

# Repensar la evaluación del aprendizaje en el contexto de la Educación Superior

Tiburcio Moreno Olivos

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO

## RESUMEN

En la actualidad, enseñar es más complejo que antes, y aprender se ha convertido en una experiencia cada vez más desafiante para los alumnos. De la capacidad para adquirir y almacenar conocimientos en la memoria, se ha transitado al desarrollo de competencias y habilidades intelectuales de orden superior. Este giro significa un cambio profundo en la concepción tanto del aprendizaje como de la enseñanza, lo que, a su vez, obliga a replantear el papel de la evaluación en el aula, de modo que esta se convierta en un mecanismo efectivo para valorar tanto los procesos como los productos del aprendizaje. En este artículo se destaca la función formativa y retroalimentadora del proceso de evaluación, así como la necesidad de proveer al profesorado herramientas conceptuales y metodológicas que le permitan implementar una evaluación del aprendizaje y para el aprendizaje. También se alude a la evaluación sostenible y a las condiciones que esta requiere para asegurar el aprendizaje presente y a largo plazo.

**Palabras clave:** evaluación del aprendizaje, evaluación sostenible, evaluación y motivación, aprendizaje a largo plazo, educación superior

## INTRODUCCIÓN

Desde una perspectiva integradora, se reconoce que aprender, enseñar y evaluar son procesos estrechamente relacionados y mutuamente influyentes. Es habitual que los profesores evalúen mientras enseñan y, en ocasiones, los alumnos aprenden mientras son evaluados. Por otra parte los mecanismos previstos para evaluar determinan de manera decisiva la forma en que los alumnos aprenden y el modo en que los profesores enseñan (Monereo y Castelló, 2009: 15). Esto significa que la evaluación puede llegar a tener un alcance y poder excesivos que le permitan condicionar intensamente el proceso de enseñanza-aprendizaje.

En el discurso educativo actual se insiste en la necesidad de que los alumnos aprendan de forma significativa y relevante, pero sabemos que lograr esta meta no es una tarea sencilla; depende de muchos factores que interactúan de forma compleja. En la educación formal la evaluación tiene un peso privilegiado y, para la mayoría de los alumnos, constituye la principal razón de sus prácticas de estudio y aprendizaje. En otras palabras, los alumnos estudian para aprobar exámenes, y, en el mejor de los casos, aprenden para la evaluación (Moreno, 2009), lo que genera un aprendizaje superficial que permanece en la memoria a corto plazo y al poco tiempo se olvida.

Según Monereo y Castelló (2009), cuando la evaluación supone repetir mecánicamente la información recibida (por ejemplo, aplicando fórmulas matemáticas para resolver ejercicios artificiales), aprender para la evaluación, sin duda, se convierte en un esfuerzo estéril, en un ardid que se evapora tan pronto como se supera dicha evaluación. Sin embargo, cuando una evaluación exige al individuo dar solución a problemas y conflictos de la vida cotidiana, debiendo argumentar convenientemente las razones de determinada respuesta o solución, y porque una es mejor que otra, entonces aprender para esa evaluación no se diferencia mucho de aprender para la vida. En este último caso el aprendizaje es *sostenible*. El alumno habrá adquirido unas competencias que con bastante frecuencia deberá emplear para afrontar las demandas del entorno social en el que vive y se desarrolla. En coherencia con este tipo de aprendizaje, Boud (2000) propone una *evaluación sostenible*.

## EVALUACIÓN Y APRENDIZAJE A LARGO PLAZO

### CONVERTIRSE EN UN ASESOR

Preparar a los estudiantes para el aprendizaje a lo largo de la vida necesariamente implica habilitarlos para formular juicios complejos acerca de su propio trabajo y el de otros, y para tomar decisiones en circunstancias impredecibles e inciertas con las que se encontrarán en el futuro. Un rasgo central de esta perspectiva es que los estudiantes son imaginados como jugadores mucho más activos en el proceso de evaluación.

Tradicionalmente los estudiantes no han tenido una participación activa en el proceso de evaluación, más bien se han limitado a responder pruebas y a recibir retroalimentación en materias que los profesores juzgan importantes. Ellos se convierten en recipientes de las acciones de otros, no son agentes activos en el proceso de evaluación. Tales concepciones de la evaluación son inapropiadas para el aprendizaje a largo plazo y también limitan el aprendizaje actual. Ni los profesores ni un currículum provocan el aprendizaje después de la graduación; es el deseo de los aprendices, las iniciativas que ellos tomen y el contexto en el que el aprendizaje tiene lugar los que ejercen una influencia determinante. Las tradicionales prácticas de evaluación pueden minar la capacidad de los estudiantes para juzgar su propio aprendizaje y, por ende, limitar la agenda del aprendizaje para toda la vida.

En los contextos del mundo real, los estudiantes deben ser capaces de juzgar o evaluar lo adecuado, lo completo o lo apropiado de su propio aprendizaje; así que cualquier práctica de evaluación empleada debe ser comprensible para ellos, de modo que pueda ser internalizada como criterio para la autoevaluación crítica. En otras palabras, si los estudiantes están siempre atendiendo los juicios de otros, podrían no adquirir un conjunto más amplio de habilidades que les permitan hacer esto para sí mismos. Además, las tareas de evaluación a menudo enfatizan la solución de problemas antes que la formulación de problemas y comúnmente utilizan ejemplos no reales y descontextualizados para evaluar el aprendizaje.

En muchas prácticas de evaluación se omite la participación del estudiante en el diseño de la evaluación y, así, son invisibles las etapas clave para juzgar el aprendizaje, tales como el establecimiento de criterios apropiados para la terminación de tareas. Además, en los cursos a menudo se insinúa —y los profesores usualmente asumen— que la colaboración es

una trampa y, por tanto, los estudiantes son desalentados para trabajar cooperativamente. En relación con el tema de la cooperación en el aula, se plantea que: “históricamente, el propósito de las pruebas de rendimiento ha sido medir la competencia individual de los estudiantes. El supuesto tradicional es que la competencia individual es mejor medida sólo por el trabajo individual sin la asistencia de otros. Desde esta perspectiva, la colaboración es vista como ‘trampa’ y las puntuaciones no son consideradas válidas porque no reflejan el trabajo ‘propio’ del estudiante” (Webb, 1997: 205).

Es frecuente que los estudiantes no tengan la oportunidad de ver cómo funciona realmente el proceso de evaluación. Esto es algo que ellos experimentan como un procedimiento al que se someten, en lugar de que sea algo que poseen. Los sistemas y prácticas de evaluación ejercen considerables limitaciones sobre la conducta de aprendizaje de los estudiantes. El conjunto altamente individualizado de supuestos sobre la evaluación de los estudiantes manifiesto en grados y certificación individual es una de tales limitaciones. Otra es la fragmentación de las tareas de evaluación —y, en efecto, del currículum— propuestas por la modularización que inhibe algunos enfoques holísticos para evaluar. La más dominante de todas las limitaciones es el tratamiento de la evaluación cuando clasifica, porque conduce a los estudiantes a centrarse en las calificaciones antes que en el aprendizaje que estas pretenden representar.

Estas preocupaciones acerca de la influencia negativa de las prácticas de evaluación no sólo aplican para los enfoques tradicionales. Ahora hay nuevas trampas de evaluación. Las estrategias que tienen un efecto positivo en el aprendizaje actual (por ejemplo, proveer a los estudiantes con criterios para la evaluación), podrían tener consecuencias no deseadas a largo plazo que aun tienen que ser identificadas. Por ejemplo, mientras que el uso de los resultados de evaluación y la especificación de los criterios podrían ser deseables, una consecuencia no prevista es representar en los estudiantes la idea de que la especificación de los criterios y resultados está determinada de antemano y que el aprendizaje sólo se logra siguiendo tales especificaciones dadas por otros.

