

Aligning LANGUAGECERT General Test Scores with Canadian Language Benchmarks: Enhancing Accessibility for Learners in Canada



- › Ofqual-regulated Awarding Organisation
- › Focusing on the assessment of language proficiency
- › Recognised by HEIs, organisations and government bodies in 90+ countries

» LANGUAGECERT portfolio

CEFR Levels	LC Global Scale	Schools/Domestic Universities			Study Abroad	Migration		Work	LC Global Scale	CEFR Levels
C2	90-100			ESOL C2	Academic/ Academic SELT		General/ General SELT	LTE	90-100	C2
C1	75-89			ESOL C1					75-89	C1
B2	60-74		ESOL B2 for schools	ESOL B2					60-74	B2
B1	40-59		ESOL B1 for schools	ESOL B1		ESOL SELT			40-59	B1
A2	20-39		ESOL A2 for schools	ESOL A2		ESOL SELT			20-39	A2
A1	10-19	YL ESOL Owl	ESOL A1 for schools	ESOL A1		ESOL SELT			10-19	A1
PreA1	0-9	YL ESOL Fox							0-9	PreA1

» LANGUAGECERT General

- › LANGUAGECERT General (LCG) and its counterpart qualification, LANGUAGECERT Academic (LCA) were conceived as a dynamic response to changing markets and stakeholder expectations.
- › Both tests are derived from the LANGUAGECERT item bank and report scores across relevant levels on the same measurement scale that is used for all LANGUAGECERT – the Global Scale.

(Jones, 2024, The LANGUAGECERT General Test: Assessing Language in the Migration and Employment Domains)

» Mapping LANGUAGECERT General to the CLB

- › Multi-level, computer-based English language test spanning CEFR A2 to C1
- › Developed from the LANGUAGECERT IESOL B2
- › Built on a bed of underpinning research and validation
- › Secure and reliable test delivery
- › Meaningful and user-friendly results reporting

» Mapping LANGUAGECERT General to the CLB

- › The test measures a range of skills and competences appropriate for personal, occupational and vocational contexts: reading and listening for gist and detailed understanding of a range of written and audio sources including adverts, articles, websites, diaries, radio programmes, and podcasts. The test includes writing in formal and informal registers, expressing viewpoints and taking part in role plays in real-life scenarios.
- › The Global Scale allows ease of interpretation for test users and provides a finely-tuned results service across all language skills. As shown, performance can be separated both by each skill and overall, so that a candidate is not only described as having “B2 ability”, but a more precise level of detail is provided on a candidate’s performance.

» Mapping LANGUAGECERT General to the CLB

- › LANGUAGECERT is committed to demonstrating test validity and fitness-for-purpose. This includes internal research analyses as well as external validation projects, such as the benchmarking of the LANGUAGECERT Global Scale against different tests and frameworks, demonstrating criterion-related validity.

» Mapping LANGUAGECERT General to the CLB

Test takers receive:

- › a score per skill on the LANGUAGECERT Global Scale (0-100)
- › an overall score
- › the CEFR level attained (A2, B1, B2, C1)
- › personalised feedback on their performance in the productive skills

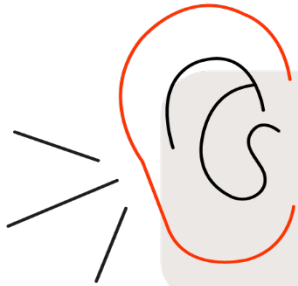
Widely accessible through the LANGUAGECERT Secure Test Centre Network operational in 110+ countries globally

- › Can also be taken online with live remote invigilation*
- › Providing test takers with their results within five business days

» Exam features



Duration
~ approx. 2'20"



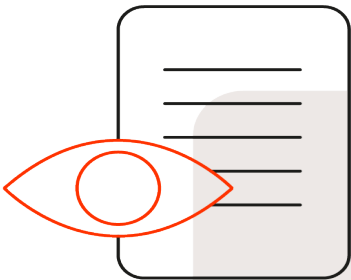
Listening

Ả Ỗ Ắ Ắ Ồ Ồ 30 Ắ Ắ Ắ Ắ Ắ Ắ
Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề
R Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ



Writing

Ả Ỗ Ắ Ắ Ồ Ồ 30 Ắ Ắ Ắ Ắ Ắ Ắ
Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề
R Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ



Reading

Ả Ỗ Ắ Ắ Ồ Ồ 30 Ắ Ắ Ắ Ắ Ắ Ắ
Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề Ề
R Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ Ầ



Speaking

4 parts
Questions/role-play/read aloud/presentation
~ 12 minutes

» LANGUAGECERT General - Listening

Part	No. of Items	Task Type	Task Description	Context
1	7	Multiple Choice (MC)	Two speakers; 7 incomplete exchanges	A variety of contexts which would be encountered in an English-speaking environment, ranging from informal to formal. Unfinished dialogues.
2	10 (5x2 items)	MC	Two speakers; 5 short conversations	Exchanges in formal and informal settings, for example an interaction between two colleagues. Two speakers.
3	7	Note-taking	Monologue; identify information to complete notes	An informational talk on a topical or work-based theme. Monologue.
4	6	MC	Group discussion (max. 3 speakers); follow-up multiple-choice questions	A discussion on a topical issue, e.g. a discussion at a conference, discussion at a meeting, a web podcast.

» LANGUAGECERT General - Reading

Part	No. of Items	Task Type	Task Description	Context
1a	6	MC	replace the highlighted word in the text with no change in meaning.	6 sentences with one word to replace and choice of 4 options.
1b	5	MC	select the correct word or phrase from a choice of three to fill each gap	Authentic source, minimal editing, range of genres.
2	6	Gapped text	select the correct sentences to complete the text	A text from which sentences have been removed. Text has a recognisable line of development which can be narrative, discursive or chronological. The text has strong cohesion and is 'standalone' with a defined beginning, middle and end.
3	7	4 texts, multiple-matching	4 texts; read the texts to match the text to the questions.	Four texts that would be seen in an academic context. Texts may have different types of source/genres.
4	6	MC	long text; answer questions to show understanding of a complex text and complex argumentation.	A long text with academic argumentation. A range of genres can be used, including article, extract from a book or report etc which presents argumentation and opinion.

» LANGUAGECERT General - Writing

Part	Task Description	Context
1	Text input; cover three required pieces of information. Emphasis on clarity of message and communicative effect.	Express opinion and feelings, describe activities, give justification, give advantages and disadvantages, giving reasons, describe purpose/cause/result, present a solution, making comparisons, needs and wants, obligation, ability and inability etc.
2	Informal writing describing experience, ideas on a topic, future plans or explaining feelings.	Describing past events and experience, future plans, hopes, wishes and regrets, likes/dislikes/preferences, narratives, describing people and events, stating opinions etc.

» LANGUAGECERT General - Writing

Assessment Criteria

Criteria	Description
Task Fulfilment	A measure of the extent to which the candidate has demonstrated the ability to manage the tasks adequately for the level.
Accuracy and Range of Grammar	A measure of the ability to vary and demonstrate control of grammatical structures as appropriate to the task.
Accuracy and Range of Vocabulary	A measure of the ability to vary and demonstrate control of lexis and register as appropriate to the task.
Text Organisation	A measure of how coherently ideas are linked together in the text and how accurately punctuation is used.

» LANGUAGECERT General - Speaking

Assessment Criteria

Part	Task Type	Approx time	Task Description
1	Warm up and interview	2-3 mins	Candidates give and spell their name and give their country of origin. They then answer up to five questions from different topic areas.
2	Role plays	2 mins	Two real-life situations are presented by the interlocutor. Candidates respond and initiate in each situation, each of which lasts approximately four turns.
3	Reading aloud and answering questions	3-4 mins	Candidates read aloud a short text and answer questions about what they have read.
4	Presentation	4 mins	Candidates prepare and deliver a short presentation in response to a stimulus and answer questions about it.

