



CLT

2019

CLT¹⁰ Norm Reference Report

2019

CLT¹⁰ Norm
Reference Report

CONTRIBUTORS

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

Technical Manual Lead

Noah Tyler

Psychometricians

Hong Jiao

Liru Zhang

Xiaobin Zhou

Writers/Reviewers

Gabriel Blanchard

Mary Trent

Leadership Reviewers

Jeremy Tate

Graphic Designer

Meg Prom

Copyright © 2018 Classic Learning Initiatives, LLC. All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of Classic Learning Initiatives, LLC.

Letter from the President

The founding fathers of our nation were formed by a rich education, one that explored profound questions of morals, politics, and philosophy. The answers they gave to these questions remain relevant today, because the questions themselves are timeless, keyed into issues of life and purpose that affect every one of us. Their education was humanist—that is, it was aimed at giving its students a fully human intellectual life.

My experience as a public school student in the 1980s was very different. Great ideas were gone, replaced by rote memorization of uninspiring material. My classmates and I were desperately bored in consequence. Neither our books nor our teachers were presenting us with sacred or lasting concerns. The notion that school could have helped us engage with questions about human happiness and purpose would have seemed laughable.

In creating the Classic Learning Test and its preparatory exams, knowing that teachers will teach towards the test, I want to invigorate students to engage with the same timeless ideas that inspired previous generations, reaching back to our country's founding and to the education that they received.

Jeremy Tate

Jeremy Tate,
CLT President





Table of Contents

CHAPTER 1. INTRODUCTION	10
1.1 Introduction to CLT ¹⁰	
1.2 Introduction to Norming	
1.3 Overview of CLT ¹⁰ Norming Samples	
1.4 Procedures	
CHAPTER 2. NORMING	14
2.1 Overview of CLT10 National Norms	
2.2 National Target Populations	
2.3 CLT10 Initial Sample	
2.4 CLT10 National Percentile Norms	
2.5 Limitations and Future Considerations	
2.6 Conclusions	
CHAPTER 3. EMPIRICAL NORMING	22
3.1 Data Preparation	
3.2 PSAT and CLT Matching and Concordance Table Development	
3.3 Data Preparation	
3.4 Concordance Between CLT10 and PSAT Scores	
3.5 Limitations and Future Considerations	
3.6 Conclusions	
CHAPTER 4. SUMMARY AND DISCUSSIONS	36
4.1 Results Summary	
4.2 Final Remarks	
APPENDICES	40
Appendix A: Alignment Study	
Appendix B: Screenshot for running the LEGS program	
Appendix C: CLT10 Score Frequencies	
Appendix D: CLT10 and PSAT Percentile Rankings	
Appendix E: Assessment of Spelling on the CLT10	

Executive Summary

THIS REPORT PROVIDES NORMATIVE INFORMATION for interpreting CLT10 scores. Two avenues are explored in this report. The first one is a norming study that develops norms for CLT10 scores based on a nationally representative sample of the CLT10 population, while the other is to provide normative information for CLT10 scores relative to two PSAT nationally representative samples, through linking the two tests. The norming study reported in Chapter 2 focuses on the nationally representative sample of the CLT10 student population; the normative information presented in Chapter 3 compares CLT10 students' performance relative to two PSAT populations.

The norming study derived a representative sample from the spring 2019 CLT10 test administration. The NCES 2016 national survey on private and homeschool students (published in the Digest of Education Statistics) was used for the population targets. Percentile rank of a given CLT10 score was created based on this national normative sample. Details on the norming study are presented in Chapter 2.

The linking study established the concordance relationship between CLT10 scores and PSAT scores using the equipercentile linking method. This mapping relates the percentile ranks of CLT10 scores to two 2018 PSAT normative samples (the PSAT nationally representative sample, and PSAT/NMSQT and PSAT10 users for the 2018). A concordance table shows how each CLT10 score is mapped onto a PSAT score or a range of PSAT scores. The percentile for a CLT10 score was obtained from the normative information of the linked PSAT score(s). This linking study is described in detail in Chapter 3.

Along the majority of the CLT10 score scale range, the percentile ranks from the national norming study are lower than those from the linking study. Such differences are as expected because the CLT10 target population is expected to perform better than the general K-12 education population, which is represented by the PSAT normative sample. Thus, for the same CLT10 score, the percentile rank is lower on the CLT10 national norming group (which uses non-district educated students as the reference group) than on the PSAT linking sample (which uses the general ed population as the reference group). Furthermore, the score distributions for the norming sample and the linking sample differ from each other. The norming and the linking studies used samples that differ in size and other demographic variables, due to the availability of the PSAT scores reported. The norming sample resembles the target CLT10 population, whereas the linking sample, limited by the availability of the PSAT scores, differed from the target CLT10 student population in terms of the key demographic variables. The linking sample also has higher performance compared with those in the norming study. Thus, the differences in target populations and sample representations and performance are consistent with the lower percentile ranks observed in the CLT10 national norm sample than in the PSAT linking samples.

The studies presented in this report are first steps in creating CLT10 national norms. Due to the availability of the sample, caution should be exercised in interpreting the norms from this report. As the CLT10 population evolves in the future, it is recommended that norm updating be conducted periodically until the sample representation of the CLT10 target population becomes stable.

<i>2019 CLT10 Scores</i>	<i>CLT10 Norming Percentiles</i>	<i>PSAT Nationally Representative Percentiles</i>	<i>2019 CLT10 Scores</i>	<i>CLT10 Norming Percentiles</i>	<i>PSAT Nationally Representative Percentiles</i>
45	4	28	83	67	86
46	4	33	84	70	87
47	6	33	85	72	88
48	6	33	86	74	89
49	7	36	87	76	89
50	8	36	88	77	90
51	8	36	89	79	91
52	9	38	90	81	91
53	11	38	91	82	91
54	12	40	92	84	91
55	13	40	93	86	92
56	15	42	94	87	93
57	16	42	95	88	93
58	19	42	96	90	94
59	19	44	97	92	95
60	20	46	98	93	95
61	21	48	99	94	96
62	23	50	100	94	96
63	25	52	101	94	97
64	27	54	102	95	97
65	28	56	103	96	97
66	31	56	104	97	98
67	34	60	105	97	98
68	36	62	106	97	98
69	38	63	107	98	98
70	40	65	108	98	99
71	42	67	109	98	99
72	44	69	110	99	99
73	46	71	111	99	99
74	47	73	112	99	99+
75	50	75	113	99	99+
76	51	76	114	99	99+
77	54	78	115	99	99+
78	56	79	116	99	99+
79	58	82	117	99	99+
80	61	82	118	99	99+
81	62	83	119	99	99+
82	65	84	120	99	99+

2019 CLT¹⁰ NATIONAL PERCENTILES



ONE

INTRODUCTION

1.1 INTRODUCTION TO THE CLT¹⁰

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

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

1.2 INTRODUCTION TO NORMING

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

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

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

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

1.3 OVERVIEW OF CLT¹⁰ NORMING SAMPLES

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

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

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

1.4 PROCEDURES

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

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

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

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

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

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

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



TWO NORMING

2.1 OVERVIEW OF CLT10 NATIONAL NORMS

THE CLT10 NATIONAL NORMS WERE developed to provide a norm referenced interpretation for CLT10 total scores. A normative sample, resembling the national demographics of the target student population, forms the peer group with whom a student's CLT10 score can be compared. A percentile rank represents the percentage of scores that are at or below it, indicating that the student performs at a higher level than that percentage of students in the population. The percentile norm also allows education legislators to identify low performing students and provide them with additional learning resources and remedial instructional help in preparation for college.

For a test that does not have widespread coverage of the target population, conducting a national norming study requires deliberately recruiting a representative sample of the target population. For this norming study, the data used was taken from the spring 2019 administration of the CLT10, which may not be fully representative of the CLT10 student population. In addition, as the CLT10 evolves in the upcoming years, it is likely that the sample used for this current study will be superseded. This initial exploration of norms based on the available sample provides preliminary information about the normative standing of CLT10 scores relative to the targeted population.

This study employed the following method to obtain a representative sample. First, a set of exclusion rules were applied in cleaning the data. Sample stratification was conducted to match the national 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 CLT10 national percentile norms.

2.2 NATIONAL TARGET POPULATIONS

The CLT10 national target population demographics were derived from the 2016 national survey by the National Center for Education Statistics (NCES) for private schools and home schooling households. Specifically, Table 206.10 (retrieved from https://nces.ed.gov/programs/digest/d17/tables/dt17_206.10.asp) and Table 206.30 (retrieved from https://nces.ed.gov/programs/digest/d18/tables/dt18_206.30.asp), published in the Digest of Education Statistics, were used. Table 206.10 reports 3.76% of grade 9-12 students were home schooled in 2016. Table 206.30 indicates grades 9-12 students in public and private school systems, including 3% in charter schools and 9.03% in private schools. Thus, the entire grade 9-12 NDE student population includes 3.76% home schooled, 2.89% (or $.03 \times (1 - 0.0376)$) in charter, and 8.69% (or $.0903 \times (1 - 0.0376)$) in private schools. That is, 15.34% (i.e., $3.76 + 2.89 + 8.69$) of grade 9-12 students in the US in 2016 make up the CLT10's target population. Within the CLT10's target population, 24.51% (or $3.76 / 15.34$) are home schooled, 18.84% (or $2.89 / 15.34$) attend charter schools, and 56.65% (or $8.69 / 15.34$) attend private schools.

