



DESARROLLO Y VALIDACIÓN DE UNA PRUEBA PARA EL DIAGNÓSTICO DE LOS APRENDIZAJES DEL IDIOMA INGLÉS EN EDUCACIÓN SECUNDARIA

Maite Rodríguez de la Vega Elu
Instituto de Investigación y Desarrollo Educativo

Juan Carlos Pérez Morán
Instituto de Investigación y Desarrollo Educativo

Ana Sofía Pérez Salido
Facultad de Idiomas

Área temática: Evaluación educativa.

Línea temática: Diseño y validación de instrumentos.

Tipo de ponencia: Reportes parciales o finales de investigación.

Resumen:

En las últimas décadas el uso de instrumentos para evaluar el dominio de una segunda lengua es una práctica común de los sistemas educativos del mundo para el diseño de políticas educativas que fortalezcan el intercambio y desarrollo económico y cultural entre las naciones (OCDE, 2014). El propósito del presente estudio es diseñar y validar una prueba diagnóstica de la expresión escrita del inglés como segunda lengua, alineada al currículum de tercero de secundaria del Sistema Educativo Nacional (SEN). El método integró las propuestas metodológicas de Nitko (1994) y Contreras (2009) para el diseño de pruebas criterioles alineadas al currículum, así como los modelos para el diseño de pruebas cognitivas propuestos por Mislevy (2007) y Embretson (2001). El procedimiento constó de cuatro etapas: (1) Definición y análisis de dominios establecidos en el currículum; (2) Desarrollo de los ítems; (3) Validación del aspecto de contenido y del proceso de respuesta de los ítems; (4) Piloteo y (5) Análisis de las propiedades psicométricas. Los resultados fueron satisfactorios mostrando el logro de los criterios y estándares de las evidencias de validez de constructo de los aspectos del contenido, del proceso de respuesta y de la estructura interna de la prueba diseñada. Se puede concluir que se lograron cabalmente los objetivos del diseño y validación de la prueba, sin embargo, es importante aplicar un piloteo con una muestra lo suficientemente grande como para realizar análisis más sofisticados que aporten a las evidencias de la calidad técnica de la prueba.

Palabras claves: evaluación diagnóstica, inglés, segunda lengua, currículum, educación secundaria.

Introducción

La enseñanza del inglés como segunda lengua se incorpora al currículum de Secundaria en la Reforma Educativa de 2009, con el propósito de acrecentar y consolidar las prácticas sociales del lenguaje. En la elaboración del Programa Nacional de inglés en Educación Básica (PNIEB) se incluyen objetivos del dominio de comprensión y producción oral y escritas en los niveles de dominio B1 del Marco Europeo de Referencia para la Lengua (Subsecretaría de educación media superior, 2014). Sin embargo, desde entonces no se ha llevado a cabo una evaluación o investigación con el fin de conocer la eficiencia del Programa (Ramírez, 2012). Por tanto, se desconoce el nivel de dominio de la lengua de los estudiantes de secundaria y la eficiencia del programa de enseñanza del inglés de acuerdo, es decir, si los estudiantes alcanzan el nivel de dominio B1 establecido en los objetivos del currículum nacional (Subsecretaría de Educación Media Superior, 2014).

Se pueden identificar dos distintos usos a las pruebas diseñadas para la evaluación del logro del aprendizaje del inglés como segunda lengua: la certificación del nivel de dominio de un individuo y la de logro de objetivos de aprendizaje. Las pruebas con fines de certificación del dominio del idioma inglés gozan de reconocimiento y aceptación internacional y son un producto ampliamente comercializado por las agencias e instituciones encargadas de su desarrollo y aplicación. Sin embargo, las pruebas alineadas al currículum no presentan el mismo caso. Los ejemplos de pruebas diseñadas para evaluar el logro de la enseñanza del idioma inglés en un contexto o programa específico con la finalidad, no de certificar, sino de conocer el avance alcanzado por los estudiantes y buscar las áreas de oportunidad de mejora a los programas, son escasos.