Antes de que el aprendizaje pueda iniciar, los estudiantes tienen la necesidad de identificar por sí mismos lo que requieren aprender, tomando en cuenta un conjunto de factores contextuales. Además, tienen que juzgar lo que cuenta como un ‘buen trabajo’. Por supuesto, la apreciación de los criterios apropiados podría necesitar apoyarse en la retroalimentación de los profesores. Sin embargo, la meta siempre debe ser que los estudian-

tes aprendan a juzgar por sí mismos lo que constituye un buen trabajo y darles oportunidades para practicar esta habilidad. Pero, sin el desarrollo de los estudiantes como asesores, como una meta clara traducida en prácticas específicas, podemos terminar inadvertidamente minando lo que estamos tratando de lograr.

#### EVALUACIÓN SOSTENIBLE

Las prácticas de evaluación deben ser juzgadas desde la perspectiva de si efectivamente equipan a los estudiantes para una evaluación de su propio aprendizaje a lo largo de la vida. Para esto se propone la adopción del concepto de *evaluación sostenible*. La visión de que la evaluación puede ser sostenible significa que “satisface las necesidades del presente sin comprometer la capacidad de los estudiantes para satisfacer sus propias necesidades futuras de aprendizaje” (Boud, 2000: 151). Es decir, las actividades de evaluación no deben dirigirse sólo a las necesidades inmediatas de certificación o retroalimentación de los estudiantes en su aprendizaje actual, sino también contribuir en alguna forma a su aprendizaje futuro.

Fue Boud (2000) quien identificó la importancia de considerar la evaluación para el aprendizaje a lo largo de la vida; pero ha habido poca discusión acerca de dónde instalar la evaluación sostenible en el panorama más amplio. Sin embargo, para la evaluación sostenible no es la preparación de los estudiantes lo que realmente importa, sino el uso de la evaluación sostenible como parte de la disposición de los estudiantes para su futuro aprendizaje y experiencias de evaluación (Ecclestone y Pryor, 2003). La evaluación sostenible no es un nuevo tipo de práctica de evaluación, sino una forma de construir sobre la evaluación formativa y sumativa para fomentar metas de aprendizaje a largo plazo.

Es claro que todos los aspectos de la enseñanza, el aprendizaje y la evaluación tienen un importante papel que jugar en el aprendizaje a largo plazo. Pero hay un número de razones importantes para centrarse en la evaluación como tal. Como ya se ha mencionado al inicio de este artículo, la evaluación tiene un poderoso efecto sobre los alumnos, los profesores y los programas educativos. La evaluación comunica la intención de los estudiantes y es un indicador para ellos de lo que se considera más importante (Moreno, 2004). Si las prioridades del aprendizaje no están contempladas en las prácticas de evaluación, particularmente en las prácticas de evaluación sumativa, pueden no ser tomadas en serio. La evaluación

influye las percepciones que los estudiantes tienen del currículum y las formas en que ellas podrían emplear el proceso para fomentar las habilidades del aprendizaje (Moreno, 2010).

Boud (2000) utilizó el análisis de Black y Wiliam (1998) de una investigación sobre evaluación formativa como punto de partida para identificar los temas a ser incorporados en un programa de reforma de la evaluación. Los puntos considerados en la revisión de la práctica de evaluación para hacerla más sostenible fueron:

1. La importancia de un marco basado en referentes o criterios claros que permita a los estudiantes ver su propio trabajo a la luz de una práctica aceptable.
2. La creencia de los profesores de que todos los estudiantes pueden tener éxito.
3. La necesidad de fomentar en los estudiantes la confianza acerca de su capacidad para aprender porque sus creencias al respecto afectan su rendimiento.
4. La necesidad de separar los comentarios de las calificaciones porque estas distraen a los estudiantes de los comentarios de retroalimentación.
5. La necesidad de centrar la evaluación en el aprendizaje en lugar del desempeño.
6. El papel vital del desarrollo de habilidades de autoevaluación.
7. El fomento de una evaluación reflexiva con pares.
8. Asegurarse que los comentarios sobre las tareas de evaluación en realidad se utilizan efectivamente para influir el aprendizaje posterior.

## EVALUACIÓN DEL APRENDIZAJE Y PARA EL APRENDIZAJE

Hay otra forma en la que la evaluación puede contribuir al desarrollo de escuelas efectivas que ha sido en gran parte ignorada en la evolución de los estándares, la evaluación y el movimiento de la rendición de cuentas. Nosotros también podemos usar la evaluación *para aprender*. De acuerdo con Black *et al.* (2004: 10), la *evaluación para el aprendizaje* es cualquier evaluación que tiene como primera prioridad, en su diseño y en su práctica, servir al propósito de promover el aprendizaje de los estudiantes. Por lo tanto, difiere de la evaluación que se diseña sobre todo para servir

a los propósitos de rendición de cuentas, clasificación o acreditación de competencias. Una actividad de evaluación puede ayudar al aprendizaje, si proporciona información que los profesores y los estudiantes pueden usar como retroalimentación para evaluarse a sí mismos y a otros, y para modificar las actividades de enseñanza y aprendizaje en que participan. Dicha evaluación se convierte en *evaluación formativa* cuando la evidencia es utilizada efectivamente para adaptar la enseñanza de modo que responda a necesidades de aprendizaje.

Si las *evaluaciones del aprendizaje* proveen evidencia del logro para informes públicos, *las evaluaciones para el aprendizaje* sirven para ayudar a los estudiantes a aprender más. La distinción crucial está entre una evaluación para determinar el estatus del aprendizaje y una evaluación para promover un aprendizaje mayor. Tanto las evaluaciones del aprendizaje como para el aprendizaje son importantes. Aunque actualmente se tienen muchas evaluaciones del aprendizaje disponibles, si se quiere balancear las dos, hay que hacer una inversión mucho más fuerte en la evaluación para el aprendizaje. Se pueden obtener beneficios sin precedentes en el rendimiento académico, si se logra convertir el actual proceso diario de evaluación en el aula en un instrumento más poderoso para el aprendizaje. Para ello, es necesario proveer a los profesores con las herramientas de evaluación que requieren para hacer mejor su trabajo.

Lo anterior es intentar equiparar la idea de evaluación para el aprendizaje con un término común entre nosotros: *evaluación formativa*. Pero no son lo mismo. La evaluación para el aprendizaje es algo que va más allá de las pruebas frecuentes y no provee a los profesores con evidencia de modo que ellos puedan revisar la enseñanza, aunque esos pasos son parte de esto. Además, la evaluación para el aprendizaje debe incluir a los estudiantes en el proceso. Cuando los profesores evalúan para el aprendizaje, usan el proceso de evaluación en el aula y el continuo flujo de información acerca del logro del alumno que esto provee, a fin de obtener beneficios, no sólo de verificar el aprendizaje adquirido.

En resumen, el efecto de la evaluación para el aprendizaje consiste en que los alumnos se mantengan aprendiendo y permanezcan confiados en que ellos pueden continuar aprendiendo productivamente. En otras palabras, que no caigan en la frustración ni en la desesperanza. Si se quiere conectar a la evaluación con la mejora de la escuela en formas significativas, debemos ver la evaluación a través de nuevos ojos. Los actuales sistemas de evaluación son nocivos para un gran número de alumnos y ese daño

surge, entre otras causas, de nuestro fracaso para balancear el uso de las pruebas estandarizadas y las evaluaciones de aula al servicio de la mejora escolar.