Figure 1. U.S. students, 9-12th grade by school type

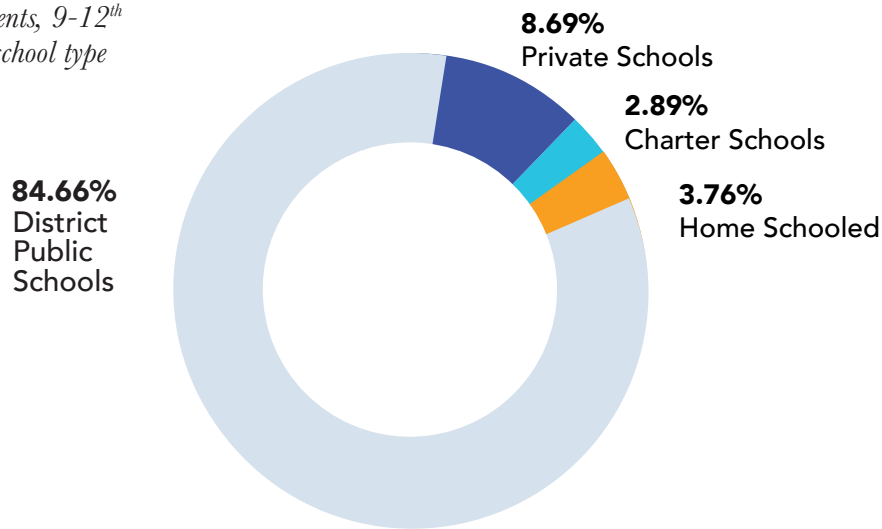


Figure 2. Non-district U.S. students, 9-12th grade by school type

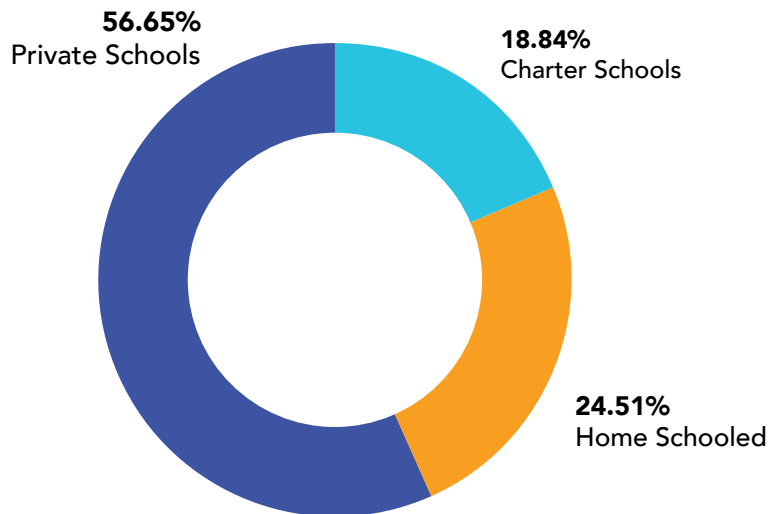


Table 2.1. Demographics of Population Targets by School Type for the CLT10

	SCHOOL TYPE			WEIGHTED CLT10 POPULATION TARGET %
	CHARTER ² (18.84%)	PRIVATE ² (56.65%)	HOMESCHOOL ¹ (24.51%)	
<i>Gender</i>				
Male	51.4	50.3	47.8	49.9
Female	48.6	49.7	52.5	50.1
<i>Ethnicity</i>				
White	30.4	62.1	59.2	55.4
Black	26.2	12.1	7.8	13.7
Hispanic	36.0	15.0	26.3	21.7
Asian/Pacific islander	4.7	6.3	2.6	5.1
Other	2.7	4.4	4.1	4.0
<i>Region</i>				
Northeast	16.8	26.1		23.8
South	16.7	23.3		21.7
Midwest	20.6	26.5		25.0
West	45.9	24.1		29.6
<i>Locale</i>				
City	57.8	37.2	29.2	39.1
Suburb	34.8	48.1	38.5	43.2
Town	3.4	3.0	10.5	4.9
Rural	4.0	11.8	21.8	12.8

The available demographics for each school type were obtained from each NCES Figure. These demographic percentages were weighted, based on the above calculated proportion of school types, to find the population demographic targets. Regional information was not available for homeschool students; hence, regional percentages were weighted based on charter and private school proportions only.

Table 2.1 presents the student demographic percentages available from NCES Tables 206.10 and 206.30 and the weighted percentages used as CLT10's target population demographics. (For example, the percentage for Male is obtained as $0.1884*51.4 + 0.5665*50.3 + 0.2451*47.8 = 49.9$.)

2.3 CLT10 INITIAL SAMPLE

The norming sample was drawn from the CLT10 spring 2019 administration. A total of 3,029 students had valid CLT10 scores. Several exclusion rules were first applied in data cleaning, including:

1. Removing all paper-and-pencil test scores.
2. Removing students from grades other than 9 and 10.
3. Removing students with missing demographic data in gender, ethnicity, or geographic area.

After data cleaning, the remaining 2,269 students were included in the initial sample. Table 2.2 presents the demographics for the initial sample. The initial sample is unbalanced in several categories, especially in ethnicity and geographic region.

Table 2.2. Student Demographics by School Type: Initial Sample vs. Target Population

	INITIAL SAMPLE %	POPULATION TARGET %	DIFFERENCE IN %
<i>School Type</i>			
Charter	4.8	18.8	-14.0
Private	56.5	56.7	-.2
Homeschool	38.7	24.5	14.2
<i>Gender</i>			
Male	45.7	49.9	-4.2
Female	54.3	50.1	4.2
<i>Ethnicity</i>			
White	79.3	55.4	23.9
Black	3.3	13.7	-10.4
Hispanic	6.6	21.7	-15.1
Asian/Pacific islander	5.2	5.1	.1
Other	5.7	4.0	1.7
<i>Region</i>			
Northeast	8.5	23.8	-15.3
South	51.4	21.7	29.7
Midwest	18.7	25.0	-6.3
West	21.4	29.6	-8.2
<i>Locale</i>			
City	19.2	39.1	-19.9
Suburb	27.1	43.2	-16.1
Town	0.9	4.9	-4.0
Rural	52.9	12.8	40.1

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

2.3.1 SAMPLE STRATIFICATION

In order to prepare a representative sample while maintaining a sufficient sample size for norming, sample stratification involved duplicating scores of students from under-represented categories and eliminating scores from over-represented categories.

Both duplication and elimination have inherent limitations. Duplication process implies the assumption that students from the same demographics would perform the same. Given the large score range in CLT10 (i.e., 0-120), this assumption would probably not hold at the individual score point being duplicated. Elimination, on the other hand, risks removing information from the score distribution.

To limit the impact of the stratification process on the test score distribution and avoid introducing systemic bias, controls were used in both steps. Duplication was strictly used 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. The stratification process triplicated 56 students, duplicated 224 students, and eliminated 1445 students, resulting in a total of 1160 students in the normative sample. Table 2.3 shows descriptive statistics for CLT10 scores from both the initial sample and the stratified sample. The score range and shape of the score distribution in the stratified sample remain similar to the original initial sample.

Table 2.3. CLT10 Score Statistics: Initial vs. Stratified Sample

SAMPLE	N	MEAN	STD	MIN	MAX
Initial	2269	75.7	16.5	22	117
Stratified	1160	74.7	16.9	22	117

2.3.2 CLT10 NORMING SAMPLE

The final national norm sample includes 1160 students. Table 2.4 shows the comparison in school type and other key demographics between the national norm sample and the population target. The national norm sample matches the target perfectly in school type distribution and resembles it closely in gender and ethnicity. The largest discrepancies are in geographic categories, with Southern and rural student proportions diverging by 10.5% and 12.9%, respectively. However, these were markedly improved from the initial sample, where the South is over-represented by 29.7% and rural by 40.1%. These national norm sample demographics and sample size are adequate for percentile norm development.

Table 2.4. CLT10 Student Demographics: Weighted Sample vs. Target Population

	NATIONAL NORM SAMPLE %	POPULATION TARGET %	DIFFERENCE %
<i>School Type</i>			
Charter	18.8	18.8	0
Private	56.7	56.7	0
Homeschool	24.5	24.5	0
<i>Gender</i>			
Male	45.8	49.9	-4.1
Female	54.2	50.1	4.1

<i>Ethnicity</i>			
White	56.5	55.4	1.1
Black	12.8	13.7	-.9
Hispanic	21.7	21.7	0
Asian/Pacific islander	5.1	5.1	0
Other	4.0	4.0	0
<i>Region</i>			
Northeast	24.5	23.8	.7
South	32.2	21.7	10.5
Midwest	19.5	25.0	-5.5
West	23.8	29.6	-5.8
<i>Locale</i>			
City	31.4	39.1	-7.7
Suburb	41.1	43.2	-2.1
Town	1.8	4.9	-3.1
Rural	25.7	12.8	12.9

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

Table 2.5 presents the percentile rank correspondence to each obtainable CLT10 score. For ease of use, scores associated with the same percentile rank are grouped together in the Table.