La enseñanza de idiomas y su evaluación son procesos necesarios y complementarios (Hussain, 2015). Mientras que la evaluación del aprendizaje de un idioma ayuda a los alumnos a conocer su avance, a los maestros aporta información de las fortalezas, debilidades de sus estudiantes (Allison, 1999). Las características de los instrumentos que midan estos aspectos deben ser definidas en relación al propósito a que se servirá. Así, la evaluación del logro de los estudiantes o de la efectividad de un programa debe ser realizada de acuerdo con los dominios establecidos por el currículum; en el caso de la escritura en inglés la evaluación debe ser acorde a los objetivos establecidos en éste (CAP, 1990).

Lo anterior lleva al planteamiento de la necesidad de contar con una evaluación del desempeño de los alumnos de secundaria en el sistema público de educación en México con el fin de conocer la eficiencia del programa de enseñanza del inglés como lengua adicional y de diagnosticar el grado de dominio de los estudiantes al ingreso, durante y al final del programa. Adicionalmente aportaría conocimiento útil para mejorar la instrucción y el desempeño.

El propósito general de este trabajo es diseñar y validar una prueba de la expresión escrita del inglés como segunda lengua, alineada al currículum de tercero de secundaria del Sistema Educativo Nacional (SEP) para lo que es necesario analizar el Plan Nacional de la Enseñanza del Inglés en Educación Básica (PNIEB), conocer sus contenidos y dimensiones y la medida en la que estos se encuentran alineados con el MCER y los modelos cognitivos relacionados, especialmente con la expresión escrita. Adicionalmente, como

aportación metodológica se propone la definición del constructo de la prueba con base en el análisis de los contenidos del Programa de Inglés de Educación básica PNIEB y la adaptación de un modelo de diseño fuerte basado en evidencias para la evaluación de la competencia de expresión escrita en el idioma inglés como segunda lengua.

Desarrollo

El Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza y evaluación (MCER) es un documento presentado por el Consejo Europeo en 2001, como resultado de la investigación durante más de 10 años por parte de especialistas en el campo de la lingüística aplicada cuya finalidad es promover una mayor unidad entre los países miembros facilitando la comunicación, la interacción y el intercambio entre los miembros del conforman el Consejo Europeo. Este marco fue diseñado con la finalidad de proporcionar una base común para la elaboración de programas de aprendizaje de lenguas, orientaciones curriculares y exámenes en toda Europa (Ministerio de Educación Cultura y Deporte, 2002); además de otros usos como son la planificación de certificados, en particular de los contenidos de los exámenes y de los criterios de evaluación en función del rendimiento positivo. El marco analiza y recopila de manera sistemática los estudios sobre el aprendizaje y enseñanza de lenguas y es utilizado como obra de consulta para el desarrollo de planes curriculares y estudios de investigación a nivel mundial (Ministerio de Educación, Cultura y Deporte, 2002).

El PNIEB incluye los programas de estudio para los niveles de escolar, primaria y secundaria, elaborados de manera alineada con estándares nacionales e internacionales, particularmente con los niveles de dominio del MCER para las Lenguas. Los contenidos se diseñaron considerando que los alumnos alcancen al final de la educación básica el nivel de logro BI; de acuerdo con esto el estudiante de secundaria debe ser capaz de participar activamente en distintos contextos sociales mediante el uso de las habilidades receptivas y productivas de la comunicación y para hacerlo poner en práctica los conocimientos adquiridos durante los niveles de educación básica previos. Según lo señalado en el programa las habilidades sociales del lenguaje se llevan a cabo en tres ambientes; las relacionadas con el lenguaje oral en el ambiente familiar y comunitario, la comprensión lectora en el lúdico y de producción escrita en el ambiente literario y académico (PNIEB, 2011).