## EVALUACIÓN Y MOTIVACIÓN

### LA EVALUACIÓN COMO FUENTE DE MOTIVACIÓN PARA APRENDER

Una de las funciones pedagógicas más importantes de la evaluación es la motivación que puede producir en los participantes. La investigación más reciente ha revelado que lo que hace más efectivo al aprendizaje es que los alumnos se involucren activamente en el proceso de evaluación (Stiggins, 2002). Las formas en las cuales la evaluación puede afectar la motivación y la autoestima de los alumnos, así como los beneficios que implica comprometerlos en la autoevaluación, sugieren que ambos merecen una cuidadosa atención.

Los alumnos sólo pueden evaluarse a sí mismos cuando tienen una imagen suficientemente clara de las metas de aprendizaje que se pretenden conseguir. Y muchos alumnos carecen de tal imagen. Si a los educandos se les da solamente notas o calificaciones, ellos no se beneficiarán de la retroalimentación de su trabajo. En este tenor se menciona que “cuando a los estudiantes se les da una nota final en una tarea junto con comentarios formativos, es altamente probable que ellos atiendan sólo la nota y no tomen en serio los comentarios formativos” (Butler, 1988, *apud* Moreno, 2007: 63).

La mayoría de nosotros hemos crecido en aulas en las cuales nuestros profesores creían que la forma de potenciar el aprendizaje era maximizar la ansiedad, y la evaluación siempre tenía que ser el intimidador más grande. A causa de sus propias experiencias exitosas en ascender a posiciones de liderazgo y autoridad, la mayoría de los políticos y líderes escolares aprendieron que el camino para el éxito —cuando se enfrentan con un fuerte desafío— es redoblar sus esfuerzos. Es decir, ellos consideran que la forma de provocar que los alumnos aprendan más es enfrentarlos a un reto más difícil. Esto provocará que dupliquen sus esfuerzos, con lo que aprenderán más, sus puntuaciones en las pruebas subirán y las escuelas serán más efectivas.

Según esta óptica, se piensa que se puede motivar a los alumnos para que realicen un esfuerzo mayor, estableciendo criterios académicos más altos, subiendo la barra e implementando más pruebas estandarizadas. Este es el fundamento de la creencia en el poder de tales pruebas, orienta-

das hacia la rendición de cuentas para la mejora de la escuela. De hecho, cuando algunos alumnos son enfrentados con un desafío más difícil, lo que ellos hacen es doblar sus esfuerzos, y aprenden más de lo que podrían haber aprendido sin el incentivo adicional. Pero hay que hacer notar, sin embargo, que esto es cierto sólo para algunos alumnos.

Otro amplio segmento de la población escolar, cuando es enfrentado con un reto todavía más difícil que el primero en el que ya ha fracasado, no redoblará sus esfuerzos —un punto que la mayoría de la gente olvida—. Estos alumnos percibirán las demandas por puntuaciones más altas como inalcanzables, y caerán en la desesperanza. Por lo tanto, es irracional construir nuestros ambientes de evaluación sobre el supuesto de que las pruebas estandarizadas tendrán el mismo efecto en todos los alumnos. Esto no sucederá. Algunos alumnos afrontan las pruebas con un historial personal académico fuerte y una expectativa de éxito. “Algunos alumnos vienen a dar muerte al dragón, mientras que otros esperan ser devorados por él” (Stiggins, 2002: 761). Como resultado, la evaluación mediante pruebas estandarizadas fortalecerá el aprendizaje de algunos mientras desmotivará a otros y causará en ellos desaliento.

Conviene aclarar que, aunque las pruebas estandarizadas suelen emplearse para la selección de candidatos a ingresar a la educación media superior y superior, o bien, con fines de certificación una vez que egresan de una determinada carrera, lo cierto es que en determinados momentos de su trayectoria escolar la mayor parte de los alumnos inevitablemente tiene que afrontar este tipo de instrumentos de evaluación (por ejemplo, para certificar el dominio de un idioma, habilidades matemáticas, habilidades de lecto escritura, etc.), cuyos resultados pretenden dar cuenta de sus logros de aprendizaje durante un periodo de tiempo determinado. Estas experiencias marcan de forma significativa la vida de muchos individuos, llegando en algunos casos a predisponerles negativamente hacia los estudios y otras experiencias de evaluación.

#### EVALUACIÓN Y CONFIANZA EN EL ALUMNO

Los estudiantes tienen éxito académicamente sólo si ellos quieren tener éxito y se sienten capaces de hacerlo. Si ellos carecen de deseo o confianza, no serán exitosos. Por lo tanto, la pregunta esencial es doble: ¿cómo ayudamos a nuestros estudiantes para que quieran aprender y se sientan capaces de aprender?

Históricamente, hemos construido nuestra respuesta a estas preguntas sobre algunos supuestos. Asumimos que podemos estimular al máximo el esfuerzo del profesor y el aprendizaje del estudiante mediante amenazas públicas que avergüenzan tanto a los estudiantes como a los profesores, si los estudiantes no tienen éxito académico. Un ejemplo es el uso que se hace de los resultados de las pruebas estandarizadas que en nuestro país el Centro Nacional para la Evaluación de la Educación Superior (CENEVAL) aplica en las universidades públicas estatales, cuyas puntuaciones son consideradas para valorar la calidad (o su ausencia) de los programas educativos, y, por extensión, el mérito de los alumnos, los profesores y la institución misma. Si los resultados no son los esperados, esto tendrá consecuencias negativas en términos no sólo de imagen pública o prestigio en la comunidad, sino que posiblemente represente una merma en el presupuesto económico asignado a la institución. Se sanciona públicamente a los usuarios, ni una sola palabra en cuanto a la calidad de las pruebas suministradas.

Se asume que la razón por la que los estudiantes no aprenden es porque ellos y sus profesores no están haciendo el suficiente esfuerzo requerido para tener éxito. Y la sabiduría convencional ha sido que la forma para estimular un mayor esfuerzo es mediante la intimidación por medio de la amenaza de consecuencias negativas por las bajas puntuaciones de las pruebas.

#### MOTIVACIÓN E IMPORTANCIA DE LA RETROALIMENTACIÓN

El aprendizaje no es sólo un ejercicio cognitivo: implica a la persona en su totalidad. La necesidad de motivar a los alumnos es evidente, pero a menudo se asume que la prestación de recompensas extrínsecas, como calificaciones, estrellitas y premios, es la mejor manera de hacerlo. Sin embargo, hay amplia evidencia para cuestionar este supuesto. Los estudiantes invertirán esfuerzos en una tarea sólo si creen que pueden lograr algo. Si un ejercicio de aprendizaje es visto como una competencia, entonces todo el mundo es consciente de que habrá ganadores y perdedores, y los que tienen un historial como perdedores verán que tiene poco sentido intentarlo. Así, el problema es motivar a todos, aunque algunos están limitados a conseguir menos que otros. Al abordar este problema, el tipo de retroalimentación dado es muy importante. Muchas investigaciones

apoyan esta afirmación (Black *et al.*, 2004). A continuación se presentan algunos ejemplos:

- Los estudiantes a los que se les dice que la retroalimentación *les ayudará a aprender* aprenden más que aquellos a quienes se dice: “como tú dices que eres listo ya sabes que calificaciones obtendrás”. La diferencia es mayor para los alumnos de bajo rendimiento.
- Los estudiantes a quienes se retroalimenta con calificaciones es probable que lo vean como una manera de compararse con los demás (la intervención del ego); en cambio, aquellos a quienes sólo se les hacen comentarios los verán como una ayuda para mejorar (implicación en la tarea). Este último grupo, supera al anterior.
- En un sistema competitivo, los estudiantes de bajo rendimiento atribuyen sus resultados a la falta de capacidad; los estudiantes de alto rendimiento, a su esfuerzo. En un sistema orientado a la tarea, todos los resultados se atribuyen al esfuerzo, y el aprendizaje es mejor, sobre todo entre alumnos de bajo rendimiento.
- Una revisión exhaustiva de la investigación encontró que la retroalimentación mejora el rendimiento en 60% de los estudios. En los casos donde la retroalimentación no fue útil, esta resultó ser sólo un juicio o calificación sin ninguna indicación de cómo mejorar. En general, la retroalimentación dada en forma de recompensas o calificaciones aumenta la implicación del ego en lugar de la participación en la tarea. Se puede centrar la atención de los estudiantes en su capacidad y no en la importancia del esfuerzo, dañando así la autoestima de los alumnos de bajo rendimiento y conduciendo a problemas de indefensión aprendida. La retroalimentación que se centra en lo que hay que hacer puede alentar a todos a pensar que pueden mejorar. Dicha retroalimentación puede mejorar el aprendizaje, tanto directamente, a través del esfuerzo que puede derivarse, e indirectamente, mediante el apoyo a la motivación para invertir en tal esfuerzo.

