

## Exámenes de aprovechamiento computarizados. Un paso más en la evolución académica del departamento de Español\*

*Eva Campos Gómez\*\**

*Aída Espinosa\*\*\**

*Martha Jurado Salidas\*\*\*\**

**Resumen:** En el primer curso de 2011 se implementó en el Centro de Enseñanza para Extranjeros (CEPE) la modalidad computarizada de una batería de siete exámenes (EC), que se administran al final de cada curso intensivo. Éstos se diseñaron con base en el perfil de egreso definido en siete de los ocho programas que conforman el programa de estudios del CEPE, el cual, a su vez, sigue los lineamientos propuestos en el Marco Común Europeo de Referencia para el Aprendizaje, la Enseñanza y la Evaluación de Lenguas (MCER), aparecido en su versión española en 2002. Esta innovación representa una etapa más en el proceso de actualización y estandarización de los cursos de español, que tuvo como punto de partida el año 2000, cuando se empezó a trabajar en un nuevo programa de estudios, el cual entró en vigor en 2007.

El objetivo de este artículo es describir el proceso de elaboración de los EC, el marco teórico-metodológico que los sustenta, así como los resultados preliminares (índices de dificultad de los reactivos) obtenidos en tres aplicaciones llevadas a cabo con 650 estudiantes en total.

**Palabras clave:** Lengua Española, evaluación, exámenes computarizados.

---

\*\* Maestra en Lingüística Aplicada por la UNAM. Ha desempeñado cargos académicos y administrativos en el CEPE y en la Escuela Permanente de Extensión en San Antonio (EPESA) desde 1988.

\*\*\* Maestra en Lingüística Aplicada por la UNAM. Profesora en los programas de español y de formación de profesores del CEPE, con especialidad en la enseñanza de la pronunciación.

\*\*\*\* Doctora en Lingüística por el Colegio de México. Profesora Titular del CEPE, donde ha impartido cursos desde 1982.

---

\* Las autoras aquí mencionadas fueron comisionadas para elaborar este artículo, el cual describe el trabajo de un equipo conformado también por las profesoras: Patricia Martínez y María de la Luz Munguía.

**Abstract:** A series of seven computer-based exams (CES) were applied at the end of the first intensive course of 2011 at CEPE. These tests were designed based on the exit profile defined in seven of eight programs that form the course curriculum.

The *Common European Framework of References for Languages* (CEFR) was used as basis for the program of studies.

These tests continue with a series of initiatives that seek to standardize and update the courses taught at CEPE, beginning in 2000 when a new program of studies began its drafting and was launched in 2007.

This article describes the building process of the CES, their theoretical framework, methodology and preliminary results (index of difficulty for the test items) obtained from three exam applications given to 650 students.

**Keywords:** Spanish, language testing, Computer-based exams.

## Introducción

En el CEPE se elaboran y aplican tres procedimientos formales de evaluación: 1. exámenes de clasificación (colocación), 2. de aprovechamiento (verificación de objetivos) y 3. de grado de dominio.<sup>1</sup> En este artículo nos referiremos en particular a los de aprovechamiento, cuya versión computarizada fue diseñada a lo largo de 2010 y se implementó en el curso de invierno (enero-febrero) de 2011, año conmemorativo del 90 aniversario del Centro de Enseñanza para Extranjeros (CEPE).

Con la puesta en práctica de la modalidad computarizada de una batería de siete exámenes finales culmina una etapa más en el desarrollo académico del departamento de Español. La elaboración de dichos instrumentos significa otro paso en el proceso de actualización y adecuación a estándares internacionales que tuvo como punto de partida el año 2000, cuando se empezó a trabajar un nuevo programa de estudios que entró en vigor en el 2007 e implicó la realización de una serie de ocho libros acordes con la nueva estructura curricular.

Como se señaló, los exámenes fueron diseñados con base en el perfil de egreso definido en siete de los ocho programas que conforman el currículo<sup>2</sup> y éste, a

---

1 Los primeros sirven para distribuir (clasificar) a los estudiantes en diferentes cursos, de acuerdo con su nivel. Para este fin, el CEPE cuenta con una batería de siete exámenes que evalúan las habilidades de expresión oral, escrita y la competencia lingüística. Esta última está diseñada con base en 40 ítems o reactivos de opción múltiple.

Los exámenes de dominio tienen como objetivo evaluar el nivel de competencia lingüística del candidato, sin considerar sus estudios o capacitación previa. Asimismo, tienen un carácter prospectivo, ya que la información que arrojan no sólo expone lo que éste puede hacer con la lengua, sino que también predice su actuación en situaciones de uso real de la lengua meta. El CEPE cuenta con un sistema de evaluación de dominio que comprende un examen tipo score, denominado Examen de Posesión de la Lengua Española (EPLÉ) y tres de dominio alcanzado (básico, intermedio y superior), denominados Certificado de Español como Lengua Adicional (CELA). El EPLÉ sigue el modelo norteamericano y evalúa de manera directa tres habilidades lingüísticas (comprensión auditiva, de lectura, y expresión oral) e, indirectamente, la expresión escrita, a través de una prueba de estructuras de la lengua y vocabulario. Los CELA, por su parte, evalúan directamente las cuatro habilidades e incluyen, además, pruebas que miden específicamente el dominio gramatical.

Cabe hacer notar que también se aplican exámenes extraordinarios a los estudiantes que no cumplen con los requisitos institucionales para acreditar los cursos. Con ellos se evalúan las cuatro habilidades y el sistema de la lengua (gramática).

2 La metodología que se utiliza en el nivel Superior se basa en el enfoque de aprendizaje por proyectos. En éste, no se aplica un examen final, únicamente se lleva a cabo una evaluación de carácter formativo: “La evaluación formativa, que se suele oponer a la sumativa, está muy relacionada con la evaluación continua, ya que se aplica a lo largo de los distintos estudios y momentos del proceso de enseñanza-aprendizaje, valiéndose de una variedad de procedimientos para recabar la información pertinente respecto a los estudiantes” (Bordón, 2006: 285).

su vez, corresponde, en general, a las pautas propuestas en el *Marco Común Europeo de Referencia para el Aprendizaje, la Enseñanza y la Evaluación de Lenguas* (MCER), publicado en su versión española en 2002.

Al igual que los dos principales productos de las etapas precedentes —el programa de estudios y la serie de libros de texto—, los exámenes computarizados (EC) tienen como sello distintivo el haber sido resultado de un trabajo colegiado, llevado a cabo por integrantes del personal académico del departamento de Español.

El objetivo de este artículo es describir el proceso de elaboración de los EC, el marco teórico-metodológico que los sustenta, así como los resultados preliminares (índices de dificultad de los reactivos) obtenidos en tres aplicaciones (invierno y primavera 1 y 2) llevadas a cabo con 650 estudiantes en total. A modo de conclusión se valorarán de manera general los logros y los pendientes.

## ANTECEDENTES

Desde la década de los ochenta un poco antes de la implantación del programa de estudios que tuvo como base la serie de textos *Pido la palabra*, se instituyeron en el CEPE los exámenes departamentales. Esta medida significó un primer paso en el camino de la estandarización de criterios de evaluación. Los exámenes eran elaborados —la mayoría de las veces de manera individual— por profesores con una incipiente formación en el campo de la evaluación. Pese a que hubo varios intentos por conjuntar un equipo de trabajo que se dedicara sistemáticamente a la realización de exámenes,<sup>3</sup> dicho objetivo no se logró a cabalidad (principalmente porque la evaluación es una disciplina relativamente nueva en el campo de la enseñanza de lenguas y más reciente aún en el de la lengua española), razón por la cual la renovación y actualización de la batería de exámenes dependía de la iniciativa individual de los maestros que se ofrecían a elaborar alguno nuevo.

3 Por ejemplo, en 1990, 1995 y 2007 se abrieron a concurso plazas destinadas específicamente a la contratación de personal dedicado a la evaluación (aunque en los hechos no se dieron las condiciones para que dicho personal cumpliera esa función). En 1996, se promovió un grupo de interés sobre el tema; en el año 2000, el *Centro Nacional de Evaluación* (CENEVAL) impartió a los profesores del CEPE un “Taller de elaboración de reactivos”. En el Primer Simposio del CEPE octubre de 2001, el invitado a la conferencia magistral ofreció un curso sobre “La evaluación formativa”; en agosto de 2007, maestros del CEPE congregados en torno al proyecto del Certificado Internacional de Español (CIE), organizaron el “Primer Curso Taller de Evaluación”, al año siguiente se llevó a cabo el segundo, de donde surgieron los criterios y ponderaciones que estuvieron vigentes hasta diciembre de 2010.

Dos circunstancias se conjuntaron para propiciar un cambio en el panorama antes descrito. En primer término, el hecho de que actualmente se cuenta con un programa de estudios estandarizado y apegado a los lineamientos internacionales establecidos en un documento de reconocido prestigio internacional, el Marco Común Europeo de Referencia para el Aprendizaje, la Enseñanza y la Evaluación de Lenguas (MCER). En segundo lugar, a raíz de la firma de un convenio interinstitucional con el Instituto Cervantes, se produjo un acercamiento a los estándares internacionales en materia de evaluación, emanados del Consejo de Europa.

La conjunción principalmente de estas dos circunstancias (aunque hubo otras coyunturales y algunas otras de índole laboral) favoreció en el año 2010 la conformación de equipos de trabajo que se dieron a la tarea de elaborar los modelos de examen, retomando la experiencia acumulada tanto en un proyecto previo con el Instituto Cervantes, como en dos Cursos Taller de Evaluación que fueron impartidos en 2007 y 2008. Finalmente, hay que apuntar que, en relación con el diseño del sistema aplicativo de los exámenes, la evolución en esta materia por parte del personal del departamento de Tecnología Educativa jugó un papel importante.

