistically significant.

2.6 CONCLUSIONS

The CLT8 national percentile user norm was created using a representative sample of 1,238 grade 7 and 8 students taking the CLT8 during Spring 2021 administration. A CLT8 percentile user norm score links the student's total score to the percentage of performance in the normative sample. Three major limitations should be cautioned:

1. The current CLT8 normative sample was derived from a convenience sample with a very unbalanced initial representation of the target population. This limitation could reduce the stability and interpretability of the current norm. A larger and more representative sample should be used for future national norms development and update.
2. Covid-19 pandemic could have impacted on student's learning experience and effectiveness in the school year 2020-2021. The effect of the impact awaits further post-pandemic research and evaluation.
3. The current user base for CLT8 is fast growing. Performance shift in future test administration may be expected.

The availability of the inaugural CLT8 national norm will provide test stake-holders such as teachers and administrators a much-needed norm-referenced interpretation for students' performance on the CLT8. It will also assist the test developer in future research of student performance and test enhancement. However, given the limitations cautioned above, particularly the unique challenge caused by the pandemic, periodical evaluation and norms update are strongly recommended to maintain the interpretability of the normative score.



3. *CONCORDANCE BETWEEN CLT8, CLT10, AND PSAT*

3.1 OVERVIEW

Chapter 2 presents the norming results based on the Spring 2021 CLT8 test-takers' performance, measured against the targeted CLT8 national population. The demographics for the CLT8 national population were derived from the 2016 national survey by the NCES for private schools and home schools published in the Digest of Education Statistics. The test scores and demographic information from the Spring 2021 CLT8 were then used to develop a representative sample targeting the national population for the CLT8. Then, this national normative sample was used for developing norms. The norming results from Chapter 2 indicate the relative ranking of a given CLT8 total test score relative to the national normative sample; these norms are often referred to as user norms.

This chapter presents the results from an empirical linking study which provides normative information for CLT8 examinees on their test performance relative to the 2018 PSAT normative sample. This linking study intends to establish a concordance relationship between the CLT8 and the PSAT. It should be noted that students that take the CLT8 in 7th and 8th grade do not yet have a PSAT score; however, some CLT8 test takers from previous administrations went on to take CLT10 approximately two years later. Through the design of this study, the concordance table in this chapter shows how each CLT8 score is mapped onto the PSAT scale. To provide some context of the meaning of the CLT8 scores, the normative information constructed for the Spring 2018 PSAT can be used to compare a given CLT8 score with a mapped PSAT score, indicating the relative ranking of CLT8 scores compared with the normative sample for the 2018 norms of the PSAT. However, given the situation of cancelled or reduced numbers of standardized tests such as PSAT due to the Covid-19 pandemic, fewer students took the PSAT than during a normal year. Consequently, this study first links the CLT8 to the CLT10, then uses the linkage hitherto set up between the CLT10 and PSAT to obtain normative information against the normative sample for the 2018 PSAT national population.

The following sections start with the content alignment results between the CLT8, CLT10, and PSAT, then present data preparation and analyses for developing a concordance relationship between CLT8 and CLT10, and ultimately between CLT8 and PSAT. Next, the technical details and the results from the empirical norming study are presented. The last section of this chapter presents a summary of the findings.

3.2 CONTENT ALIGNMENT STUDY

The goal of this empirical norming study is to derive normative information against the PSAT national population. To make valid inferences based on the results of this linking study, it is necessary to evaluate the alignment of the content between the CLT8 and the CLT10, the CLT10 and the PSAT, and, ultimately, the CLT8 and the PSAT. This section summarizes the results from a content analysis for the CLT8 in comparison with a content summary for CLT10 and PSAT.

The CLT8 is the first test in the suite of CLT series. It was designed for students in grades 7 and 8. It consists of three sections: Verbal Reasoning, Grammar/Writing, and Quantitative Reasoning (a structure employed in all CLT exam-suite tests). The PSAT/NMSQT consists of two sections: Evidence-Based Reading and Writing (EBRW) and Mathematics. The EBRW section consists of Reading and Writing & Language sections, while the Math section consists of Math without Calculator and Math with Calculator sections.

As with the CLT10, there are 40 questions in each CLT8 section. The test takes approximately 2 hours and 15 minutes (135 minutes total), and students are given 45, 40, and 50 minutes for the Verbal Reasoning, Grammar/Writing, and Quantitative Reasoning sections, respectively. The PSAT EBRW section has 91 questions to be completed in 95 minutes and the PSAT Math section has 48 questions to be completed in 65 minutes. Students are allotted 1 minute on average for CLT8 and CLT10 items, while students are allotted an average of 1.15 minutes for each PSAT question. (Refer to Table A.1, A.2, A.3 in the Appendix for details.)

Verbal Reasoning (CLT8, CLT10) and Reading (PSAT)

On both the CLT8 and the CLT10 Verbal Reasoning sections, questions are broken down into two domains: Comprehension and Analysis. Comprehension questions include the subdomains “Passage as a Whole”, “Passage Details”, and “Passage Relationships.” Analysis questions include the subdomains “Textual Analysis” and “Interpretation of Evidence.” The CLT8 Verbal Reasoning section has a total word count between 1,700-1,750, for an average of 1,725 words total. The CLT10 Verbal Reasoning section has a total word count between 2,175 and 2,225, for an average of 2,200 words total. The CLT8 Verbal Reasoning section consists of four passages, including one paired-set, and is broken into 40 questions to be completed in 45 minutes. The CLT10 Verbal Reasoning section consists of four passages, also including one paired-set, and is broken into 40 questions to be completed in 40 minutes.

The CLT8 Verbal Reasoning section most closely corresponds to the CLT10 Verbal Reasoning section and to the PSAT Reading test from the Evidence-Based Reading and Writing section. The range of the passage length on the PSAT is 500-750, not to exceed 3,000 words total. The PSAT Reading test consists of five passages, including one paired-set, and contains 47 questions to be completed in 60 minutes.

Table 3.2.1 The Blueprint for the CLT8 Verbal Reasoning Section

SECTION	DOMAIN	SUBDOMAIN
Verbal Reasoning (40 questions)	Comprehension (27 questions)	Passage as a Whole (8 questions)
		Passage Details (11 questions)
		Passage Relationships (8 questions)
	Analysis (13 questions)	Textual Analysis (8 questions)

		Interpretation of Evidence (5 questions)
--	--	---------------------------------------------

Table 3.2.2 The Blueprint for the CLT10 Verbal Reasoning Section

SECTION	DOMAIN	SUBDOMAIN
Verbal Reasoning (40 questions)	Comprehension (27 questions)	Passage as a Whole (8 questions)
		Passage Details (11 questions)
		Passage Relationships (8 questions)
	Analysis (13 questions)	Textual Analysis (8 questions)
		Interpretation of Evidence (5 questions)

Table 3.2.3 The Blueprint for the PSAT Reading Test (Averages) ¹

SECTION	DOMAIN	SUBDOMAIN
Verbal Reasoning (47 questions)	Comprehension (26.25 questions)	Passage as a Whole (10.75 questions)
		Passage Details (15.5 questions)
		Passage Relationships (0 questions)
	Analysis (20.75 questions)	Textual Analysis (6.75 questions)
		Interpretation of Evidence (14 questions)

Tables 3.2.1-3.2.3 contain the Verbal Reasoning and PSAT Reading Test blueprints for the CLT8, CLT10, and PSAT.

The CLT8 and the CLT10 each contain 11.65% more Comprehension questions than the PSAT, while the PSAT contains 11.65% more Analysis questions than the CLT tests. Within Comprehension, the PSAT contains more questions about both the subdomains “Passage as a Whole” and “Passage Details” than the CLT8 and CLT10. No PSAT items assess “Passage Relationships.” In Analysis, the PSAT includes 5.64% fewer questions on “Textual Analysis” and 17.29% more questions on the “Interpretation of Evidence” than the CLT8 and CLT10.

Both the CLT8 and CLT10 include Comprehension questions involving analogies in the subdomain entitled “Passage Relationships”—a subdomain which is absent on the PSAT since College Board removed analogies from the SAT and PSAT in 2005. The CLT8 and CLT10 include analogies to test logical reasoning and synthesis; the analogies always pertain to the reading passages.

¹ Data available upon request

Grammar/Writing (CLT8, CLT10) and Writing and Language (PSAT)

On the Grammar/Writing section for both the CLT8 and CLT10, questions are broken down into two domains: Grammar and Writing. Grammar questions include the subdomains “Agreement” and “Punctuation and Sentence Structure.” Writing questions include the subdomains “Structure”, “Style”, and “Word Choice.” In test form assembly for CLT8, the Grammar/Writing section has a total word count between 1,550-1,600 words, for an average of 1,575 words total. The CLT8 Grammar/Writing section consists of four passages and is broken into 40 questions to be completed in 40 minutes. In test form assembly for the CLT10, each Grammar/Writing passage fits narrowly within a word count range of 440-560 words; the total must be between 1,975 and 2025 words, for an average of 2,000 words total. The CLT10 Grammar/Writing section consists of four passages and is broken into 40 questions to be completed in 35 minutes.

The CLT8 and CLT10 Grammar/Writing sections most closely correspond to the PSAT Writing and Language test from the Evidence-Based Reading and Writing section. The passage length on the PSAT ranges 400-450 words, not to exceed 1,700 words total. The PSAT Writing and Language test consists of four passages and contains 44 questions to be completed in 35 minutes.

