

PARAGON RESEARCH REPORTS

Language Proficiency Index (LPI) Test 2016 Report



TESTING ENTERPRISES
Paragon



a subsidiary of
THE UNIVERSITY OF BRITISH COLUMBIA

Contact Information

Paragon Testing Enterprises

110-2925 Virtual Way, Vancouver BC V5M 4X5

research@paragontesting.ca

Contents

1 – Description of the Test.....	1
1.1 General Description	1
1.2 Test Format.....	1
2 – Scoring and Reporting of Results.....	2
2.1 Explanation of Scoring.....	2
2.2 Procedures for Reporting Scores.....	2
3 – Interpreting LPI Results.....	2
4 – Test-Taking Population.....	3
5 – Test Statistics	4
5.1 Descriptive Statistics and Distribution of Essay Scores.....	4
5.2 Measurement Consistency.....	5
5.3 Rater Agreement for the Writing Component.....	6

List of Tables

Table 1: Format and Content of the LPI	1
Table 2: Proposed Interpretation of Writing Band Scores	3
Table 3: Distribution (in %) of LPI Test Takers by Age	4
Table 4: Distribution (in %) of LPI Test Takers by Purpose for Taking the Test	4
Table 5: Descriptive Statistics for the LPI Test	4
Table 6: Distribution (in %) of Test Takers at each Essay Band Score	4
Table 7: Reliability Estimates for the Multiple-Choice Items on the LPI Test	5
Table 8: LPI Test-Retest Reliability	6
Table 9: LPI Test-Retest Writing Band Score Changes	6

1 — Description of the Test

1.1 General Description

The Language Proficiency Index® (LPI) is a standardized test of English language reading and writing skills. The LPI, originally developed at The University of British Columbia, is currently administered by Paragon Testing Enterprises Inc., a subsidiary of The University of British Columbia.

The purpose of the LPI is to provide a consistent and standardized measurement of a test taker’s competency in English. The LPI is used by a number of post-secondary institutions and professional associations as proof of English language proficiency for academic and professional purposes.

Paragon is committed to upholding the highest standards in educational measurement. All parts of the LPI are written following specified guidelines, and results are closely monitored to ensure that they are accurate and informative. Paragon works closely with test centres to make certain that the LPI is administered in a way that is fair and accessible but also secure, ensuring that the LPI is available to all individuals who wish to take the exam and that the scores issued are defensible.

1.2 Test Format

The LPI is a paper-based test of reading and writing skills. The test duration is 2 ½ hours. Table 1 describes the format and content of the LPI.

Table 1: Format and Content of the LPI			
Component	Time	Item Description	Number of Items
Reading	Suggested: 1 hour	Sentence Structure (multiple choice) Test takers are asked to identify common errors in sentence structure.	10
		English Usage (multiple choice) Test takers are asked to identify common errors in English usage.	10
		Reading Comprehension Section A (multiple choice) Test takers read two short reading passages and answer comprehension questions.	8
		Reading Comprehension Section B (summary writing) Test takers read and summarize the main idea(s) of three prose passages.	3 tasks
Writing	Suggested: 1 ½ hours	Test takers write a 300–400-word, opinion-based essay on one of the three topics provided.	1 task

2 — Scoring and Reporting of Results

2.1 Explanation of Scoring

The multiple-choice items are scored by computer. Each correct answer contributes proportionately to the final score, and there are no points deducted for wrong answers. Reading Comprehension Section B is evaluated by trained raters using a 4-point scale. The Writing section is graded according to a scale established by Paragon and is assessed by at least two Paragon-certified raters.

2.2 Procedures for Reporting Scores

Test takers receive a score report that provides section results with information about the maximum possible score. The result for the Reading component is reported as a raw score on a 0–40-point scale. The Writing component is reported both as a raw score on a 0–40-point scale and as a band score (0–6) with a brief descriptor for each level.