La introducción de la nueva modalidad supuso modificaciones de distinta índole respecto al sistema anterior. En el terreno metodológico, por una parte, la definición del contenido de los EC está basada en tablas de especificaciones construidas con base en los perfiles de egreso plasmados en el programa de estudios, los cuales, a su vez, derivan de los niveles de referencia propuestos en el MCER. Por otro lado, la calidad metodológica de los instrumentos de evaluación se vio fortalecida por el trabajo colegiado que subyace tanto en su diseño, como en la aplicación de las pruebas productivas (expresiones escrita y oral). En relación con este último punto, cabe señalar que a diferencia de lo que ocurría en el antiguo sistema en el que un solo profesor evaluaba las cuatro habilidades<sup>4</sup> y el sistema de la lengua (gramática), el protocolo de aplicación de los EC establece que el maestro titular del curso aplica la prueba de expresión escrita, pero su cotejo o verificación<sup>5</sup> se lleva a cabo mediante un intercambio entre pares. En

4 En el programa de estudios del CEPE, así como en los documentos derivados de él se utiliza el término tradicional habilidades, el cual no coincide con el utilizado en el MCER (actividades comunicativas); sin embargo, ambos designan el mismo referente. En los EC se evalúan las habilidades de comprensión auditiva y de lectura, expresión escrita y expresión oral; mientras que en el MCER, se habla, respectivamente, de producción escrita y producción hablada), además de interacción oral. En el programa de estudios del CEPE no se abordan de manera específica las actividades de mediación.

5 En la literatura especializada sobre evaluación se distingue entre *marking* y *grading*. En español no existe tal distinción; sin embargo, el primero podría traducirse como cotejo o corrección, y el segundo como calificación o puntuación. En este artículo se optó por los términos "cotejar" vs. "calificar". En *The Multilingual Glossary of Language Testing Terms* aparecen las siguientes definiciones: grade: A test score may be reported to the candidate as a grade, for example on a scale of A to E,

cuanto a la prueba de expresión oral, es administrada aleatoriamente a estudiantes de distintos grupos y niveles, por un profesor que entrevista y otro que coteja o verifica; ambos deciden por consenso la calificación final. Finalmente, cabe señalar que el proceso de cotejo de los dos tipos de pruebas subjetivas se lleva a cabo mediante escalas o rúbricas que evalúan distintos componentes de la competencia comunicativa (subcompetencia lingüística, sociolingüística, pragmático discursiva y estratégica), concepto que constituye el fundamento teórico-metodológico del programa de estudios.

En el terreno práctico, con la aplicación de un solo examen final, se incrementaron las horas de clase de los cursos, ya que se suprimieron las destinadas a los exámenes parciales. Este reacomodo se produjo gracias a que los nuevos instrumentos miden un perfil de egreso concebido como la puesta en práctica de un conjunto de habilidades que tienen como sustento la lengua más que los conocimientos —principalmente de índole gramatical—, apegados a la secuencia de algún manual de enseñanza.

En el terreno administrativo, la vinculación de los sistemas computarizados de aplicación, cotejo y expedición de calificaciones minimiza los errores en la verificación de las pruebas objetivas (comprensión auditiva, comprensión de lectura y sistema de la lengua), así como en el cálculo de la calificación final, además de incrementar la eficiencia de los procesos administrativos.

## METODOLOGÍA DE CONSTRUCCIÓN DE LOS INSTRUMENTOS

En la bibliografía sobre evaluación se establece que un examen es bueno cuando es válido y es válido cuando mide adecuadamente lo que intenta medir. Atendiendo a este prurito, la confección de buenos exámenes comprende varias etapas. Con base en Bordón (2006) y en relación, específicamente, con la construcción de los EC, describiremos a continuación el desarrollo de las fases previa, de diseño, de experimentación y de recapitulación.

---

where A is the highest grade available, B is a good pass, C a pass and D and E are failing grades.  
grading: The process of converting test scores or marks into grades.

marking: Assigning a mark to a candidate's responses to a test. This may involve professional judgement, or the application of a mark scheme which lists all acceptable responses.

mark: The outcome of an examination, often expressed as a percentage. Because of adjustments such as heavier weighting for some items, the mark is not always the same as the total score. (The multilingual Glossary of Language Testing Terms, 2011.)

### *Fase previa*

Según Bordón (2006), la fase previa consiste en la determinación del tipo de examen y su objetivo; la precisión del perfil de los candidatos; la selección de las tareas de examen en relación con el uso de la lengua; así como la estimación de los recursos necesarios.

Respecto a la selección del tipo de examen, los equipos de trabajo tuvieron claro desde un principio que se trataba de exámenes de aprovechamiento, ligados a procesos de enseñanza-aprendizaje, que se aplicarían al final de los cursos intensivos de seis semanas. Los EC tienen como finalidad medir el logro de objetivos definidos en los programas de estudio de cada nivel (así como en el perfil de egreso correspondiente). Asimismo, las calificaciones obtenidas sirven para expedir certificados de un programa conformado por 90 horas de instrucción (en el caso de los cuatro niveles básicos) y de 78 horas (en el de los tres niveles intermedios).

En cuanto al perfil de los candidatos, se cuenta con algunos datos recabados en los últimos diez años.<sup>6</sup> Según estos registros, el CEPE atiende mayoritariamente a estudiantes de origen asiático. Los japoneses son los más numerosos (22.6 por ciento en 2000 y 19.6 por ciento en 2010). En segundo lugar están los coreanos, cuyo porcentaje decreció sensiblemente en 2010 (4.6 por ciento), después de haberse mantenido estable por arriba del 20 por ciento a lo largo de la década. Los estudiantes chinos, por su parte, han triplicado su presencia en los últimos diez años (de un 3.65 por ciento en 2000 a un 10.9 por ciento en 2011). Las estadísticas muestran que, aunque ha habido una disminución de 13 puntos en el periodo de 2000 a 2011 (del 47.8 por ciento al 34.7), la comunidad asiática aún siguen constituyendo el grueso de nuestra población.

Los estudiantes norteamericanos representan un segundo grupo importante que se ha mantenido entre 11 y 13 por ciento. El resto son, principalmente, europeos; en su mayoría, alemanes.<sup>7</sup>

Respecto a las razones por las que se encuentran estudiando español en México, todavía no se ha hecho un estudio amplio; contamos sólo con el análisis

6 Estadística proporcionada por la Jefa de la oficina de Servicios Escolares el 27 de julio de 2011.

7 Cabe señalar que, a pesar de los factores adversos que aquejan a México, la matrícula del CEPE se ha incrementado en la última década (de 1 834 a 2 930 estudiantes por año). Independientemente de que aprender español esté de moda en muchos países, podríamos arriesgar la hipótesis de que este aumento está ligado al perfil predominante de estudiantes. El hecho de que se trate de jóvenes asiáticos que ven en su aprendizaje una inversión para su desarrollo personal (principalmente de índole económica) en un mundo globalizado y competitivo, nos conduce a esforzarnos en mejorar nuestros servicios —incluyendo el de evaluación— para continuar atrayendo a este tipo de público que demuestra anteponer sus intereses instrumentales a otro tipo de factores coyunturales.

de una encuesta aplicada en 2008 y con los comentarios que se recogen en el aula. La motivación mencionada por los estudiantes asiáticos es principalmente de tipo instrumental, ya sea “laboral” (su compañía u orden religiosa los manda porque de aquí serán enviados a otros países de Hispanoamérica), o de estudios (becas otorgadas por Conacyt y convenios con universidades como Ulsan y Takushoku). En el caso de los estadounidenses, las razones son de índole familiar (matrimonio o familiares mexicanos). Por lo que respecta a los del continente europeo, también se aducen causas de tipo familiar, como el noviazgo, matrimonio o acompañamiento de un cónyuge que ha sido asignado a nuestro país por motivos laborales o diplomáticos. Finalmente, muchos otros admiten que están en nuestro país por el simple gusto de aprender la lengua española.<sup>8</sup>

Por lo que respecta al nivel de escolaridad, se tienen datos de los japoneses que estuvieron inscritos de 2001 a 2007: el 70 por ciento de ellos contaba con estudios de nivel superior.

Según un análisis hecho en 2007, entre la población asiática y la estadounidense (498 estudiantes) la edad promedio oscila entre los 20 y 34 años (81 por ciento).

En suma, los EC están dirigidos a estudiantes mayoritariamente asiáticos, de entre 20 y 34 años, con un nivel de escolaridad universitario, que pasan en México un promedio de seis meses con el objetivo de perfeccionar su nivel de lengua, por razones eminentemente instrumentales (de estudio o laborales).

La selección de las tareas estuvo determinada, en primera instancia, por el perfil de egreso especificado en los programas (que a su vez se relaciona con el uso de la lengua, según los niveles de referencia del MCER); pero también, en gran medida, por el recurso tecnológico elegido para la aplicación de las pruebas objetivas y por el factor tiempo, en el caso de las subjetivas. Con base en estos elementos, se determinó que los exámenes estarían conformados por tres pruebas objetivas aplicadas en computadora (comprensión auditiva, comprensión de lectura y sistema de la lengua) y dos subjetivas (expresión escrita y expresión oral). Se trataría de una medición aislada de habilidades —es decir, éstas no estarían integradas entre sí—, aunque habría cierta relación temática entre los textos de comprensión de lectura y los ítems de la sección de sistema de la lengua.

<sup>8</sup> La encuesta se aplicó en julio de 2008 a 77 estudiantes. Los resultados fueron los siguientes: el 37 por ciento estudia español porque es necesario para su actividad; el 28 por ciento porque es importante comunicarse en español; el 14 por ciento por razones familiares; y el 17 por ciento porque vive en México. Según Bachman (1996), la factibilidad o viabilidad es una de las seis características que competen a la “utilidad” de un examen. Consiste en la relación entre los recursos requeridos para el diseño, desarrollo y uso del examen y los recursos de los que se dispondrá para estas actividades. Se define como la relación entre los recursos (humanos, académicos, estructurales y económicos) que se necesitan para el diseño, desarrollo y uso del examen y la disposición de los mismos, es decir, la factibilidad es la relación que existe entre los recursos necesarios y los recursos disponibles. Un examen es viable cuando sus exigencias, a nivel de recursos tanto humanos como técnicos o económicos, se pueden satisfacer.