## REFLEXIONES FINALES

Hoy en día sabemos cómo construir sanos ambientes de evaluación que puedan satisfacer las necesidades de información de todos los que toman decisiones de enseñanza, ayudar a los alumnos a que quieran aprender

y se sientan capaces de aprender y así apoyar un incremento significativo en sus logros de aprendizaje. Pero para alcanzar esta meta se deben establecer mecanismos que hagan posible una evaluación sana, lo cual requerirá que empecemos por ver la evaluación de un modo distinto. El bienestar de nuestros alumnos depende de nuestra buena voluntad para hacerlo así. Si deseamos maximizar el logro de los educandos, debemos poner una mayor atención a la mejora de la evaluación de aula.

## REFERENCIAS

- Black, P., C. Harrison, C. Lee, B. Marshall, y D. Wiliam. (2004). "Working inside the black box: Assessment for learning in the classroom". *Phi Delta Kappan*. 96(1): 9-21.
- Black, P. y D. Wiliam. (1998). "Inside the black box: raising standards through classroom assessment". *Phi Delta Kappan*. 80(2): 139-148.
- Boud, D. (2000). "Sustainable assessment: Rethinking assessment for the learning society". *Studies in Continuing Education*. 22(2): 151-167.
- Ecclestone, K. y J. Pryor. (2003). "Learning careers or assessment careers? The impact of assessment systems on learning". *British Educational Research Journal*. 29(4): 471-488.
- Monereo, C., y M. Castelló. (2009). "La evaluación como herramienta de cambio educativo: evaluar las evaluaciones". En C. Monereo (coord.). *Pisa como excusa. Repensar la evaluación para cambiar la enseñanza* (15-31). Barcelona: GRAÓ.
- Moreno, T. (2004). "Evaluación cualitativa del aprendizaje: enfoques y tendencias". *Revista de la Educación Superior*. 33(3): 93-110.
- Moreno, T. (2007). "La evaluación del aprendizaje en educación superior: El caso de la carrera de Derecho". *Revista Reencuentro. Análisis de problemas universitarios*. 48: 61-67.
- Moreno, T. (2009). "La evaluación del aprendizaje en la universidad: tensiones, contradicciones y desafíos". *Revista Mexicana de Investigación Educativa*. 24(41): 563-591.
- Moreno, T. (2010). "El currículum por competencias en la universidad: más ruido que nueces". *Revista de la Educación Superior*. 39(2): 77-90.
- Stiggins, R. (2002). "Assessment crisis: The absence of assessment for learning". *Phi Delta Kappan*. 83(10): 758-765.
- Webb, N. M. (1997). "Assessing students in small collaborative groups". *Theory into Practice*. 36(4): 205-213.

# Exámenes de dominio del CELE y el MCRE: análisis comparativo

Alma Ortiz Provenzal

Elsa María Esther López del Hierro

María Teresa Mallén Estebaranz

Karen Beth Lusnia

María Aurora Marrón Orozco

Barbara Byer Clark

CENTRO DE ENSEÑANZA DE LENGUAS EXTRANJERAS

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

## RESUMEN

El impacto que el Marco Común de Referencia Europeo (MCRE) ha tenido en el campo de la enseñanza de lenguas extranjeras a nivel internacional es evidente. Aunque dicho documento no fue creado especialmente para la evaluación, ha tenido también mucha influencia en este ámbito. El Centro de Enseñanza de Lenguas Extranjeras (CELE) de la Universidad Nacional Autónoma de México (UNAM), centro certificador de conocimientos de idiomas, no se abstrae de las tendencias internacionales actuales, y cuestiona la pertinencia y el impacto real o construido de diversos parámetros internacionales para tomar decisiones, en este caso concreto, en relación a la certificación de idiomas, pero sin dejar de tomar en cuenta a la vez las variables del entorno educativo propio. Este trabajo muestra avances del equipo que está involucrado en el proyecto correspondiente adscrito al Departamento de Lingüística Aplicada del CELE-UNAM. Se presenta el acercamiento a los documentos del MCRE, especialmente al manual, *Relating Language Examinations to the CEF for Languages: Learning, Teaching and Assessment, 2003*, y a los Anexos publicados en el documento del MCRE, con la finalidad de confrontarlos con los criterios para la elaboración de los exámenes de dominio de la Coordinación de Evaluación y Certificación del CELE-UNAM.

**Palabras clave:** certificación, evaluación, dominio, MCRE, exámenes

## INTRODUCCIÓN

El presente artículo hace un recuento del avance del proyecto “Los exámenes de certificación de Dominio de francés e inglés del Centro de Enseñanza de Lenguas Extranjeras (CELE). Un análisis comparativo con el Marco Común de Referencia Europeo (MCRE) y una propuesta de lineamientos para la actualización de exámenes”. En este proyecto participamos un equipo de profesoras de tiempo completo del Departamento de Lingüística Aplicada (DLA).

El objetivo propuesto, como su título lo indica, consiste en llevar a cabo un análisis comparativo de los exámenes hasta ahora utilizados en la Coordinación de Evaluación y Certificación (CEC) del CELE. Dichos exámenes llamados de ‘dominio’ o de ‘posesión’ del idioma certifican los conocimientos de lengua extranjera de los alumnos de ciertas carreras del campus Ciudad Universitaria (comunicación, relaciones internacionales y actuaría, por ejemplo) de la Universidad Nacional Autónoma de México (UNAM), de los profesores de nivel medio-superior y superior del Sistema Incorporado a la UNAM, así como también de los aspirantes a obtener la acreditación como guías de turistas por la Secretaría de Turismo. La comparación se planteó en relación a los parámetros que el Marco Común de Referencia Europeo (MCRE) ha propuesto, por considerarlo uno de los estándares internacionales más difundidos y conocidos durante los últimos años. La inquietud de realizar esta comparación emana, no de una justificación en el diseño de los exámenes ni de su conceptualización, sino de la necesidad de ampliar sus alcances al establecer equivalencias con estándares internacionales y de la conveniencia de que el universitario sepa qué nivel de dominio de lengua extranjera maneja y lo que significa en dichos términos comparativos.

Las funciones de certificación del CELE han sido una de las tareas primordiales del Centro desde su fundación. En ese transcurrir, el reto del diseño de instrumentos de evaluación se ha realizado de manera seria e incluyente en cuanto a modalidades de evaluación, integrando estilos, conocimientos, técnicas y autenticidad en la selección tanto de textos escritos como auditivos y evaluando el desempeño de tareas reales de estos conocimientos de lenguas. El CELE no se abstrae de las tendencias educativas internacionales actuales y plantea cuestionamientos académicos sobre

la pertinencia y el impacto real o construido de parámetros de evaluación internacionales para poder tomar decisiones, considerando las variables del entorno educativo propio, en este caso las necesidades de certificación de los universitarios de la UNAM.

