Pearson

Bridging Policy, Practice, and Lived Experience: Evaluating a High-stakes Language Test for Canadian Migration

Languages Canada 2026

Sarah Hughes, Director of Assessment Research and Validity

Ellen Barrow, Senior Researcher

Agenda

01   **PTE Core**

Context and Background

02   **Establishing Score Meaning**

Design, Development and Test Validation

03   **Understanding Impact**

Test Taker Perspectives

PTE Core Overview

The PTE Portfolio: A World of Testing

The PTE Portfolio offers a range of English language tests tailored to different global needs. From academic study to migration and family visas, there's a PTE test to support every journey.

PTE Academic

PTE Academic • PTE Academic UKVI

A globally recognised English test trusted by universities and governments for study and migration. Fast, fair, and entirely computer-based.

PTE Core

Designed for Canadian permanent residency and work visa applications, it assesses real-life English skills in everyday and workplace contexts.

PTE Home

PTE Home A1 • PTE Home A2 • PTE Home B1

A simple, secure English test accepted for UK family, settlement and citizenship visas, with results typically available in just 2 days.

How do PTE Core and PTE Academic compare?

PTE Core

PTE Academic

Purpose

PTE Core tests **general English** through tasks related to everyday life and workplace scenarios to be used to support immigration decisions in Canada.

PTE Academic tests **academic English** with tasks related to study and professional skill to be used to support higher education and skilled migration decisions worldwide.

Skills

Assesses four skills: listening, reading, speaking, writing

Designed to align to the **Canadian Language Benchmarks**

Assesses four skills: listening, reading, speaking, writing

Designed to align to the **Common European Framework of Reference**

Test Content

General content relevant to everyday life, workplaces and some study situations (non-academic). Content covers subjects that would be familiar to everyone.

Authentic source content such as videos/podcasts on the internet, announcements, automated telephone helplines, work meetings, and broadcasts.

Academic content relevant to study and professional situations such as university lectures, interviews, and academic articles.

Authentic source content such as reading and audio texts from books and articles covering broad topics such as social sciences, humanities, business, technology, and health & medicine.

Test Structure

Part	Minutes	Minutes
Pre – test	Technical check	Not timed
Part 1: **Speaking & Writing**	• Personal Introduction • Read aloud (6–7) • Repeat sentence (10–12) • Describe image (3–4) • **Respond to a situation (2–4)** • Answer short question (5–6) • Summarize written text (1–2) • **Write email (2–3)**	46 - 67 minutes
Part 2: **Reading**	• Fill in the blanks (R&W) (5–6) • Multiple choice, choose multiple answers (1–2) • Reorder paragraph (2–3) • Fill in the blanks (R) (4–5) • Multiple choice, choose single answer (1–2)	27 - 38 minutes
Part 3: **Listening**	• Summarize spoken text (1–2) • Multiple choice, choose multiple answers (1–2) • Fill in the blanks (2–3) • Multiple choice, choose single answer (1–2) • Select missing word (1–2) • Highlight incorrect words (1–2) • Write from dictation (3–4)	30 - 37 minutes

Establishing Score Meaning

Design, Development and Test Validation

Forming the Blueprint

Iterative Phases of Research

Initial assessment design and Item writing and development

Feedback from Canadian Test Review Group

Initial mapping exercise by Canadian experts to map items to the CLB

Review/Amendment of Items

Sample test forms piloted with test takers in focus groups and interviews

Large scale field tests and item and test analyses

PTE Core Alignment to Canadian Language Benchmark (CLB)

Do the CLB levels assigned by PTE Core genuinely reflect what Canadian experts recognise as CLB-level performance?

Initial Standard Setting (2021)

- To identify score thresholds that correspond to CLB levels 3–10, based on expert evaluation of field-test responses.

- 14 independent judges with extensive experience in teaching and assessing English language, strong familiarity with the CLB.

- Majority had over 15 years experience in language learning and assessment in Canada.

- An item-centred modified Angoff method was used for receptive skills.

- A test-taker modified Analytic judgement method was used for productive skills.

Establishing alignment between PTE Core and the Canadian Language Benchmarks (CLB)

Research Report

PTE Core Alignment to Canadian Language Benchmark (CLB)

Do the CLB levels assigned by PTE Core genuinely reflect what Canadian experts recognise as CLB-level performance?