Por razones de viabilidad o factibilidad,<sup>9</sup> se decidió que las pruebas de expresión escrita fueran aplicadas el último día de clases por los profesores titulares a sus respectivos grupos; sin embargo, para el cotejo o verificación de las mismas, se estableció un procedimiento que consiste en la recolección y repartición aleatoria de los escritos entre los maestros que imparten cada nivel. Dos de ellos cotejan los textos con base en escalas o rúbricas que evalúan distintos aspectos de las tareas escritas. De esta manera, cada escrito recibe una doble verificación (el puntaje final de cada uno de los aspectos de la escala se determina sacando el promedio de la calificación otorgada por ambos evaluadores).

Para las pruebas computarizadas (después de una primera aplicación del periodo de invierno, que sirvió como experimentación), se definió una logística que consiste en la distribución de un promedio de 16 estudiantes en cada uno de los tres espacios habilitados, en dos turnos diarios, durante tres días. En la aplicación del curso de invierno se asignaron dos horas para la resolución de los exámenes computarizados tanto básicos como intermedios. Se observó que para los básicos, hora y media era suficiente; mientras que para los intermedios se requería de dos horas.

Las pruebas de expresión oral significaron el mayor reto, en primer lugar, por la novedad que conllevaba para los profesores la evaluación de tareas de distinto nivel a estudiantes que no habían tomado el curso con ellos. Asimismo, para los estudiantes también fue un desafío el desempeñar las tareas orales frente a maestros distintos al propio.

La organización de un total de 100 parejas de estudiantes para ser evaluados cada media hora por nueve pares de profesores representó el otro reto importante planteado por la nueva modalidad de pruebas orales; principalmente, por la logística que implicó y por la dificultad de diseñar tareas en las que los dos participantes ejecutaran equilibradamente actos de habla y funciones comunicativas semejantes (en términos de complejidad del idioma y de tiempo de posesión del turno de palabra).

En cuanto a los recursos, aspecto relacionado también con el concepto de *factibilidad* o *viabilidad*, el desarrollo de los EC requirió los siguientes:

---

<sup>9</sup> Según Bachman (1996), la factibilidad o viabilidad es una de las seis características que competen a la "utilidad" de un examen. Consiste en la relación entre los recursos requeridos para el diseño, desarrollo y uso del examen y los recursos de los que se dispondrá para estas actividades. Se define como la relación entre los recursos (humanos, académicos, estructurales y económicos) que se necesitan para el diseño, desarrollo y uso del examen y la disposición de los mismos, es decir, la factibilidad es la relación que existe entre los recursos necesarios y los recursos disponibles. Un examen es viable cuando sus exigencias, a nivel de recursos tanto humanos como técnicos o económicos, se pueden satisfacer.

a) Humanos. Dos coordinadoras supervisaron el desarrollo de los exámenes de los bloques básico e intermedio —respectivamente— desde la definición de las especificaciones, hasta el análisis y la revisión de los resultados. Hubo también sendos equipos de redactores<sup>10</sup> que seleccionaron los textos y formularon colegiadamente los ítems. El grupo dedicado al bloque básico estuvo conformado por cuatro profesoras de asignatura que trabajaron durante seis semanas; en el grupo abocado a la elaboración de los exámenes del bloque intermedio participaron cinco profesoras de tiempo completo que trabajaron de manera consistente a lo largo de seis meses. Los profesores que imparten los cursos (algunos de ellos ya con experiencia en los exámenes de dominio del CEPE<sup>11</sup>) fueron los encargados de llevar a cabo las pruebas escritas y orales. El desarrollo tecnológico del sistema aplicativo de los EC corrió a cargo de un equipo conformado por un coordinador y dos programadores.

b) Materiales. El CEPE contaba con dos espacios equipados con computadoras, el primero con quince y el segundo con veinte. Para expandir la capacidad, se instalaron otros 15 equipos portátiles en una sala de usos múltiples inaugurada el año 2010.

#### *Fase de diseño*

La fase de diseño, consistió en la especificación de los contenidos (áreas temáticas, nociones y funciones pragmáticas, exponentes lingüísticos —gramática y vocabulario—) para cada uno de los siete niveles, así como en el establecimiento de los tipos de pruebas y su duración, la redacción de los ítems y, finalmente, la definición de criterios de evaluación.

El contenido de los exámenes —como ya se mencionó—se definió con base en los niveles comunes de referencia para medir el dominio de la lengua establecidos en el MCER así como en los programas por nivel y en el perfil de egreso<sup>12</sup> asen-

10 La primera versión de los cuatro exámenes básicos se debe al trabajo de las profesoras de asignatura Berta Berzunza, Lilia Castellanos, Claudia Cárdenas y María Elena Sánchez, quienes estuvieron coordinadas por la Mtra. Eva Campos. En cuanto a los intermedios, fueron desarrollados por los profesores de tiempo completo José Carlos Escobar, Aída Espinosa, Ma. de la Luz Munguía, Patricia Martínez y Eva Campos. En distintas etapas participaron también los maestros Ricardo Ancira, Rosa Esther Delgadillo, Rosario Gutiérrez, Ma. del Carmen Koleff y Emma Jiménez. Este equipo fue coordinado por Martha Jurado. El grupo de programadores, integrado por Maribel Carmona y Roberto Sierra, fue supervisado por Rodrigo Galindo.

11 Debido al factor tiempo, la capacitación que recibieron los profesores habilitados como correctores y como administradores de las pruebas subjetivas fue mínima, no obstante, se apeló y se confió en su experiencia; además, se tiene previsto subsanar esta carencia con un entrenamiento programado para año 2012.

12 Véase anexo 1.

tados en el programa general de estudios, versión 2010. La descripción que se hace en los perfiles de egreso establece las funciones comunicativas que el alumno tendrá que ejecutar, además del tipo y la extensión de los textos orales y escritos que deberá ser capaz de comprender y producir.

Para convertirlo en un contenido operativo susceptible de ser evaluado en un examen, el caudal de información proveniente de estas fuentes fue sistematizado en dos tipos de instrumentos: tablas de especificaciones,<sup>13</sup> para organizar los exponentes lingüísticos y el contenido de las pruebas receptivas (comprensión auditiva y de lectura); y escalas o rúbricas, para sistematizar la medición de las habilidades productivas (expresión escrita, oral e interacción).

En las tablas de especificaciones se determinan las operaciones (Hughes, 1989) o microdestrezas (Alderson, 1995) a evaluar en cada habilidad, así como el nivel cognoscitivo de las mismas y el número suficiente de ítems para medirlas adecuadamente. Las operaciones están enunciadas con verbos que denotan conductas observables organizadas en tres grandes categorías cognoscitivas: identificación/conocimiento, comprensión y aplicación.

El nivel de identificación/conocimiento es el más elemental, ya que implica el uso de la memoria para recordar o reconocer ideas. En él se da mayor énfasis a los procesos psicológicos del recuerdo y la relación. Para evaluar las operaciones incluidas en este nivel cognitivo se debe exponer la información y solicitar su identificación o reconocimiento. Así pues, se pide al candidato ubicar datos que aparecen de manera explícita y literal en el texto. Estas preguntas operan a nivel local del texto.<sup>14</sup>

El nivel de comprensión incumbe al proceso psicológico de formación de conceptos o grado en que se éstos han sido aprendidos, captados o incorporados en una estructura cognoscitiva previamente existente, tomando en cuenta su relación significativa con ideas previas. En el modelo de los EC, el nivel de comprensión incluye las operaciones de análisis y síntesis. En él se requiere recuperar, seleccionar o relacionar información que aparece de manera explícita en el texto (aunque pueda estar parafraseada). Apunta hacia aspectos locales o globales del escrito.<sup>15</sup>

13 Las tablas se trabajaron en una primera versión en el Curso Taller de Evaluación realizado en 2007 y se depuraron para elaborar los EC. Obsérvese un ejemplo en el anexo 2.

14 Los verbos utilizados en las tablas de especificaciones para evaluar este nivel son: reconocer, recordar, evocar, distinguir, seleccionar, enlistar, señalar, identificar, indicar, localizar, marcar, registrar, subrayar; apuntar, nombrar, enumerar, citar, contar, definir, describir, enunciar, exponer, mencionar, relatar, repetir, escribir.

15 Los verbos utilizados para medir este nivel son: categorizar, diferenciar, distinguir, analizar, asociar, relacionar, contrastar, comparar, generalizar, analizar, jerarquizar, descomponer, fraccionar, separar, ordenar, organizar, transformar, deducir, concluir, ilustrar, pronosticar, resumir, sintetizar, explicar, expresar, narrar, definir, describir, parafrasear, predecir, discutir, debatir, criticar, ejemplificar, traducir, transcribir.

El nivel de aplicación implica la utilización de abstracciones en situaciones particulares y concretas, así como la transferencia de conocimiento, es decir, capacidad de generalización.<sup>16</sup> En este nivel cognitivo es necesario evaluar o interpretar la información a partir tanto del contenido como del contexto. Incluye la capacidad de sintetizar y de hacer explícitos algunos efectos del enunciado que adquieren sentido en contextos particulares. Para responder a ítems de este nivel, el candidato debe recurrir a información tanto explícita como implícita, así como a su propia experiencia comunicativa.<sup>17</sup>

Los exponentes lingüísticos también se especificaron en tablas. Éstos corresponden a los contenidos gramaticales establecidos en los programas de estudio de los cuatro niveles básicos y los tres intermedios. A todos ellos se les adjudicó el nivel cognoscitivo tres, es decir, el nivel de aplicación, puesto que se consideró que el uso de la lengua implica la puesta en práctica de las reglas gramaticales que el aprendiz va adquiriendo y aprendiendo.