Table 3.2.4 The Blueprint for the CLT8 Grammar/Writing Section

SECTION	DOMAIN	SUBDOMAIN
Grammar/Writing (40 questions)	Grammar (20 questions)	Agreement (10 questions)
		Punctuation and Sentence Structure (10 questions)
	Writing (20 questions)	Structure (8 questions)
		Style (8 questions)
		Word Choice (4 questions)
		Quantitative Analysis (0 questions)

Table 3.2.5 The Blueprint for the CLT10 Grammar/Writing Section

SECTION	DOMAIN	SUBDOMAIN
Grammar/Writing (40 questions)	Grammar (20 questions)	Agreement (10 questions)
		Punctuation and Sentence Structure (10 questions)
	Writing (20 questions)	Structure (8 questions)
		Style (8 questions)

		Word Choice (4 questions)
		Quantitative Analysis (0 questions)

Table 3.2.6 The Blueprint for the PSAT Writing and Language Test (Average)

SECTION	DOMAIN	SUBDOMAIN
Grammar/Writing (44 questions)	Grammar (15.5 questions)	Agreement (4.5 questions)
		Punctuation and Sentence Structure (11 questions)
	Writing (28.5 questions)	Structure (11.25 questions)
		Style (8.75 questions)
		Word Choice (7 questions)
		Quantitative Analysis (1.5 questions)

Tables 3.2.4-3.2.6 contain the Grammar/Writing and Writing and Language blueprints for the CLT8, CLT10, and PSAT. The CLT8 and CLT10 contain 14.78% more Grammar questions than the PSAT, while the PSAT contains 14.77% more Writing questions. Within Grammar, the CLT8 and CLT10 contain 14.77% more questions about the subdomain “Agreement,” and all three tests have an equal percentage of questions about “Punctuation and Sentence Structure.” Within the domain of “Writing,” the PSAT contains 5.57% more questions about “Structure,” all three tests contain the same percent of questions about “Style,” and the PSAT contains 5.91% more questions about “Word Choice.”

The PSAT also includes questions about “Quantitative Analysis,” in which the test taker is asked to select the sentence option that accurately interprets data from a graph. Neither the CLT8 nor the CLT10 include questions about data interpretation of this kind. Moreover, the Science passages in the Grammar/Writing section do not include visual representations of data.

Text Evaluation

The difficulty of the passages in both verbal tests is determined, in part, by the Educational Testing Service (ETS) Text Evaluation (TE) grade level score. The relevant TE data are as follows:

- » CLT8
 - » TE Grade Level 8
 - » TE Range 6-10
- » CLT10
 - » TE Grade Level 10.5
 - » TE Range 9-12
- » PSAT
 - » TE Grade Level 10.8
 - » TE Range 8-12.

Quantitative Reasoning (CLT8, CLT10) and Math (PSAT)

The CLT8 Quantitative Reasoning section consists of questions in three domains: Pre-Algebra and Algebra, Geometrical Reasoning, and Mathematical Reasoning. Pre-Algebra and Algebra questions include the subdomains “Arithmetic and Operations” and “Algebraic Expressions and Equations.” Geometry questions include the subdomains “Plane Geometry” and “Properties of Shapes.” Mathematical Reasoning questions include the subdomains “Logic” and “Word Problems.”

On the CLT10 Quantitative Reasoning section, questions are broken into three domains: Algebra, Geometry, and Mathematical Reasoning. Algebra questions include the subdomains “Arithmetic and Operations” and “Algebraic Expressions and Equations.” Geometry questions include the subdomains “Plane Geometry” and “Properties of Shapes.” Mathematical Reasoning questions include the subdomains “Logic” and “Word Problems.”

The CLT8 and CLT10 Quantitative Reasoning sections encompass the content of both tests in the PSAT Math section (Math with Calculator and Math without Calculator), though the CLT8 and CLT10 are taken entirely without a calculator. Given this, it can be concluded that the CLT8 and CLT10 do not measure Mathematics/Quantitative Reasoning in the same way that PSAT measures mathematical skills. Tables 3.2.7-3.2.9 contain the Quantitative Reasoning and Math Section for the CLT8, CLT10, and PSAT. Both the CLT8 and PSAT include 35% Algebra questions, about 10% more items than the CLT10. The CLT10 contains 10% more Geometry Questions than the CLT8 and 18.86% more items than the PSAT, while the PSAT contains 8.44% more Mathematical Reasoning questions than the CLT8 and CLT10. Within Algebra, the CLT8 and CLT10 contain 14.9% and 10.4% more questions about “Arithmetic and Operations” than the PSAT does, while the PSAT contains 13.75% and 18.25% more questions about “Algebraic Expressions and Equations” than the CLT8 and CLT10, respectively. Under Geometry, the PSAT contains 2.4% more questions about “Plane Geometry” than the CLT8 and about the same amount as the CLT10; the CLT10 contains 7.5% and 18.75% more questions about “Properties of Shapes” than the CLT8 and PSAT, respectively. In Mathematical Reasoning, both the CLT8 and CLT10 contain 18.4% more “Logic” problems, and the PSAT contains 6.56% more “Word Problems” items.

Table 3.2.7 The Blueprint for the CLT8 Quantitative Reasoning Section

SECTION	DOMAIN	SUBDOMAIN
Quantitative Reasoning (40 questions)	Algebra (14 questions)	Arithmetic and Operations (7 questions)
		Algebraic Expressions and Equations (7 questions)
	Geometry (10 questions)	Plane Geometry (3 questions)
		Properties of Shapes (7 questions)
	Mathematical Reasoning (16 questions)	Logic (8 questions)
		Word Problems (8 questions)
		Statistics (0 questions)

Table 3.2.8 The Blueprint for the CLT10 Quantitative Reasoning Section

SECTION	DOMAIN	SUBDOMAIN
Quantitative Reasoning (40 questions)	Algebra (10 questions)	Arithmetic and Operations (5 questions)
		Algebraic Expressions and Equations (5 questions)
	Geometry (14 questions)	Plane Geometry (4 questions)
		Properties of Shapes (10 questions)
	Mathematical Reasoning (16 questions)	Logic (8 questions)
Word Problems (8 questions)		
		Statistics (0 questions)

Table 3.2.9 The Blueprint for the PSAT Math Section (Average)

SECTION	DOMAIN	SUBDOMAIN
Quantitative Reasoning (48 questions)	Algebra (16.75 questions)	Arithmetic and Operations (1.25 questions)
		Algebraic Expressions and Equations (15 questions)
	Geometry (7.75 questions)	Plane Geometry (4.75 questions)
		Properties of Shapes (3 questions)
	Mathematical Reasoning (23.25 questions)	Logic (<1 question)
Word Problems (12.75 questions)		
		Statistics (10.25 questions)

In addition, the PSAT Math section is comprised of 21.35% questions about Statistics, which are not included on the CLT8 and CLT10. These questions concern descriptive statistics, interpretation of data sets, and interpretation of visual data. The CLT8 and CLT10 do contain questions about probability, but these are classified under “Arithmetic and Operations.”

Additionally, the PSAT includes questions in which the student is given a mathematical word problem and asked to determine which function in the answers matches the question, a type of question the CLT8 and CLT10 do not ask. These questions have been classified as “Word Problems,” though the CLT8 and CLT10 tests the same skills of function notation in “Algebraic Expressions and Equations.”

The alignment study results are summarized and presented graphically in Figures 1 and 2, which follow below. In general, content domains and content coverage differ among the CLT8, the CLT10, and PSAT. Though the domain level coverage is similar to some extent, the coverage and the distribution at the subdomain level can be quite different in some subdomains.

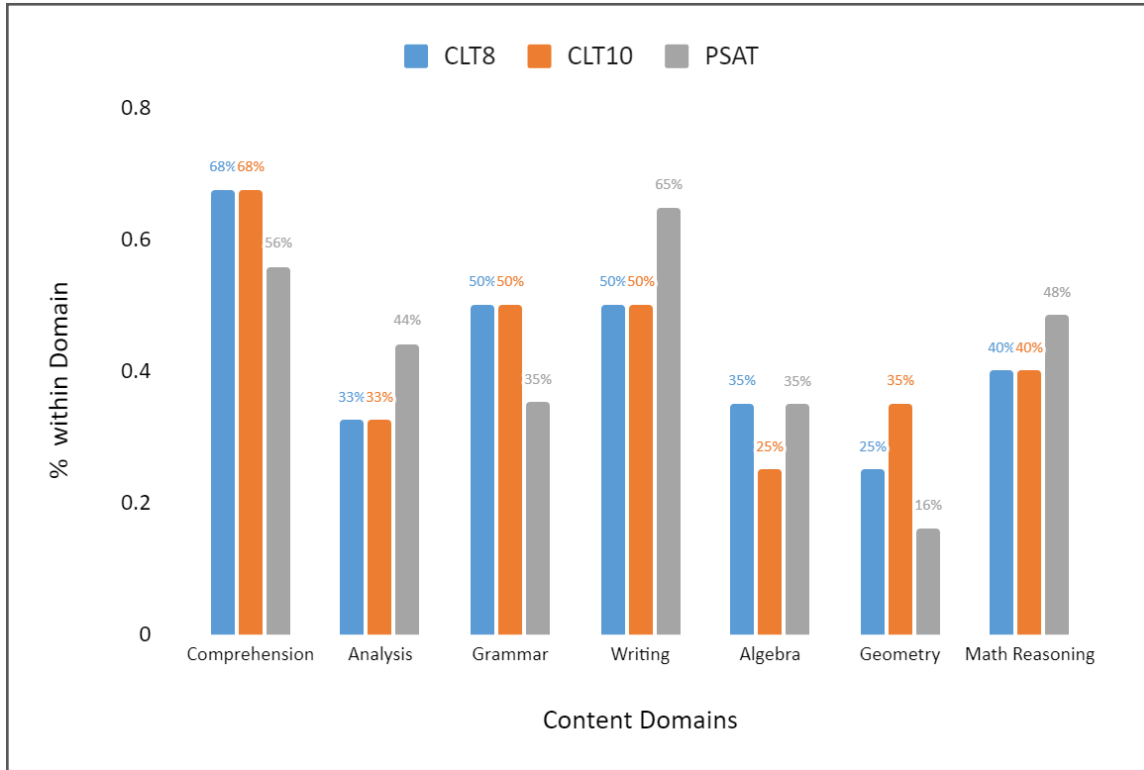


Figure 1. Content Domain Distributions for CLT8, CLT10, and PSAT.