Quienes trabajamos en el campo de la enseñanza de lenguas, estamos conscientes del gran impacto que el MCRE ha tenido en nuestra profesión y reconocemos sus aportaciones. Sin embargo, en el presente artículo destacamos la necesidad de acercarse a él de manera crítica, para rescatar lo que pueda ser aplicado a nuestro propio contexto y necesidades, y descartar, o quizás adaptar, lo que esté alejado de nuestra realidad.

## OBJETIVO

El objetivo del proyecto de trabajo que aquí presentamos es analizar qué tanto se relacionan los exámenes que existen actualmente en la Coordinación de Evaluación y Certificación (CEC) con el MCRE y cómo se pueden establecer lineamientos para el diseño de exámenes a futuro que tomen en cuenta tanto estos estándares internacionales como las particularidades de nuestra institución y de sus estudiantes.

## METODOLOGÍA DE TRABAJO

En el seno de las discusiones del proyecto, se plantearon varias etapas para el logro del objetivo antes mencionado. En este artículo describiremos las primeras tres etapas: *a)* el análisis de los exámenes de dominio de la CEC de inglés y francés. *b)* El estudio, traducción y adaptación del manual *Relating language examinations to the common European framework of reference for languages: Learning, teaching, and assessment* (CEF) (primera versión, septiembre 2003), a las necesidades de nuestro proyecto y de la CEC. *c)* Estudio y análisis del MCRE y de las escalas y parámetros de la Association of Language Testers in Europe (ALTE) y del proyecto DIALANG, creado por la Unión Europea para diagnosticar el nivel de dominio de lengua (14 lenguas europeas) de los hablantes de lenguas extranjeras. Dicho análisis se hizo para determinar su posible adaptación y uso al contexto de la CEC del CELE y de la UNAM.

## PRIMERA ETAPA

En la primera etapa se realizó la descripción de los exámenes de dominio y posesión de la CEC de inglés y francés. Estos fueron analizados en cuanto a estructuras gramaticales, funciones, vocabulario, estrategias y grado de dificultad de los reactivos, de las secciones y de cada examen en su conjunto.

**Figura 1**  
Exámenes de dominio y posesión de la CEC de inglés y francés

EXAMEN	POBLACIÓN
Examen de dominio de inglés.	Estudiantes de licenciatura: Ciencias Políticas y Sociales, FES-Aragón (licenciatura en Matemáticas, Ciencias Genómicas, Ciencias Ambientales).
Examen de dominio de inglés (guía de turistas).	Aspirantes a ser guía de turistas, certificados por la Secretaría de Turismo.
Examen de dominio de inglés (profesores de nivel medio superior, escuelas incorporadas a la UNAM-DGIRE).	Profesores en servicio en el bachillerato de escuelas del sistema incorporado a la UNAM.
Examen de dominio de francés.	Estudiantes de licenciatura (Ciencias Políticas y Sociales).

Al iniciar el análisis, las descripciones fueron realizadas de manera independiente y cada profesora comentó lo que, a su juicio, parecía interesante o causaba alguna inquietud. Cada miembro del equipo contó con una copia del examen asignado (incluidos videos), y no se predeterminaron criterios para el análisis; es decir, se partió de una descripción 'impresionista' de cada examen, con base en los conocimientos que cada académica aporta a este proyecto por su trayectoria en el diseño de exámenes y en el área de evaluación.

## SEGUNDA ETAPA

Una vez concluidas estas descripciones y comentados los hallazgos, se decidió trabajar con un solo referente para lograr una descripción que pudiera servir a futuro, no únicamente para el proyecto, sino también para compartirlo con otros colegas. Se utilizó la adaptación del manual *Relating language examinations to the common European framework of reference for languages: Learning, teaching and assessment* (CEF) (primera versión, septiembre 2003), el cual se pensó que podría brindar un buen parámetro referencial.

Este primer acercamiento arrojó datos y/o carencias de información en algunos rubros marcados por el manual. Se constató que los exámenes analizados cubren muchos de los aspectos contemplados en dicho documento; sin embargo hay otros que convendría revisar y de los cuales se debería ponderar su importancia. Algunos de estos aspectos son de fondo, mientras que algunos otros son sólo de forma.

La lectura del manual sugiere dos fases en el proceso de especificación de cualquier examen. La primera es una descripción general del instrumento y la segunda consiste en una descripción detallada de este. Para la primera fase se proporcionan las formas de la A1 a la A7 (CEF, 2003: 35-40) que sirven para describir los diferentes aspectos que aparecen a continuación:

- A1: Descripción general del examen.
- A2: Desarrollo del examen.
- A3: La corrección del examen.
- A4: La calificación del examen.
- A5: Reporte de resultados.
- A6: Análisis de datos.
- A7: Fundamentos para toma de decisiones.

Por último, la forma A8 permite registrar, con base en los aspectos anteriores, una impresión general del nivel del examen analizado.

El hecho de utilizar los formatos contenidos en este manual para referenciar exámenes nos permitió utilizar una base común para analizar los diferentes instrumentos, ver las coincidencias y las diferencias que había entre ellos, así como tener un desglose de elementos a analizar en las diferentes etapas del examen (diseño, presentación, aplicación, calificación y

reporte de resultados). Este manual sólo está disponible en inglés y, dado que nos pareció muy útil al proporcionarnos una forma consistente y sistemática de analizar y de establecer concordancias y diferencias entre los exámenes, el grupo de trabajo decidió traducirlo para que pueda servir como un referente para cualquier idioma, y, en un futuro, adecuarlo a la propuesta de este proyecto al tomarlo como punto de partida.

A manera de ejemplos, incluimos nuestra traducción de la forma A3: La corrección del examen y de la forma A5: Reporte de los resultados.

**Figura 2**  
Forma A3: La corrección del examen

CORRECCIÓN DEL EXAMEN: SECCIÓN _____	COMPLETE UNA COPIA DE ESTA FORMA PARA CADA SECCIÓN DEL EXAMEN. DESCRIPCIÓN CORTA Y/O REFERENCIA
¿Cómo se corrige cada actividad de la prueba?	<input type="checkbox"/> Para actividades receptivas: <input type="checkbox"/> Lector óptico <input type="checkbox"/> Corrección por expertos locales/profesores <input type="checkbox"/> Para actividades productivas o integradas: <input type="checkbox"/> Examinadores entrenados <input type="checkbox"/> Profesores
¿Dónde se corrigen las actividades de la prueba?	<input type="checkbox"/> En el organismo que maneja el examen <input type="checkbox"/> En la institución que aplica el examen <input type="checkbox"/> Por equipos locales <input type="checkbox"/> Por examinadores individuales
¿Qué criterios se usan para seleccionar al personal que corrige?	
¿Cómo se asegura la exactitud al corregir?	<input type="checkbox"/> Revisiones periódicas del coordinador <input type="checkbox"/> Entrenamiento para el personal que corrige y/o evalúa <input type="checkbox"/> Sesiones de discusión para estandarizar criterios <input type="checkbox"/> Uso de ejemplos estandarizados de actividades de la prueba <input type="checkbox"/> Calibrados en relación al MCRE <input type="checkbox"/> Calibrados a otra escala de descripción <input type="checkbox"/> No calibrados