Para las habilidades receptivas (comprensión auditiva y de lectura), las operaciones especificadas en los niveles básicos se enfocan en la evaluación de la comprensión global del texto y en la localización de datos específicos. En los intermedios, se evalúan la discriminación entre ideas principales y secundarias; la identificación de información y opiniones presentes en el texto, así como el reconocimiento de elementos implícitos.<sup>18</sup>

En los instrumentos diseñados para sistematizar la evaluación de las pruebas productivas, es decir, en las escalas, rúbricas o baremos,<sup>19</sup> se desglosan los criterios para cotejar las características del discurso y los escritos de los estudiantes. Con las escalas para las pruebas orales se verifican la fluidez, el dominio del léxico y de la gramática, la adecuación fonética, así como distintos componentes de las subcompetencias lingüística, sociolingüística, pragmático-discursiva y estratégica. Asimismo, con las escalas para las pruebas escritas se cotejan, además del dominio del léxico y de la gramática, aspectos textuales como el contenido, la estructura, la coherencia y la cohesión; así como la organización de las ideas y el uso de los marcadores discursivos.

16 En el modelo de los EC, este nivel incluye la operación cognitiva de evaluación.

17 Los verbos utilizados para medir este nivel son: evaluar, contrastar, verificar, concluir, comparar resultados, estimar, juzgar, inferir, discernir, interpretar, aplicar, construir, crear, desarrollar, formular, componer, planear, demostrar, organizar, relacionar, emplear, combinar, sintetizar, encontrar, producir, representar, resolver, transferir, usar, utilizar, dramatizar, criticar, ejemplificar, explicar, describir, acordar, refutar.

18 En los exámenes de Intermedio 3, a manera de indicadores de nivel de logro, se incluyen uno o dos reactivos que evalúan operaciones, tales como la identificación de las alusiones, humor, metáforas o ironías, así como la deducción de las alusiones y referencias implícitas en los textos, las cuales propiamente pertenecen al grado superior.

19 Véase, anexo 3.

Por cuestiones de factibilidad o viabilidad, se determinó que, tanto la ponderación como el orden, fueran homogéneos para todos los niveles:

Pruebas	Porcentaje
Expresión escrita	15
Comprensión auditiva	15
Comprensión de lectura	15
Sistematización lingüística (gramática)	15
Expresión oral	20
Total	80

Cuadro 1. Orden y ponderación.

Se determinó una ponderación homogénea del 15 por ciento para tres de las habilidades y para la sistematización lingüística (gramática). Esta decisión se adoptó tomando en cuenta que el grado de desarrollo de las destrezas es diferenciado (al igual que en los hablantes nativos). Al asignar un mismo porcentaje, los estudiantes tienen la posibilidad de equilibrar su actuación; por ejemplo, si alguno de ellos tiene problemas en la audición, puede compensar una baja calificación en esta prueba desarrollando una buena tarea escrita y obtener así un mejor resultado global. A la producción oral (objetivo principal de un programa basado en el enfoque comunicativo) y a la evaluación formativa<sup>20</sup> se les asignó un porcentaje de 20 por ciento.

El número de ítems que conforman cada prueba se determinó con base en el juicio de los expertos creadores, el cual atendió a la importancia relativa y al dominio exigido para cada componente de la tabla de especificaciones. Aunque sufrió leves ajustes en las distintas secuelas, en general, el modelo quedó definido desde la primera versión:

<sup>20</sup> La evaluación formativa “se aplica a lo largo de los distintos estadios y momentos del proceso de enseñanza aprendizaje, valiéndose de una variedad de procedimientos para recabar la información pertinente respecto a los estudiantes” (Bordón, 2006: 285).

Nivel	Versión 1			Versión 2			Versión 3		
	CA*	CL	SL	CA	CL	SL	CA	CL	SL
Básico 1	13 (15 pts)	10	30	15	15	30			
Básico 2	15	15	40	15	18 (15 ptos.)	38 (30 ptos.)			
Básico 3	15	15	40	15	15	40			
Básico 4	15	15	40	15	15	40	15	15	45
Intermedio 1	29	20	30	29	20	30	15	20	30
Intermedio 2	23	20	25	24	20	25			
Intermedio 3	20	22	20	20	22	20	20	25	20

\* CA (Comprensión auditiva), CL (Comprensión de Lectura), SL(Sistema de la Lengua)

Cuadro 2. Número de ítems por prueba

### Fase de experimentación

La experimentación discriminada de ciertos ejercicios incluidos en los exámenes fue llevada a cabo por algunas profesoras, en sus clases de los cursos de otoño 2010. Los modelos completos, ya montados en el sistema, se experimentaron únicamente con el cuerpo académico del área de español, el cual los resolvió e hizo llegar sus observaciones a las coordinadoras.

Debemos admitir que la experimentación representa el punto débil en la construcción de los instrumentos, ya que nunca se probaron los modelos completos con un grupo similar al que estaba destinado el examen. Dos factores nos impidieron proceder de esa manera. En primer lugar, la dificultad que implica probar reactivos con una población que toma cursos intensivos de seis semanas y que paga por ellos. En estas condiciones no hay disponibilidad para que los estudiantes dediquen tiempo a resolver ítems. En segundo término, el factor tiempo fue decisivo. La confección en un año de una batería de siete exámenes y de un sistema computarizado *ad hoc* para aplicarlos representó un *tour* de force que casi superó nuestras posibilidades. Nos propusimos aplicar la batería completa en el primer periodo de 2011 y así lo hicimos; apostamos a la experiencia de los creadores de las pruebas, pero sobre todo, a la fortaleza que nos daba el haber trabajado de manera colegiada. Confiados en estos dos pilares, decidimos aplicar los exámenes sin la fase de la experimentación de los modelos completos con los estudiantes. Estábamos conscientes de los riesgos, pero también de la posibilidad de mejorar los instrumentos posteriormente.

Una vez definido (y probado) el modelo en las aplicaciones sucesivas de los cursos de Invierno y Primavera 1, durante el periodo de Primavera 2 (abril- mayo de 2011), se elaboraron y aplicaron segundas versiones. Así pues, se cambió el 100

por ciento del contenido de los exámenes básicos y un 30 por ciento (una de las pruebas) de los intermedios. Sobre los índices de validez de los reactivos y fiabilidad de las dos versiones de los instrumentos hablaremos en la siguiente sección.

### *Fase de recapitulación*

La fase de recapitulación comprende la comprobación de la validez de las pruebas y la fiabilidad de los resultados, con el fin de generar un efecto retroactivo o de rebote sobre la enseñanza. Esta etapa se encuentra todavía en proceso. No obstante, contamos hasta el momento con una primera sistematización de las calificaciones obtenidas por los estudiantes y con el análisis de los índices de dificultad de los exámenes aplicados a una muestra de 650 estudiantes que cursaron estudios durante los periodos de Invierno y Primavera 1 y 2.

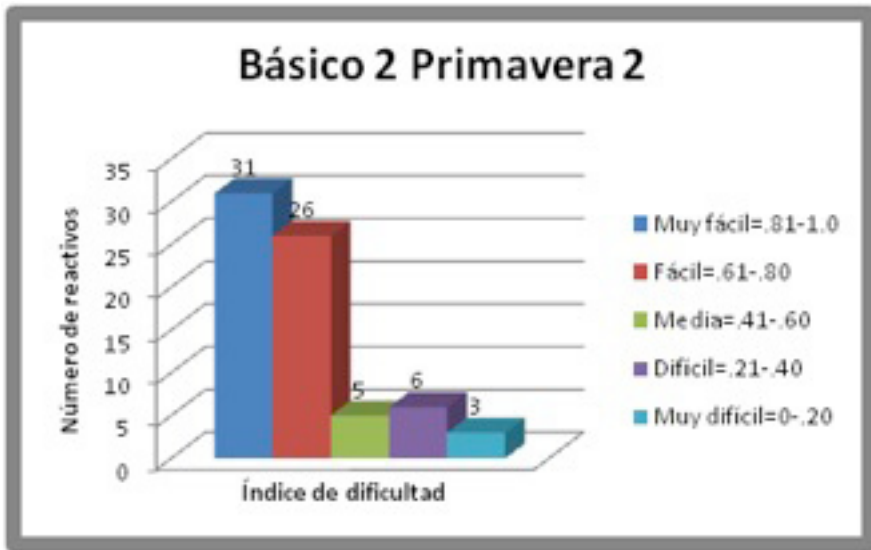
Nivel	Invierno	Primavera 1	Primavera 2	Total
Básico 1	35	47	30	112
Básico 2	37	41	42	120
Básico 3	42	35	32	109
Básico 4	30	50	31	111
Intermedio 1	36	28	40	104
Intermedio 2	14	18	18	50
Intermedio 3	7	15	22	44
Total				650

Cuadro 3. Número de estudiantes que han presentado los EC.

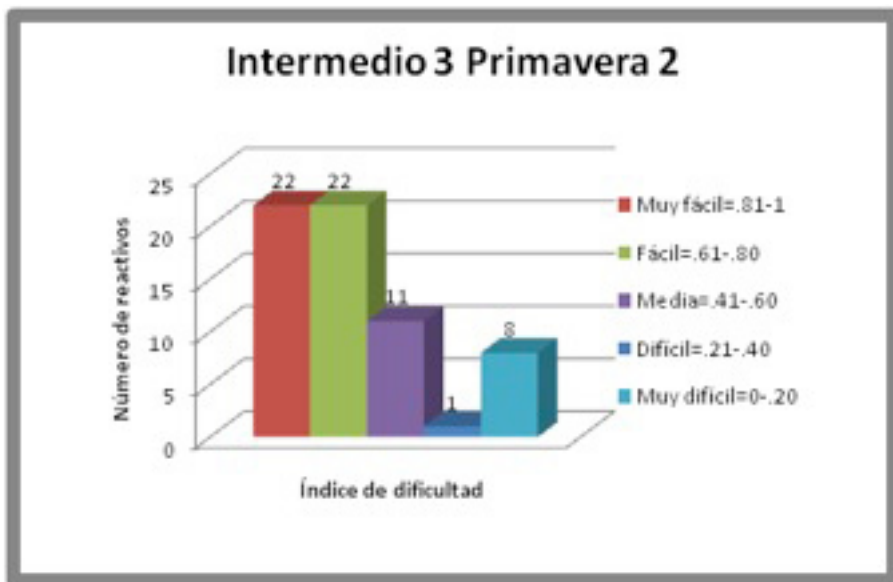
En cuanto a los resultados obtenidos, se observa que, con excepción del examen de Básico 4 (3ª versión) y el de Intermedio 1 (1ª versión), la mayoría de los estudiantes obtiene una calificación superior a siete en las pruebas computarizadas. Además, puede advertirse una gran consistencia entre los resultados de las diferentes pruebas y la calificación asignada por los profesores (Anexo 6).