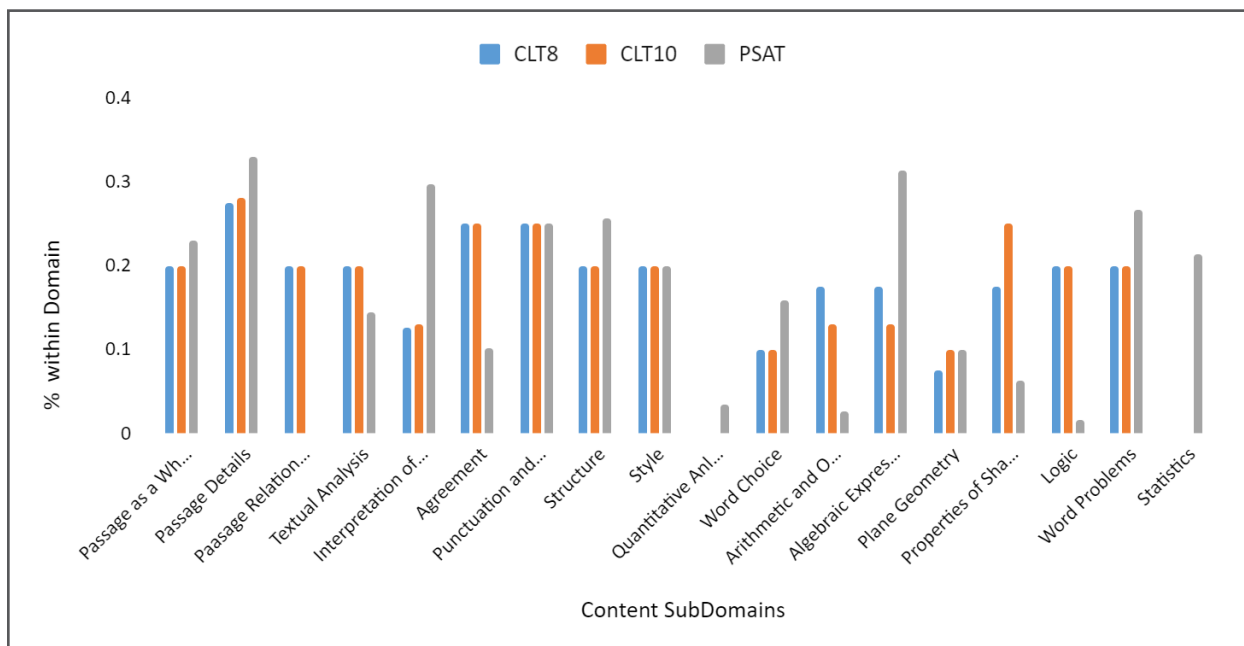


Figure 2. Content Sub-Domain Distributions for CLT8, CLT10, and PSAT.

3.3 DATA PREPARATION

To map CLT8 scores to PSAT scores, CLT8 test scores are first mapped onto CLT10 scores. Then, based on the concordance relationship established between the CLT10 and PSAT, CLT8 scores are mapped to the PSAT. The following sections elaborate the details related to the mapping of CLT8 scores to CLT10 scores first, then to the PSAT scores and the national percentile ranks for the normative sample for the 2018 PSAT score interpretation.

The data used for the statistical linkage is slightly different from that used in Chapter 2. Chapter 2 uses valid CLT8 scores from Spring 2021 test administration with exclusion rules applied, in order to generate CLT8 national user norms; similar exclusion rules were applied in cleaning data for this empirical norming study bridging CLT8 and CLT10 scores. Students who took the CLT8 in Spring 2021 are mostly in grades 7 and 8. In terms of timeline, most of them have not taken either the PSAT or the CLT10. Thus, this study used CLT10 test-takers who just took the CLT10 and have taken the CLT8 in the past. However, some students might have taken the CLT8 or CLT10 in grades other than they were intended for. Test records for these students were excluded from further analyses. The exclusion rules applied in data cleaning for this concordance study are summarized here:

1. Remove students from public schools.
2. Remove students from graduation years other than 2023 and 2024.
3. Remove students who do not have either CLT8 or CLT10 scores reported.

Test records for 1,054 students were retained after the application of these exclusion rules and a reasonableness check of the data. Table 3.1 summarizes the demographics for this sample compared with the target population and the constructed national sample for norming in Chapter 2. The concordance sample is unbalanced in a number of categories, especially in school type, ethnicity, and geography.

Table 3.1. Percentages of Students by School Type and Demographics: Concordance Sample vs. the Target CLT8 Student Population

	CONCORDANCE SAMPLE CLT8 & CLT10	NATIONAL NORM SAMPLE% CLT8	POPULATION TARGET % CLT8	NATIONAL NORM SAMPLE % CLT10	POPULATION TARGET % CLT10	DIFFERENCE % IN NORM- ING SAM- PLE-CLT8	DIFFER- ENCE % IN CONCOR- DANCE SAM- PLE-CLT8
SCHOOL TYPE							
Charter	1.4	8.2	29.6	18.8	18.8	-6.80	-28.20
Private	47.7	60.9	51.2	56.7	56.7	-13.20	-3.50
Homeschool	49	30.9	19.3	24.5	24.5	18.10	29.70
Others	1.9	0	0			1.90	1.90
Gender							
Male	41.8	49.6	50.1	45.8	49.9	-7.80	-8.30
Female	46.0	50.4	49.9	54.2	50.1	-4.40	-3.90
Not Provided	12.1					12.10	12.10
Ethnicity							
White	65.0	75.4	52.2	56.5	55.4	-10.40	12.80
Black	2.0	3.0	15.4	12.8	13.7	-1.00	-13.40
Hispanic	4.3	10.2	23.4	21.7	21.7	-5.90	-19.10
Asian/Pacific islander	3.7	5.4	5.1	5.1	5.1	-1.70	-1.40
Other	5.3	6.0	3.8	4.0	4.0	-0.70	1.50
Not Provided	19.8					19.80	19.80
Region							
Northeast	6.1	4.2	22.7	24.5	23.8	1.90	-16.60
South	42.1	46.4	20.9	32.2	21.7	-4.30	21.20
Midwest	26.4	22.2	24.3	19.5	25.0	4.20	2.10
West	22.5	27.2	32.1	23.8	29.6	-4.70	-9.60
Missing	2.9					2.90	2.90
Locale							
City	15.6	17.5	41.7	31.4	39.1	-1.90	-26.10
Suburb	25.8	29.9	42.3	41.1	43.2	-4.10	-16.50
Town	0.2	0.2	4.6	1.8	4.9	0.00	-4.40
Rural	55.5	52.5	11.4	25.7	12.8	3.00	44.10
Missing	2.9					2.90	2.90

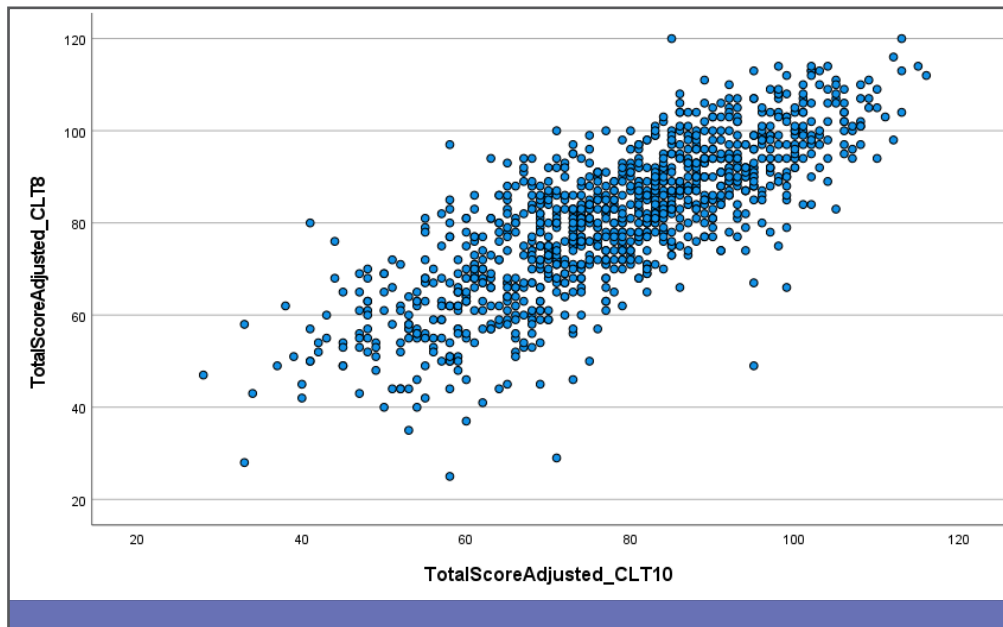
Note: Total percentages may not add up to 100% due to rounding.

Table 3.2 presents the descriptive statistics for the CLT8 and CLT10 total scores from the stratified sample used for norming conducted in Chapter 2, and the sample for the concordance study in this chapter. It should be noted that the score distribution of the CLT8 for the concordance study differs from that from the norming sample used in Chapter 2. In general, the students who took the CLT10 performed better than those students in the norming sample, with a mean CLT8 score of 80.86 and a standard deviation of 15.9. Their CLT10 scores ranged from 28 to 116, with a mean of 78.15 and a standard deviation of 15.7. The sample used for developing the concordance relationship between the CLT8 and the CLT10 is a relatively high-performing group of students. This is not a surprising finding, since students that did not perform well on CLT8 in 7th or 8th grade may have been discouraged from taking the CLT10 in 9th and 10th grades. Figure 3 presents the scatterplot of the CLT8 and CLT10 scores.

Table 3.2. Descriptive statistics for the CLT10 total scores for the initial and norming samples in chapter 2 and the concordance sample in chapter 3.

SAMPLE	N	MEAN	STD	MIN	MAX
CLT8 Norming Sample	1,238	68.6	19.3	18	116
CLT8-Concordance	1,054	80.86	15.9	25	120
CLT10-Concordance	1,054	78.15	15.7	28	116
CLT10-2019	220	85.3	15.1	38	117
PSAT-2019	220	1144.2	138.7	720	1520

Figure 3. Scatterplot for CLT8 and CLT10 Scores.