Table 2.5. CLT10 National Percentile Norms (Target Population)

CLT10 Total Score	National Percentile Rank	CLT10 Total Score	National Percentile Rank
0-38	1	76	51
39-41	2	77	54
42-44	3	78	56
45-46	4	79	58
47-48	6	80	61
49	7	81	62
50-51	8	82	65
52	9	83	67
53	11	84	70
54	12	85	72
55	13	86	74
56	15	87	76
57	16	88	77
58-59	19	89	79
60	20	90	81
61	21	91	82
62	23	92	84
63	25	93	86
64	27	94	87
65	28	95	88
66	31	96	90
67	34	97	92
68	36	98	93
69	38	99-101	94
70	40	102	95
71	42	103	96
72	44	104-106	97
73	46	107-109	98
74	47	110-120	99
75	50		

2.4 CLT10 NATIONAL PERCENTILE NORMS (TARGET POPULATION)

Percentile norms are established for the CLT10. First, the 1st-99th percentile ranks for each total score in the normative sample were calculated. Then, the correspondence between each unique total score and percentile was obtained. If a score was associated with more than one percentile rank, the higher percentile value was used. For example, both the 10th percentile and 11th percentile were associated with total score 53, thus, 11 was chosen as the percentile rank for 53. Scores not directly associated with a percentile rank were placed in the adjacent rank down: for example, a total score of 50 was associated with the 8th percentile and a score of 52 was associated with the 9th percentile, so a score of 51 was placed in the score category of 50, i.e., the 8th percentile.

Percentile norms map each performance level directly to the distribution of CLT10 scores for the normative sample. This norm score has a straightforward interpretation. A norm score indicates the percentage of students in the sample with a score at or above it. For example, a CLT10 score of 79 is normed at the 58th percentile, i.e. a student who gets a score of 79 performed as well as or better than 58% of students in the normative sample. This interpretation of performance with reference to peers makes it easy to identify students in need of academic assistance, and provides legislators empirical data for allocating education resources to them.

2.5 LIMITATIONS AND FUTURE CONSIDERATIONS

A limitation in the current norming study lies in the sample size and imbalance of the CLT10 user base as of spring 2019. The initial sample, as presented in Table 2.2, was unbalanced in several demographic categories. Most of the imbalance was smoothed out through stratification and weighting, but at the cost of rather substantial sample size reduction. Although random selection was used during sample stratification and the stratified normative sample is of sufficient size with close resemblance to the national targets for school type, gender, and ethnicity, and displays satisfactory resemblance to geographic targets, a more representative initial sample would be desirable.

In addition, it is expected that the CLT10's user base is in the growing phase. As a result, a shift in performance could happen. Such a shift is often upward, as more information and resources become available to students taking the exam. A shift downward could also take place, as a broader demographic range of students gain access to the test. During the period of rapid user growth for a test like CLT10, norm updating studies should be conducted, at least once per year, to track any significant performance shifts among examinees and to ensure that the national norms are referencing the most current CLT10 population performance. Norm updates could become less frequent once the CLT10 user base reaches equilibrium.

2.6 CONCLUSIONS

This study is the first step in creating CLT10 national norms, based on the sample available from the spring 2019 test administration. A CLT10 percentile norm score links the student's total score to the percentile rank in the normative sample. Allowing for the limitations discussed above, this provides serviceable norms for present use.



THREE

CLT10-PSAT LINKING STUDY

3.1 LINKING STUDY OVERVIEW

IN CHAPTER 2, THE TARGET demographics for the CLT10 national population were derived from the 2016 national survey by the NCES. Then, the students who took the spring 2019 CLT10 test were used to establish a representative sample of the national target population for the CLT10. Last, this national normative sample was used for developing norms for the CLT10. The results from Chapter 2 indicate the relative ranking of a given CLT10 score compared with the national normative sample.

This chapter carries out an empirical linking study which provides normative information on CLT10 test scores relative to the 2018 PSAT normative sample. This study establishes a concordance relationship through the equipercentile linking method. The concordance table shows how each CLT10 score is mapped onto the PSAT scale. Then, the normative information constructed for the spring 2018 PSAT can be related to a given CLT10 score with a mapped PSAT score, indicating the relative ranking of a CLT10 score compared with the normative sample for the 2018 PSAT norms.

The following sections start with the content alignment results between the CLT10 and the PSAT; data preparation follows. The technical details and results from the empirical norming study are presented next. Lastly, a summary of the findings from this study is presented.

3.2 CONTENT ALIGNMENT STUDY

The validity of the interpretation of this study's results is closely related to the content alignment between the CLT 10 and PSAT. A high correlation between tests is only one indication of similarity, to be confirmed by an examination of the exams themselves. This chapter summarizes the results from this content alignment study by the content experts for the CLT 10.

The CLT10 (first launched in the spring of 2017) is comprised of three sections: Verbal Reasoning, Grammar/Writing, and Quantitative Reasoning. The PSAT/NMSQT consisted of two sections: Evidence-Based Reading and Writing (EBRW) and Math. The EBRW portion consists of two subsections, one on Reading and one on Writing and Language, while Math consists of Math without Calculator and Math with Calculator.

Each CLT10 section contains 40 questions, and students are given 40, 35, and 45 minutes for Verbal Reasoning, Grammar/Writing, and Quantitative Reasoning respectively. The PSAT contains 91 questions and 95 minutes for Evidence-Based Reading and Writing, and 48 questions and 65 minutes for Math. Students are allotted one minute on average for each question on the CLT10, and students are allotted 1.15 minutes on average for each PSAT question. (See Figures A.1 and A.2 in Appendix A for details.)

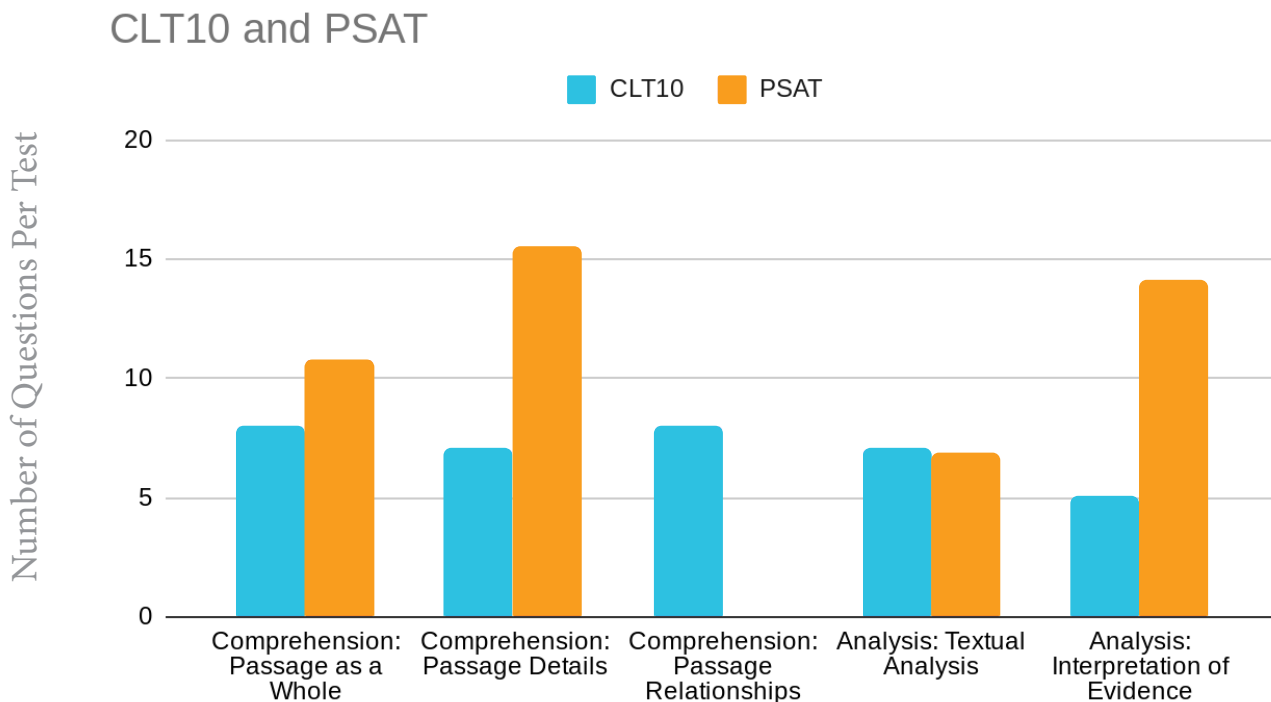


Figure 3: Subdomains Relationships in Verbal Reasoning

3.2.1 VERBAL REASONING (CLT10) AND READING (PSAT)

On the CLT10 Verbal Reasoning section, 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.” Tests are calibrated so that each Verbal Reasoning passage fits narrowly within a word count range of 475-625 words. The total must be between 2,175 and 2,225, for an average of 2,200 words total. The CLT10 Verbal Reasoning section consists of four passages—including one paired-set—and is broken into 40 questions to be completed in 40 minutes.

