PARAGON RESEARCH REPORTS

CELPIP Data Report of 2024 Test Takers

research@prometric.com

CELPIP® | by PROMETRIC

1 Description of the Test

1.1 General Description

The Canadian English Language Proficiency Index Program (CELPIP) is a standardized test of general English language proficiency. It is administered by Paragon Testing Enterprises Inc, a Prometric company.

The CELPIP test evaluates a test taker's ability to use English in everyday situations such as communicating in the workplace, interacting socially, understanding news broadcasts, and responding to written materials. The CELPIP is officially recognized by Immigration, Refugees and Citizenship Canada (IRCC) for applications related to permanent residency and Canadian citizenship. It is also accepted by a number of post-secondary institutions and professional associations as evidence of English language proficiency for academic and professional purposes.

Paragon is committed to upholding the highest standards in educational measurement. The CELPIP test is developed following established guidelines, and results are rigorously monitored to ensure they are accurate, informative, and defensible. Paragon also works closely with test centres to ensure that the CELPIP test is administered fairly, securely, and is accessible to all individuals who wish to take the exam.

1.2 Test Format

There are two versions of the CELPIP Test: the CELPIP-General Test and the CELPIP-General LS Test. Individuals who take the CELPIP-General Test are assessed on on four components: Listening, Reading, Writing, Speaking. Individuals who take the CELPIP-General LS Test are assessed on the Listening and Speaking components. Table 1 summarizes the format and content of each test component.

Table 1: Format and Content of the CELPIP Test			
Component	Duration (Minutes)	Item Description	Number of Items
Listening	47 - 55	Test takers listen to seven passages and answer comprehension questions. The listening passages cover topics in daily conversation, problem-solving, news items, discussions, and viewpoints.	38
Reading	55 – 60	Test takers read several passages and answer comprehension questions. The reading passages engage the test takers in	38

Table 1: Format and Content of the CELPIP Test

		understanding correspondence, interpreting a diagram, and reading for viewpoints.	
Writing	53 – 60	Test takers write an email and write a response to survey questions.	2
Speaking	15-20	Test takers speak to give advice, talk about personal experiences, describe scenes, make predictions, compare and persuade, deal with difficult situations, express opinions, and to describe an unusual situation.	8

Note: The Listening and Reading components may contain additional unscored items. These unscored items are indistinguishable from scored items.

1.3 Scoring and Score Reporting

The CELPIP Test assesses English language proficiency in general social, educational, and workplace contexts. Each test taker receives a score report that provides a score for each component. Multiple-choice items (Listening and Reading components) are scored by computer. Each correct answer contributes proportionately to the final score, and no points are deducted for wrong answers. The Writing and Speaking components are each evaluated by at least three trained and certified raters according to standardized scoring rubrics.

CELPIP scores are reported in bands ranging from 0 to 12. These scores are aligned with the Canadian Language Benchmarks (CLB). Table 2 able 2 provides a description of each CELPIP level, alongside its CLB equivalent.

Table 2: CELPIP Score Interpretation and CLB Equivalencies

Descriptor	CELPIP Level	CLB Level
Advanced proficiency in workplace and community contexts	12	12
Advanced proficiency in workplace and community contexts	11	11
Highly effective proficiency in workplace and community contexts	10	10
Effective proficiency in workplace and community contexts	9	9
Good proficiency in workplace and community contexts	8	8

Table 2: CELPIP Score Interpretation and CLB Equivalencies		
Adequate proficiency in workplace and community contexts	7	7
Developing proficiency in workplace and community contexts	6	6
Acquiring proficiency in workplace and community contexts	5	5
Adequate proficiency for daily life activities	4	4
Some proficiency in limited contexts of personal relevance	3	3
Limited ability in contexts related to immediate needs	2	1, 2
Insufficient information to assess	1	/
Insufficient information to assess	0	/
Not Administered: test taker did not receive this test component	NA	/

Note: CELPIP levels 0, 1, and 2 were previously reported as level "M"

When interpreting a CELPIP Test score report, it is important to remember that the CELPIP Test estimates test takers' true proficiency by approximating the kinds of tasks they may encounter during their daily lives, study, or at work. There are, inevitably, small mismatches between the tasks that test takers complete as part of the test and the tasks that they perform in a specific context. Also, temporary factors unrelated to test takers' true proficiency, such as fatigue, anxiety, or illness, may affect their CELPIP test results.