CORRECCIÓN DEL EXAMEN: SECCIÓN _____	COMPLETE UNA COPIA DE ESTA FORMA PARA CADA SECCIÓN DEL EXAMEN. DESCRIPCIÓN CORTA Y/O REFERENCIA
<p>Describa las especificaciones de los criterios de evaluación de las actividades productivas y/o integrativas de la prueba.</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Un resultado holístico para cada actividad</li> <li><input type="checkbox"/> Calificaciones para diferentes aspectos de cada actividad</li> <li><input type="checkbox"/> Una escala de calificación para el desempeño general en el examen</li> <li><input type="checkbox"/> Una tabla de calificaciones para aspectos del desempeño en el examen</li> <li><input type="checkbox"/> Una escala de calificaciones para cada actividad</li> <li><input type="checkbox"/> Una tabla de calificaciones para aspectos de cada actividad</li> <li><input type="checkbox"/> Se definen bandas en una escala de calificaciones. No de acuerdo al MCRE</li> <li><input type="checkbox"/> Se definen bandas en una escala de calificaciones en relación al MCRE</li> </ul>
<p>¿Las actividades productivas o integrativas de la prueba se califican por uno o dos evaluadores?</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Un solo evaluador</li> <li><input type="checkbox"/> Dos evaluadores simultáneamente</li> <li><input type="checkbox"/> Doble corrección de escritos o de grabaciones</li> <li><input type="checkbox"/> Otro: (especifique) _____</li> </ul>
<p>¿Si intervienen dos evaluadores, qué procedimientos se usan cuando existen diferencias entre evaluadores?</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Se usa un tercer evaluador y la calificación que este otorga es la que se sostiene</li> <li><input type="checkbox"/> Se solicita una tercera calificación y se utilizan las dos calificaciones más cercanas</li> <li><input type="checkbox"/> Se promedian las dos calificaciones</li> <li><input type="checkbox"/> Los dos evaluadores discuten y llegan a un acuerdo</li> <li><input type="checkbox"/> Otro: (especifique) _____</li> </ul>
<p>¿Se calcula la confiabilidad entre evaluadores?</p>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Sí</li> <li><input type="checkbox"/> No</li> </ul>

Figura 3  
Forma A5: Reporte de los resultados

RESULTADOS	DESCRIPCIÓN CORTA Y/O REFERENCIA
¿Qué resultados se reportan a los candidatos?	<input type="checkbox"/> Una calificación global o aprobado/no aprobado <input type="checkbox"/> Calificación o aprobado/no aprobado por cada sección de la prueba <input type="checkbox"/> Calificación global más un perfil por cada sección de la prueba <input type="checkbox"/> Perfil de aspectos de su desempeño por cada sección de la prueba
¿En qué forma se reportan los resultados?	<input type="checkbox"/> Puntaje <input type="checkbox"/> Calificaciones no definidas <input type="checkbox"/> Un nivel en una escala definida <input type="checkbox"/> Un diagnóstico personalizado
¿En qué documento se reportan los resultados?	<input type="checkbox"/> Carta o correo electrónico <input type="checkbox"/> Boleta de calificaciones <input type="checkbox"/> Certificado/diploma
¿Se da información para ayudar a los candidatos a interpretar sus resultados? Dar detalles.	
¿Tienen los candidatos el derecho de ver sus exámenes corregidos y calificados?	
¿Tienen los candidatos el derecho de pedir que su examen vuelva a corregirse/calificarse?	

### TERCERA ETAPA

El hecho de plantear una comparación con base en un parámetro establecido y utilizado mundialmente implica conocerlo lo más a fondo posible, saberlo analizar y tener claros sus propósitos, para, en su caso, adaptarlo al contexto propio. Este equipo de investigación se abocó por tanto a la lectura y análisis del *Common European framework of reference for languages: Learning, teaching, assessment* (Council of Europe, 2001) para lograr la meta comparativa. Aunque parezca extraño, en esta fase inicial nos concentra-

mos en los Anexos del MCRE, pues en ellos se encuentra información relevante sobre los orígenes de este documento, así como sobre los criterios que fueron tomados en cuenta para su elaboración, lo cual revela mucho de su fundamentación y de sus metas en el espacio y tiempo en el que fue desarrollado. Aunque en el MCRE en sí estos anexos aparecen en un orden diferente, para la presentación en este artículo decidimos seguir un orden más bien cronológico, considerando que da mayor claridad a la evolución y a las aportaciones del documento en sus distintas fases. A continuación presentaremos un resumen de los contenidos del anexo, así como una reflexión sobre las implicaciones que puede (o no) tener para nuestro propio contexto.

Las escalas ilustrativas de los descriptores

Empezaremos por el anexo que describe los orígenes del MCRE y detalla el proceso de su creación. Este aparece como *Appendix B: The illustrative scales of descriptors en el documento original*.

El uso de los descriptores del MCRE parece ahora muy común y se consideran de utilidad ya que describen con cierto detalle lo que los alumnos pueden hacer en la lengua meta. Sin embargo, en algunas ocasiones los profesores sentimos que lo que plantean algunos de esos descriptores, especialmente en los niveles de B2 a C2, es difícil de lograr por muchos de nuestros alumnos. Entre los profesores hemos comentado en ocasiones que algunos de nuestros estudiantes no alcanzan esos niveles ni siquiera en su lengua materna (por ejemplo, el dominio de la redacción). Esto nos lleva a cuestionar qué tan factible es que un alumno de lengua extranjera en nuestro contexto logre alcanzar un nivel C2, que describe más bien el perfil de un hablante nativo educado.

El apéndice B da cuenta de la historia de los descriptores, lo cual es de utilidad para entenderlos mejor y ubicarlos en su propio contexto. Estos descriptores son el producto de un proyecto de investigación suizo que se llevó a cabo entre 1993 y 1996, y cuyo propósito fue desarrollar descriptores transparentes sobre habilidades o competencias descritos en diferentes aspectos del esquema del MCRE y que pudieran contribuir al desarrollo del Portafolio Europeo.

Los descriptores se hicieron con base en encuestas y evaluaciones realizadas por profesores de lengua extranjera (primero de dominio de inglés, luego de francés y alemán). A las evaluaciones de los profesores se agre-

garon autoevaluaciones e información de exámenes de la Universidad de Cambridge, del Instituto Goethe y del Ministerio de Educación Francés. Participaron casi 300 profesores y 2800 aprendientes de unos 500 grupos de niveles secundaria, bachillerato, vocacional y educación de adultos, de las regiones en Suiza de habla alemana, francesa, italiana y romansch (los dialectos utilizados en Grisons, Suiza).

Para esta investigación se tomaron algunas escalas ya existentes, por ejemplo, las del American Council of the Teaching of Foreign Languages (ACTFL *Proficiency Guidelines*), algunas holísticas, otras de cuatro habilidades, otras de evaluación oral, de actividades comunicativas y algunos marcos de evaluación de sílabos. Estas escalas fueron analizadas y de-construidas para ubicarlas en categorías. Se hicieron grabaciones de profesores para corroborar que el metalenguaje usado por ellos estuviera representado. Posteriormente, se organizaron talleres con profesores donde clasificaron los descriptores en las diferentes categorías, juzgaron su claridad, relevancia y precisión, y los colocaron en bandas o rangos.

En 1994 los profesores evaluaron a alumnos de distintos niveles (como mínimo 80 horas de inglés hasta avanzados) utilizando siete cuestionarios con 50 descriptores. En 1995 se repitió el estudio pero con cinco cuestionarios diferentes. En ambas situaciones se utilizó el mismo material para analizar la interacción oral.

Los resultados se analizaron con el modelo Rasch para establecer de manera matemática un valor de dificultad por descriptor e identificar la variación estadísticamente significativa en la interpretación de descriptores. Durante este proceso se descartaron ciertas categorías. Se encontró que los descriptores expresados en términos negativos funcionaban mejor con una redacción afirmativa. Es interesante destacar que la escala de lectura se hizo posteriormente de forma independiente y después se relacionó con los demás descriptores. Los descriptores de producción escrita fueron desarrollados a partir de la producción oral. Finalmente produjeron las escalas de seis niveles que conocemos actualmente (A1, A2, B1, B2, C1 y C2).