Initial Standard Setting (2021)

CLB	PTE Core Listening	PTE Core Reading	PTE Core Speaking	PTE Core Writing
10	89–90	88–90	89–90	90
9	82–88	78–87	84–88	88–89
8	71–81	69–77	76–83	79–87
7	60–70	60–68	68–75	69–78
6	50–59	51–59	59–67	60–68
5	39–49	42–50	51–58	51–59
4	28–38	33–41	42–50	41–50
3	18–27	24–32	34–41	32–40

Full Research Report
Establishing alignment between PTE Core and the Canadian Language Benchmarks (CLB)

Establishing alignment between PTE Core and the Canadian Language Benchmarks (CLB)

Research Report

PTE Core Alignment to Canadian Language Benchmark (CLB)

Do the CLB levels assigned by PTE Core genuinely reflect what Canadian experts recognise as CLB-level performance?

Follow up Standard Verification Activities

2022
Verification

- To check the appropriacy of the 2021 thresholds and confirm that the original alignment decisions still held up

- Found good levels of agreement and support for the initial CLB alignment.

2025 (post-launch)
Monitoring

- Showed high levels of agreement with the PTE Core classifications
- Judge commentaries showed a nuanced analysis, referencing elements of specific test taker performances in relation to CLB standards.

Pearson | PTE

Full Research Report
Establishing alignment between PTE Core and the Canadian Language Benchmarks (CLB)

Establishing alignment between PTE Core and the Canadian Language Benchmarks (CLB)

Research Report

Construct Comparison

? Are we measuring the right kind of English for Canadian migration?

Independent researchers examined tasks, skills, and scoring criteria in PTE Core and compared them with CELPIP General.

Results indicate that PTE Core effectively evaluates overall English proficiency, with tasks that reflect real-life communication.
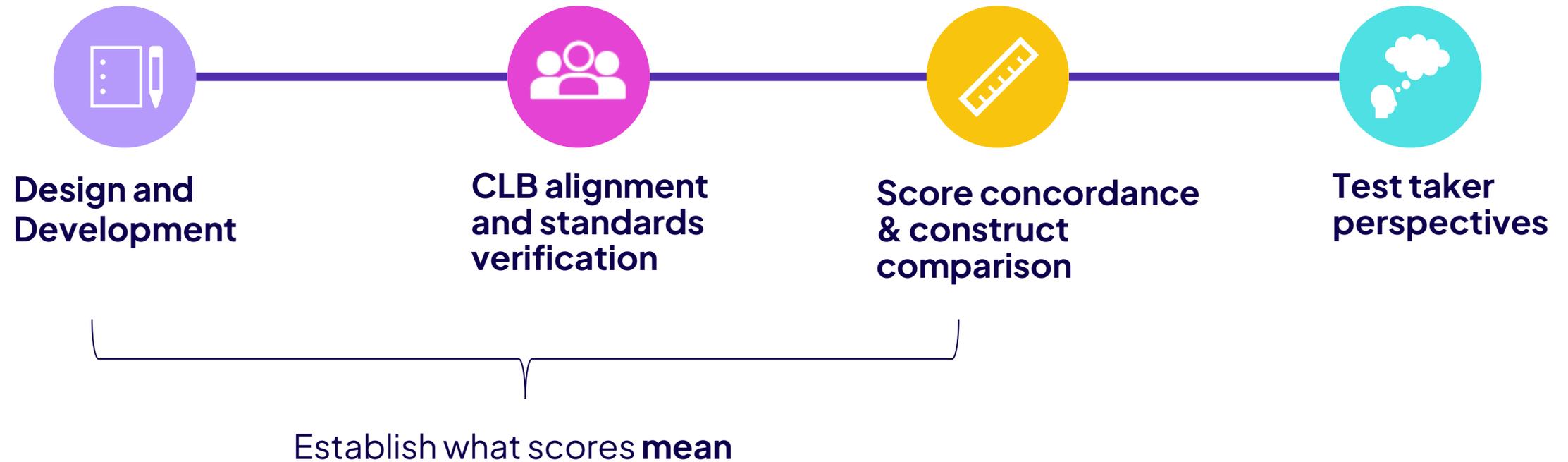
Concordance with CELPIP

? Do different accepted tests lead to comparable outcomes in practice?

Uses empirical data to find equivalent scores between two different tests.

Adds complementary perspectives to standards and alignment research

Preliminary analysis conducted last year, in progress...