3.4 THE CONCORDANCE RELATIONSHIP BETWEEN THE CLT8 AND CLT10 SCORES

Equipercentile linking was conducted to construct the linkage of scores on the CLT8 and CLT10 using a single group sample. The common-group design was used in linking based on the raw scores for the cleaned sample for the concordance relationship development. Equipercentile linking based on the matched sample of the CLT8 and the CLT10 was carried out using the software program, Linking with Equivalent Group or the Single Group Design, abbreviated as LEGS (Kolen & Brennan, 2004). The raw CLT8 scores and the corresponding CLT10 scores were used to link the CLT8 and CLT10. With the proper specification of the format of the input data, subgroup information, input data file names, smoothing values, the score range for the CLT8, and the truncation choice, the program conducts the equipercentile linking and calculates the results. In Appendix B, a screenshot captures the input window for linking CLT8 and CLT10 scores.

Two smoothing values were compared in post-linking: 0.3 and 1. The choice of using smoothing values is supported by the results from simulation studies, which indicate that the smoothed results outperform the non-smoothed method in reducing linking errors when the population test scores are in fact smooth

(Cui & Kolen, 2009; Hanson et al., 1994). LEGS output results for different linking methods including mean, linear, parallel-linear, and equipercentile methods with and without post-smoothing. The results with a smoothing value of 1 has the smallest root mean squared error for linking (RMSEL) as shown in Table 3.3. Table 3.4 presents the mapped CLT10 scores and the PSAT scores from the LEGS output based on the equipercentile linking with a smoothing value of 1.

Table 3.3. RMSEL for different linking methods

METHOD	MEAN	LINEAR	PARALLEL LINKING	EQUIPERCENTILE	SMOOTHING=0.30	SMOOTHING=1.00
RMSEL	10.4116	10.2844	10.2844	10.2636	10.2553	10.2547

Table 3.4. The Mapped CLT8 and CLT10 Scores from the Equipercentile Linking with a Smoothing Value of 1

CLT8	CLT10	CLT8	CLT10	CLT8	CLT10	CLT8	CLT10
25	28	49	50	73	73	97	97
26	29	50	51	74	74	98	98
27	30	51	52	75	75	99	99
28	31	52	53	76	76	100	100
29	32	53	54	77	77	101	101
30	33	54	55	78	78	102	102
31	33	55	56	79	79	103	103
32	34	56	57	80	80	104	104
33	35	57	58	81	81	105	105
34	36	58	59	82	82	106	106
35	37	59	60	83	83	107	107
36	38	60	61	84	84	108	108
37	39	61	62	85	85	109	109
38	40	62	63	86	86	110	110
39	41	63	64	87	87	111	111
40	42	64	65	88	88	112	111
41	43	65	66	89	89	113	112
42	44	66	66	90	90	114	113
43	45	67	67	91	91	115	113
44	46	68	68	92	92	116	114
45	47	69	69	93	93	117	114
46	48	70	70	94	94	118	115
47	48	71	71	95	95	119	116
48	49	72	72	96	96	120	116

As seen in Table 3.4, the range for the CLT8 score is only from 25 to 120, which are the highest and lowest possible CLT8 scores in the cleaned sample for the concordance study. The LEGS program has conducted interpolation in filling the blanks in the CLT8 scores. Appendix C presents the frequency for the CLT8 scores for equipercentile linking. Some CLT8 score points, such as 26, 27, and 30-34 (see Appendix C), were missing in the data for the concordance relationship development. As the valid CLT8 scores range from 0 to 120, scores below 25 need to be extrapolated onto the CLT10 scores (see Table 3.4).

As the CLT8 scores at the lower end of the scale were not present in the matched samples, extrapolation is required. Several models, including power, linear, exponential, and polynomial (with different orders going from 2 to 6), were fitted with the matched scores based on equipercentile linking. A scatterplot (see Figure 3) was generated to examine the relationship between the CLT8 and the mapped CLT10 scores output from LEGS, as presented in the concordance table in Table 3.4. A prediction equation was developed using Excel's adding the trendline function to find the best fitting model. The equations for the fitted models and the R-squares which indicates the total variance explained by the fitted models are summarized in Table 3.5.

Figure 4. Scatterplot for CLT8 and CLT10_Mapped Scores from LEGS.

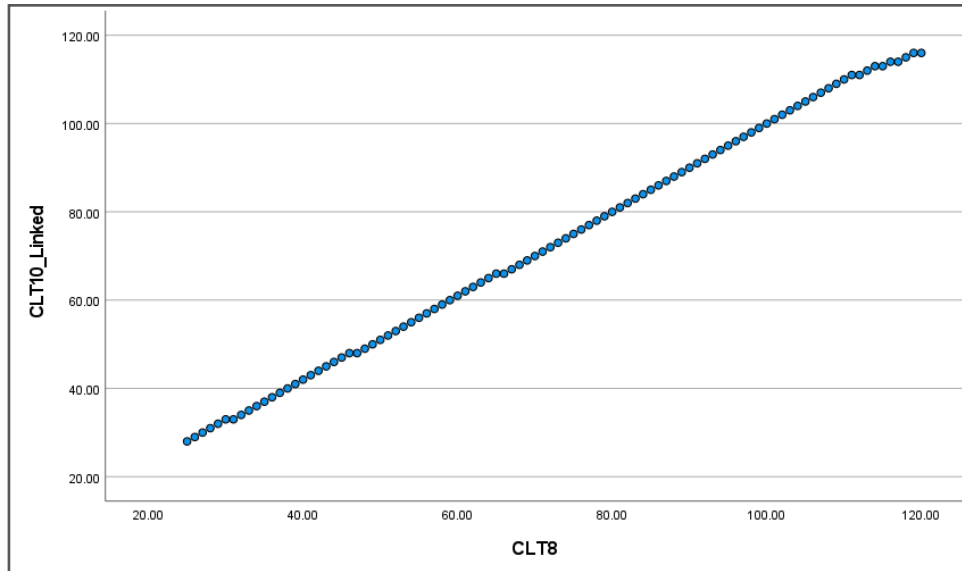


Table 3.5. The Prediction Equations for the CLT8 to CLT10 scores

TRENDLINE OPTIONS	R-SQUARE	PREDICTION EQUATION
Power	0.9992	■
Linear	0.9995	
Exponential	0.9689	■
Polynomial 2	0.9995	■
Polynomial 3	0.9997	■
Polynomial 4	0.9998	■
Polynomial 5	0.9999	■
Polynomial 6	0.9999	■

Note: X represents the CLT8 scores while Y represents the CLT10 scores.

Based on the total variance explained (R-square), the reasonableness of the extrapolated values for the CLT10 and the differences between actual CLT10 scores and the predicted CLT10 scores based on the CLT8 scores, the linear function was identified as the best fitting model with the best reasonableness.

Though all the polynomial functions displayed better fit in terms of the R-squares when compared with the linear model, some of the predicted CLT10 values were out of boundaries or deviated too much from the original CLT10 scores used for developing the best fitting model. Thus, for the lower end of the CLT8 scale, the linear extrapolation equation was used to fill in the CLT10 scores. There is no need for extrapolation at the upper end of the scale. The complete concordance table is presented in Table 3.6, with rounding of the mapped CLT10 scores.

Table 3.6. The Mapped CLT8 and the CLT10 scores.

CLT8	CLT10	CLT8	CLT10	CLT8	CLT10	CLT8	CLT10
0	4	31	33	61	62	91	91
1	5	32	34	62	63	92	92
2	5	33	35	63	64	93	93
3	6	34	36	64	65	94	94
4	7	35	37	65	66	95	95
5	8	36	38	66	66	96	96
6	9	37	39	67	67	97	97
7	10	38	40	68	68	98	98
8	11	39	41	69	69	99	99
9	12	40	42	70	70	100	100
10	13	41	43	71	71	101	101
11	14	42	44	72	72	102	102
12	15	43	45	73	73	103	103
13	16	44	46	74	74	104	104
14	17	45	47	75	75	105	105
15	18	46	48	76	76	106	106
16	19	47	48	77	77	107	107
17	20	48	49	78	78	108	108
18	21	49	50	79	79	109	109
19	22	50	51	80	80	110	110
20	23	51	52	81	81	111	111
21	24	52	53	82	82	112	111
22	25	53	54	83	83	113	112
23	26	54	55	84	84	114	113
24	27	55	56	85	85	115	113
25	28	56	57	86	86	116	114
26	29	57	58	87	87	117	114
27	30	58	59	88	88	118	115
28	31	59	60	89	89	119	116
29	32	60	61	90	90	120	116
30	33						

Based on the concordance table developed for the conversion of CLT10 and PSAT scores (See Appendix D), the mapped score for the CLT8 and PSAT can be obtained as shown in Table 3.7. For reasonableness, the CLT8 score of 0 is mapped to the lowest PSAT score of 320 even though the corresponding CLT10 score is 4, mapping to a PSAT score of 4.

Table 3.7. The Mapped CLT8 and the PSAT scores.