The CLT10 Verbal Reasoning section most closely corresponds to the PSAT Reading subsection from EBRW. The passage lengths on the PSAT fall in a range of 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.

The CLT10 contains 21% more Comprehension questions, while the PSAT contains 36% more Analysis questions than the CLT10. Within Comprehension, the PSAT contains more questions about both the

subdomains “Passage as a Whole” and “Passage Details” than the CLT10. In Analysis, the PSAT includes 38.9% more questions about “Textual Analysis” and 16.9% more questions about the “Interpretation of Evidence” than the CLT10. These subdomain proportions are summarized in Figure 3 below.

The CLT10 also includes 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 CLT10 includes analogies to test logical reasoning and synthesis, and the analogies are always attached to the reading passages.

3.2.2 GRAMMAR/Writing (CLT10) AND Writing AND LANGUAGE (PSAT)

On the Grammar/Writing section, 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.” Tests are calibrated so that 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.

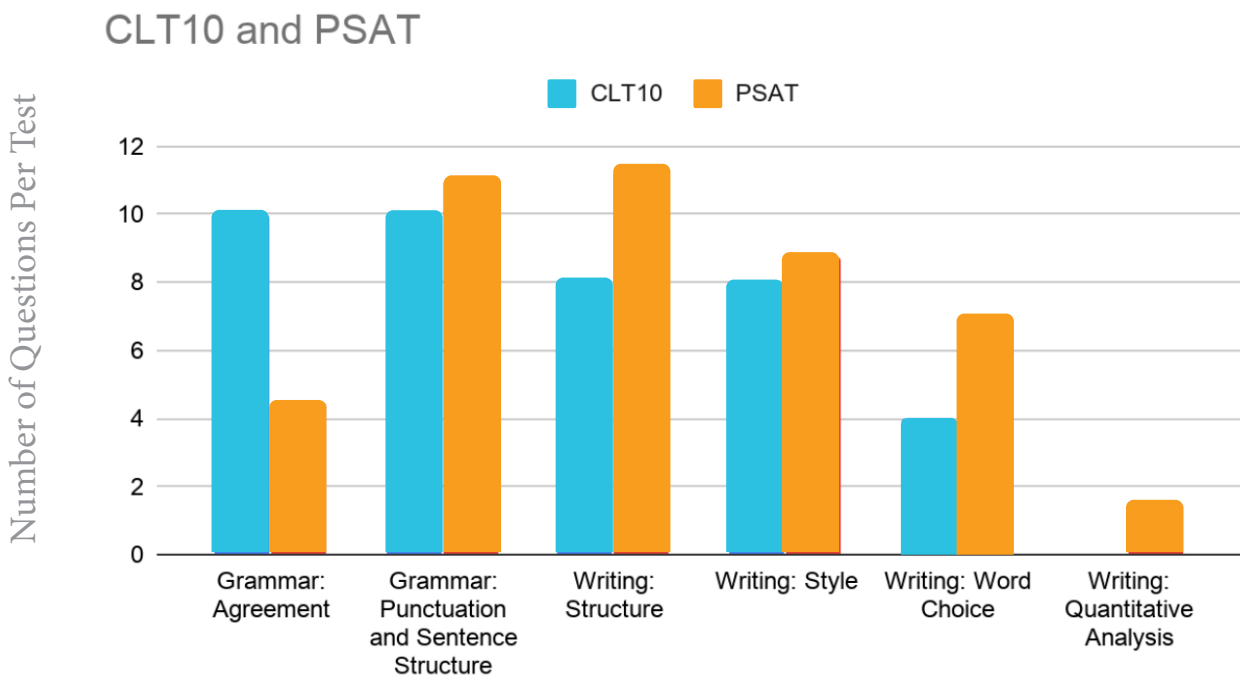


Figure 4: Domain Relationships in Grammar and Writing

The CLT10 Grammar/Writing section most closely corresponds to the PSAT Writing and Language subsection from EBRW. The passage lengths on the PSAT fall into a range of 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.

The CLT10 contains 42% more Grammar questions than the PSAT, and the PSAT contains 29.6% more Writing questions. Within Grammar, the CLT10 contains 145.1% more questions under the subdomain “Agreement,” and both tests have an equal percentage of questions about “Punctuation and Sentence Structure.” Within the domain of “Writing,” the PSAT contains 28% more questions about “Structure,”

the same percentage of questions about “Style,” and the PSAT Contains 59% more questions about “Word Choice.” These subdomain proportions are summarized in Figure 2 below.

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. The CLT10 does not include questions about data interpretation, as the Science passages in the Grammar/Writing section do not include visual representations of data.

3.2.3 TEXT EVALUATION

The difficulty of the passages in both verbal tests is determined in part by the Educational Testing Service’s Text Evaluation (TE) grade level score. The relevant TE data, showing a close similarity between the two exams, are as follows:

CLT10

TE Grade Level: 10.5

TE Range: 9-12

PSAT

TE Grade Level: 10.8

TE Range: 8-12

3.2.4 QUANTITATIVE REASONING (CLT) AND MATH (PSAT)

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 CLT10 Quantitative Reasoning section encompasses the content of both tests in the PSAT Math section (Math with Calculator and Math without Calculator), though the CLT10 is taken entirely without a calculator. The PSAT includes 40% more Algebra questions, the CLT10 112.1% more Geometry Questions, and the PSAT 23.75% more Mathematical Reasoning questions. Within Algebra, the CLT10 contains 363% more questions about Arithmetic and Operations, and the PSAT contains 155.25% more questions about Algebraic Expressions and Equations. Under Geometry, the CLT10 contains approximately the same number of questions about Plane Geometry, and 56% more questions about Properties of Shapes. In Mathematical Reasoning, the CLT10 contains 1,150% more Logic problems, and the PSAT contains 38.5% more Word Problems. These subdomain relationships are summarized in Figure 5 below.

On average, the PSAT Math section is comprised of 21.3% questions about Statistics, which are not included on the CLT10. These questions concern descriptive statistics, interpretation of data sets, and interpretation of visual data. The CLT10 does include questions about probability, 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 not included on the CLT10. These questions have been classified as “Word Problems”; the CLT10, however, tests the same skills of function notation in “Algebraic Expressions and Equations.”

The alignment study results for domains as a whole are summarized and presented graphically in Figure 4 below. In general, content domains and coverage do not dramatically differ between the CLT 10 and PSAT. Both exams are broadly covering aptitude and achievement in the fundamental subjects of reading, writing and mathematics. The divergence of the two tests at the subdomain level indicates stronger emphasis on aptitude in the case of the CLT10 (logic and analogy questions for example) and a stronger emphasis on achievement on the PSAT (quantitative analysis and statistics, for example).