Test Validation: Building a picture

Design and Development

CLB alignment and standards verification

Score concordance & construct comparison

Test taker perspectives

Establish what scores **mean**

Test Validation: Building a picture

Test taker perspectives

Alignment with policy standards is a critical foundation for test validity, it is only **one part of a bigger picture.**

The degree to which test takers perceive the test as **relevant**, **appropriate**, and **representative of real-world language use** or the skills it claims to assess.

Domain relevance of the item types

High stakes, big **consequences**

Understanding Impact

Test Taker Perspectives of PTE Core

Research Questions

01 To what extent do test takers perceive the tasks in PTE Core to be authentic representations of real-life language use in Canadian society?

02 How relevant do test takers find the PTE Core item types to their daily lives?

03 What challenges do test takers face in integrating into Canadian society, and how might these challenges inform the design and support structures of PTE Core?

Areas of Focus

Test Taker Background

- Demographic information
- Motivations for taking PTE Core
- Long-term goals
- Confidence in digital skills

Feedback on the Test

- General feedback on experiences of taking PTE Core
- Task authenticity /relevance, demand and user experience of test items

Test Preparation

- Familiarity with PTE Core
- What resources are test takers using?

Attitudes to AI

- General AI usage
- Perceptions of AI scoring systems: trust, accuracy, objectivity

Data Collection

| Research Method	Sample Size
Test Taker Survey	280
Test Taker Interviews	9
Focus Groups with Language Instruction for Newcomers to Canada (LINC) Students	10

Key Themes that emerged from analysis:

- Feedback on test experience and domain relevance of the tasks and items

- Contextual insights into test taker experiences and personal language learning journeys

Feedback on Test Experience

92% of survey respondents agree/strongly agree that the tasks in PTE Core are useful in measuring English language proficiency.

I think the tasks in the PTE Core test are useful in measuring English language proficiency (n=280)

■ Strongly agree　■ Agree　■ Disagree　■ Strongly disagree
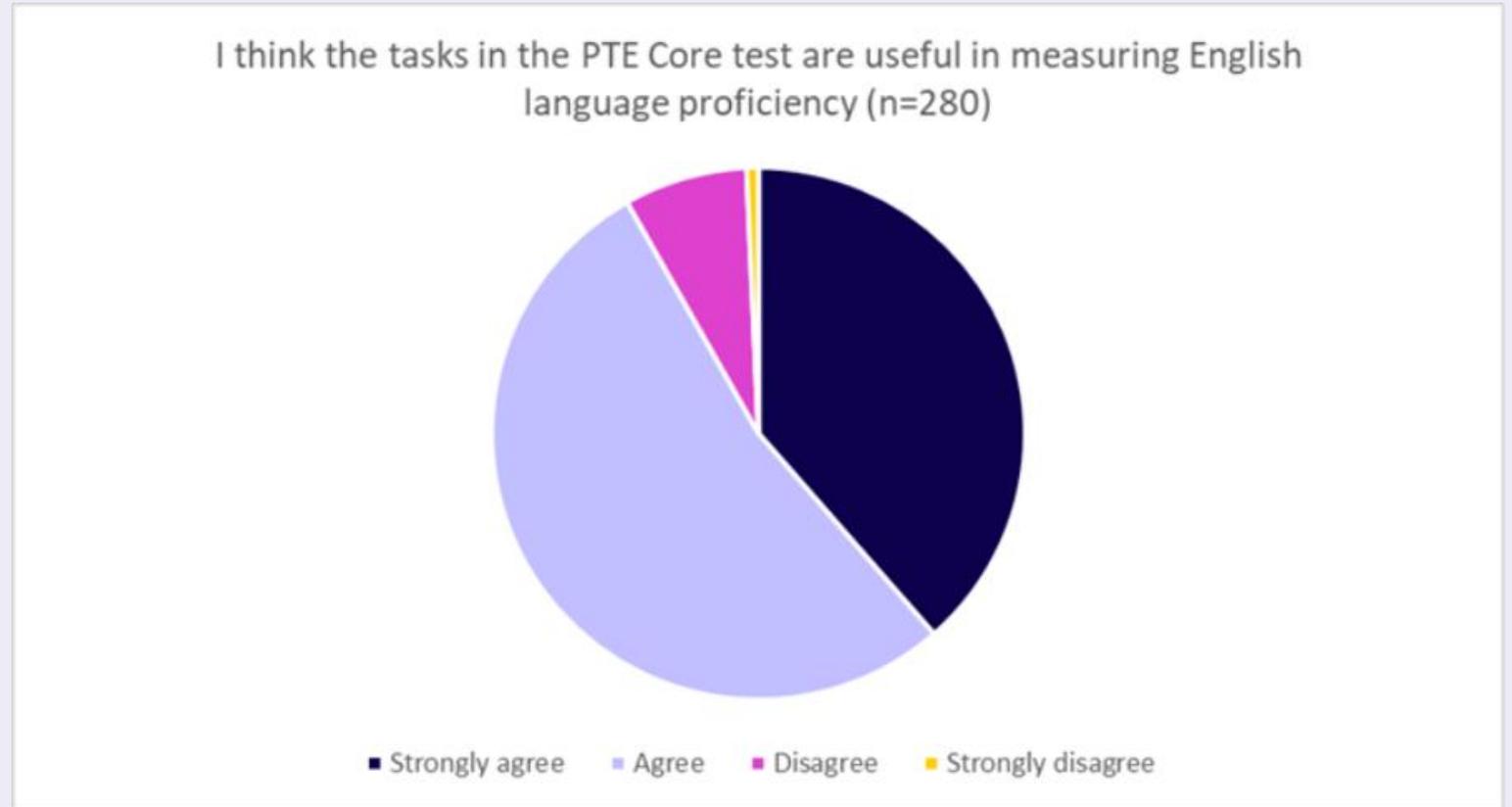
85% of all survey participants said that they were either generally relevant or very relevant to use in daily life.