Entender un poco la historia de la elaboración de los descriptores nos ayuda a tener otra perspectiva del MCRE en México. No enseñamos a alumnos suizos. Ni somos profesores suizos. Nuestro contexto es muy diferente. Por mencionar sólo un aspecto, en nuestro país no hay tanta movilidad de alumnos a los países donde se habla la lengua meta, como la hay en Europa. Justamente así como los participantes en el proyecto

suizo se apoyaron en una variedad de escalas anteriores para luego crear las propias, nosotros debemos conocer lo que el MCRE ofrece, analizar otras opciones y crear lo que mejor sirva, acorde con nuestra propia realidad educativa.

Elaboración de los descriptores de dominio de la lengua (competencia)

El segundo anexo al que haremos referencia, llamado *Appendix A: Developing proficiency descriptors* en el MCRE, se centra en la formulación de los descriptores y en las metodologías que se siguieron para el desarrollo de las escalas en términos: intuitivo, cualitativo y cuantitativo.

En esta sección se describen diversas características importantes tomadas en cuenta para la redacción de los descriptores. A continuación mencionamos algunas de ellas.

- *Formulación positiva.* Cuando se requiere describir niveles de dominio bajo o insuficiente, es común utilizar formulaciones negativas o restrictivas, ya que resulta más fácil referirse a lo que no se sabe o no se puede hacer que a lo que sí se es capaz de hacer en esos niveles iniciales. Por tanto, si se quiere que las escalas de competencias sean útiles no sólo para definir objetivos, sino además para seleccionar a los examinados, es necesario formular las competencias de manera positiva. En ciertos casos se puede formular el mismo enunciado de manera positiva y negativa.

Ejemplo:

- a) Formulación positiva: “Expresa y reconoce un conjunto de palabras y expresiones breves aprendidas de memoria” (Trim 1978, nivel 1, *apud* Saville, 2005).
- b) Formulación negativa: “Únicamente es capaz de expresarse por medio de enumeraciones de enunciados convencionales” (Principiante ACTFL).

Otra de las dificultades que existen al redactar formulaciones negativas es que algunas características del dominio lingüístico no son acumulativas; por ejemplo, el grado en que el alumno depende de algo o alguien para comunicarse es difícil de expresar en términos negativos: “Interactúa con razonable facilidad en situaciones

estructuradas y conversaciones breves siempre que el interlocutor ayude, si es necesario”.

- *Precisión.* Otro de los factores tomados en cuenta al redactar en el MCRE fue la descripción de las características concretas de la actuación, de las tareas y/o niveles precisos de destreza al realizar tareas (*performance*). Por ello se evitaron las vaguedades; por ejemplo en “Puede utilizar una gama de *estrategias adecuadas*” las partes en cursivas no serían acertadas, ya que se considerarían ambiguas. Asimismo, este anexo sugiere que para distinguir los grados entre las escalas no se debe sustituir únicamente el uso de un cuantificador por otro. Ejemplos: *unos pocos-muchos, algún-la mayoría, bastante amplio-muy amplio-moderado*. Se señala que las distinciones deben ser “reales” no sólo verbales; sin embargo se reconoce que las diferencias concretas y significativas no siempre se pueden enunciar con facilidad.
- *Claridad.* Una característica importante es que los descriptores deben ser transparentes y no estar dominados por la “jerga”. La sintaxis debe ser sencilla y la estructura lógica y explícita.
- *Brevedad.* En cuanto a este aspecto particular de la redacción de los descriptores se hace referencia a dos escuelas o tendencias. Por un lado, la holística, que sugiere escribir un párrafo extenso que trate de cubrir de manera exhaustiva los aspectos que se consideran esenciales. Esto tendría como objetivo hacer un “retrato detallado del alumno tipo para un nivel X”. En relación a esta propuesta se mencionan dos inconvenientes: 1) ningún individuo es realmente típico y 2) los profesores, en general, prefieren manejar descriptores cortos.

Por otro lado, la otra tendencia propone la independencia entre los descriptores. Los descriptores cortos presentan dos ventajas: 1) es más probable que describan un comportamiento específico y fundamental: si la persona es o no capaz de hacer algo; y 2) son independientes unos de otros, por lo que pueden servir para formular cuestionarios de evaluación continua, de autoevaluación y de objetivos.

Además de describir las características particulares de los descriptores, en ese anexo se mencionan las metodologías que se siguieron para el desarrollo de las escalas.

Para determinar qué descriptores se pueden poner en un nivel y no en otro y cuáles descripciones, con cierto grado concreto de destreza, pertenecen a un nivel y no a otro, lo más recomendable, según este anexo, es combinar tres enfoques: el intuitivo, el cualitativo y el cuantitativo en un proceso complementario y acumulativo.

En algunas ocasiones sentimos que en ciertos contextos el MCRE se trata de imponer como una camisa de fuerza, cuando en sus orígenes no fue concebido así, sino, como su nombre lo dice, como un marco de referencia, cuyo carácter es más bien descriptivo que prescriptivo. Este anexo del MCRE nos pareció de utilidad, pues nos da pautas de los aspectos que fueron considerados al redactar los descriptores para poder ubicarlos en las distintas escalas, y nos ayudará a formular descriptores particulares de acuerdo a nuestro propio contexto y necesidades, sin olvidar tomar en cuenta las diferentes características mencionadas en este anexo para la redacción de los descriptores y para hacerlo de forma más fundamentada y consistente.

#### El proyecto DIALANG

En este apartado nos referiremos al proyecto DIALANG (*Appendix C: The DIALANG Project*), el cual es una aplicación del MCRE para fines de diagnóstico. DIALANG fue iniciado en 1996 explícitamente para desarrollar pruebas en 14 lenguas europeas y está pensado para alumnos de idiomas que se interesan en diagnosticar el dominio de la lengua que estudian. Es decir, se dirige a los adultos que quieren conocer su nivel de dominio de una lengua y obtener retroalimentación sobre las cualidades y las insuficiencias de dicho dominio. El proyecto se llevó a cabo con la ayuda financiera de la Unión Europea, entre otros organismos.

El sistema DIALANG se compone de autoevaluación, pruebas lingüísticas y retroalimentación que están disponibles en cada una de las lenguas europeas que ofrece. Se puede acceder a este sistema a través de Internet de forma gratuita. No concede certificados. El marco de evaluación y las escalas descriptivas utilizadas para informar de los resultados a los usuarios se basan directamente en el MCRE. La mayoría de las especificaciones de autoevaluación empleadas también se ha extraído del MCRE y se ha adaptado para satisfacer las necesidades específicas del sistema.

DIALANG tiene pruebas en cinco aspectos del lenguaje y de su uso: lectura, comprensión auditiva (indirecta), escritura, gramática y vocabulario. Los resultados en los seis niveles del MCRE se reportan en cada una de estas pruebas de manera separada y no se suman para dar un resultado que abarque todas las secciones. Individualmente cada reactivo está diseñado para evaluar aspectos particulares de cada habilidad. Los resultados de estas sub-habilidades se reportan de manera separada para que los usuarios puedan ver sus respuestas a cada reactivo.

Además de los resultados de sus pruebas se les da retroalimentación de dos tipos: el primer tipo de retroalimentación es sobre la diferencia entre su autoevaluación y los resultados de sus pruebas, y se les invita a explorar si hay discrepancias en los resultados y las posibles razones. Esto se hace con el fin de aumentar su “conciencia lingüística” (*language awareness*) y ayudarlos a diagnosticar mejor sus fortalezas y debilidades. El segundo tipo es lo que DIALANG llama “*advisory feedback*” que son consejos que pueden ayudar al aprendiente a progresar del nivel en el que está al siguiente rango señalado en el MCRE.

Se dice que es un sistema de diagnóstico a diferentes niveles. A nivel macro porque en las diferentes habilidades puede ayudar a los usuarios a decidir si están listos para tomar algún examen comercial o público de determinado nivel o si aún necesitan más trabajo o tomar un examen de nivel más bajo en el MCRE. A nivel micro porque el usuario podrá conocer sus áreas débiles y concentrarse en trabajar en ellas.