3 — Interpreting LPI Results

The LPI has been designed to assess the readiness of test takers for entry into first-year, post-secondary academic programs or courses or for entry into professional contexts. Table 2 describes the proposed interpretation of each band score for the Writing component.

When interpreting an LPI score report, it is important to remember that the LPI estimates a test taker's true proficiency by approximating the kinds of tasks that they may encounter in the course of their academic or professional work. There are, inevitably, small mismatches between the tasks that test takers complete as part of the test and the tasks that they have to perform in a specific academic or professional context. Also, temporary factors unrelated to a test taker's true proficiency, such as fatigue, anxiety, or illness, may affect their LPI test results.

When using test scores for decision-making, check the date the test was taken. Language ability changes over time; it can improve with active use and further study of the language, but may also diminish if an individual does not continue to study or use English on a regular basis. It is also important to remember that test performance is only one aspect to be considered. Communicative language ability consists of both knowledge of language and knowledge of the world. Therefore, it is important to consider how factors other than language affect how well someone can communicate. For example, in a post-secondary context, the ability to function effectively involves not only knowledge of English but also other knowledge and skills such as content knowledge and academic study skills.

Table 2: Proposed Interpretation of Writing Band Scores

Band Score	Proficiency	Description
6	Advanced Proficiency	Demonstrates exceptional fluency marked by a wide range of skills, including excellent organizational abilities and original insights. Advanced proficiency is indicated in the clear articulation of both complex and straightforward concepts.
5	Effective Proficiency	Demonstrates fluent competency. For the most part, the writing is clear and controlled. Occasional errors in expression and structure do not significantly detract from the coherent articulation of ideas. There is clarity in development and organization.
4	Adequate Proficiency	Demonstrates adequate competency with satisfactory organization and structure although expression errors are evident, particularly occurring in occasional clusters. This is writing that requires some revision.
3	Developing Proficiency	Demonstrates some familiarity with written communication, using simple tools of expression. Major problems in diction, sentence structure, and organization, however, are evident. Numerous second language errors may also be present. This is writing that requires considerable revision.
2	Minimal Proficiency	Demonstrates only limited ability in written communication. Second language expression errors are the dominant feature of this writing.
1	No Proficiency	Demonstrates very limited knowledge of written English. The writer does not have the language skills to write a series of statements that can be understood by the reader.
0	Essay cannot be evaluated	Essays are placed at Level 0 when the component has not been attempted, not enough material has been written to allow for a fair evaluation, or the essay does not directly develop the topic chosen.

4 — Test-Taking Population

This section presents an overview of the test takers who took the LPI in 2016, providing demographic information for the test population. Every test taker completes a registration form which asks for their date of birth and purpose for taking the test. Cases where information has not been given have been treated as missing data.

Tables 3 and 4 present the distribution of test takers by age and purpose for taking the test. Table 3 shows that the majority of LPI test takers were 17–25 years old (64.4%). This is in line with the main purpose for taking the test, i.e., for entrance into post-secondary education. Table 4 shows that almost all of the test takers reported that they took the test for academic purposes (94.9%).

Table 3: Distribution (in %) of LPI Test Takers by Age

Age	% of Test Population
< 12	0.2
13–16	2.0
17–19	39.7
20–25	24.7
26–30	7.2
31–35	6.2
36–40	5.9
> 40	14.1
Missing Data	0.2

Table 4: Distribution (in %) of LPI Test Takers by Purpose for Taking the Test

Purpose	% of Test Population
Academic	94.9
Professional Certification	5.1

5 – Test Statistics

5.1 Descriptive Statistics and Distribution of Essay Scores

Table 5 presents descriptive statistics for each LPI section, essay raw score, and essay band score. Table 6 shows the percentage of test takers at each essay band score. The data indicates that the majority of test takers receive an essay band score 4 (55.32%).