To what extent are the tasks in the PTE Core test relevant to use in your daily life? (n=280)
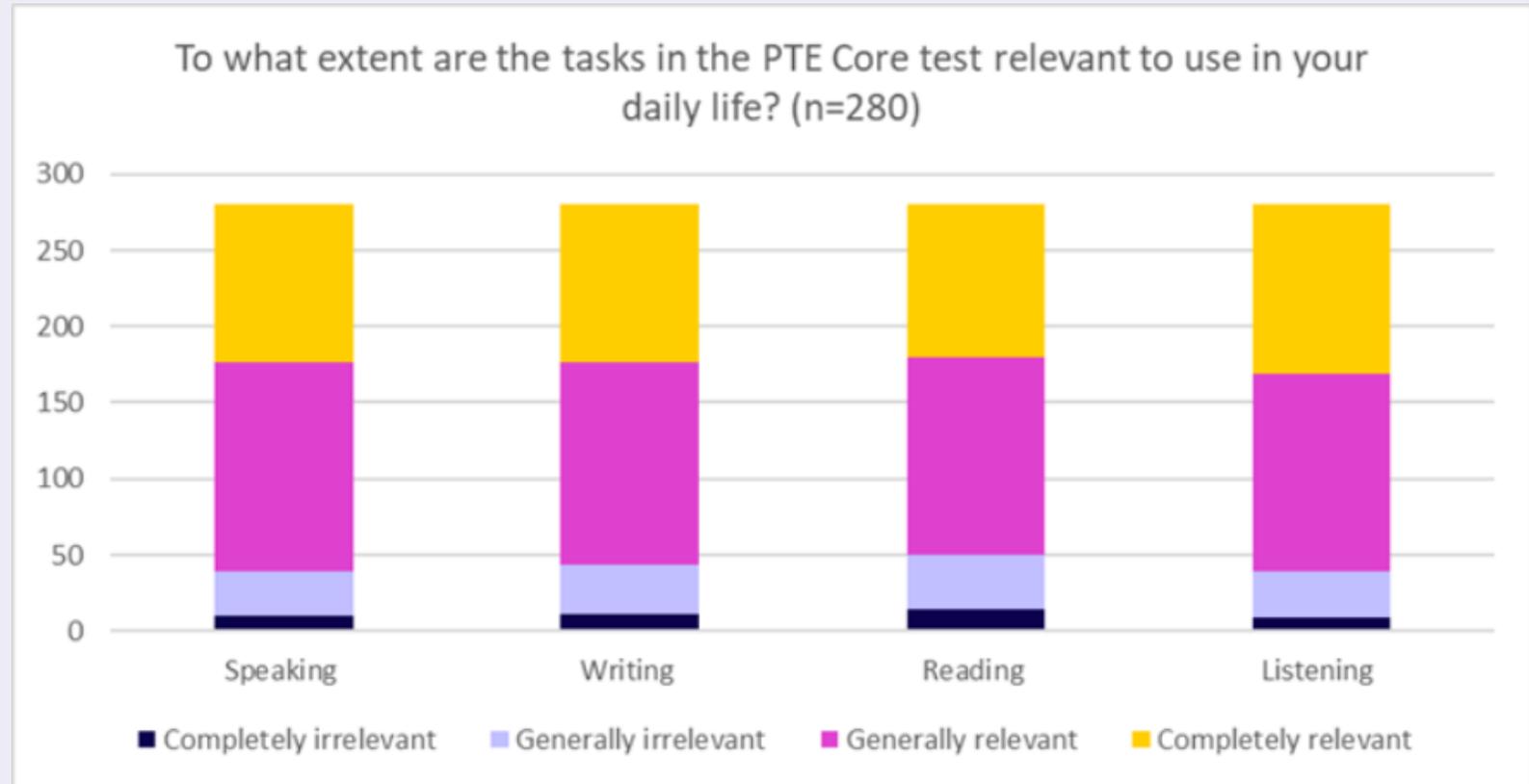
Respond to a Situation