CLT8	PSAT	CLT8	PSAT	CLT8	PSAT	CLT8	PSAT
0	320*	31	680	61	920	91	1180
1	370	32	690	62	930	92	1190
2	370	33	700	63	940	93	1200
3	390	34	710	64	950	94	1210
4	400	35	720	65	950	95	1210
5	410	36	730	66	950	96	1230
6	420	37	750	67	970	97	1240
7	430	38	780	68	980	98	1250
8	440	39	800	69	990	99	1260
9	450	40	800	70	1000	100	1270
10	460	41	810	71	1010	101	1280
11	470	42	820	72	1020	102	1290
12	480	43	820	73	1030	103	1310
13	490	44	840	74	1040	104	1320
14	510	45	840	75	1050	105	1330
15	520	46	840	76	1060	106	1340
16	530	47	840	77	1070	107	1350
17	540	48	850	78	1080	108	1360
18	550	49	850	79	1100	109	1370
19	560	50	850	80	1100	110	1380
20	570	51	860	81	1110	111	1400
21	580	52	860	82	1120	112	1400
22	590	53	870	83	1130	113	1430
23	600	54	870	84	1140	114	1460
24	610	55	880	85	1150	115	1460
25	620	56	880	86	1160	116	1470
26	630	57	880	87	1160	117	1470
27	640	58	890	88	1170	118	1480
28	660	59	900	89	1180	119	1490
29	670	60	910	90	1180	120	1490
30	680						

Once the concordance relationship between the CLT8 and the PSAT scores is constructed, normative information for the PSAT can be used as a reference to compare CLT8 scores with PSAT norm groups. The normative information for the 2018 PSAT is presented in Appendix E. The table below provides the percentile ranks of each PSAT score in reference to two norm groups: one is the nationally representative sample of 10th graders taking the PSAT, and the other is the PSAT/NMSQT and PSAT 10 users. Through the mapping of the CLT8 and PSAT scores and the norms developed for the PSAT, normative comparison of the CLT8 scores in reference to the nationally representative sample of 10th graders taking the PSAT and the PSAT/NMSQT and PSAT 10 users can be established empirically. Table 3.8 presents the mapped CLT8 and PSAT scores, and the percentile ranks for the PSAT scores based on different norm groups.

Table 3.8. The percentile ranks of the CLT8 scores in reference to the PSAT 10th grade nationally representative sample and the PSAT/NMSQT and PSAT 10 users.

CLT8	PSAT	NATION- ALLY REP- RESEN- TATIVE SAMPLE	PSAT/ NMSQT AND PSAT 10 USER	CLT8	PSAT	NATION- ALLY REP- RESEN- TATIVE SAMPLE	PSAT/ NMSQT AND PSAT 10 USER
0	360	1-	1-	61	920	50	51
1	370	1-	1-	62	930	52	53
2	370	1-	1-	63	940	54	54
3	390	1-	1-	64	950	56	56
4	400	1-	1-	65	950	56	56
5	410	1-	1-	66	950	56	56
6	420	1-	1-	67	970	60	60
7	430	1-	1-	68	980	62	62
8	440	1-	1-	69	990	63	64
9	450	1-	1-	70	1000	65	66
10	460	1-	1-	71	1010	67	67
11	470	1-	1-	72	1020	69	69
12	480	1-	1-	73	1030	71	71
13	490	1-	1-	74	1040	73	72
14	510	1-	1-	75	1050	75	74
15	520	1-	1-	76	1060	76	76
16	530	1-	1-	77	1070	78	77
17	540	1-	1-	78	1080	79	78
18	550	1-	1	79	1100	82	81
19	560	1-	1	80	1100	82	81
20	570	1-	1	81	1110	83	83
21	580	1-	1	82	1120	84	84
22	590	1-	1	83	1130	86	85
23	600	1-	1	84	1140	87	86
24	610	1-	1	85	1150	88	87
25	620	1	2	86	1160	89	88
26	630	1	2	87	1160	89	88
27	640	1	2	88	1170	90	89
28	660	2	4	89	1180	91	90
29	670	2	4	90	1180	91	90
30	680	3	5	91	1180	91	90
31	680	3	5	92	1190	91	91
32	690	4	6	93	1200	92	92
33	700	5	7	94	1210	93	93
34	710	6	9	95	1210	93	93
35	720	8	10	96	1230	94	94
36	730	9	12	97	1240	95	94
37	750	13	16	98	1250	95	95
38	780	19	22	99	1260	96	95

39	800	24	26	100	1270	96	96
40	800	24	26	101	1280	97	96
41	810	26	28	102	1290	97	97
42	820	28	30	103	1310	97	97
43	820	28	30	104	1320	98	98
44	840	33	34	105	1330	98	98
45	840	33	34	106	1340	98	98
46	840	33	34	107	1350	98	98
47	840	33	34	108	1360	99	98
48	850	36	36	109	1370	99	99
49	850	36	36	110	1380	99	99
50	850	36	36	111	1400	99	99
51	860	38	38	112	1400	99	99
52	860	38	38	113	1430	99+	99
53	870	40	41	114	1460	99+	99+
54	870	40	41	115	1460	99+	99+
55	880	42	43	116	1470	99+	99+
56	880	42	43	117	1470	99+	99+
57	880	42	43	118	1480	99+	99+
58	890	44	45	119	1490	99+	99+
59	900	46	47	120	1490	99+	99+
60	910	48	49				

The interpretation of the percentile ranks in Table 3.8 for each of the CLT8 scores in reference to the PSAT norm groups is the same as that explained in Chapter 2. Based on the norming study presented in Chapter 2, a CLT8 score of 91 was in the 85th percentile against the nationally representative sample of the CLT8 student population. Based on the method used in this chapter, a CLT8 score of 91 was the 91st percentile, indicating that a student who gets a CLT8 score of 91 performed the same or better than 91% of the PSAT national representative sample, and the same or better than 90% of the PSAT/NMSQT and PSAT 10 users. These results suggest that the students used for this exploratory empirical norming study were higher-performing students compared with the general 10th grade PSAT students and PSAT/NMSQT and PSAT 10 users.

3.5 LIMITATIONS AND FUTURE CONSIDERATIONS

One of the limitations of the linking study presented in this chapter is the availability of related test scores. This initial exploratory report intends to provide normative information about CLT8 scores against different nationally representative norming groups. The CLT8 is designed for students in grades 7 and 8. When these students took the CLT8, most of them had not yet taken the PSAT. Moreover, the Covid-19 pandemic caused multiple cancellations of PSAT test administrations. Thus, PSAT scores became even more sparse.

Most students who have taken the CLT8 have not yet taken the CLT10. This current study used students' test scores on both the CLT8 and CLT10 to set up the linkage with the PSAT. This is a retrospective linkage. Students included in this study took the CLT10 recently, but likely took the CLT8 approximately two years ago. Though we could assume that the CLT8 student population stayed relatively stable, caution

should be exercised when interpreting the results. Anecdotally based on feedback from interviews with CLT8 customers, it seems that some CLT8 test takers do not go on to take the CLT10, since the CLT8 assessment is quite challenging for students and, until now, educators, students, and parents did not have enough context in the student analytics to understand what the scores mean. The goal of this study is to put those scores into context so that they are more interpretable to educators and parents.

Further, as noted, the students with both valid CLT8 and CLT10 scores did not match the national target population well; some targets were unbalanced in many variables compared with the national target population. This is not surprising given the CLT test taking population. The limitation from the study linking the CLT10 and PSAT scores also deserves note. The study used students' self-reported PSAT scores; it was observed that higher-performing students were more likely to report their PSAT scores. Thus, the students that participated in that empirical norming study were likely higher-performing students than the CLT10 test taker pipeline as a whole. Given this, the percentile ranks for the same CLT10 score could be different from the norming study reported in Chapter 2 and that for the empirical norming study reported in this chapter. In consequence, over-interpretation or use of the normative information of the CLT8 scores should be avoided. As more representative CLT8 and CLT10 test taker data becomes available through our pipeline, CLT will conduct additional empirical norming studies using a nationally representative sample and/or representative CLT8 user sample.

As highlighted in Chapter 2, CLT8's user base is expected to be changing, especially due to the impact of the COVID-19 pandemic and due to a growing CLT8 population. CLT is committed to conducting additional studies on the empirical norms related to the PSAT norm groups, to ensure that the empirical norms are referencing the most up-to-date CLT8 population performance. The normative information presented in this chapter will have improved validity and generalizability when CLT8 students who participate in the empirical norm study become stable and representative of the CLT8 student population.

3.6 CONCLUSIONS

This chapter presented a linking study to develop a concordance relationship between the CLT8 and PSAT. This relationship helps to contextualize CLT8 scores for educators and parents. Given the challenge of collecting representative data, this study should be considered as a preliminary exploration of the comparison of CLT8 students' performance in reference to the PSAT national representative sample and the PSAT/NMSQT and PSAT 10 users. As most CLT8 test takers have not taken the PSAT, and considering the widespread cancellations of PSAT test administrations in 2020, this study utilized a two-step linkage approach to set up the concordance relationship between CLT8 and PSAT scores. First, the concordance relationship was established between CLT8 and CLT10 scores; based on the concordance relationship previously established between CLT10 and PSAT scores, the CLT8 and PSAT scores were then mapped to one another. Utilizing the normative information developed for the PSAT score in reference to the PSAT national representative sample and the PSAT/NMSQT and PSAT 10 users, the empirical percentile ranks for CLT8 scores in reference to the two norm groups for the PSAT score interpretation were obtained through these mapped scores between the CLT8 and the PSAT.

Two major highlights are summarized below:

1. The CLT8 and CLT10 sample used in this empirical norming study was a convenience sample, and may not be appropriately representative of the CLT8 student population, especially under pandemic conditions. Even if the sample of students is adequately representative, their CLT8 performance may be negatively impacted by the pandemic, which could reduce the generalizability

of the empirical norms developed in this chapter. Moreover, when the concordance relationship was developed between the CLT10 and PSAT, the PSAT scores used were self-reported. The integrity of the data will be reviewed in future replication studies of this work. A larger and more representative sample will be used for future empirical norm development and updating.

2. As highlighted in Chapter 2, the current user base for the CLT8 is likely to change in the future. Given the need for stakeholders to use the empirical norms in comparing CLT8 students with the PSAT national representative sample and PSAT/NMSQT and PSAT 10 users, CLT will conduct additional norming studies in the future to maintain the integrity of the normative scores obtained in such studies.



4. SUMMARY AND DISCUSSIONS

4.1 SUMMARY

The purpose of the studies reported in Chapters 2 and 3 is to provide normative information about the performance of a CLT8 student who earns a specific test score as compared to a target population. Different target populations have been used as a reference for interpreting the CLT8 test scores within a normative framework. In Chapter 2, normative information about the CLT8 is developed by identifying a national representative sample of the CLT8 student population. In Chapter 3, CLT8 students' performance is compared to two PSAT populations: namely, the nationally representative sample of the PSAT and the PSAT/NMSQT and PSAT10 user group. This was done through a longitudinal linking of CLT8 scores to CLT10 scores, and then linking the CLT10 scores to the PSAT samples utilizing self-reported PSAT scores.