» LANGUAGECERT General - Speaking

Criteria	Description
Task Fulfilment and Communicative Effect	A measure of the ability to manage the tasks adequately for the level and link the utterances into coherent speech.
Coherence	A measure of the ability to provide coherent responses, particularly over extended speech, and the linking of ideas and contributions.
Accuracy and Range of Grammar	A measure of the ability to vary and demonstrate control of grammatical structures as appropriate to the task.
Accuracy and Range of Vocabulary	A measure of the ability to vary and demonstrate control of lexis and register as appropriate to the task.

» LANGUAGECERT General

Feedback on performance

- › All candidates receive feedback on their performance in the Writing and Speaking tests, on the basis of the marking criteria:

Writing

- › Task Fulfilment
- › Accuracy and Range of Grammar
- › Accuracy and Range of Vocabulary
- › Organisation and Coherence

Speaking

- › Task Fulfilment and Communicative Effect
- › Coherence
- › Accuracy and Range of Grammar
- › Accuracy and Range of Vocabulary
- › Pronunciation, Intonation and Fluency

» LANGUAGECERT General

Delivery Methods

› **Secure Test Centre Network**

All test centres have controlled access and are under live video and audio surveillance (and direct test invigilation), with recordings securely stored for audit and investigation purposes. Test centre security check-in procedures include the use of hand-held metal detectors to scan candidates for mobile phones, cameras, or other prohibited devices.

» LANGUAGECERT General

Delivery Methods

› Online Proctoring

Security features include:

- › Synchronous, human invigilation
- › Data encryption
- › KYC software
- › Database matching
- › Device lockdown software
- › Unauthorised hardware detection software
- › Option shuffling
- › Second Camera
- › Data analysis
- › Similarity detection software
- › Machine key log & Geolocation
- › Video recording

» Mapping LANGUAGECERT General to the CLB

Aim:

- › To conduct a standard-setting investigation into how the Canadian Language Benchmarks (CLB) correspond with LANGUAGECERT General (LCG) and the LanguageCert Global Scale

Previous work has

- › validated the LANGUAGECERT Global Scale
- › mapped LCG to the CEFR
- › as well as the correspondence between the CLB and the CEFR,

but a direct link between LCG and CLB has yet to be established.

» The LANGUAGECERT Global Scale

- › ‘The utility of the scaled score comes from allowing for meaningful score interpretations and, at the same time, minimizing misinterpretations and inappropriate inferences.’ (Kolen & Brennan, 2004).
- › For the recently introduced LANGUAGECERT Academic and General tests, results are reported on the LANGUAGECERT Global Scale. The Global Scale score (which is provided by language skill and overall result) gives finer gradations of performance within the CEFR levels but is also a standalone measure that can be aligned with any relevant external scale.

The LANGUAGECERT Global Scale

Design and Development

CEFR level	LID scale range
C2	151-170
C1	131-150
B2	111-130
B1	91-110
A2	71-90
A1	50-70

- › Based on the LANGUAGECERT Item Difficulty scale (0-200, left)
- › Created for item-banking and test construction purposes
- › Common across all products to ensure comparable difficulty

» The LANGUAGECERT Global Scale

Internal Validation

Calibrated through four studies (2019-2022):

1. Calibration (Rasch) of four multi-level linear (PB) tests (282 items, 3,000 cand.s., 2020)
2. Calibration (Rasch) of multi-level, adaptive item bank (827 items, 6,000 cand.s. 2021)
3. Full adaptive dataset (827 items) simulated to 50,000 cand.s., 2022)
4. Final calibration of dataset (827 items, 48,000 cand.s., 2022)

