



CANADIAN ENGLISH LANGUAGE  
BENCHMARK ASSESSMENT FOR NURSES

# FACTS & FIGURES

ISSUE 6



Prepared by:

**TOUGHSTONE**  
INSTITUTE  
COMPETENCY EVALUATION EXPERTS

# CELBAN

# WRITING TEST RENEWAL



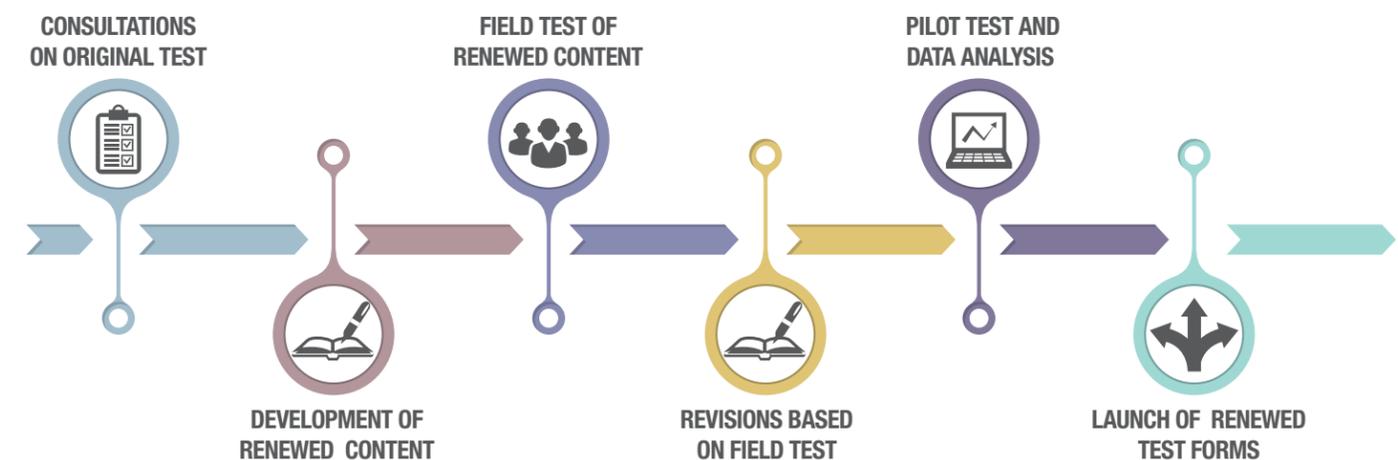
CELBAN is built on communication tasks essential to nursing that represent authentic language use, and it is one of two tests accepted by nursing licensing bodies across Canada as evidence of English language proficiency for licensure. CELBAN includes a separate test component for each of the four language skills (listening, speaking, reading and writing). The assessment is task-based and embeds the use of language within a profession-specific framework with a focus on contexts that are fundamental to nursing practice. All test tasks have been designed to be authentic and engaging to test takers, to facilitate demonstration of communicative competence.

A CELBAN test renewal project was launched in 2014, when the CELBAN Centre was first established at Touchstone Institute. As the original CELBAN test forms had been in use for more than a decade, the purpose of the renewal project was to develop additional forms of the test that would retain the salient features of the original model and introduce some new task types and fresh content.

The renewal process for the CELBAN Writing Test involved review by the test renewal lead, consultations with subject-matter experts, development of renewed tasks and scoring procedures, field testing, revisions, pilot testing, data analysis, construction of renewed test forms and a launch of the renewed procedures.

## Background

The Canadian English Language Benchmark Assessment for Nurses (CELBAN) is an occupation-specific language test that offers Internationally Educated Nurses (IENs) the opportunity to demonstrate their proficiency within a real-world context that is familiar and relevant. Touchstone Institute is the national CELBAN test administrator, overseeing its delivery at test sites across the country and managing a quality assurance framework that involves test maintenance and development. CELBAN test results are referenced to the Canadian Language Benchmarks (CLB), a descriptive scale of communicative ability in English as a Second Language (ESL). The test proprietor, the Centre for Canadian Language Benchmarks (CCLB), has oversight over CELBAN.



This issue of *CELBAN Facts & Figures* presents a summary of the work that was carried out to produce renewed forms of the CELBAN Writing Test. Information on the renewal of other CELBAN skill tests has been published in previous issues of *Facts & Figures*.

# WRITING TEST RENEWAL PROCESS

## Consultations

The renewal process began with a review of the current Writing Test by a test development specialist. Responses to each task were examined, and consideration was given to the range of writing abilities represented across the test. Following this review, an approach was designed for consultation with a small group of nurses serving as subject-matter experts (SMEs). The SMEs were brought together in a facilitated workshop, during which they were invited to share their feedback on the authenticity of the task genres and the currency and relevance of the writing topics.

The original CELBAN Writing Test consisted of two tasks, the first of which was related to a listening passage. To complete that task, examinees recorded information in point form in a chart format while viewing a video. While this was an authentic activity reflecting many real-life situations in which nurses need to take notes while listening, the heavy reliance on a skill other than writing introduced challenges for accurately evaluating writing performance.

Accordingly, the main change that was made to the test model was to eliminate the task that relied on listening and to introduce another task of a different genre. This change provided examinees with a fuller opportunity to demonstrate the extent of their writing ability.



## Development of Tasks and Scoring Procedures

Content development began with the identification of a range of relevant genres, task types and topics to be tried out in a small field test. Instructions and prompts were prepared for six different task types, and booklets were compiled so that field test examinees could attempt a combination of tasks. Following the field test administration, a group of trained examiners evaluated and scored each examinee's performance on each task.