The norming results obtained relative to a representative sample from the Spring 2021 CLT8 test administration was reported in Chapter 2. The targets of the population characteristics of the demographics for the CLT8 national population were derived from the 2016 NCES national survey for private and home schools, published in the Digest of Education Statistics. The normative information from Chapter 2 indicates the relative ranking of a specific CLT8 score relative to the national normative sample developed from the Spring 2021 CLT8 test data.

The empirical normative information about the ranking of CLT8 scores relative to two 2018 PSAT normative samples was presented in Chapter 3.¹ Such information was obtained based on an empirical linking study to develop the concordance relationship between the CLT8 scores and the PSAT scores. The linking study developed a concordance relationship between the CLT8 and CLT10 scores through the equipercentile linking method first, then the concordance relationship between the CLT10 scores and the PSAT scores from the previous concordance study was utilized to map the CLT8 scores onto the PSAT scores. In the end, a concordance table was developed to show how each CLT8 score is mapped onto the PSAT scale. Then, the normative information constructed for the Spring 2018 PSAT is used to compare a CLT8 score with a mapped PSAT score, indicating the relative ranking of a CLT8 score

¹ As discussed in Chapter 3, the sample used to derive the Predicted PSAT score was a different sample than the sample to generate the CLT8 User Percentiles in Chapter 2. The sample used in Chapter 3 required us to link the CLT8 to the PSAT through the relationship of the CLT8 to the CLT10. Therefore the sample used in Chapter 3 required us to select students that had both CLT8 and CLT10 scores.

compared with the two normative sample for the 2018 norm development of the PSAT: the nationally representative sample of the PSAT, and the PSAT/NMSQT and PSAT10 user group.

The percentile ranks corresponding to each CLT8 score based on the norming study for the CLT8 nationally representative sample and the empirical norming investigation based on the PSAT nationally representative sample and the PSAT/NMSQT and PSAT10 users respectively are summarized in Table 4.1. In general, the percentile ranks for the CLT8 scores for the CLT8 norming sample are lower than those for the PSAT nationally representative sample and the PSAT/NMSQT and PSAT10 User sample. This is consistent with expectations, given what is presented in Table 3.2, that is, the mean score for the CLT8 norming sample based on the Spring 2021 CLT8 test administrations was lower than that for the sample used in the concordance study. This sample difference in their mean CLT8 scores is worthy of note.

Table 4.1 The Summary of the Norming Results

CLT8 SCORES	CLT8 NORMING SAMPLE	PSAT NATIONALLY REPRESENTATIVE SAMPLE	PSAT/NMSQT AND PSAT 10 USER	CLT SCORES	CLT8 NORMING SAMPLE	PSAT NATIONALLY REPRESENTATIVE SAMPLE	PSAT/NMSQT AND PSAT 10 USER
0	1	1-	1-	61	36	50	51
1	1	1-	1-	62	39	52	53
2	1	1-	1-	63	41	54	54
3	1	1-	1-	64	43	56	56
4	1	1-	1-	65	45	56	56
5	1	1-	1-	66	48	56	56
6	1	1-	1-	67	50	60	60
7	1	1-	1-	68	51	62	62
8	1	1-	1-	69	53	63	64
9	1	1-	1-	70	54	65	66
10	1	1-	1-	71	56	67	67
11	1	1-	1-	72	58	69	69
12	1	1-	1-	73	61	71	71
13	1	1-	1-	74	62	73	72
14	1	1-	1-	75	64	75	74
15	1	1-	1-	76	66	76	76
16	1	1-	1-	77	67	78	77
17	1	1-	1-	78	69	79	78
18	1	1-	1	79	70	82	81
19	1	1-	1	80	72	82	81
20	1	1-	1	81	73	83	83
21	1	1-	1	82	75	84	84
22	1	1-	1	83	76	86	85
23	1	1-	1	84	77	87	86
24	1	1-	1	85	78	88	87

25	1	1	2	86	80	89	88
26	1	1	2	87	81	89	88
27	1	1	2	88	82	90	89
28	1	2	4	89	83	91	90
29	1	2	4	90	85	91	90
30	1	3	5	91	85	91	90
31	2	3	5	92	86	91	91
32	2	4	6	93	87	92	92
33	2	5	7	94	88	93	93
34	3	6	9	95	89	93	93
35	3	8	10	96	90	94	94
36	4	9	12	97	91	95	94
37	5	13	16	98	92	95	95
38	5	19	22	99	93	96	95
39	5	24	26	100	94	96	96
40	7	24	26	101	94	97	96
41	8	26	28	102	95	97	97
42	8	28	30	103	96	97	97
43	9	28	30	104	97	98	98
44	10	33	34	105	97	98	98
45	11	33	34	106	98	98	98
46	12	33	34	107	98	98	98
47	13	33	34	108	98	99	98
48	14	36	36	109	99	99	99
49	16	36	36	110	99	99	99
50	17	36	36	111	99	99	99
51	19	38	38	112	99	99	99
52	21	38	38	113	99	99+	99
53	23	40	41	114	99	99+	99+
54	24	40	41	115	99	99+	99+
55	26	42	43	116	99	99+	99+
56	28	42	43	117	99	99+	99+
57	30	42	43	118	99	99+	99+
58	31	44	45	119	99	99+	99+
59	33	46	47	120	99	99+	99+
60	35	48	49				

The change patterns of the percentile ranks along the CLT8 score scale are graphically presented in Figure 4.1. Along the vast majority of the CLT8 scale range, except at the ends of the scale, the percentile ranks corresponding to each CLT8 score relative to the CLT8 nationally representative sample developed in Chapter 2 are lower than those in reference to the PSAT nationally representative sample

and to the PSAT/NMSQT and PSAT10 users developed in Chapter 3. The differences are larger for the score range from 38 to 64, with the largest difference of 24 and the smallest difference of 13 within this interval. The differences in the two PSAT norming samples are much smaller within this scale range, with the largest difference of 3 and the smallest of 0. It is expected that such differences are present in the percentile ranks. This is because the norming results are sample dependent. Figures 4.2 and 4.3 present the score distributions for the norming sample used in Chapter 2 and the concordance sample used in Chapter 3, respectively. They differ from each other especially in terms of the mean, the score range, the minimum score, and the standard deviation. The norming and the concordance studies used different samples, which diverge in sample size, school type, gender, ethnicity, region, and locale. The norming sample used in Chapter 2 resembles the target CLT8 population while the sample used in Chapter 3 in the concordance study differed more from the target CLT8 student population in terms of the key demographic variables as listed above. Further, the sample in the concordance study is a relatively high-performing group of students compared with those in the norming study in Chapter 2. Thus, it is expected that the percentile ranks associated with the same CLT8 score would be lower for the CLT8 nationally representative sample compared with those from the CLT8 concordance sample who reported their PSAT scores.

Figure 4.1 The percentile ranks for each CLT8 score relative to different target populations.