Table 5: Descriptive Statistics for the LPI Test

Component	Section	Minimum Score	Maximum Score	Mean	Standard Deviation
Reading	Sentence Structure	0	10	3.99	2.42
	English Usage	0	10	4.19	2.79
	Reading Comprehension	0	19	6.30	4.18
Writing	Essay Scores	0	37	25.77	3.69
	Essay Levels	0	6	4.00	0.71

Table 6: Distribution (in %) of Test Takers at each Essay Band Score

Essay Band Score	% of Test Population
6	0.11
5	22.88
4	55.32
3	21.16
2	0.21
1	0.02
0	0.31

5.2 Measurement Consistency

Test scores are a numerical measure of a test taker’s ability. In theory, a test taker’s score should be the same each time the test is taken. In practice, however, even when the test conditions are carefully controlled, an individual’s performance on a set of test items will vary from one administration to another due to variation in the items across different versions of the same test or due to variability in individual performance. Among the reasons for variations in individual performance are temporary factors unrelated to a test taker’s English language proficiency. These include fatigue, anxiety, or illness. Therefore, test scores always contain a small amount of measurement error. The aim is to keep this error to a minimum. For standardized examinations such as the LPI Test, a reliability figure of 0.80 and above is expected and acceptable.

Paragon monitors two facets of measurement consistency: internal consistency and test-retest reliability. Internal consistency is a measure of whether items in the test that are intended to measure the same construct produce similar results. Internal consistency estimates (Cronbach’s alpha) are calculated for the multiple-choice items on each LPI test form. Three sections of the Reading component comprise multiple-choice items: Sentence Structure, English Usage, and Reading Comprehension Section A. Table 7 shows the overall reliability estimate for each form administered in 2016. It shows that the average reliability estimate for all the forms administered in 2016 was .89 and the lowest reliability estimate was .87. These values suggest excellent internal consistency of measurement for the LPI multiple-choice items.

Table 7: Reliability Estimates for the Multiple-Choice Items on the LPI Test	
Test Form	Overall
DO	.897
DP	.908
DZ	.865
DV	.891
DU	.877
DR	.875
DW	.893
DS	.879
DT	.898
Average	.887

Test-retest reliability refers to the stability of a test taker’s scores if they take the test on multiple occasions over a period of time. Under experimental conditions, test takers will take the same test form each time. Under naturalistic conditions, test takers will take different, but equivalent, test forms. Paragon monitors test-retest reliability under naturalistic conditions. It is not uncommon for LPI test takers to attempt the test more than once. Paragon has accumulated retest records dating back to April 2004. Changes in test taker performance were measured between subsequent attempts (rather than between first attempt and final attempt). Since a test taker’s central purpose for taking the test again is to

get a higher score, Paragon encourages test takers to undertake a period of study between test attempts. Therefore, changes in test scores and a moderate to low correlation between test attempts are expected. Table 8 presents the test-retest correlations (Pearson’s r) for the Reading and Writing component raw scores. The correlation figures for both components are moderate.

Table 8: LPI Test-Retest Reliability	
Component	Test-Retest Correlation (Pearson’s r)
Reading	0.60
Writing	0.48

Since the Writing component score is also reported as a band score, Table 9 presents patterns in band changes when test takers retake the LPI. Though test takers can expect to see changes in their raw score for the Writing component, changes in the band score awarded are less common; almost 62% of retests result in the same Writing band score.

Table 9: LPI Test-Retest Writing Band Score Changes	
Band Score Change	% of Test Takers
2	0.98
1	26.26
0	61.92
-1	10.75
-2	0.09

5.3 Rater Agreement for the Writing Component

The raters for the Writing component are highly proficient in English and are fully trained and certified by Paragon. Each essay is rated independently by two accredited raters using a 20-point rating scale. The essay is evaluated by a third rater if scores awarded by the original two raters fall outside a tolerance of ± 2 points.

Paragon monitors rater agreement for quality control purposes. The percentage of within tolerance agreement (± 2 score points) between the raters is monitored. Across the 2016 LPI test administrations, the overall within tolerance agreement percentage was 98.2%. This figure suggests excellent agreement among raters.