El análisis del índice de dificultad (ID) de los ítems de los siete EC muestra un patrón bastante regular. En general, son pruebas cuyos ítems son respondidos de manera correcta por aproximadamente el 70 de los estudiantes que asistieron regularmente al curso y cumplieron con las actividades encomendadas por los profesores.<sup>21</sup>

<sup>21</sup> Los índices de dificultad se calculan a partir del número de los examinados que respondieron al reactivo de manera correcta sobre el total de aciertos. Al igual que en otras instituciones de la UNAM, se tomaron los siguientes rangos: si un ID es igual a 1.0 es muy fácil (todos lo responden correctamente); un ID=0 indica que el reactivo es muy difícil (ninguno lo contesta); y un ID=0.5 tiene un índice de dificultad media (la mitad de los candidatos lo responde correctamente). Un buen ítem tiene un índice entre .5 y .75



Gráfica 1



Gráfica 2

El examen más fácil resultó ser el básico 2, con un 80 por ciento de ítems con más de .8 de ID; en los niveles intermedios, este porcentaje es menor, por ejemplo, en el caso del Intermedio 3 es de 68 por ciento.



Si bien es cierto que los datos evidencian la necesidad de elevar el grado de dificultad de los reactivos, también es justo considerar que son muestra del buen desempeño de los docentes en el aula. Esta afirmación se apoya en el hecho de que existe una gran consistencia entre la calificación otorgada por ellos (equivalente al 20 por ciento de la calificación total) y los resultados de las pruebas objetivas y subjetivas.

En relación con el punto de vista de los estudiantes, se aplicó una pequeña encuesta de opinión a finales del periodo de invierno, para valorar la validez de apariencia de los instrumentos según la percepción de los reinscritos en el periodo de Primavera 2. Los datos mostraron entre un 60 y un 70 por ciento de aprobación (además de 42 comentarios adicionales positivos, frente a 23 negativos, relativos en su mayoría a dificultades no de índole académica, sino práctica, como el tamaño de la letra y la incomodidad de subir y bajar la pantalla para leer los textos). La sistematización lingüística es percibida como la prueba más fácil (70 por ciento) y la comprensión auditiva como la más difícil (60 por ciento).

A reserva de seguir analizando los datos en relación con los índices de discriminación y de fiabilidad, podemos decir que las primeras versiones de los exámenes cumplieron su cometido de constituirse en modelos válidos para evaluar el perfil de egreso de nuestro programa de estudios.

## CONCLUSIONES

La construcción de los EC fue un proceso en el que cristalizó la experiencia acumulada por el departamento de Español en relación con dos de los principales proyectos institucionales que se desarrollaron en la última década: el diseño de un programa de estudios y la elaboración de instrumentos de evaluación, ambos apegados a estándares internacionales. La plataforma que conformaron estos dos factores posibilitó la confección, en un tiempo relativamente corto, de siete instrumentos razonablemente válidos, cuya implementación, además de contribuir al incremento de la especialización de los productos y servicios de evaluación del CEPE, ha tenido repercusiones positivas en otros ámbitos de nuestra labor profesional. A manera de conclusión, enumeraremos los logros y los pendientes en los que tendremos que seguir trabajando.

El logro más destacado, en relación con el trabajo académico, fue la reafirmación de que la vía imprescindible para las transformaciones educativas es el trabajo colegiado, y de que el cambio y el compromiso para realizarlo se alcanzan confiriendo al docente el rol central.

En cuanto al avance en materia metodológica, se conformó al menos un equipo experto en la redacción de ítems, que se apropió de un procedimiento estandarizado para la elaboración de instrumentos. Por otra parte, la evaluación colegiada de las habilidades productivas (expresión escrita y oral) favoreció procesos de homologación de criterios entre el personal docente del área de español.

Gracias al efecto retroactivo (o de rebote) que sobre el proceso de enseñanza ejerce la evaluación, la definición de los siete modelos de examen contribuyó a la estandarización de los cursos. Los profesores se preocuparon desde un principio por conocerlos y por dar un mayor énfasis a la práctica de las habilidades productivas, las cuales habían estado un tanto supeditadas a la enseñanza de la gramática. Asimismo, el hecho de que los índices de rendimiento estén ahora expuestos a la comparación entre grupos del mismo nivel propició que los maestros pusieran mayor atención en el perfil de egreso y en el proceso individual de aprendizaje de cada uno de sus estudiantes. Ahora, se cuenta con mayor certeza y control sobre el nivel de actuación que alcanzan los estudiantes respecto al perfil definido en el programa de estudios del CEPE. Este control, a su vez, favorece el trabajo en el aula con grupos más homogéneos.

Es importante destacar el efecto positivo de la incorporación de la tecnología, ya que ahora se cuenta con una plataforma que posibilita la conformación de una base de reactivos suficientemente amplia para integrar una batería de exámenes equivalentes (misma forma y mismo contenido) para cada nivel. La tecnología contribuye también a minimizar el error en el cotejo o verificación de las pruebas objetivas (comprensión auditiva, comprensión de lectura y sistematización lingüística). En relación con los estudiantes, el uso de la computadora les permite integrar las competencias generales (uso de dicha herramienta) con las comunicativas.

Hay que señalar además que la vinculación tecnológica de los sistemas de aplicación de exámenes y de servicios escolares proporciona un mayor control administrativo en el proceso de expedición de certificados, lo cual, a su vez, contribuye a mejorar el servicio que se brinda a los estudiantes, ya que se acortan los tiempos de entrega de resultados e incluso se les facilita el acceso directo a su calificación por vía Internet y sin necesidad de acudir personalmente a la oficina de Servicios Escolares.

Finalmente, debe apuntarse que a lo largo de estas cuatro aplicaciones hemos detectado algunas debilidades en las que será necesario trabajar el resto del año. La más urgente es mejorar los índices de dificultad de los ítems. Es necesario también profundizar en la investigación para definir con mayor precisión —entre otras cuestiones— el perfil de los estudiantes extranjeros del CEPE y en qué medida el resultado de la evaluación refleja su aprovechamiento académico.

Otros pendientes son la elaboración de los documentos descriptivos de las características de los exámenes y de los protocolos de aplicación de las pruebas orales ; así como la estandarización de la tipología de procedimientos empleados (portafolios, autoevaluación, evaluación entre iguales) para asignar el 20 por ciento de la calificación, correspondiente a la evaluación formativa que llevan a cabo los profesores.

Avanzar en la resolución de estos pendientes requiere de un plan de gestión de calidad que nos lleve a contar con un sistema de evaluación ligado a la enseñanza, el cual cumpla los requisitos para ser certificado por algún organismo internacional. Es un objetivo que nos hemos trazado para un horizonte de dos años. Los resultados obtenidos en este año de celebración de nuestro 90 aniversario son un indicio que nos da una gran confianza en que el CEPE podrá lograrlo, reafirmando así su estatus como institución líder en la enseñanza del español como segunda lengua a nivel mundial.

#### BIBLIOGRAFÍA

- ALDERSON, J. Ch., C. Claphman & D. Wall, (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- ALDERSON, J. Ch., C. Claphman & D. Wall, (1998) *Exámenes de idiomas. Elaboración y evaluación*. Madrid: CUP.
- BACHMAN, L. & A. Palmer. (1996) *Language Testing in Practice: designing and developing useful language tests*. Oxford: OUP.
- BORDÓN, T., (2004) Panorama histórico del desarrollo de algunas de las cuestiones fundamentales en la evaluación de segundas lenguas. *Carabela*, (55), pp. 5-29.
- \_\_\_\_\_, (2006). *La evaluación de la lengua en el marco de E/L2: Bases y procedimientos*. Madrid: Arco/Libros.
- CAMPOS, E. (2006). *Lineamientos generales para la elaboración de reactivos y consignas del Certificado Internacional de Español (CIE)*. M. Jurado (coord.), México: Centro de Enseñanza para Extranjeros, Universidad Nacional Autónoma de México. [Documento electrónico s/r].
- Directrices de ALTE para los creadores de reactivos* (1995) actualizado en julio de 2005. (Versión española: Instituto Cervantes, 2006. Trad. José Ramón Parrondo).
- Exámenes computarizados del CEPE* [en línea] <<http://escolar.cepe.unam.mx/examenes/login.php>> México: Centro de Enseñanza para Extranjeros, Universidad Nacional Autónoma de México.

- ERES, G., (2010) *La enseñanza de lenguas extranjeras y la evaluación*, Madrid: Arco/Libros.
- FIGUERAS, N., (2004) Docencia, evaluación y contexto de uso real de la lengua oral. *Carabela* (55), pp. 63-84.
- GRANDA, B., (2008) *Lineamientos para la elaboración de reactivos del examen CELA Básico*. México: Centro de Enseñanza para Extranjeros, Universidad Nacional Autónoma de México. [Documento electrónico s/r].
- HUGHES, A., (1989) *Testing for language teachers*, Cambridge: CUP.
- PRATI, S., (2007) *La evaluación en español lengua extranjera*, Buenos Aires: Libros de la Araucaria.
- PRIETO J. M., La elaboración de una prueba de nivel: la reforma de los DELE. *Carabela*, (55), pp. 85-140.
- Programa de estudios del CEPE*. Versión 2010, México: Centro de Enseñanza para Extranjeros, Universidad Nacional Autónoma de México. [Documento electrónico s/r].
- Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual* (2009) [en línea]. <[http://www.coe.int/t/dg4/linguistic/manual1\\_en.asp](http://www.coe.int/t/dg4/linguistic/manual1_en.asp)>, Council of Europe Language Policy Divisions. Recuperado 19 julio 2011.
- SANTA CECILIA, A., (1999) *El currículo de español como lengua extranjera*. Madrid: Edelsa.
- Studies in language testing 6 multilingual glossary of assessment terms*. [en línea] <<http://www.alte.org/projects/glossary.php>> Cambridge: Cambridge University Press. Recuperado: 19 julio 2011.