CLT10 and PSAT

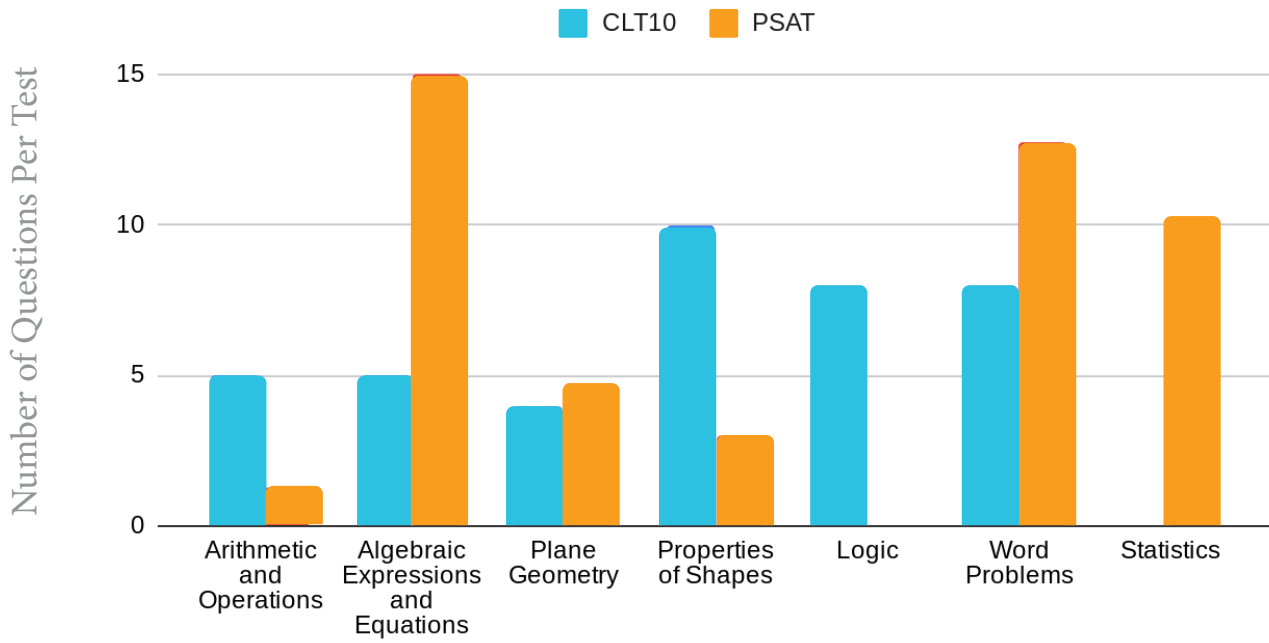


Figure 5. Subdomain Relationships in Quantitative Reasoning

CLT10 and PSAT

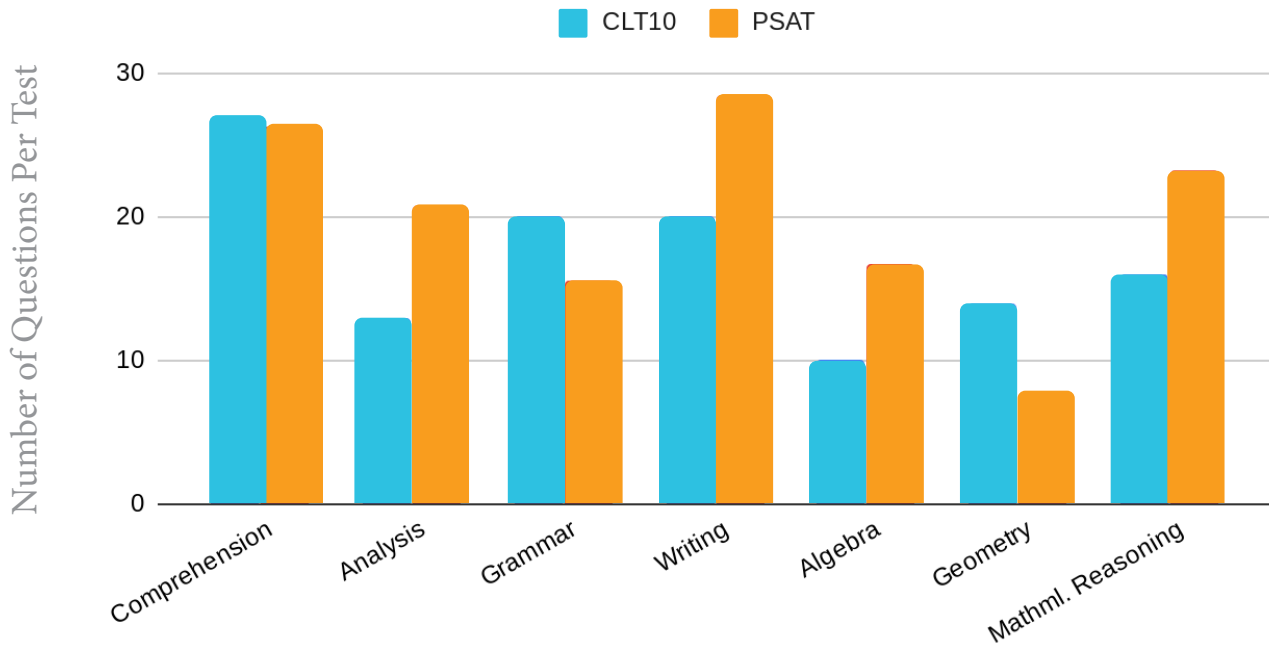


Figure 6. Domain Relationships in General

3.3 DATA PREPARATION

The data used for the statistical linkage differs slightly from that used in Chapter 2. Chapter 2 uses students with valid CLT10 scores after exclusion rules have been applied; similar exclusion rules were applied in cleaning data for this study. In addition, the data used here only contains students who took both the spring 2019 CLT10 and the PSAT (in any test administration before spring 2019). The exclusion rules applied in data cleaning for this concordance study are summarized below.

1. Remove all paper-and-pencil test scores.
2. Remove students from grades other than 9 and 10.
3. Remove students from public schools.
4. Remove students who do not have the PSAT scores reported.

After applying these exclusion rules, data for 228 students was retained. However, among these 228 student records, 8 had invalid PSAT scores like 822, 999, 1125 (valid PSAT scores are in multiples of ten). After removal of these records, the final sample size is 220 for this concordance study. Table 3.1 summarizes the demographics for this sample as compared with the target population and with the constructed national sample for norming in Chapter 2. The concordance sample is unbalanced in several variables, particularly school type, ethnicity, and geographic region.

Table 3.1. Student Demographics: PSAT Norming Sample vs. CLT10 Norming Sample

	SAMPLE FOR THE CONCORDANCE STUDY	NATIONAL NORM SAMPLE %	POPULATION TARGET %	DIFFERENCE % IN NORMING SAMPLE	DIFFERENCE % IN CONCORDANCE SAMPLE
<i>School Type</i>					
Charter	3.2	18.8	18.8	0	-15.6
Private	70.0	56.7	56.7	0	13.3
Homeschool	26.4	24.5	24.5	0	1.9
Others	0.5				0.5
<i>Gender</i>					
Male	51.4	45.8	49.9	-4.1	1.5
Female	48.6	54.2	50.1	4.1	-1.5
<i>Ethnicity</i>					
White	83.6	56.5	55.4	1.1	28.2
Black	2.7	12.8	13.7	-0.9	-11.0
Hispanic	3.2	21.7	21.7	0	-18.5
Asian/Pacific islander	4.1	5.1	5.1	0	-1
Other	6.4	4.0	4.0	0	2.4
<i>Region</i>					
Northeast	11.4	24.5	23.8	0.7	-12.4
South	50.5	32.2	21.7	10.5	28.8
Midwest	12.7	19.5	25.0	-5.5	-12.3
West	25.5	23.8	29.6	-5.8	-4.1

Table 3.1. (cont'd) Student Demographics: PSAT Norming Sample vs. CLT10 Norming Sample

<i>Locale</i>					
City	19.5	31.4	39.1	-7.7	-19.6
Suburb	32.3	41.1	43.2	-2.1	-10.9
Town	0.5	1.8	4.9	-3.1	-4.4
Rural	47.7	25.7	12.8	12.9	34.5

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

Table 3.2 presents descriptive statistics for CLT10 scores from the initial sample and the stratified sample used for norming reported in Chapter 2, as well as the sample used for the concordance study in this chapter. The score distribution in the stratified sample closely resembles the original initial sample for the norming study in chapter 2; however, the score distribution for the concordance study differs noticeably from the distribution obtained from the norming sample. In general, those students who reported the PSAT scores performed better than those students in the initial and the stratified norm samples from the spring 2019 CLT 10 administration. The linking sample has a mean CLT10 score of 85.3 and standard deviation of 15.1. Their reported PSAT scores ranged from 720 to 1520, with a mean of 1144.2 and standard deviation of 138.7. The sample used for developing the concordance relationship between the CLT10 and the PSAT is a high-performing group of students.

Table 3.2. CLT10 and PSAT Score Statistics Across Samples

SAMPLE	N	MEAN	STD	MIN	MAX
Initial Sample	2269	75.7	16.5	22	117
Stratified Norming Sample	1160	74.7	16.9	22	117
Concordance Sample	220	85.3	15.1	38	117
PSAT	220	1144.2	138.7	720	1520

3.4 CONCORDANCE BETWEEN CLT10 AND PSAT SCORES

To establish score correspondences between the CLT10 and the PSAT, equipercentile linking was conducted using the matched sample whose CLT10 scores and PSAT scores were available. This linking was carried out using the software program, Linking with Equivalent Group or the Single Group Design, abbreviated as LEGS (Kolen & Brennan, 2004). Raw CLT10 scores and corresponding PSAT scores were used to link the CLT10 and the PSAT. With the proper specification of the format of the input data, subgroup information, input data file names, smoothing values, the score range for the CLT10, and the truncation choice, this program conducts the equipercentile linking and outputs the results in the window. (Appendix B provides a screenshot of the input window for linking the CLT10 and PSAT.)

Two smoothing values were compared in post-linking, 0.3 and 1. The choice of using smoothing values is supported by results from simulation studies: 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 include mean, linear, parallel-linear, and equipercentile methods (with and without post-smoothing). The results with a smoothing value of 1 has the smallest RMSEL, as shown in Table 3.3. Table 3.4 presents the mapped CLT10 and PSAT scores from the LEGS output, based on equipercentile linking with a smoothing value of 1.