2 Test-Taking Population

This section presents an overview of the test takers who took the CELPIP test in 2024. As shown in Figure 1, the vast majority (96.0%) of CELPIP test takers reported their primary purpose as applying for Canadian immigration or citizenship (IRCC).
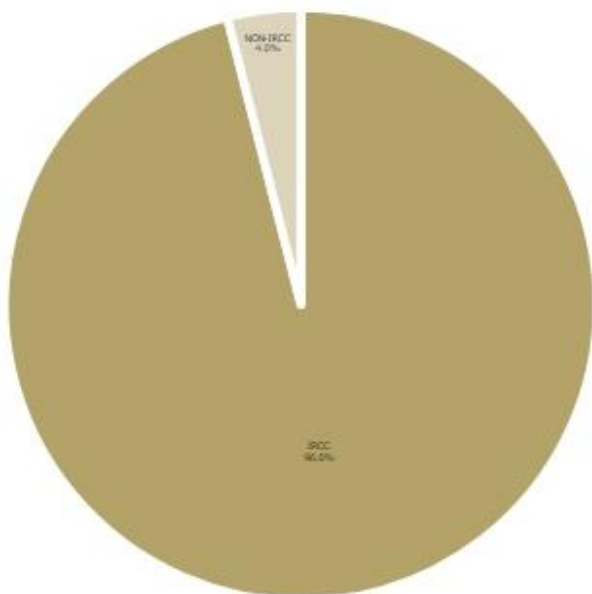
Figure 1: Distribution of CELPIP Test Takers by Test-Taking Purpose

Figure 2 displays the proportion of test takers by test version (CELPIP-General vs CELPIP-General LS).
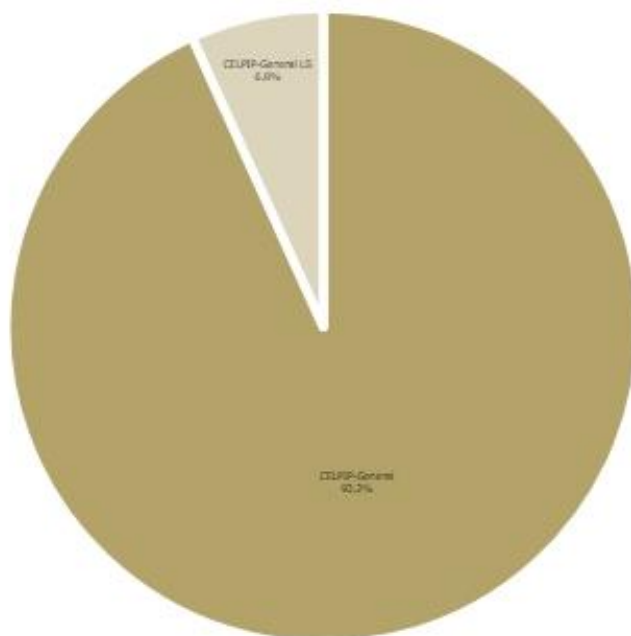

Figure 2: Distribution (in %) of CELPIP Test Takers by Test

Table 3 shows the age group distribution of test takers by test type. The majority of test takers (80.5%) were between 20 and 40 years old. CELPIP-General attracted a younger demographic, while CELPIP-General LS had relatively more test takers in the 36–55 age range.

Table 3: Distribution of CELPIP Test Takers by Age Group (%)

Age Group	Overall	CELPIP-General	CELPIP-General LS
< 20	0.20	0.19	0.36
20-25	13.96	14.66	4.48
26-30	29.74	30.98	12.94
31-35	21.37	21.33	21.88
36-40	15.67	15.10	23.51
41-45	10.68	10.06	19.14
46-50	5.48	5.00	12.01
51-55	2.21	1.96	5.58
> 55	0.68	0.73	0.08

3 Test Statistics

3.1 Score Distributions

Tables 4 presents the mean scores and standard deviations for the CELPIP-General and CELPIP-General LS component scores. The mean score is the simple average of all test takers' scores, and the standard deviation quantifies how scores are spread out from the mean.

Table 4: CELPIP-General and CELPIP-LS Test Scores

Test	Component	Mean	Standard Deviation
CELPIP-General	Listening	8.22	2.29
	Reading	7.21	2.39
	Writing	7.91	1.93
	Speaking	7.76	2.08

Test	Component	Mean	Standard Deviation
CELPIP-LS	Listening	6.70	2.57
	Speaking	6.67	2.48

Tables 5a and 5b show the percentage breakdown of test takers who received each component band score on the CELPIP-General and CELPIP-General LS Tests.

Table 5a: Score Distributions (%) on CELPIP-General Test

Band	Listening	Reading	Speaking	Writing
0-2	0.3	1.6	0.5	0.3
3	1.4	3.6	1.4	0.8
4	4.3	7.5	3.5	2.3
5	8.7	13.6	7.4	5.3
6	10.9	15.7	14.5	10.9
7	12.1	14.4	17.5	23.3
8	10.4	10.3	22.5	27.3
9	18.2	14.2	13.1	10.2
10	18.0	8.1	7.1	8.0
11	9.3	8.5	7.9	5.4
12	6.5	2.5	4.5	6.2

Table 5b: Score Distributions (%) on CELPIP-General LS Test

Band	Listening	Speaking
0-2	2.2	3.4
3	6.1	6.2
4	13.6	10.0
5	17.9	14.6
6	14.1	17.2
7	10.3	13.5
8	7.3	13.1

Band	Listening	Speaking
9	10.0	7.6
10	9.6	4.7
11	4.9	6.0
12	4.0	3.8

The above tables (Tables 4, 5a, and 5b) suggest that a larger proportion of test takers achieve lower bands on the CELPIP-General LS Test than on the CELPIP-General Test. This reflects the differing purposes and required proficiency levels: CELPIP-General is typically used by applicants for permanent residency, whereas CELPIP-General LS is used for citizenship applications, which have lower language requirements.