<b>ANEXO 1</b>	
Básico 3 (Usuario independiente Umbral)	<p>En este nivel el estudiante</p> <ul style="list-style-type: none"><li>• Usará las estructuras de la lengua para ejecutar las siguientes funciones comunicativas: expresar posesión y pertenencia, expresar negación, expresar acciones y estados que acontecieron y se desarrollaron en el pasado, pedir/dar información para ubicar el origen temporal de una acción, expresar acciones y estados que se desarrollaron con anterioridad o con posterioridad a otra acción pasada, solicitar algo de manera cortés, establecer comparaciones y expresar cantidad.</li><li>• Escuchará diálogos (dos minutos aprox.) entre dos interlocutores, charlas, entrevistas. Escuchará también monólogos (90 segundos). Programas de radio y televisión, noticias, reportajes, lecturas cortas (narraciones) Discurso claro y a velocidad normal.</li><li>• Leerá textos de máximo 250 palabras: noticias, entrevistas, biografías, leyendas, cuentos y narraciones breves.</li><li>• Escribirá anuncios, cartas, currículos, comunicados, recados, mensajes, informes, (auto)biografías (entre 100-150 palabras).</li></ul>

**Anexo 2**  
**Tabla de especificaciones para la comprensión auditiva**  
**de los niveles básicos 3 y 4**

COMPRENSIÓN AUDITIVA			
Operaciones	Nivel cognoscitivo y No. de reactivos		
	1	2	3
Básicos 3 y 4	1	2	3
1. Generaliza el tema.		2	
2. Identificar el propósito.	2		
3. Identifica las ideas principales en las intervenciones de cada uno de los participantes.	2		
5. Localiza información concreta y específica.	4		
6. Sigue instrucciones.		1	
7. Sigue puntos principales de un debate, charla o conferencia.		1	
8. Resume el argumento		1	
No. total de reactivos	15		

**ANEXO 3**  
**Ejemplo de una escala para evaluar la producción escrita**  
**de los niveles básicos**

Calificación	
Vocabulario	/4
Gramática	/4
Coherencia y cohesión	/4
Ortografía	/3
TOTAL	/15

Fecha: \_\_\_\_\_

Tarea: \_\_\_\_\_

Nombre \_\_\_\_\_ Nivel \_\_\_\_\_

Nombre y firma del profesor 1 \_\_\_\_\_

Nombre y firma del profesor 2 \_\_\_\_\_

Nota: Si el candidato ejecuta una tarea distinta a la solicitada en las instrucciones, la calificación no podrá ser mayor a 4 puntos.

VOCABULARIO

	Ptos.	Evaluador 1	Evaluador 2	Final
Maneja el léxico suficiente y adecuado para el nivel.	4			
Maneja un vocabulario limitado para el nivel.	2			
Carece del vocabulario necesario.	0			

## GRAMÁTICA

	Ptos.	E. 1	E. 2	Final
Emplea la sintaxis y la morfología adecuadas para el nivel. Usa algunas oraciones complejas.	4			
Maneja un repertorio básico de estructuras gramaticales. Usa mayormente estructuras simples.	2			
Su desconocimiento de las estructuras de la lengua hace poco comprensible el texto.	0			

## COHERENCIA Y COHESIÓN

	Ptos.	E. 1	E. 2	Final
Produce un texto bien estructurado. Usa elementos de conexión entre las ideas de manera adecuada, aunque limitada.	4			
Produce un texto aceptable aunque con un limitado uso de conectores.	2			
Produce un texto poco estructurado, sin conectores o con conectores usados inadecuadamente.	1			

## ORTOGRAFÍA

	Ptos.	E. 1	E. 2	Final
La mayoría de las palabras están escritas y acentuadas correctamente.	4			
Tiene menos de 4 distintos tipos de errores.	2			
Tiene más de 4 distintos tipos de errores.	1			



## ANEXO 4

### Especificaciones del examen de nivel Intermedio 1

#### 1. Propósito del examen

Evaluar el aprovechamiento de los estudiantes del CEPE que terminan el nivel intermedio 1.

#### 2. Perfil de los estudiantes a los cuales va dirigido

Jóvenes y adultos mayores de 18 años que han cursado 78 horas de instrucción formal correspondiente al nivel Intermedio 1 del CEPE.

#### 3. Marco de referencia del examen

Se evalúa el nivel B2.1 del Marco Europeo Común de Referencia (MECR). Las operaciones y los contenidos se encuentran detallados en la tabla de especificaciones.

#### 4. Estructura del examen escrito

Se evalúan las cuatro habilidades y la sistematización lingüística. El porcentaje asignado a cada sección es la siguiente:

Secciones	Duración en minutos	No. de reactivos	Porcentaje
Expresión escrita	40 a 60		15%
Comprensión auditiva	15	20	15%
Comprensión de lectura	30	20	15%
Sistematización lingüística	20	30	15%
Expresión oral	15		20%
Total	3hrs. (Aprox.)		80%

Expresión escrita	
Duración	De 40 a 60 minutos.
Tarea(s) (Cf. instrumento 4)	Artículos de opinión, cartas, cuentos, memorias, reseñas, resúmenes (Aprox. 250 palabras).
Criterios para evaluar la tarea	Cf. Rúbrica

<b>Expresión escrita</b>	
Mecánica de aplicación	<p>El profesor entrega a los estudiantes la hoja con las instrucciones.</p> <p>Explica de manera general la tarea.</p> <p>Pregunta si hay dudas (puede hacer aclaraciones de cualquier tipo, incluyendo significado de vocabulario).</p> <p>Indica que no está permitido consultar ninguna fuente externa.</p> <p>Solicita apagar los teléfonos celulares.</p> <p>Da comienzo a la tarea y mide el tiempo.</p> <p>Diez minutos antes de concluir el plazo, señala que quedan 10 minutos.</p> <p>Una vez cumplido el plazo, recoge los textos producidos.</p>

<b>Comprensión auditiva</b>	
Duración	12 minutos aprox.
Total de reactivos	30
Pruebas (Cada texto sonoro se escucha 3 veces)	<p>Monólogo de 3 min. aprox. Programas de radio y televisión, noticias, reportajes, declaraciones, mensajes, conferencias, informes, lecturas cortas, instrucciones sobre el funcionamiento de aparatos.</p> <p>Temas concretos. Discurso articulado con claridad y a velocidad normal.</p> <p>Materiales auténticos o ex profeso (con alto grado de verosimilitud).</p>
Tipo de pregunta	<p>Falso verdadero.</p> <p>Opción múltiple.</p> <p>Relación con elementos visuales.</p>

<b>Comprensión de lectura</b>	
Duración	20 minutos
Total de reactivos	20
Pruebas	<p>2 textos de entre 450 y 700 palabras. Noticias, entrevistas, textos dialogados, textos periodísticos de opinión con estructura argumentativa clara, instrucciones, cuentos y narraciones breves con secuencias dialogadas.</p>
Tipo de pregunta	Opción múltiple

<b>Sistematización lingüística</b>	
Duración	30 minutos aprox.
Total de reactivos	30
Pruebas	Integración temática con la comprensión de lectura.
Tipo de preguntas.	Opción múltiple Cloze.

<b>Expresión oral</b>	
Duración	15 minutos (por candidato; si las tareas se ejecutan en parejas, el tiempo se duplica).
Tareas	Tarea 1. Entrevista entre pares (2 minutos aprox.). Tarea 2. Producción (3 minutos aprox.). Tarea 3. Interacción (3 minutos aprox.). Tiempo de preparación de las tareas (3 minutos aprox.). TOTAL 11-15 minutos (Aprox.).
Criterios para evaluar la tarea	Cf. Rúbrica
Mecánica de aplicación	Cf. Instrucciones para los profesores.

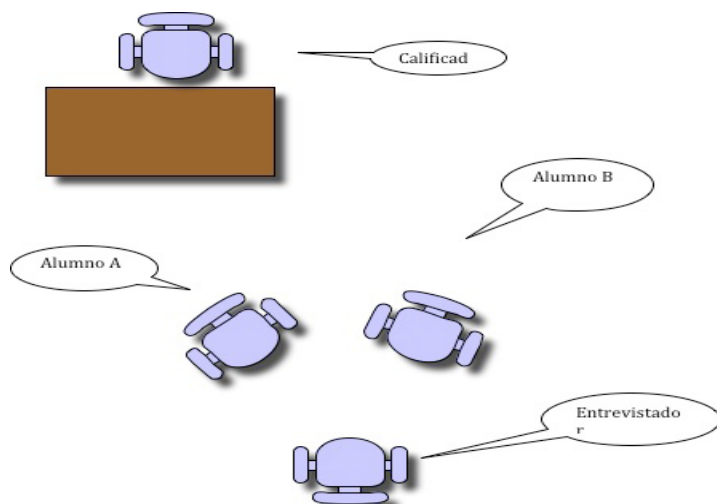
## ANEXO 5

### Instrucciones para los profesores

#### Examen final de expresión oral

#### Nivel Básico 1

#### I. Mecánica de aplicación de la prueba (Disposición de los participantes).



#### II. Protocolo de aplicación de la prueba

Entrevistador:

1. Recibe a los 2 estudiantes y les indica que tomen asiento.
2. Les informa que la prueba de expresión oral del nivel básico 1 consta de 3 partes y que él les indicará lo que tienen que hacer en cada una de ellas.
3. Indica que comiencen la primera parte de la prueba y les explica en qué consiste.

#### Tarea 1. Entrevista entre pares

Caja con papelitos que contienen preguntas que serán insaculadas por los estudiantes

4. Se dirige al alumno A y le pide que saque de la caja una pregunta y la formule al estudiante B.
5. Indica al alumno B que tiene 1 minuto para planear su respuesta y 1 para ejecutarla.
6. Repite el procedimiento (pasos 4 y 5) con el alumno B.
7. Dirige la acción marcando tiempos.
8. Recoge los papelitos, da por concluida la tarea y efectúa la transición a la segunda.