Table 3.3: RMSEL for Score Linking Methods

MEAN - 126.89 **PARALLEL LINKING** - 90.72 **SMOOTHING=0.30** - 90.44
LINEAR - 90.72 **EQUIPERCENTILE** - 90.69 **SMOOTHING=1.00** - 90.18

Table 3.4. Mapped CLT10 and PSAT Scores (Equipercentile Method, Smoothing: 1)

CLT 10	PSAT	CLT 10	PSAT	CLT 10	PSAT	CLT 10	PSAT
38	731	58	880	78	1078	98	1253
39	753	59	885	79	1095	99	1263
40	775	60	896	80	1102	100	1271
41	797	61	909	81	1111	101	1280
42	804	62	920	82	1121	102	1292
43	810	63	932	83	1134	103	1306
44	817	64	943	84	1144	104	1318
45	823	65	948	85	1153	105	1329
46	839	66	953	86	1157	106	1342
47	842	67	965	87	1160	107	1349
48	844	68	983	88	1168	108	1355
49	847	69	994	89	1177	109	1366
50	849	70	1000	90	1180	110	1381
51	852	71	1009	91	1184	111	1401
52	856	72	1020	92	1192	112	1432
53	862	73	1030	93	1201	113	1458
54	868	74	1039	94	1205	114	1469
55	872	75	1046	95	1214	115	1476
56	875	76	1055	96	1228	116	1488
57	877	77	1065	97	1241	117	1510

The available score range for the CLT10 is only from 38 to 117 in the sample for the concordance linking study. The LEGS program has conducted interpolation to fill in the blanks among CLT10 scores. Appendix C presents the frequency of CLT10 scores for equipercentile linking. Some CLT10 score points, such as 39 and 40, were missing in the data for the concordance relationship development. As valid CLT10 scores range from 0 to 120, these missing scores and scores outside 38 and 117 had to be extrapolated to establish equivalent PSAT scores (see Table 3.4).

Several models, including power, linear, exponential, and polynomial (at different orders, from 2 to 6), were fitted with the matched scores based on equipercentile linking. A scatterplot was generated to examine the relationship between CLT10 and PSAT scores in the concordance Table shown in Table 3.4. A prediction equation was developed using Excel’s trendline function to find the best fitting model. The equations for the fitted models and the R-squares, which indicate the total variance explained by the fitted models, are summarized in Table 3.5.

Table 3.5. Prediction Equations for Mapping CLT10 Scores to PSAT Scores

TRENDLINE OPTIONS	R-SQUARE	PREDICTION EQUATION
Power	0.9732	$Y=80.319x^{0.6004}$
Linear	0.9869	$Y=9.0292x+383.35$
Exponential	0.9949	$Y=554.65e^{0.0084x}$
Polynomial 2	0.9953	$Y=0.0404x^2+2.7652x+604.53$
Polynomial 3	0.9955	$Y=0.0003x^3-0.0364x^2+8.4048x+475.23$
Polynomial 4	0.9964	$Y=0.00003x^4-0.0098x^3+1.0958x^2-45.471x+1391.8$
Polynomial 5	0.9981	$Y=0.000002x^5-0.0008x^4+0.1168x^3-8.1138x^2+277.99x-2985.8$
Polynomial 6	0.9993	$Y=-0.0000001x^6+0.00005x^5-0.0092x^4+0.9402x^3$ $52.346x^2+1508.2x-16812$

Note: x represents the CLT 10 scores while Y represents the PSAT scores.

Based on the total variance explained (R-square) and the reasonably extrapolated values for the PSAT based on actual CLT10 scores, the exponential function was identified as the best fitting model with the best reasonableness (extrapolated PSAT scores corresponding to each CLT10 score for each of the models listed in Table 3.5 are available upon request.). Though all the polynomial functions displayed better fit in terms of R-squares when compared with the exponential model, some extrapolated PSAT values were out of boundaries, or deviated too much from the original PSAT scores used to develop the best-fitting model. Using the exponential prediction equation, PSAT scores for the values at the upper end of the CLT10 scale in the matched sample were computed.

On the other hand, for the lower end of the CLT10 scale, a linear extrapolation equation was developed by fixing the lowest possible scores for the CLT10 and the PSAT, and the lowest scores in the concordance Table (Table 3.4) were developed based on the matched sample. The linear equation developed by fixing two data points (0, 320) and (38, 731) is as follows, with x representing CLT10 scores while Y represents PSAT scores.

Based on the equipercentile linking using the matched sample, the identified best-fitting exponential model for extrapolation for the upper end of the CLT10 scale, and the linear extrapolation equation for the lower end of the CLT10 scale, the complete concordance Table is presented in Table 3.6 (with rounding of the mapped PSAT scores to multiples of ten, to best approximate real possible PSAT scores).

Once the concordance relationship between CLT10 and PSAT scores is established, the normative information for the PSAT can be used as a reference to compare CLT10 scores with the norm groups. (Normative information for the PSAT from 2018 is presented in Appendix D.) This Table provides the percentile ranks of each PSAT score in reference to two norm groups: one is the nationally representative sample of the 10th grades taking the PSAT, and the other is the PSAT/NMSQT and PSAT 10 users. By mapping CLT10 and PSAT scores and the norms developed for the PSAT, a normative comparison of CLT10 scores to the nationally representative sample of the 10th graders taking the PSAT and the PSAT/NMSQT and PSAT 10 users can be established empirically. Table 3.7 presents the mapped CLT10 and PSAT scores and the percentile ranks for PSAT scores based on different norm groups.

Table 3.6. Score Concordance Table (Based on Spring 2019 Testing Data)

CLT 10	PSAT	CLT 10	PSAT	CLT 10	PSAT	CLT 10	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						

Table 3.7. CLT10-PSAT Percentile Concordance Table

CLT 10	PSAT	NATIONALLY REPRESENTATIVE SAMPLE	PSAT/NMSQT AND PSAT 10 USER	CLT 10	PSAT	NATIONALLY REPRESENTATIVE SAMPLE	PSAT/NMSQT AND PSAT 10 USER
0	320	<1	<1	61	910	48	49
1	330	<1	<1	62	920	50	51
2	340	<1	<1	63	930	52	53
3	350	<1	<1	64	940	54	54
4	360	<1	<1	65	950	56	56
5	370	<1	<1	66	950	56	56
6	390	<1	<1	67	970	60	60
7	400	<1	<1	68	980	62	62
8	410	<1	<1	69	990	63	64
9	420	<1	<1	70	1000	65	66
10	430	<1	<1	71	1010	67	67
11	440	<1	<1	72	1020	69	69
12	450	<1	<1	73	1030	71	71
13	460	<1	<1	74	1040	73	72
14	470	<1	<1	75	1050	75	74
15	480	<1	<1	76	1060	76	76
16	490	<1	<1	77	1070	78	77
17	510	<1	<1	78	1080	79	78
18	520	<1	<1	79	1100	82	81
19	530	<1	<1	80	1100	82	81
20	540	<1	<1	81	1110	83	83
21	550	<1	1	82	1120	84	84
22	560	<1	1	83	1130	86	85
23	570	<1	1	84	1140	87	86
24	580	<1	1	85	1150	88	87
25	590	<1	1	86	1160	89	88
26	600	<1	1	87	1160	89	88
27	610	<1	1	88	1170	90	89
28	620	1	2	89	1180	91	90
29	630	1	2	90	1180	91	90
30	640	1	2	91	1180	91	90
31	660	2	4	92	1190	91	91
32	670	2	4	93	1200	92	92

33	680	3	5	94	1210	93	93
34	690	4	6	95	1210	93	93
35	700	5	7	96	1230	94	94
36	710	6	9	97	1240	95	94
37	720	8	10	98	1250	95	95
38	730	9	12	99	1260	96	95
39	750	13	16	100	1270	96	96
40	780	19	22	101	1280	97	96
41	800	24	26	102	1290	97	97
42	800	24	26	103	1310	97	97
43	810	26	28	104	1320	98	98
44	820	28	30	105	1330	98	98
45	820	28	30	106	1340	98	98
46	840	33	34	107	1350	98	98
47	840	33	34	108	1360	99	98
48	840	33	34	109	1370	99	99
49	850	36	36	110	1380	99	99
50	850	36	36	111	1400	99	99
51	850	36	36	112	1430	99+	99
52	860	38	38	113	1460	99+	99+
53	860	38	38	114	1470	99+	99+
54	870	40	41	115	1480	99+	99+
55	870	40	41	116	1490	99+	99+
56	880	42	43	117	1510	99+	99+
57	880	42	43	118	1500	99+	99+
58	880	42	43	119	1510	99+	99+
59	890	44	45	120	1520	99+	99+
60	900	46	47				

The interpretation of the percentile ranks in Table 3.7 for each CLT10 score relative to the PSAT norm groups is the same as that explained in chapter 2. Based on the norming study presented there, a CLT10 score of 80, for example, fell in the 61st percentile of the CLT10 target population. Based on the method used in this chapter, a CLT10 score of 80 was in the 82nd percentile nationwide, indicating a student who gets a CLT10 score of 80 performed the same or better than 82% of the PSAT national representative sample. This indicates that the students used for this empirical norming study were higher-performing students compared with the general 10th grade PSAT students and users of the PSAT/NMSQT and PSAT 10. This is supported by the mean PSAT score for the students who were in the sample for this study (1144.2 in Table 3.2) and the mean PSAT scores of the norm groups for the PSAT (939 and 934 in Appendix D).