When evaluating the field test writing responses, examiners worked with draft scoring grids that retained key criteria established in the original CELBAN Writing Test along with some additional CLB-based indicators of successful performance on the new genres. During and after the scoring, examiners were invited to comment on the content and quality of the scoring grids, with the following questions guiding their consultation:

1. What are the strongest features of the draft scoring grids?
2. What are the limitations of the draft grids, if any?
3. Did you experience any challenges in working with the grids? Please specify.
4. Do you have suggestions for refining or improving the grids? Please specify.

Following the scoring, each examiner independently provided feedback on the field test task types, indicating which genres and topics were most effective in eliciting the target writing skills and strategies. Based on this feedback, the two most appropriate task types were selected for the renewed test model.

The model was informed by the following considerations:

- Like the original CELBAN, the renewed test model would consist of two tasks.
- One task would be prompted by a chart that provides standardized information.
- The other task would have an open-ended prompt to elicit more original discourse.

With the renewed model in place, the team moved on to the process of developing topics and prompts for four pilot versions of the test. A group of nurse experts and examiners worked with the renewal development lead to generate and refine the new content. Prompts were reviewed for parallelism and standardization. Instructions were carefully worded to stress the importance of original content, complete sentences and connected discourse.

## Pilot Testing

Four pilot forms (A, B, C, D) were administered in a design that called for 160 examinees in four order cohorts. Each cohort was to include the following:

- 10 low intermediate examinees
- 10 intermediate examinees
- 10 high intermediate examinees
- 10 advanced examinees

The order cohorts were intended to ensure that data analysis would not be affected by a task's placement within the timing of the administration. Tasks encountered early in a pilot session may be influenced by nervousness or distraction, while those encountered at a later point can be affected by fatigue or restlessness. The order of administration for the four cohorts was AB, BC, CD and DA. This counterbalanced pilot testing approach ensured that each pilot form appeared in the initial and the final position.

## Refinements to Scoring Procedures

Before the pilot scoring began, the field test scoring grids were revised based on examiner feedback from the field test scoring. Then, the renewal development lead analyzed a sample of responses written by the pilot examinees and further shaped the scoring criteria to accurately describe the discourse that writers produced when responding to each of the pilot tasks. This was followed by a comparison of the scoring grid criteria to descriptors in the CLB document, and adjustments to reflect target levels were made. The revised grids were prepared for workshops with trained examiners.

In the first workshop, examiners were oriented to the renewed tasks and trained on how to score responses using the criteria in the revised grids. A selection of pilot test papers was then used as the basis for a group scoring exercise. The examiners rated the tasks and provided general feedback on the grids. The examiners' comments were taken back by the renewal development lead to inform further refinements to the grids and scoring procedures.



The second workshop served as preparation and training for scoring the pilot test responses. The final refined scoring grids were introduced, and examiners worked in one large group to score a selection of papers, discussing their ratings to develop a common understanding of how to apply the criteria in the grids. When the group was sufficiently calibrated and confident, the examiners worked in pairs to practice rating, and finally they began scoring pilot test papers independently.

The renewal development lead kept track of the agreement among independent ratings. At the end of this workshop it was determined that the grids were functioning effectively and that the examiners were ready to move on to the pilot scoring, which required every examiner to independently score the same sample of test papers. Analysis of data from this scoring exercise confirmed that the scoring procedures were sufficiently reliable, and a double-rating system was put in place to further ensure reliability in the early stages of implementation.



## RESULTS OF THE WRITING TEST RENEWAL

As a result of the renewal process, there are several new parallel forms of the CELBAN Writing Test. These new forms and their accompanying scoring procedures have been designed and constructed based on a process of consultation, field testing, pilot testing, examiner feedback and scoring workshops. Each final form contains two tasks that reflect themes and topics relevant to the health-care field. The renewed test versions preserve the original CELBAN model, representing complementary tasks that emphasize different writing skills and strategies. The new forms are equivalent in difficulty.

The CELBAN renewal process has focussed on preserving and building on the excellent foundation that was established in the original CCLB test development project. The content of the new forms is current and relevant to nursing professionals, and the task genres address a range of writing ability at an appropriate level of difficulty. As always, specificity to the nursing domain has been accomplished by a team of developers working closely with nurse experts, so that CELBAN can continue to provide examinees with the opportunity to demonstrate their writing capabilities in an authentic, meaningful and relevant context.

Under the stewardship of Touchstone Institute, with input from language specialists and consulting nurses, the CELBAN content and procedures have been evolving. With the renewed test forms launched and in current use, research and development remains an ongoing element of quality control. As well as tracking examinee test results, the Touchstone team conducts regular psychometric analysis of CELBAN test content to ensure quality, integrity, and fairness. This practice will continue as one component of Touchstone Institute's commitment to ensuring the validity and reliability of each skill test.



Facts & Figures Reports are developed and published by The CELBAN Centre to support transparency in test research and development and to document our experiences and activities as administrators of CELBAN. The publication is available from The CELBAN Centre's website. For more information, contact [celban@tsin.ca](mailto:celban@tsin.ca).





CELBAN Facts & Figures is a biannual report series prepared by the Communication Competency Program at Touchstone Institute.

**TOUCHSTONE**  
INSTITUTE  
COMPETENCY EVALUATION EXPERTS

The CELBAN Centre is the national administrative centre for the Canadian English Language Benchmark Assessment for Nurses (CELBAN).

  
**CELBAN CENTRE**



Centre for  
Canadian Language  
Benchmarks

Centre des niveaux de  
compétence linguistique  
canadiens

CELBAN was developed by, and remains property of, the Centre for Canadian Language Benchmarks (CCLB).