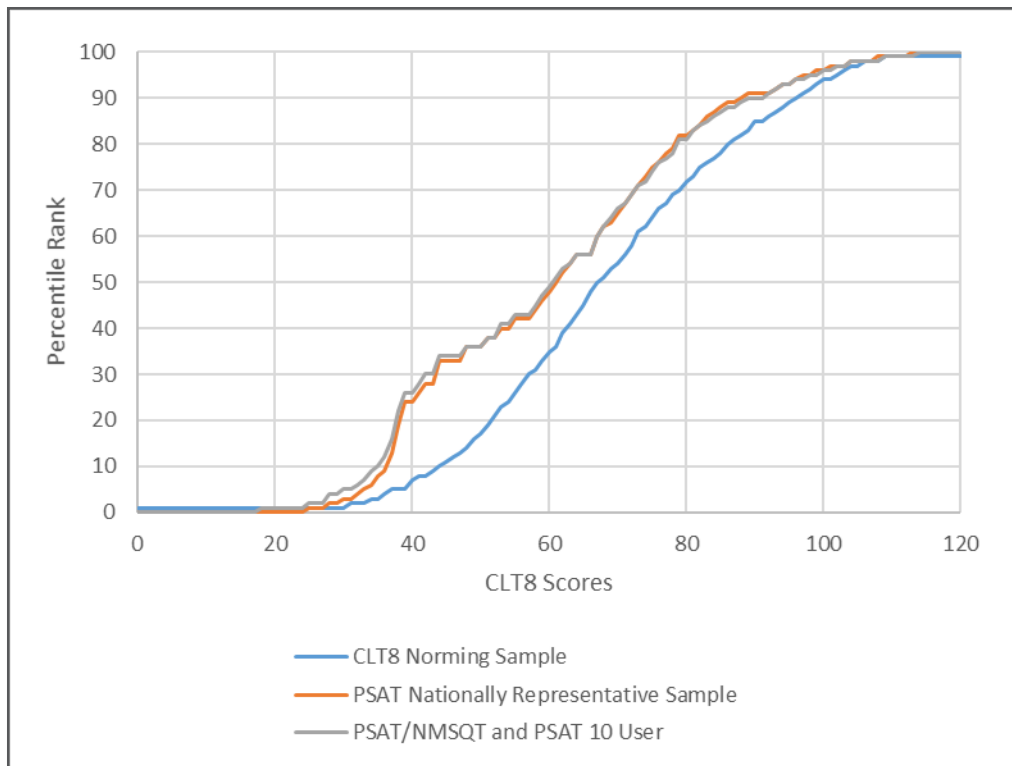


Figure 4.2 Histogram for the CLT8 scores for the concordance sample

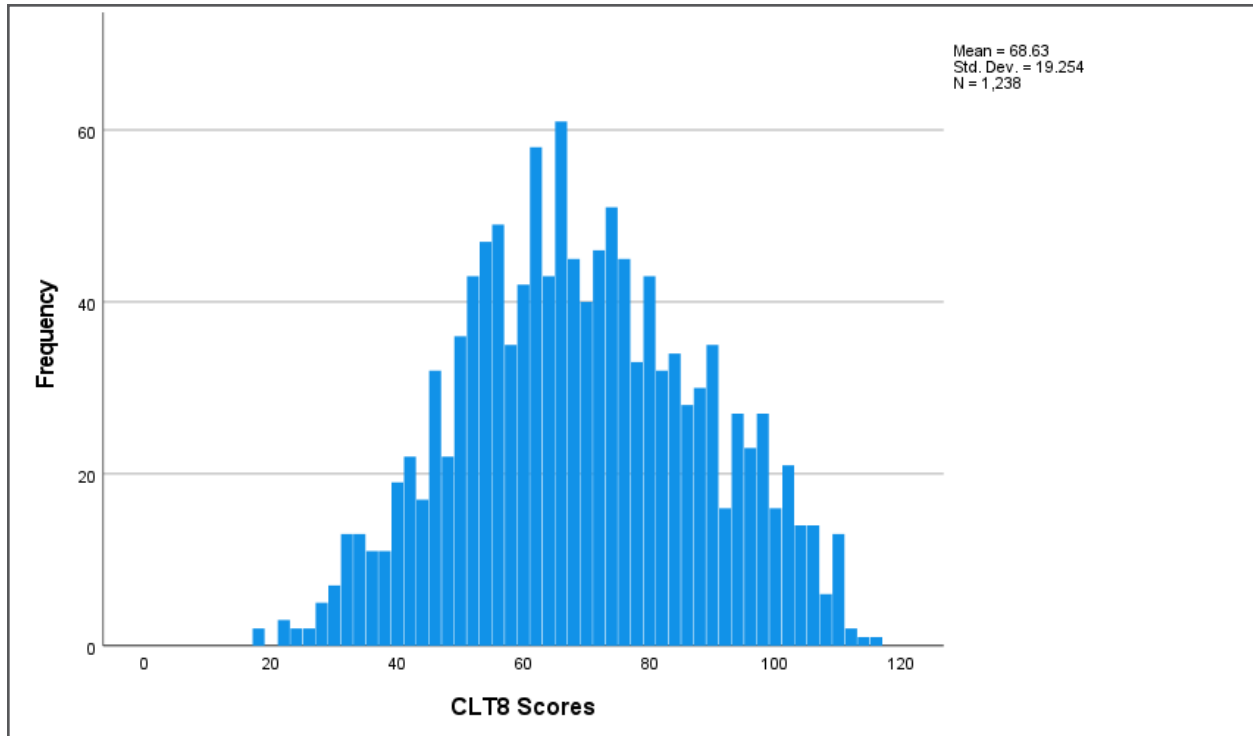
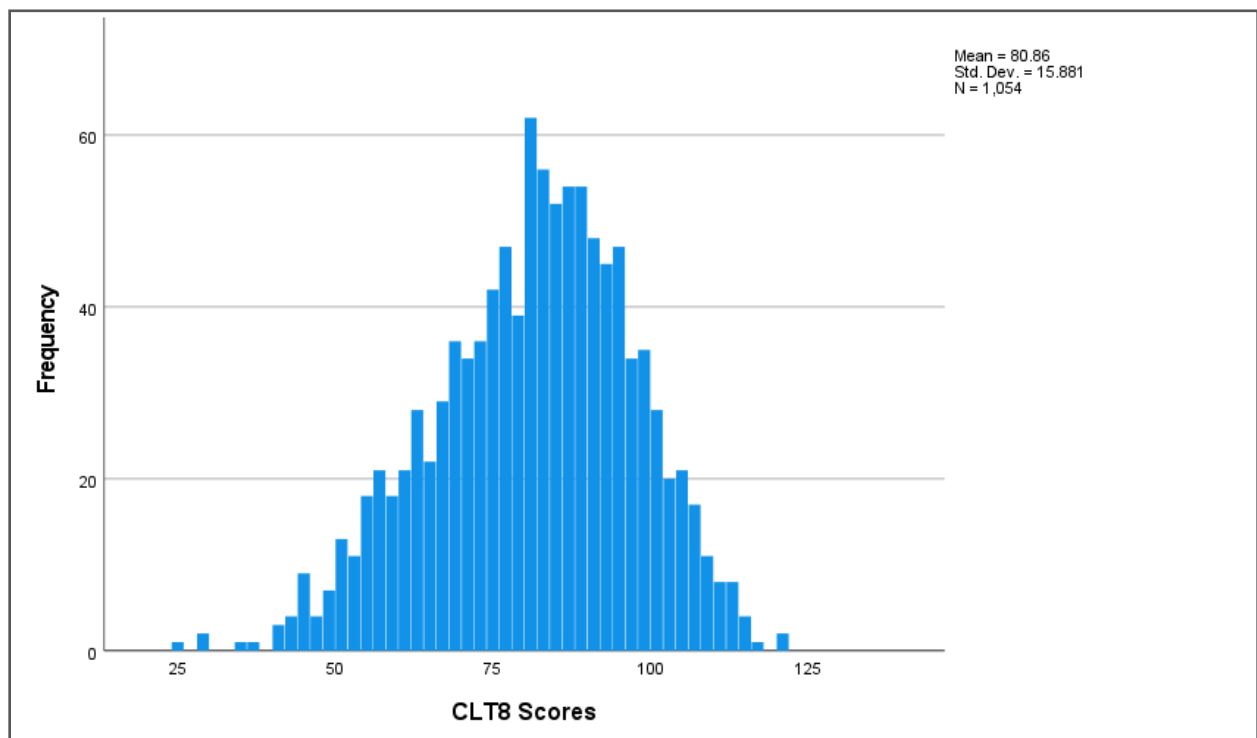


Figure 4.3 Histogram for the CLT8 scores for the national norming sample



In terms of the representativeness of the two study samples used in Chapters 2 and 3 of the target national population of the CLT8, the norming sample constructed in Chapter 2 is a more nationally representative sample of the target CLT8 population, which is composed of primarily private schooled and homeschooled students. However, the percentile ranks produced in mapping the relationship between the CLT8 and PSAT scores provide normative information about the CLT8 students' performance relative to the PSAT nationally representative sample and to the PSAT/NMSQT and PSAT10 users, using the concordance relationship developed between the CLT10 scores and the PSAT scores from the Spring 2019 CLT10 data. The populations used in the PSAT samples are primarily public school students, whereas most of the students that take CLT8 are private schooled and homeschooled students. Given the limitations of the samples obtained for the CLT8 study, caution should be exercised in interpreting the normative information from both studies.

4.2 DISCUSSIONS

Chapters 2 and 3 report two studies intended to provide normative information about CLT8 scores. These two studies are preliminary first steps in developing national norms and provide normative information related to interpreting CLT8 scores. The CLT8 national percentile norms were developed using a representative sample of 1,238 7th and 8th grade students who took the CLT8 during a Spring 2021 administration. In general, the norms developed from this study provide information about the interpretation of the CLT8 scores compared to the national sample of the CLT8 population. As highlighted in Chapter 2, caution should be taken in interpreting the norms reported from the norming study.

1. The current CLT8 normative sample was derived from the Spring 2021 CLT8 sample with an unbalanced representation of the target population. This sample only contains scores ranging from 18 to 116, which does not cover the whole score range of the CLT8. Sample stratification and weighting balanced the demographics, at the cost of reduction of sample size and potential loss of information. In addition, the difference in region and locale between the norming sample and the population targets was larger compared with the matching on other variables. This limitation could reduce the validity of the developed norms. A larger and more representative sample with the full ability range and distributions on key demographic variables should be used for future development and updating of the national norms for the CLT8.
2. The CLT8 sample used in this norming study is from the Spring 2021 test administration. The impact of COVID-19 is unknown at the moment related to students' test performance and any potential changes in key demographic variables. Further, it is expected that the current user base for the CLT8 may change as time goes by when more students from the target population take the test. This will likely lead to a distribution shift in the CLT8 scores. It is strongly recommended that timely evaluation and norm updating be conducted to maintain the validity of the normative scores for the CLT8.

The establishment of the concordance relationship between CLT8 scores and PSAT scores provides another perspective in interpreting CLT8 scores within the norm-referenced framework. The concordance between CLT8 scores and PSAT scores helps to compare the relative performance of CLT8 test takers against two different norming groups of PSAT scores. However, caution should be taken when

interpreting the normative information from the concordance relationship for the following reasons.

1. The representativeness of the sample used in this concordance relationship study is a concern. The sample size was 1,054, with a score range from 25 to 120. This score range is wider than that for the norming sample used in Chapter 2; however, it still did not cover the full range of the score scale. In addition, the concordance sample deviated further from the national target population in some key demographic variables. In general, the sample in the concordance study was higher performing, which led to the mapping of a higher percentile rank for the same CLT8 score, compared to that from the norming study. More importantly, the content alignment for the CLT8, CLT10 and the PSAT was not completely overlapping though CLT8 and CLT10 content domains and subdomains were aligned more closely. The content coverage and distribution of the content domains differ between the two tests, and score interpretations are different. Thus, the normative comparison across different content and different populations should be treated with caution.
2. In terms of representativeness, the impact of the COVID-19 on student performance on CLT8 and CLT10 remains unclear and needs to be investigated. Further, the CLT8's user base is evolving. It is expected that empirical normative information compared to the PSAT norm groups presented in this concordance study may change and should be updated in a timely fashion to ensure that such normative information is compared to the most up-to-date CLT8 population performance. Though the percentile ranks/norms are sample dependent, when CLT8 students who participate in the concordance study become stable and representative of the CLT8 national population, the normative information will become more valid and generalizable.
3. The concordance relationship between the CLT8 scores and the PSAT scores was developed using CLT10 scores as a bridge: that is, CLT8 scores were mapped onto the CLT10 score scale. The concordance between CLT10 scores and PSAT scores established in 2019 was used to identify the mapping between CLT8 scores and PSAT scores. It is noted that the sample size of 220 for the concordance study between CLT10 scores and PSAT scores was small. It would be ideal to update the concordance relationship between CLT10 scores and PSAT scores using additional data. However, due to the cancellation of the recent PSAT test administration due to COVID, the number of students who reported their PSAT scores was scarce (N=30, minimum=790, maximum=1460). CLT plans to replicate the CLT10-to-PSAT concordance study, and update the concordance relationship between CLT10 scores and PSAT scores as needed, to reflect the relationship between the CLT10 and PSAT as populations change.