3.5 LIMITATIONS AND FUTURE CONSIDERATIONS

A limitation of this current linking study lies in the representativeness of the CLT10-PSAT linking sample. As noted in Chapter 2, the initial sample, as presented in Table 2.2, was unbalanced in a number of demographics. After retaining the students with both valid CLT10 and PSAT scores, the sample size was reduced to 220, and was unbalanced in certain variables compared with the national target population. It is also speculated that high-performing students were more likely to report their PSAT scores. Thus, the sample used in the linking study contained high performing students, and the percentile ranks for the same CLT10 score could diverge from both the norming study reported in Chapter 2 and the empirical norming study reported in this chapter. The results here should accordingly be interpreted with some caution, and further study is recommended.

As mentioned in Chapter 2, CLT10's user base is growing. It is also recommended that empirical norms related to the PSAT norm groups be updated in the near future, to ensure that the empirical norms are referencing the most up-to-date CLT10 population performance. Though percentile ranks are sample dependent, when students who participate in the empirical norm study become stable and representative of the CLT10 student population, the normative information will be more valid and generalizable.

3.6 CONCLUSIONS

This linking study is another initial exploration comparing CLT10 examinees' performance to the PSAT national representative sample and the norms for PSAT/NMSQT and PSAT 10 users. The sample used is a subset of students from the spring 2019 CLT10 test administration, who reported their PSAT scores. Two major highlights are summarized below:

1. The CLT10 sample used in this empirical norming study was a convenience sample, smaller and less representative of the CLT10 student population than is desirable. This limitation could reduce the generalizability of the norms developed in this chapter. Further, the PSAT scores were self-reported; the integrity of the data should be further checked.
2. As highlighted in Chapter 2, the current user base for the CLT10 is expanding. If stake-holders are interested in using the empirical norms to compare CLT10 students with the PSAT national representative sample and the PSAT/NMSQT and PSAT 10 users, it is recommended that updates be conducted annually at least, to maintain the integrity of the norms obtained in these studies.



FOUR

SUMMARY DISCUSSIONS

4.1 RESULTS SUMMARY

THE CURRENT NORMING STUDY INTENDS to provide normative information relative to a reference population. The norming study reported in Chapter 2 focuses on the national representative sample of the CLT10 target population of non-district educated students, while the information presented in Chapter 3 compares CLT10 students' performance to two PSAT populations.

Chapter 2 presents the norming results based on a representative sample obtained from the spring 2019 CLT10. The national population demographic targets were derived from the 2016 national survey by the NCES for private schools and home schools. The norming study from Chapter 2 presents the relative ranking of a given CLT10 score compared with the national normative sample created based on the spring 2019 test data.

Chapter 3 presents the empirical normative information through the linking of CLT10 scores and PSAT scores, relative to two 2018 PSAT normative samples. The linking study develops a concordance relationship between the CLT10 and the PSAT through the equipercentile linking method. This concordance table shows how each CLT10 score is mapped onto the PSAT scale. Then, the normative information constructed for the spring 2018 PSAT is used to compare a CLT10 score with a mapped PSAT score, indicating the relative ranking of a CLT10 score to the normative sample for the 2018 norm development of the PSAT.

Table 4.1 summarizes the percentile ranks corresponding to each CLT10 score, based on the norming study for the CLT10 national sample and the linking study based on the PSAT national sample and the national sample of PSAT/NMSQT and PSAT 10 users. Table 4.1 illustrates the change pattern of percentile ranks across the full CLT10 score range. Along the majority of the CLT10 scale range, the percentile ranks of each CLT10 score relative to the CLT10 nationally representative sample are lower than those relative to the PSAT nationally representative sample and the PSAT/NMSQT and PSAT 10 users. Such differences are larger for the score range from 40 to 60. (The differences in the two PSAT norming samples are not large, and lie within

expectations.) This is because norming results are sample dependent. The score distributions for the norming sample and the linking sample differ from each other in score range, especially the minimum score, the mean, and the standard deviation (see Figure 7 and 8). The norming and the linking studies used samples that differ in size, school type, gender, ethnicity, geographic region, and locale as well. The norming sample used in Chapter 2 more closely resembles the target CLT10 population, while the sample used in Chapter 3 in the concordance study diverged further from the target CLT10 student population in certain key demographic variables. Further, the sample in the concordance study is a 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 CLT10 score would be lower for the CLT10 national sample, compared with percentile rankings for students from the CLT10 sample who reported their PSAT scores.

Figure 7. Percentile Ranks for CLT10 Scores Relative to PSAT Samples

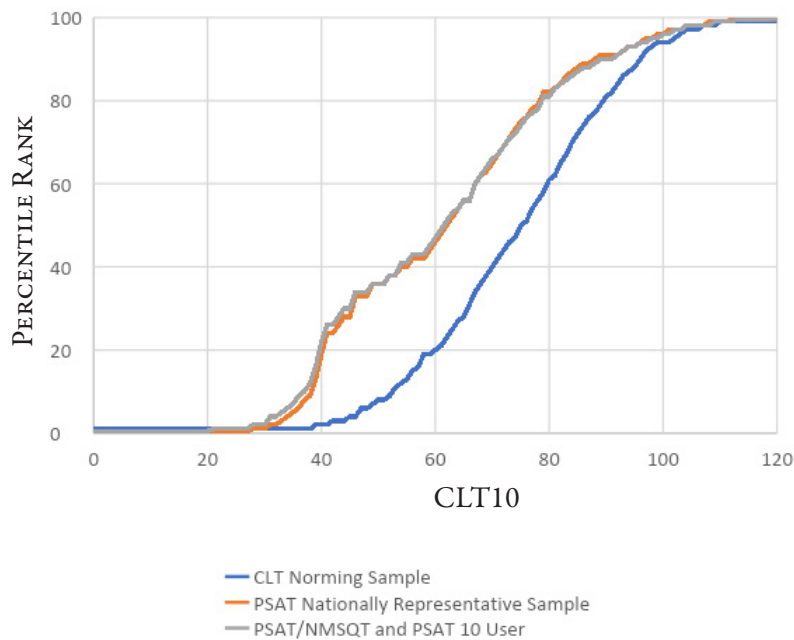


Figure 8. Histogram of CLT10 Scores from the Linking Sample

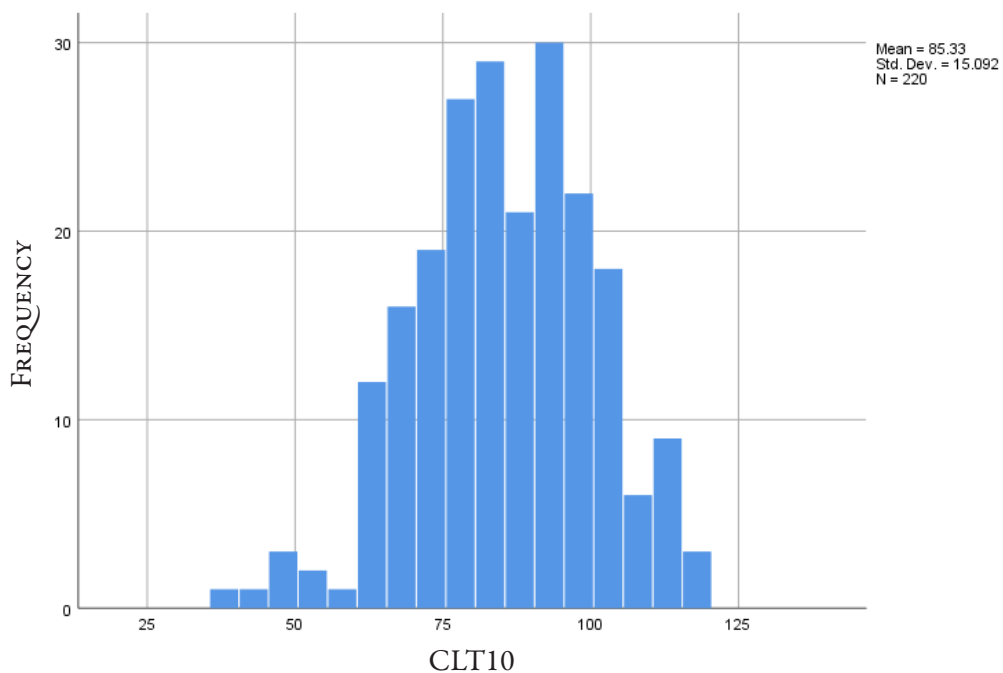


Table 4.1. Summary of Norming Results (Percentiles)