» Ecctis CEFR Referencing – LCA & LCG

- › CEFR referencing involves comparing and analysing the level of a language qualification against the Common European Framework of Reference for Languages (CEFR), which describes language proficiency on a scale from A1 (beginner) to C2 (proficient).
- › *“The results of Ecctis’ two-staged independent review demonstrated alignment between LANGUAGECERT’s reported CEFR levels for both tests and Ecctis’ findings. In addition, Ecctis’ findings corroborated the cut scores per skill reported for both tests on the LANGUAGECERT Global Scale.”*

Test name	CEFR Level(s)	LanguageCert Global Scale Cut-Scores per skill			
		Reading	Writing	Speaking	Listening
LanguageCert General	A2	20	20	20	20
	B1	40	40	40	40
	B2	60	60	60	60
	C1	75	75	75	75

» Mapping LANGUAGECERT General to the CLB

- › As part of this plan, LANGUAGECERT conducted a study to map LANGUAGECERT General and the LANGUAGECERT Global Scale against the Canadian Language Benchmarks (CLB). The aim of the study was to determine the alignment and cut-off points between the Canadian Language Benchmarks (CLB) and LANGUAGECERT General scores on the Global Scale for CLB levels 4 through 10 across the four skill areas: Listening, Reading, Writing, and Speaking, and a corresponding overall score.

» Mapping LANGUAGECERT General to the CLB

Project consists of following analyses:

- › Establishing cut scores for CLB levels 4-10
- › Exploring the validity of the panel's judgments and the extent to which results, with measurement error, can be considered a definitive standard of reference

» Mapping LANGUAGECERT General to the CLB

Study components/stages include:

- › Familiarisation: Training Activities to ensure participants' detailed knowledge of the CLB;
- › Specification: Self-audit of coverage of the examination (content, task types, etc.) and completion of relevant forms by panel participants;
- › Standardisation Training and Benchmarking: Benchmarking performance samples and tasks/items;
- › Analysis of Panel input and corroboration of Test Design and Specification claims

» Mapping LANGUAGECERT General to the CLB

- › CRELLA, a research institute on language assessment and validation at the University of Bedfordshire, experienced in conducting standard setting initiatives for high-stakes contexts was commissioned to undertake the study.

» Mapping Concordance Studies Panel



Prof. Anthony Green
Professor of Language Assessment
Centre for Research in English Language
Learning and Assessment (CRELLA)
Director, University of Bedfordshire, UK



Prof. Jason Fan
Language Testing Research
Centre Deputy Director,
University of Melbourne,
Australia



Prof. Sara Cushing
Professor of Applied Linguistics
Georgia State University, USA



Prof. Jim Tognolini
Centre for Educational Measurement
and Assessment Director, University of
Sydney, Australia



Prof. Liying Cheng
Emeritus Professor of Language
Education and Assessment,
Assessment and Evaluation Group
(AEG) Director, Queen's University,
Canada



Dr Johnathan Jones
Lecturer of Language
Learning & Assessment
CRELLA, UK



Dr Yiannis Papargyris
Assessment Development
Director
LANGUAGECER



Leda Lampropoulou
Head of Assessment
LANGUAGECER

» Mapping LANGUAGECERT General to the CLB

Following best practices in standard setting (Aera et al., 2014; Figueras et al., 2009), a series of analyses were conducted to establish cut scores and explore the validity or trustworthiness of the panel's judgments, focusing on

- › the consistency of rater judgments,
- › the accuracy of rater judgments, namely whether the allocation of items and examinee scores to a given CLB level is appropriate and
- › the extent to which aggregated results, with measurement error, can be considered a definitive standard of reference.

» Mapping LANGUAGECERT General to the CLB

- › The connection between LCG and the CLB was explored in two phases, each examining Speaking, Listening, Reading, and Writing.
- › The first phase was an exploratory range-finding phase which identified approximate areas for cut scores with relatively few productions (Speaking and Writing) and items (Reading and Listening).
- › The second phase incorporated data from additional test forms to more accurately and reliably pin-point precise divisions between proficiency levels.