Write Email

"

Every day we want to organise a daily time for the training of our kids, organising the private lessons for them**… It really happen[s] every day**. *(TT4 Interview)*

I was going to [the] airport**. Suddenly I got a message that the flight got cancelled. So I want to explain to the driver that I want to go back to my home.** So I told him that please go back to my home. I'll pay you the extra money. Please kindly change the route because I got a message that flight has been cancelled. *(TT6 Interview)*

I have to do like tens, if not hundreds of emails, like, every single day… this is again **aligning with…the skills, what is used in the day to day life**. *(TT7 Interview)*

I think it's really important to know how to write. And in the email, it's a daily life situation. You can do both in your personal life and your professional life **to like book a trip or to write an email to your director.** *(TT8 Interview)*

"

Contextual Insights

Integration Challenges

- Focus groups with 10 LINC students at College in Toronto area.

- Spoke of the difficulties they experienced in cultural, social and professional contexts.

> "
>
> My neighbour wants to help me to improve my English. So when they have a barbecue party they asked me to take part in. But you know, they speak so fast and [there are] **a lot of local joke when sometimes all of them laugh at something, but I don't know what happened.**
>
> [It's] challenging because [its a] **different culture** and the brand new here and just **my friends are almost all Chinese**.
>
> I have to try to find a job. I used to work in China. I do some marketing communication. But in Canada... **It's impossible for me to find a job because I don't know the English culture**, local culture, and especially for marketing job, the English language, how to use the vocabulary. How to use the sentence to influence your customer. It's really very hard even for local people.
>
> [I want to be a] member of here, **not the just the immigrant**
>
> "

Understanding Context = Understanding Impact

- Comparative analysis of PTE Academic and PTE Core test taker survey data showed that **PTE Core test takers consistently reported lower levels of confidence** in completing **digital tasks** than PTE Academic test takers.

- By recognising these differences and challenges, we can **reduce unintended obstacles** and support more inclusive and fair assessment results.

- Different tests serve different purposes: understanding context is an important part of understanding impact.

Purpose of the test versus use of the test

We monitor how the test operates in the real-world to ensure that the evidence we have supports the way test scores are being used in practice.

o   Are the people taking the test the **expected cohort**?

o   Is the test being used for the **intended decisions**?

o   Are the decisions linked to the **intended standards**?

o   How does the test influence English language learning (**washback**)?

o   How does success on the test relate to **meaningful outcomes** in integration, employment, and confidence?

These questions guide our ongoing validation research agenda, alongside continuous monitoring to ensure results are accurate, fair, and reliable.

Conclusions

- PTE Core is underpinned by robust validation evidence

- These evidence types address different questions and work together to establish score meaning

- **BUT** Technical evidence is not the whole story- high-stakes scores affect real people.

- Engaging stakeholders in the research & development is not just a box to tick – it is a vital part of ensuring language tests actually serve the people they impact.

Any questions?

Sarah.hughes1@pearson.com

Ellen.barrow@pearson.com

Discussion

01

Where do you see the biggest mismatch between test tasks and real-world needs?

02

What's one barrier your learners face before they even sit the test?

03

What evidence would increase your trust in score use decisions?