Appendix A

CONTENT ALIGNMENT

Table A.1 The CLT8 Test Structure

Section	Number of Questions	Time
Verbal Reasoning	40	45 min
Grammar/Writing	40	40 min
Quantitative Reasoning	40	50 min

Table A.2 The CLT10 Test Structure

Section	Number of Questions	Time
Verbal Reasoning	40	40 min
Grammar/Writing	40	35 min
Quantitative Reasoning	40	45 min

Table A.3 The PSAT Test Structure

Section	Number of Questions	Time
Reading	47	60 min
Writing and Language	44	35 min
Math (No Calculator)	17	25 min
Math (With Calculator)	31	40 min

Appendix B

SCREENSHOT FOR RUNNING THE LEGS PROGRAM

The screenshot shows the 'Equating Program - LEGS 2.0.1' window. It has a menu bar with 'File' and 'Window'. Below the menu bar are two tabs: 'Raw Scores as Input' (selected) and 'Frequency Distributions as Input'. The main area is divided into several sections:

- X Scores :**
 - Alphanumeric Identifier :
 - Column for X Score :
 - Lowest Score :
 - Highest Score :
 - File ... :
- Y Scores :**
 - Alphanumeric Identifier :
 - Column for Y Score :
 - Lowest Score :
 - Highest Score :
- Subgroups :**
 - Number of Subgroups :
 - Names of Subgroups :
 - Number of examinees in each subgroup :
- Equipercntile Smoothing**
 - slim :
 - number of Smoothing Values :
 - List of Smoothing Values :
- Truncation**
 - Lowest Valid Score(Y) :
 - Highest Valid Score(Y) :
 - Truncation :
- Options**
 - remsd_wts :

At the bottom, there are two buttons: 'Read Parameters from Control File' and 'OK'.

Appendix C

CLT8 SCORE FREQUENCY FOR LINKING

SCORE	FRE- QUENCY	PERCENT	VALID PERCENT	CUMU- LATIVE PERCENT	SCORE	FRE- QUENCY	PERCENT	VALID PERCENT	CUMU- LATIVE PERCENT
25	1	0.1	0.1	0.1	77	28	2.7	2.7	38.7
28	1	0.1	0.1	0.2	78	24	2.3	2.3	41.0
29	1	0.1	0.1	0.3	79	15	1.4	1.4	42.4
35	1	0.1	0.1	0.4	80	28	2.7	2.7	45.1
37	1	0.1	0.1	0.5	81	34	3.2	3.2	48.3
40	2	0.2	0.2	0.7	82	23	2.2	2.2	50.5
41	1	0.1	0.1	0.8	83	33	3.1	3.1	53.6
42	2	0.2	0.2	0.9	84	27	2.6	2.6	56.2
43	2	0.2	0.2	1.1	85	25	2.4	2.4	58.5
44	6	0.6	0.6	1.7	86	29	2.8	2.8	61.3
45	3	0.3	0.3	2.0	87	25	2.4	2.4	63.7
46	3	0.3	0.3	2.3	88	25	2.4	2.4	66.0
47	1	0.1	0.1	2.4	89	29	2.8	2.8	68.8
48	2	0.2	0.2	2.6	90	27	2.6	2.6	71.3
49	5	0.5	0.5	3.0	91	21	2.0	2.0	73.3
50	7	0.7	0.7	3.7	92	28	2.7	2.7	76.0
51	6	0.6	0.6	4.3	93	17	1.6	1.6	77.6
52	5	0.5	0.5	4.7	94	34	3.2	3.2	80.8
53	6	0.6	0.6	5.3	95	13	1.2	1.2	82.1
54	8	0.8	0.8	6.1	96	14	1.3	1.3	83.4
55	10	0.9	0.9	7.0	97	20	1.9	1.9	85.3
56	9	0.9	0.9	7.9	98	17	1.6	1.6	86.9
57	12	1.1	1.1	9.0	99	18	1.7	1.7	88.6
58	8	0.8	0.8	9.8	100	18	1.7	1.7	90.3
59	10	0.9	0.9	10.7	101	10	0.9	0.9	91.3
60	8	0.8	0.8	11.5	102	10	0.9	0.9	92.2
61	13	1.2	1.2	12.7	103	10	0.9	0.9	93.2
62	12	1.1	1.1	13.9	104	14	1.3	1.3	94.5
63	16	1.5	1.5	15.4	105	7	0.7	0.7	95.2
64	9	0.9	0.9	16.2	106	8	0.8	0.8	95.9
65	13	1.2	1.2	17.5	107	9	0.9	0.9	96.8
66	15	1.4	1.4	18.9	108	6	0.6	0.6	97.3
67	14	1.3	1.3	20.2	109	5	0.5	0.5	97.8
68	24	2.3	2.3	22.5	110	5	0.5	0.5	98.3
69	12	1.1	1.1	23.6	111	3	0.3	0.3	98.6
70	16	1.5	1.5	25.1	112	3	0.3	0.3	98.9
71	18	1.7	1.7	26.9	113	5	0.5	0.5	99.3
72	22	2.1	2.1	28.9	114	4	0.4	0.4	99.7
73	14	1.3	1.3	30.3	116	1	0.1	0.1	99.8
74	21	2.0	2.0	32.3	120	2	0.2	0.2	100.0
75	21	2.0	2.0	34.3	Total	1054	100.0	100.0	
76	19	1.8	1.8	36.1					

Appendix D

THE MAPPED CLT10 SCORES AND THE PSAT SCORES

CLT10	PSAT	CLT10	PSAT	CLT10	PSAT	CLT10	PSAT
0	320	31	660	61	910	91	1180
1	330	32	670	62	920	92	1190
2	340	33	680	63	930	93	1200
3	350	34	690	64	940	94	1210
4	360	35	700	65	950	95	1210
5	370	36	710	66	950	96	1230
6	390	37	720	67	970	97	1240
7	400	38	730	68	980	98	1250
8	410	39	750	69	990	99	1260
9	420	40	780	70	1000	100	1270
10	430	41	800	71	1010	101	1280
11	440	42	800	72	1020	102	1290
12	450	43	810	73	1030	103	1310
13	460	44	820	74	1040	104	1320
14	470	45	820	75	1050	105	1330
15	480	46	840	76	1060	106	1340
16	490	47	840	77	1070	107	1350
17	510	48	840	78	1080	108	1360
18	520	49	850	79	1100	109	1370
19	530	50	850	80	1100	110	1380
20	540	51	850	81	1110	111	1400
21	550	52	860	82	1120	112	1430
22	560	53	860	83	1130	113	1460
23	570	54	870	84	1140	114	1470
24	580	55	870	85	1150	115	1480
25	590	56	880	86	1160	116	1490
26	600	57	880	87	1160	117	1510
27	610	58	880	88	1170	118	1500
28	620	59	890	89	1180	119	1510
29	630	60	900	90	1180	120	1520
30	640						

Appendix E

10TH GRADE PERCENTILES

TOTAL SCORE	NATIONALLY REPRESENTATIVE SAMPLE	PSAT/NMSQT AND PSAT 10 USER	TOTAL SCORE	NATIONALLY REPRESENTATIVE SAMPLE	PSAT/NMSQT AND PSAT 10 USER	TOTAL SCORE	NATIONALLY REPRESENTATIVE SAMPLE	PSAT/NMSQT AND PSAT 10 USER
1520	99+	99+	1100	82	81	680	3	5
1510	99+	99+	1090	81	80	670	2	4
1500	99+	99+	1080	79	78	660	2	4
1490	99+	99+	1070	78	77	650	1	3
1480	99+	99+	1060	76	76	640	1	2
1470	99+	99+	1050	75	74	630	1	2
1460	99+	99+	1040	73	72	620	1	2
1450	99+	99+	1030	71	71	610	1-	1
1440	99+	99+	1020	69	69	600	1-	1
1430	99+	99	1010	67	67	590	1-	1
1420	99	99	1000	65	66	580	1-	1
1410	99	99	990	63	64	570	1-	1
1400	99	99	980	62	62	560	1-	1
1390	99	99	970	60	60	550	1-	1
1380	99	99	960	58	58	540	1-	1-
1370	99	99	950	56	56	530	1-	1-
1360	99	98	940	54	54	520	1-	1-
1350	98	98	930	52	53	510	1-	1-
1340	98	98	920	50	51	500	1-	1-
1330	98	98	910	48	49	490	1-	1-
1320	98	98	900	46	47	480	1-	1-
1310	97	97	890	44	45	470	1-	1-
1300	97	97	880	42	43	460	1-	1-
1290	97	97	870	40	41	450	1-	1-
1280	97	96	860	38	38	440	1-	1-
1270	96	96	850	36	36	430	1-	1-
1260	96	95	840	33	34	420	1-	1-
1250	95	95	830	31	32	410	1-	1-
1240	95	94	820	28	30	400	1-	1-
1230	94	94	810	26	28	390	1-	1-
1220	94	93	800	24	26	380	1-	1-
1210	93	93	790	22	24	370	1-	1-
1200	92	92	780	19	22	360	1-	1-
1190	91	91	770	17	20	350	1-	1-
1180	91	90	760	15	18	340	1-	1-
1170	90	89	750	13	16	330	1-	1-
1160	89	88	740	11	14	320	1-	1-
1150	88	87	730	9				
1140	87	86	720	8				
1130	86	85	710	6				
1120	84	84	700	5				
1110	83	83	690	4				