CLT Scores	CLT Norming Sample	PSAT Nationally Representative Sample	PSAT/NMSQT and PSAT 10 User	CLT Scores	CLT Norming Sample	PSAT Nationally Representative Sample	PSAT/NMSQT and PSAT 10 User
0	1	<1	<1	34	1	4	6
1	1	<1	<1	35	1	5	7
2	1	<1	<1	36	1	6	9
3	1	<1	<1	37	1	8	10
4	1	<1	<1	38	1	9	12
5	1	<1	<1	39	2	13	16
6	1	<1	<1	40	2	19	22
7	1	<1	<1	41	2	24	26
8	1	<1	<1	42	3	24	26
9	1	<1	<1	43	3	26	28
10	1	<1	<1	44	3	28	30
11	1	<1	<1	45	4	28	30
12	1	<1	<1	46	4	33	34
13	1	<1	<1	47	6	33	34
14	1	<1	<1	48	6	33	34
15	1	<1	<1	49	7	36	36
16	1	<1	<1	50	8	36	36
17	1	<1	<1	51	8	36	36
18	1	<1	<1	52	9	38	38
19	1	<1	<1	53	11	38	38
20	1	<1	<1	54	12	40	41
21	1	<1	1	55	13	40	41
22	1	<1	1	56	15	42	43
23	1	<1	1	57	16	42	43
24	1	<1	1	58	19	42	43
25	1	<1	1	59	19	44	45
26	1	<1	1	60	20	46	47
27	1	<1	1	61	21	48	49
28	1	1	2	62	23	50	51
29	1	1	2	63	25	52	53
30	1	1	2	64	27	54	54
31	1	2	4	65	28	56	56
32	1	2	4	66	31	56	56
33	1	3	5	67	34	60	60

Table 4.1. (cont'd) Summary of Norming Results

68	36	62	62	95	88	93	93
69	38	63	64	96	90	94	94
70	40	65	66	97	92	95	94
71	42	67	67	98	93	95	95
72	44	69	69	99	94	96	95
73	46	71	71	100	94	96	96
74	47	73	72	101	94	97	96
75	50	75	74	102	95	97	97
76	51	76	76	103	96	97	97
77	54	78	77	104	97	98	98
78	56	79	78	105	97	98	98
79	58	82	81	106	97	98	98
80	61	82	81	107	98	98	98
81	62	83	83	108	98	99	98
82	65	84	84	109	98	99	99
83	67	86	85	110	99	99	99
84	70	87	86	111	99	99	99
85	72	88	87	112	99	99+	99
86	74	89	88	113	99	99+	99+
87	76	89	88	114	99	99+	99+
88	77	90	89	115	99	99+	99+
89	79	91	90	116	99	99+	99+
90	81	91	90	117	99	99+	99+
91	82	91	90	118	99	99+	99+
92	84	91	91	119	99	99+	99+
93	86	92	92	120	99	99+	99+
94	87	93	93				

In general, the norming sample constructed in Chapter 2 is a more nationally representative sample of the target CLT10 population. Nonetheless, the percentile ranks produced in mapping the relationship between the CLT10 and PSAT scores do provide information about the CLT10 students' performance relative to the PSAT representative sample and the PSAT/NMSQT and PSAT 10 users. Care should be used in interpreting normative information from either study.

4.2 FINAL REMARKS

The studies presented in this report are an initial exploration in creating national norms for the CLT10. While the norms presented here are serviceable for present use, the reservations expressed in Chapters 2 and 3 about their future applicability and recommendations for future studies should be kept in mind.

Appendix A: Alignment Study

Table A.1 The CLT 10 Test Structure

SECTION	NUMBER OF QUESTIONS	TIME
Verbal Reasoning	40	40 min
Grammar/Writing	40	35 min
Quantitative Reasoning	40	45 min

Table A.2: 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

Table A.3 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 (7 questions)
		Passage Relationships (8 questions)
	Analysis (13 questions)	Textual Analysis (8 questions)
		Interpretation of Evidence (5 questions)

Table A.4 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)

Table A.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)

1 Data available upon request

Table A.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)

Table A.7 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 A.8 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)

Appendix B: Screenshot for running the LEGS program

Equating Program - LEGS 2.0.1

File Window

Raw Scores as Input | Frequency Distributions as Input

X Scores :

Alphanumeric Identifier : CLT10

Column for X Score : 1

Lowest Score : 38

Highest Score : 117

File ... : 9/LEGS_wg_v20/clt-raw.txt

Y Scores :

Alphanumeric Identifier : PSAT

Column for Y Score : 2

Lowest Score : 720

Highest Score : 1520

Subgroups :

Number of Subgroups : 1

Names of Subgroups : CLT

Number of examinees in each subgroup : 220

Equipercntile Smoothing

slim : 0.5

number of Smoothing Values : 2

List of Smoothing Values : 0.3
1

Truncation

Lowest Valid Score(Y) : 20

Highest Valid Score(Y) : 20

Truncation : 1

Options

remsd_wts : Comb. groups

Read Parameters from Control File

OK

Appendix C: CLT10 Score Frequencies

The CLT 10 score frequency for linking

SCORE	FREQUENCY	PERCENT	VALID PERCENT	CUMULATIVE PERCENT	SCORE	FREQUENCY	PERCENT	VALID PERCENT	CUMULATIVE PERCENT
38	1	0.5	0.5	0.5	84	4	1.8	1.8	48.6
41	1	0.5	0.5	0.9	85	4	1.8	1.8	50.5
46	1	0.5	0.5	1.4	86	5	2.3	2.3	52.7
50	2	0.9	0.9	2.3	87	3	1.4	1.4	54.1
53	1	0.5	0.5	2.7	88	3	1.4	1.4	55.5
54	1	0.5	0.5	3.2	89	5	2.3	2.3	57.7
60	1	0.5	0.5	3.6	90	5	2.3	2.3	60.0
61	5	2.3	2.3	5.9	91	6	2.7	2.7	62.7
63	4	1.8	1.8	7.7	92	4	1.8	1.8	64.5
64	2	0.9	0.9	8.6	93	8	3.6	3.6	68.2
65	1	0.5	0.5	9.1	94	5	2.3	2.3	70.5
66	3	1.4	1.4	10.5	95	7	3.2	3.2	73.6
67	3	1.4	1.4	11.8	96	9	4.1	4.1	77.7
68	4	1.8	1.8	13.6	97	6	2.7	2.7	80.5
69	5	2.3	2.3	15.9	98	3	1.4	1.4	81.8
70	1	0.5	0.5	16.4	99	2	0.9	0.9	82.7
71	3	1.4	1.4	17.7	100	2	0.9	0.9	83.6
72	4	1.8	1.8	19.5	101	3	1.4	1.4	85.0
73	5	2.3	2.3	21.8	102	3	1.4	1.4	86.4
74	3	1.4	1.4	23.2	103	4	1.8	1.8	88.2
75	4	1.8	1.8	25.0	104	2	0.9	0.9	89.1
76	5	2.3	2.3	27.3	105	6	2.7	2.7	91.8
77	4	1.8	1.8	29.1	107	1	0.5	0.5	92.3
78	5	2.3	2.3	31.4	109	5	2.3	2.3	94.5
79	7	3.2	3.2	34.5	111	7	3.2	3.2	97.7
80	6	2.7	2.7	37.3	112	2	0.9	0.9	98.6
81	4	1.8	1.8	39.1	116	2	0.9	0.9	99.5
82	8	3.6	3.6	42.7	117	1	0.5	0.5	100.0
83	9	4.1	4.1	46.8	Total	220	100.0	100.0	

Appendix D: CLT10 and PSAT Percentile Rankings

10TH GRADE PERCENTILES			10TH GRADE PERCENTILES			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	12			
1140	87	86	720	8	10	Mean Score	939	934
1130	86	85	710	6	9			
1120	84	84	700	5	7	Standard Deviation	170	180
1110	83	83	690	4	6			

Appendix E: Assessment of Spelling on the CLT10

SEVERAL U.S. STATES INCLUDE A requirement that students be tested on spelling for their education, whether public, charter, private, at-home, or parochial, to be legally recognized. North Carolina’s legal formulation is characteristic of such requirements: “The nationally standardized test or other equivalent measurement selected must measure achievement in the areas of English grammar, reading, spelling, and mathematics. Each school shall make and maintain records of the results achieved by its students. For one year after testing, all records shall be made available, subject to G.S. 115C-174.13, at the principal office of such school, at all reasonable times, for annual inspection by a duly authorized representative of the State of North Carolina.” (Quoted from G.S. 115C-549.)

The Grammar/Writing section of each CLT-suite exam normally contains one or more questions requiring students to distinguish the correct answer based on spelling, though there is not a dedicated section of the exam testing spelling alone. The following examples, taken from the CLT8 practice test available on the main CLT site, are typical of the variety and extent of spelling examination on all three tests in the CLT exam suite.

- 49. witch
 - A) NO CHANGE
 - B) whether
 - C) wich
 - D) which

- 68. Kenyas
 - A) NO CHANGE
 - B) Kenya’s
 - C) Kenyas’
 - D) Kenya

According to the blueprint for CLT-suite exams, all tests (including the CLT10) contain one to three questions that distinguish the correct answer partly or exclusively via spelling. Additional questions may combine spelling, meaning, and context as indicators of the correct answer.