La investigación cognitiva tiene como objeto la comprensión de la actuación, el aprendizaje y el desarrollo humanos enfocándose en las diferencias individuales mediante el análisis de los procesos del pensamiento. Dos de los modelos conitivos más revisados del proceso de escritura son los propuestos por Hayes y Flower de 1980, posteriormente modificado por Hayes en 1996 (MacArthur, 2017); y el de Bereiter y Scardamalia de 1987 (Cushing, 2002). El modelo de Hayes define dos aspectos; el ambiente de la tarea y el individuo. El primero se compone del ambiente social o la audiencia, estableciendo una relación recíproca entre éste y los procesos cognitivos y el físico que incluye el texto ya escrito, los esfuerzos posteriores y los medios

de composición como son la escritura a mano o el uso del procesador. El segundo aspecto, el individual es central en el modelo Hayes e involucra la interacción entre cuatro componentes la memoria de trabajo, el motivacional y afectivo, los procesos cognitivos y la memoria a largo plazo (MacArthur, 2017).

Bereiter y Scardamalia proponen dos modelos para describir el proceso de escritura distinguiendo entre la narración de conocimiento y la transformación de conocimiento. El primero requiere de poca planeación y revisión por lo que lo llaman natural; los niños y adolescentes realizan este tipo de narración (Cushing, 2002). La transformación de conocimiento involucra más esfuerzo incluye la utilización de la escritura para crear conocimiento nuevo, la transformación de la opinión del propio escritor y lo que desea comunicar (Mac Arthur, 2017). Bereiter y Scardamalia clasifican a los escritores en novatos y expertos. Los escritores novatos utilizan una versión simplificada de los procesos propuestos por Hayes (Mac Arthur, 2017) superan la falta de recursos lingüísticos por medio de tres componentes; la representación mental o el tema dentro de un contexto escolar; el conocimiento previo o esquema mental del tema extraído de la memoria a largo plazo y el texto producido durante el mismo proceso de narración para producir contenido adicional.

En contraste, el escritor experto utiliza estrategias de escritura sustancialmente diferentes y más refinadas para la elaboración de textos en los que transforman el conocimiento. Esto explica el hecho de que las tareas de escritura difieren en grado de dificultad aún entre el segundo tipo de escritores lo que es un concepto importante para la enseñanza y la evaluación de la expresión escrita (Cushing, 2002.) El paso entre relato o narración de conocimiento al de transformación de conocimiento es un proceso largo que pasa por diferentes etapas intermedias (Mac Arthur, 2017).

Para la aplicación de los modelos cognitivos en la evaluación, se debe desarrollar y validar un modelo mediante cuatro pasos: el planteamiento de la teoría; la generación de contextos en los que se observe y compruebe la teoría; el análisis de los resultados de acuerdo con los resultados predichos por el modelo y la elaboración de conclusiones basadas en los resultados. Adicionalmente, el diseño y desarrollo de evaluación en procesos cognitivos consiste en la creación y ejecución de una secuencia de acciones de evaluación con un objetivo específico. Este proceso involucra decisiones con respecto a la teoría o procesos cognitivos, el contenido y la interpretación. La teoría contiene las suposiciones sobre el modo en el que los estudiantes adquieren una competencia; el contenido o la observación que se utiliza para obtener evidencias sobre un aspecto del aprendizaje o cognición y el método o interpretación utilizado para hacer inferencias basadas en la evidencia obtenida (Nichols, 2017). La validez de las inferencias hechas a partir de los resultados de la prueba depende de la conexión explícita entre la teoría o procesos cognitivos, la observación o contenido y la interpretación o inferencias realizadas.

Este concepto de validez coincide con los revisados en la literatura sobre diseño de pruebas; los Estándares para el uso de pruebas psicológicas y educativas, 1999, definen validez como: “el grado en que la evidencia y la teoría apoyan las interpretaciones de las puntuaciones en las pruebas” (Pedrosa, 2013). La función de la validez ha cambiado, de enfocarse en la predicción a hacerlo en la explicación ya que las predicciones de

utilidad y relevancia requieren de la interpretación empíricamente sustentada de los resultados (Messick, 1993). Es decir, no existen pruebas válidas, sino que las pruebas son válidas para algo por lo que es imperante comunicar a los usuarios de la prueba las limitaciones de su uso (Pedrosa, 2013).