3.2 Test Score by Subgroups

The following tables offer a more detailed breakdown of test taker performance by gender, first language, and country of citizenship.

Table 6: CELPIP Scores by Test Purpose

Test Purpose	Listening		Reading		Writing		Speaking	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
IRCC	8.10	2.34	7.21	2.39	7.91	1.93	7.66	2.12
NON-IRCC	8.34	2.23	7.24	2.34	8.02	1.96	8.19	2.33

Table 7a: CELPIP-General Scores by the Top 10 Declared Nationality

Country	Listening		Reading		Writing		Speaking	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
India	8.59	1.93	7.41	2.08	8.01	1.61	7.93	1.63
Philippines	7.31	2.25	6.18	2.27	7.35	1.89	6.88	1.88
Nigeria	8.28	2.00	7.18	2.16	8.59	1.76	9.22	1.62
China	7.62	2.63	7.08	2.59	7.42	1.88	6.63	1.96
Ukraine	7.02	2.47	6.08	2.48	6.74	1.89	6.61	1.89

Country	Listening		Reading		Writing		Speaking	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Hong Kong, China	8.11	2.32	7.54	2.4	7.87	1.73	7.03	1.77
South Korea	7.03	2.41	6.26	2.36	6.7	1.81	5.95	2.05
Iran	8.37	2.24	7.31	2.28	7.93	1.78	7.62	1.76
Mexico	8.37	2.42	7.48	2.52	7.64	2.07	7.7	2.09
Brazil	8.86	2.33	8.17	2.45	8.14	1.95	7.97	2.03

Table 7b: CELPIP-LS Scores by the Top 10 Declared Nations/Regions

Country	Listening		Speaking	
	Mean	SD	Mean	SD
Iran	6.12	2.26	6.16	1.98
India	6.38	2.39	6.46	2.40
Brazil	8.28	2.33	7.69	2.00
Mexico	7.63	2.47	7.34	2.15
Philippines	6.86	2.20	6.62	2.01
Afghanistan	5.32	1.93	5.70	2.02
Germany	9.77	2.07	10.12	2.03
Pakistan	5.97	2.18	6.29	2.26
Ukraine	6.75	2.59	6.48	2.31
Vietnam	5.27	1.75	4.92	1.57

Table 8a: CELPIP-General Scores by the Top 10 Declared First Languages

Language	Listening		Reading		Writing		Speaking	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
English	8.84	2.11	7.79	2.38	8.88	2.01	9.33	2.01
Tagalog	7.42	2.24	6.30	2.28	7.43	1.89	6.93	1.86
Chinese	7.67	2.60	7.09	2.57	7.44	1.88	6.67	1.95

Language	Listening		Reading		Writing		Speaking	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Panjabi	8.07	2.10	6.80	2.15	7.53	1.62	7.55	1.70
Hindi	8.91	1.86	7.74	2.07	8.30	1.66	8.36	1.61
Spanish	8.53	2.36	7.75	2.44	7.90	1.96	7.80	1.99
Cantonese (Chinese)	8.15	2.30	7.57	2.37	7.88	1.70	7.03	1.72
Ukrainian	6.99	2.46	6.03	2.47	6.71	1.89	6.58	1.88
Gujarati	8.66	1.96	7.46	2.08	7.87	1.62	7.74	1.64
Malayalam	8.72	1.74	7.59	1.90	8.08	1.39	7.89	1.41

Table 8b: CELPIP-General LS Scores by the Top 10 Declared First Languages

Language	Listening		Speaking	
	Mean	SD	Mean	SD
Spanish	7.40	2.51	7.13	2.24
Farsi	5.94	2.21	6.03	2.00
Arabic	5.94	2.34	6.07	2.25
English	7.54	2.61	8.44	2.73
Portuguese	8.19	2.32	7.68	2.04
German	9.41	2.39	9.78	2.32
Panjabi	5.63	2.09	5.69	2.17
Chinese	6.51	2.29	5.85	1.73
Vietnamese	5.28	1.75	4.93	1.58
Urdu	5.99	2.22	6.30	2.26

These tables provide insights into how demographic factors correlate with performance across CELPIP components.

4 Closing Remarks

We remain committed to maintaining rigorous psychometric standards in test development, scoring, and administration of the CELPIP test. Ongoing enhancements ensure that the CELPIP test continues to be a fair, reliable, and valid measure of real-world English language proficiency.

The CELPIP test is fully computer-delivered and widely accessible, with over 140 testing centers across Canada and globally. To support test takers from diverse linguistic and educational backgrounds, a broad range of preparation resources is available—both free and paid—including official study guides, practice tests, and instructional programs.

CELPIP will continue to evolve to meet the needs of individuals, institutions, and regulatory bodies. We extend our sincere thanks to all test takers and stakeholders for their continued trust in CELPIP.

For test registration, preparation support, or further information, please visit the CELPIP website at https://www.celpip.ca/.