» Mapping LANGUAGECERT General to the CLB

- › English language experts were recruited to help map LANGUAGECERT General items and performances to CLB levels
- › Each panelist filled one of three roles: language instructor, scholar, or director.
- › Panelists had multiple years of experience working with the CLB.

» Mapping phase: Listening and Speaking Panel

Table 1. Listening and Speaking Panel Constitution

Descriptor	Group	Group size	%
Gender	Female	10	59
	Male	7	41
Role	LINC instructor	11	65
	Scholar	4	24
	Programme director	2	12
Region	British Columbia	4	24
	Alberta	2	12
	Saskatchewan	2	12
	Ontario	8	47
	Nova Scotia	1	6
Institution	University	8	47
	College	2	12
	LINC	5	29
	Public school	1	6
	Private language institute	1	6

- › Years of experience in current role: M = 14
- › Years of experience in ESL: M = 20

» Mapping phase: Listening and Speaking Panel

Table 2. Reading and Writing Panel Constitution

Descriptor	Group	Group size	%
Gender	Female	14	70
	Male	6	30
Role	LINC instructor	9	45
	Scholar	8	40
	Programme director	3	15
Region	British Columbia	5	25
	Alberta	3	15
	Saskatchewan	2	10
	Ontario	8	40
	Quebec	1	5
	Nova Scotia	1	5
Institution	University	9	45
	College	3	15
	LINC	6	30
	Public school	1	5
	Private language institute	1	5

- › Years of experience in current role: M = 14
- › Years of experience in ESL: M = 20

» Mapping LANGUAGECERT General to the CLB

- › Decision consistency is the degree to which decisions about test takers are consistent across different administrations of the test (Cizek et al., 2004). Decision accuracy refers to the extent to which the decisions made based on a test accurately reflect the true performance or ability of the test-takers (Lathrop, 2015).

DA/DC				
CLB	Speaking	Writing	Reading	Listening
4	0.99 / 0.98	0.97 / 0.96	0.99 / 0.98	0.98 / 0.97
5	0.97 / 0.96	0.96 / 0.94	0.96 / 0.94	0.95 / 0.93
6	0.88 / 0.93	0.93 / 0.91	0.93 / 0.90	0.94 / 0.91
7	0.93 / 0.90	0.92 / 0.88	0.93 / 0.90	0.93 / 0.91
8	0.90 / 0.87	0.91 / 0.88	0.92 / 0.90	0.92 / 0.89
9	0.95 / 0.93	0.94 / 0.93	0.94 / 0.92	0.94 / 0.91
10	0.98 / 0.98	0.99 / 0.99	0.95 / 0.93	0.96 / 0.94

» Mapping LANGUAGECERT General to the CLB

- › Cut scores for LCG, established by the standard-setting panels, define the performance standards required for a test taker to be considered to satisfy a given level of English language proficiency as represented by the CLB.