*[Esta tarea sirve para establecer un ambiente de equidad entre los estudiantes, romper el hielo y darles confianza. Sirve también para dejar claro que la interacción con el entrevistador será mínima. Además, es indicador para empezar a evaluar en las escalas].*

## Tarea 2. Producción

Tarjeta "Tarea 1. Producción.  
Descripción de una persona"

9. Se dirige al alumno A y le entrega la tarjeta y la imagen, cuidando que el alumno B no la vea.
10. Le indica que tiene 3 minutos máximo para leer la tarjeta, observar la fotografía y preparar su intervención
11. Indica al alumno A que debe describir a la persona que aparece en la fotografía y que debe tener cuidado de no mostrarla al alumno B.
12. Se dirige al alumno B y le informa que tendrá que poner atención a la información del alumno A, porque después tendrá que formularle una pregunta.
13. Da la instrucción al alumno A para que comience a hablar.
14. Le informa que tiene de 1 a 3 minutos para ejecutar la tarea.
15. Una vez concluida la intervención del alumno A, pide al alumno B que formule una pregunta al alumno A y le pide a éste que responda.
16. Retira al alumno A la tarjeta y la fotografía.
17. Repite el procedimiento (pasos 9 a 16) con el alumno B, cuidando de hacerlo con una fotografía distinta.
18. Pregunta a ambos estudiantes si quieren agregar algo.
19. Indica que ha terminado la tarea número 2 y efectúa la transición a la tercera y última tarea.

*[En esta tarea se evalúa fundamentalmente al estudiante que produce la descripción conforme a los criterios que aparecen en la escala, vocabulario relacionado con la descripción física, colores, ropa, etc. Se observa también la interacción de ambos estudiantes con las preguntas que se formulan entre sí].*

## Tarea 3. Interacción (Juego de roles)

Tarjeta "Tarea 2. Interacción:  
Diálogo entre un médico y un paciente"

20. Se dirige a ambos estudiantes y les entrega a cada uno una tarjeta. Les informa que tienen **1 minuto para leerla y preparar** su participación. Explica (utilizando un nivel de lengua adecuado) que se trata de actuar qué dirían si estuvieran en una situación así. Si están listos antes de ese tiempo, pueden empezar.

21. Informa a ambos que tienen de **2 a 3 minutos** para completar la tarea.
22. Da la instrucción para que el alumno que tiene la opción “médico” comience el diálogo.
23. Dirige la acción marcando tiempos.
24. Al finalizar el diálogo, recoge las tarjetas e indica que la prueba ha terminado.

[Esta tarea evalúa la capacidad del alumno para interactuar, tomar el turno de palabra y cederlo, formular el tema y utilizar vocabulario relacionado con el estado de salud].

25. El calificador que se encuentra en el escritorio entrega a cada alumno la ficha de comprobación de presentación del examen, la cual recortará de la parte inferior de las escalas.
26. Entrevistador y evaluador tendrán 3 minutos para ponerse de acuerdo sobre los puntajes otorgados a ambos estudiantes.

### III. Recomendaciones generales para el entrevistador

27. Procurar crear un clima cordial.
28. Minimizar su propia actuación, aun cuando exista alguna asimetría entre los niveles de lengua de los estudiantes. En estos casos, se debe tomar en cuenta la actuación de ambos en cuanto al uso de estrategias para compensar deficiencias en la comunicación, el manejo de la interacción y el componente actitudinal del aprendizaje (elementos como el principio de cooperación, la capacidad de respuesta ética y moral en situaciones de conflicto, de manera individual y/o grupal, el respeto, la tolerancia, etc.).
29. Llevar un estricto control del tiempo.
30. En caso de que alguno de los estudiantes exprese su insatisfacción por el tema o la tarea, se le hará la observación de que en la comunicación real no siempre nos enfrentamos a circunstancias que nos gustan, pues se trata de mostrar su nivel de lengua, su control e incluso su capacidad para ser creativo y sacar adelante cualquier situación comunicativa.

## ANEXO 6

### Resultados obtenidos por los estudiantes en los ec

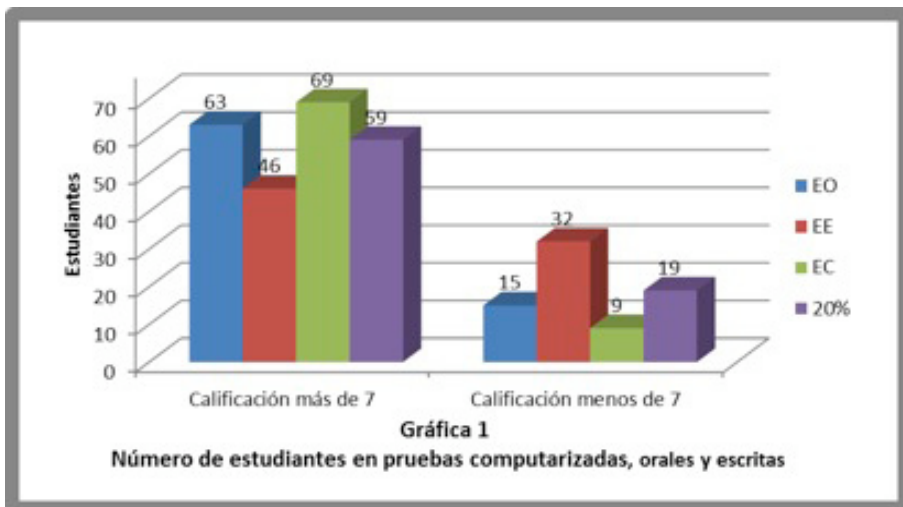
La información está organizada en relación con el número de estudiantes que obtuvieron más y menos de siete de calificación —medida estándar que nos parece un buen índice de aprovechamiento académico— de los cursos Básico 2, Básico 4 e Intermedio 3, durante los períodos de otoño, primavera 1 y primavera 2.

Básico 2. Primera versión de examen

ESTUDIANTES BÁSICO 2	Eo*	EE	EC	20%
CON MÁS DE 7	63	46	69	59
CON MENOS DE 7	15	32	9	19
TOTAL	78	78	78	78

Tabla 1. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.

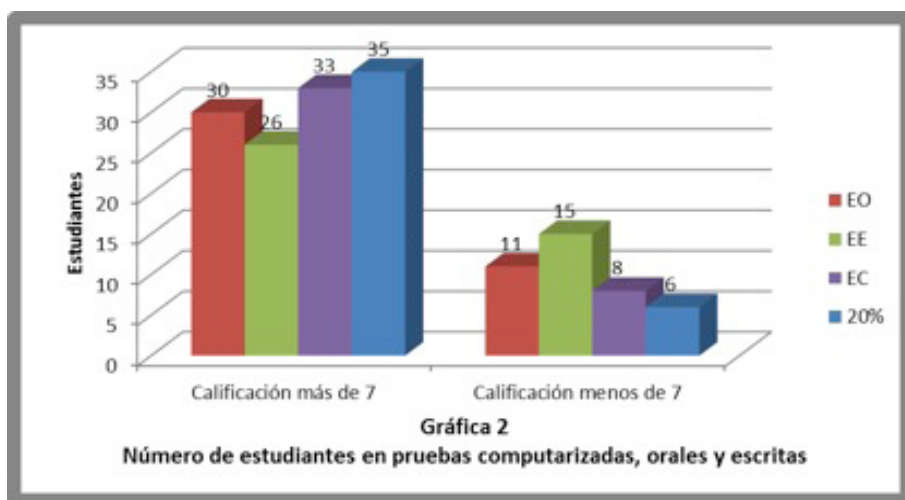


## Básico 2. Segunda versión

ESTUDIANTES BÁSICO 2	EO*	EE	EC	20%
CON MÁS DE 7	30	26	33	35
CON MENOS DE 7	11	15	8	6
TOTAL	41	41	41	41

Tabla 2. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.



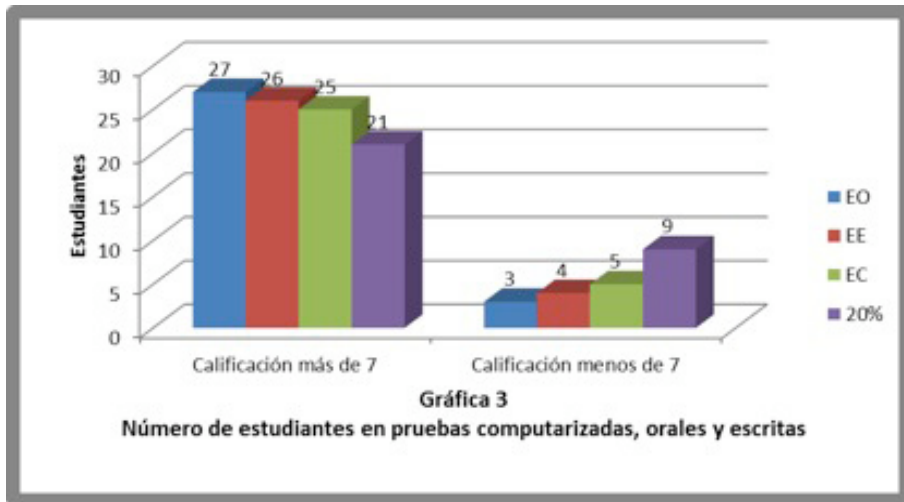
## Básico 4. Primera versión

ESTUDIANTES BÁSICO 4	EO*	EE	EC	20%
CALIFICACIÓN MÁS DE 7	27	26	25	21
CALIFICACIÓN MENOS DE 7	3	4	5	9
TOTAL	30	30	30	30

Tabla 3. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.