Este anexo nos parece de gran utilidad pues, a pesar de que el MCRE ha sido ampliamente difundido en nuestro campo de trabajo y es conocido por gran cantidad de profesores, no ocurre lo mismo con DIALANG. Lo que muchos docentes conocemos son las tablas que contienen las escalas y los descriptores del MCRE. Saber que DIALANG es una aplicación práctica del Marco con fines de diagnóstico puede ser de gran utilidad al permitirnos conocer ejemplos concretos y específicos de exámenes que están cuidadosamente calibrados con respecto al MCRE, así como a los alumnos quienes no requieren de un conocimiento teórico del MCRE pero a quienes este instrumento de diagnóstico y de autoevaluación puede darles mucha orientación en términos sencillos y prácticos.

Los niveles ALTE y los enunciados “*can do*”

El último anexo del MCRE está dedicado a las escalas realizadas por la Asociación de Evaluadores de Lengua en Europa (ALTE) y más particularmente a los *can do statements*. Por cierto, vale la pena recordar que el trabajo de esta asociación en cuanto a la descripción de los distintos niveles de lenguas extranjeras puede ser considerado como precursor del MCRE, pues empezaron a definir los “niveles claves” desde hace ya muchos años como parte de un proyecto de investigación a largo plazo que incluía 12 lenguas representadas en la asociación. La finalidad principal en cuanto a una clara definición de los niveles, y de una descripción de lo que los hablantes pueden hacer en cada uno de estos, estaba orientada principalmente a los usuarios de los distintos exámenes (empleadores, personal de la universidad, por ejemplo), quienes no sabían cómo interpretar los resultados de exámenes de certificación, y a los mismos examinados, al darles una idea clara de lo que ya pueden hacer y lo que aún les falta por lograr en la lengua extranjera.

Es importante destacar que los descriptores de los niveles dentro de ALTE son parámetros para medir el nivel de idioma en tres áreas: *a)* social y turismo, *b)* trabajo y *c)* estudio. Incluye además descriptores por habilidades generales en el idioma.

Estas habilidades están redactadas en cuanto a tareas específicas y prácticas que el hablante puede realizar en las tres áreas antes mencionadas en: comprensión auditiva, producción oral, comprensión de lectura y producción escrita. Actualmente existen alrededor de 400 enunciados (*can do statements*), lo cual da una buena idea de su especificidad.

Los organismos que administran exámenes han tratado de homologar o alinear sus exámenes con el MCRE para tener un punto de referencia para la comparación de exámenes buscando demostrar una equivalencia que les pueda dar una validez ante el público que toma estos exámenes y los necesitan para trabajos, escuelas, etc. Por lo tanto, también los niveles ALTE tienen sus equivalencias con los niveles del Consejo de Europa. Esta equivalencia, si bien no es exacta, ni tiene por qué serlo (por ejemplo, en el ALTE sólo existen cinco niveles pues no hay una equivalencia explícita con el nivel A1 del MCRE), da al usuario y a la institución aplicadora un punto de

referencia en cuanto a dónde están ubicados sus exámenes en este contexto de evaluación y certificación internacional.

Ahora bien, para la tarea práctica de certificar a los alumnos, por ejemplo, en nuestro caso en el CELE, las escalas ALTE nos servirían de esta misma forma. Los descriptores de nivel en estas tres áreas: social y turismo, trabajo y estudio se podrían comparar de una forma práctica con los niveles que alcanzan los alumnos y profesionistas que se presentan en nuestro centro para los diferentes propósitos para los cuáles necesitan la certificación. Si bien la certificación en el CELE solamente sirve para propósitos nacionales y locales, comparar los descriptores de cada nivel y los rangos de reactivos correctos con los niveles ALTE y MCRE haría más explícita la correlación que existe entre los exámenes del CELE y aquellos que se aplican en un contexto internacional.

Otra reflexión importante es que los descriptores de ALTE son prácticos y concretos para contextos específicos, lo que los hace útiles para la elaboración de unos niveles o estándares para contextos con necesidades similares en las tres áreas ya mencionadas. Tener un punto de comparación con el ALTE, MCRE u otros exámenes internacionales serviría para que los alumnos supieran dónde están ubicados en un mundo globalizado en cuanto a su dominio de la lengua extranjera, por si viajan a otros países por su trabajo o planean estudiar en el extranjero.

Tener descriptores accesibles y fáciles de entender sobre tareas prácticas similares a las representadas en el ALTE nos asegura que estos descriptores serán tomados en cuenta en la elaboración de los exámenes y que los resultados de estos exámenes nos darán información sobre el nivel de dominio de los usuarios que los presenten.

## CONCLUSIÓN

A lo largo de este artículo hemos presentado y comentado los diversos anexos que se encuentran en el MCRE, resaltando algunos aspectos que nos parecen útiles y cuestionando algunos otros que no se adecuan del todo a nuestra realidad y necesidades. Asimismo, hemos apuntado, aunque de manera breve, que existen otras escalas (como las de ALTE) que también es necesario considerar. El estudio de estos anexos nos ha permitido tener un panorama más amplio del MCRE en sí, conocer sus orígenes, sus propósitos, sus aportaciones y aplicaciones. Sin embargo, recalcamos la necesidad de

tener un acercamiento crítico hacia él, al reconocer también sus limitaciones, debido a las diferencias entre el contexto europeo y el nuestro, entre otras.

Las siguientes etapas del proyecto deberán contemplar el análisis profundo de las tablas de niveles de referencia común: *Table 2. Common reference levels: Self-assessment grid*, (CEF, 2003: 26-27) que son las que comúnmente se revisan o consultan sin mayor documentación. De igual manera el capítulo cuatro es clave para una futura adaptación o comparación de lo que ofrece el trabajo del MCRE al trabajo planteado en este proyecto. Con base en la información ya mencionada, elaboraremos nuestro propio constructo de dominio del idioma (de acuerdo con las necesidades particulares de los alumnos de la UNAM), propondremos nuevos lineamientos para los exámenes y diseñaremos formatos de resultados que ayuden a los alumnos a entender lo que significa su resultado no sólo a nivel local sino también en relación con otros estándares internacionales. Vemos pues el MCRE no como una camisa de fuerza sino como un documento que nos ayuda a ubicarnos en el contexto internacional, pero al que hay que acercarse de forma crítica y selectiva, con base en nuestras propias necesidades y realidades.

## REFERENCIAS

- Association of Language Testers in Europe. (2008). Consultado el 7 de agosto de 2008 en <http://www.alte.org/>
- Centro Virtual Cervantes. (2008). *Marco de Referencia Europeo*. Consultado el 2 de junio de 2008 en <http://cvc.cervantes.es/obref/marco/indice.htm>
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Council of Europe. (2009). *Manual for Relating Language Examinations to the Common European framework of reference for languages: Learning, teaching and assessment*. Consultado el 7 de Julio de 2009 en [http://www.coe.int/t/dg4/linguistic/Manuel1\\_EN.asp#TopOfPage](http://www.coe.int/t/dg4/linguistic/Manuel1_EN.asp#TopOfPage)
- DIALANG. Consultado el 13 de octubre de 2011 en <http://www.lancs.ac.uk/researchenterprise/dialang/about.htm>

*Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, teaching and assessment.* (2003). DGIV/EDU/LANG (2003) 5 rev.1

Saville, N. (2005). "Interview: An Interview with John Trim at 80", *Language Assessment Quarterly*. 2 (4), 263-288. Consultado el 30 de mayo de 2008 en <http://www.informaworld.com/smpp/title~content=t775653669~db=all~tab=isse.slist~branches=2 - v2>