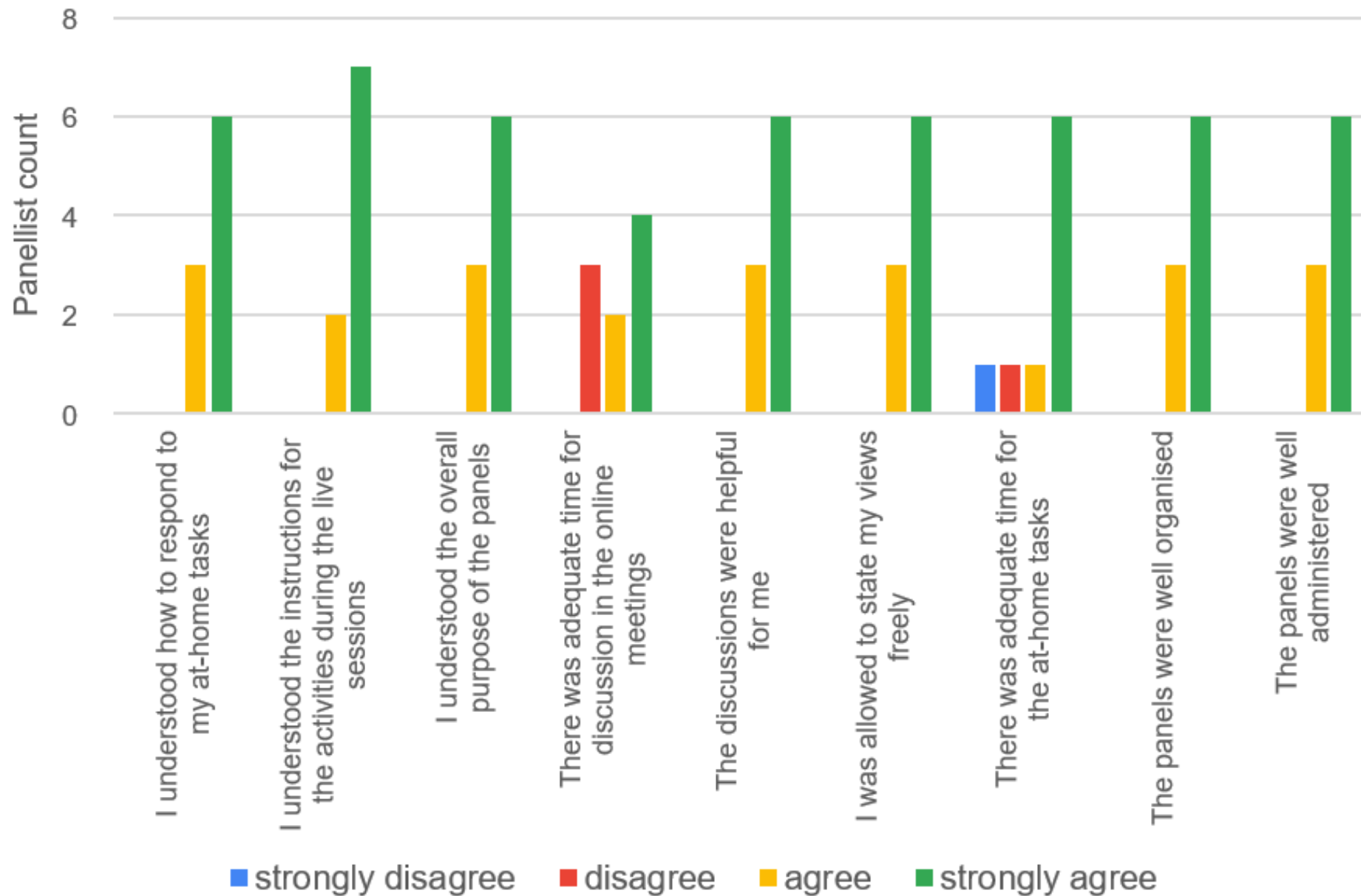
CLB	Writing	Speaking	Reading	Listening	Overall
4	35	31	29	27	31
5	44	40	40	39	41
6	54	49	51	53	52
7	63	58	60	66	62
8	72	67	71	77	72
9	81	76	77	82	79
10	90	85	83	87	86

» Mapping LANGUAGECERT General to the CLB

- › The conditional standard error of measurement CSEM represents the precision of test scores at a specified level reflecting how much error is associated with the estimate.

CLB	Speaking	Writing	Reading	Listening
4	0.51	0.39	0.46	0.44
5	0.49	0.42	0.43	0.42
6	0.52	0.44	0.42	0.42
7	0.56	0.51	0.75	0.45
8	0.61	0.53	0.47	0.53
9	0.65	0.60	0.53	0.64
10	1.00	1.60	0.64	1.04

Questionnaire Summary



» Research links



1. LANGUAGECERT Research & Validation
website: List of Publications (incl. full report)



2. Survey on the Use of
Concordance Studies

3. yiannis.papargyris@peoplecert.org