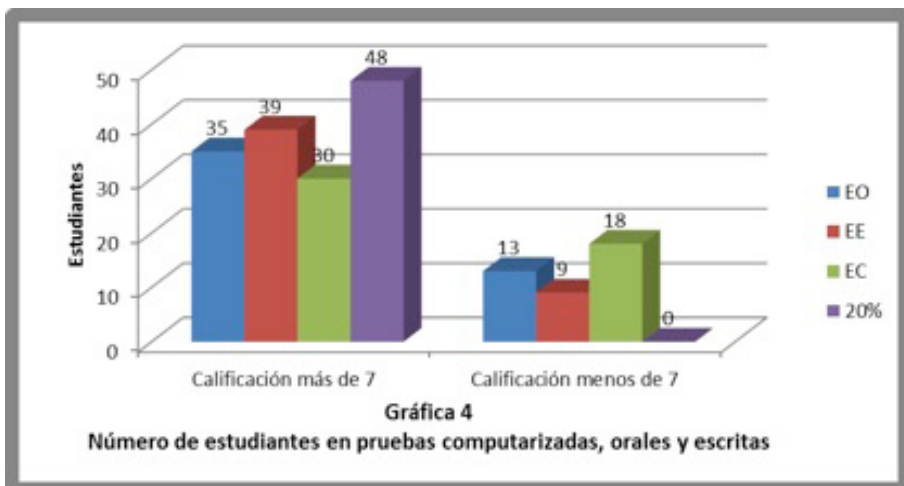


## Básico 4. Segunda versión

ESTUDIANTES BÁSICO 4	EO	EE	EC	20%
CALIFICACIÓN MÁS DE 7	35	39	30	48
CALIFICACIÓN MENOS DE 7	13	9	18	0
TOTAL	48	48	48	48

Tabla 4. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.

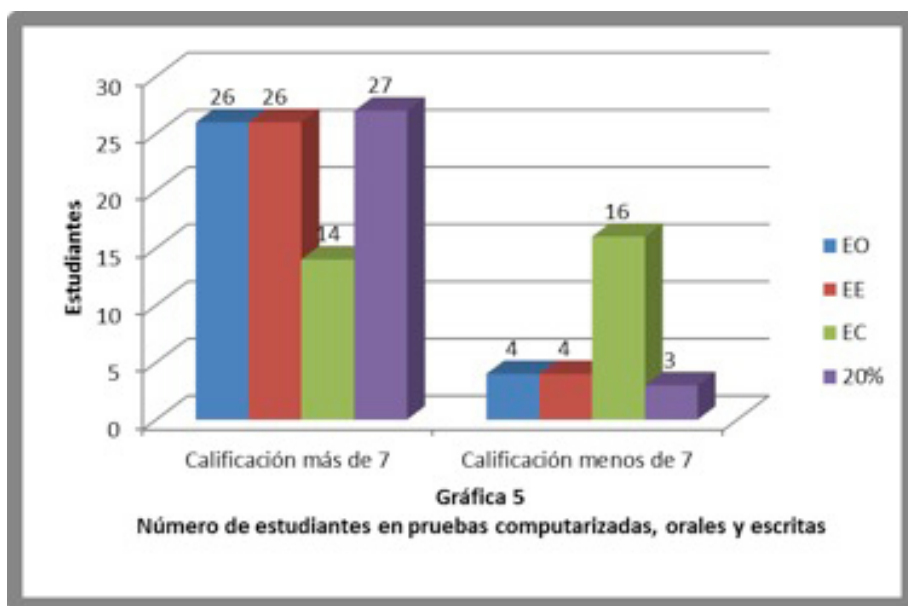


## Básico 4. Tercera versión

ESTUDIANTES BÁSICO 4	EO	EE	Ec	20%
CON MÁS DE 7	26	26	14	27
CON MENOS DE 7	4	4	16	3
TOTAL	30	30	30	30

Tabla 5. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.

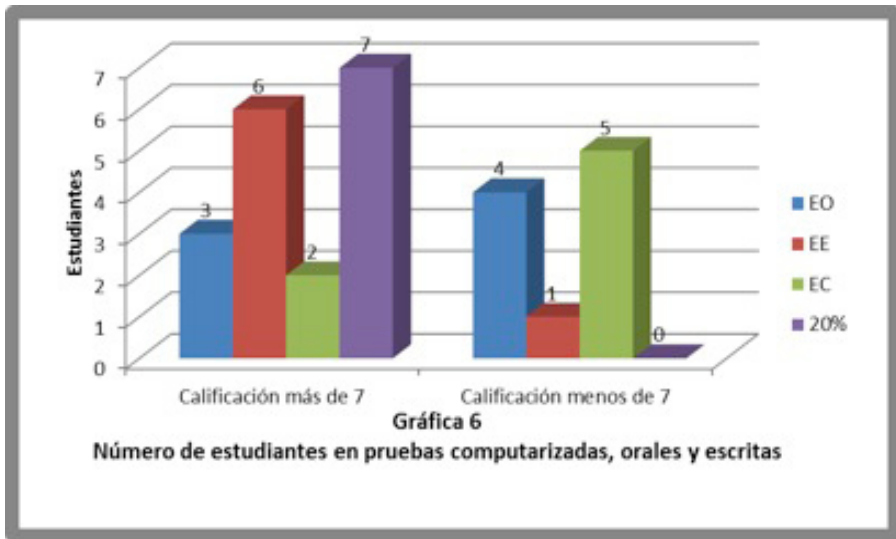


## Intermedio 3. Primera versión

ESTUDIANTES INTERMEDIO 3	EO	EE	EC	20%
CON MÁS DE 7	3	6	2	7
CON MENOS DE 7	4	1	5	0
TOTAL	7	7	7	7

Tabla 6. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.

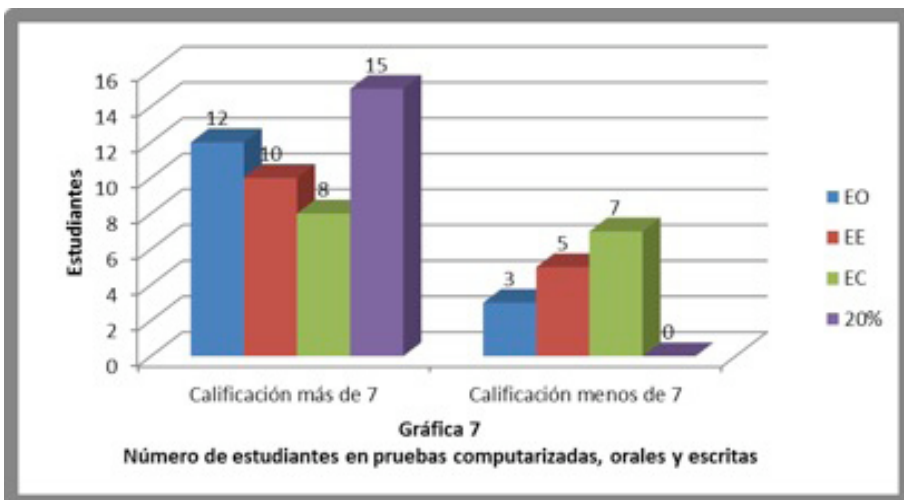


## Intermedio 3. Segunda versión

ESTUDIANTES INTERMEDIO 3	EO	EE	EC	20%
CON MÁS DE 7	12	10	8	15
CON MENOS DE 7	3	5	7	0
TOTAL	15	15	15	15

Tabla 7. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

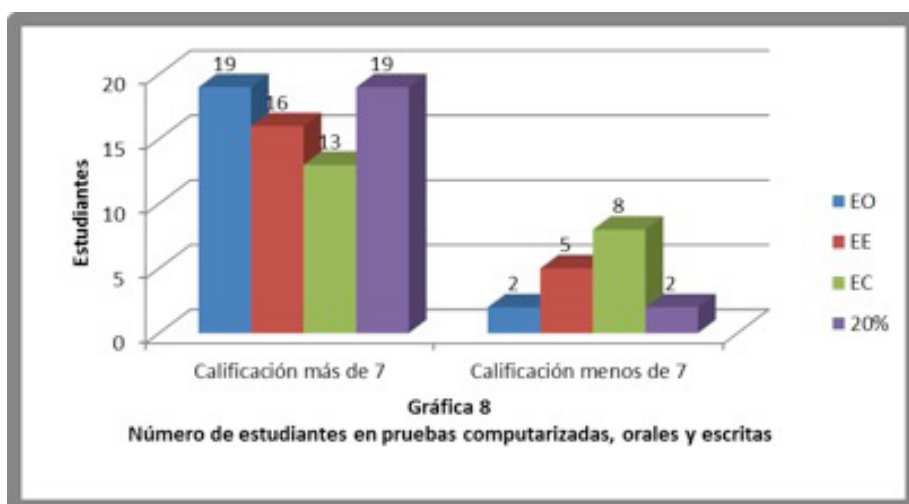
\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.



ESTUDIANTES INTERMEDIO 3	EO	EE	EC	20%
CON MÁS DE 7	19	16	13	19
CON MENOS DE 7	2	5	8	2
TOTAL	21	21	21	21

Tabla 8. Número de estudiantes a partir de sus calificaciones en una escala de 0 a 10.

\*EO: Prueba de expresión oral; EE: Prueba de expresión escrita; EC: Exámenes computarizados.



## ANEXO 7

## Encuesta de opinión aplicada a los estudiantes que se reinscribieron al curso de Primavera 2

Resultados de las encuestas - Resumen  
Encuestados 207

Opinión del examen final de español

	5/Muy fácil, Bueno.	4/Bueno	SUMA	%	3/Regular	2/Difícil	1/Muy Difícil	suma de 2 y 1	%	TOTAL respuestas
Examen Computarizado (EC)										
1. Comprensión auditiva	50	74	124	60	65	11	5	16	8	205
2. Comprensión de lectura	54	84	138	67	49	12	7	19	9	206
3. Gramática	63	81	144	70	52	7	4	11	5	207
4. Del examen de redacción	51	78	129	63	62	11	4	15	7	206
<b>TOTAL EC</b>	<b>218</b>	<b>317</b>	<b>535</b>	<b>65</b>	<b>228</b>	<b>41</b>	<b>20</b>	<b>61</b>	<b>7</b>	<b>824</b>
5. Del examen oral	69	71	140	68	55	6	4	10	5	205
6. De la evaluación formativa	56	80	136	66	59	9	3	12	6	207