Mislevy, considera a la representación del constructo como el argumento adecuado para la validez del mismo, propone el diseño cognitivo de las pruebas considerando el uso de la prueba como elemento esencial en dicho proceso. Su modelo intenta aplicar el juicio aprobatorio a través de la evidencia por lo que se conoce a este modelo como centrado en la evidencia (Mendoza, 2015). Mislevy declara que la validez surge desde el diseño de la prueba indicando cuatro pasos fundamentales en el proceso: el análisis del dominio, el modelado del dominio, el marco conceptual de evaluación y la arquitectura de la prueba.

Embretson propone el uso de los principios cognitivos psicológicos en el diseño de pruebas para aportar validez a la prueba mediante la definición de las habilidades y la selección de los ítems de acuerdo con ellas. Además de mejorar la validez del constructo los diseños basados en principios cognitivos proveen bases para la interpretación diagnóstica de los resultados, principios definidos para la calificación automatizada y estructura para la generación algorítmica de ítems. Otras ventajas del diseño con un enfoque cognitivo son la comprensión de la validez del constructo mediante la investigación de los procesos estrategias y conocimientos involucrados en el desempeño, la comprensión de la validez del constructo desde el nivel del ítem y la determinación de la complejidad cognitiva de cada uno mediante el análisis de modelos matemáticos además de hacer posible una interpretación de los resultados con la posibilidad de describir lo que un individuo es capaz o no de resolver mediante la aplicación de un ajuste de la TRI (Embretson, Gorin 2001).

El sistema de diseño cognitivo contiene un marco conceptual y uno procedimental. El marco conceptual propone la separación de la representación del constructo y la red nomotética como aspectos de la validez de constructo; el primero corresponde al significado del constructo y la segunda a la significancia de este. La red nomotética se refiere a la relación de los resultados de la prueba con otras mediciones mediante el establecimiento de correlaciones entre variables cumpliendo con las características de estar construida bajo la guía de un sistema de hipótesis generado de la representación del constructo y de tratar con la significancia y no con el significado (Embretson, Gorin 2001).

El tipo de estudio es de corte psicométrico y de diseño, desarrollo y validación de una prueba de expresión escrita del idioma inglés como segunda lengua alineada al currículum nacional de tercero de secundaria. Para ello se integraron las propuestas metodológicas de Nitko (1994) y de Contreras (2009) para el diseño de pruebas criterioles alineadas al currículum, así como los modelos para el diseño de pruebas cognitivas propuestos por Mislevy (2007) y Embretson (2001). En la Tabla 1 se pueden observar las etapas y actividades realizadas a lo largo del estudio. Nótese que el estudio se realizó en cinco etapas: (1) Definición y análisis de dominios establecidos en el currículum; (2) Desarrollo de los ítems; (3) Validación del aspecto de contenido y del proceso de respuesta de los ítems; (4) Piloteo y (5) Análisis de las propiedades psicométricas.

Tabla 1: Modelo de diseño, desarrollo y validación de la prueba

ETAPAS	ACTIVIDADES
I. DEFINICIÓN Y ANÁLISIS DE DOMINIOS ESTABLECIDOS EN EL CURRÍCULO.	CONFORMAR Y CAPACITAR A LOS INTEGRANTES DEL COMITÉ DEL DISEÑO DE LA EVALUACIÓN (CDE).
	ANALIZAR EL CURRÍCULO DE LA HABILIDAD DE EXPRESIÓN ESCRITA EN EL PNIEB DE EDUCACIÓN SECUNDARIA.
	ANALIZAR LAS HABILIDADES Y DOMINIOS ESTABLECIDOS EN EL MCER.
	ANALIZAR LAS TEORÍAS COGNITIVAS DE LA HABILIDAD DE ESCRITURA DEL INGLÉS COMO SEGUNDA LENGUA.
	ANALIZAR LA CONCORDANCIA ENTRE CONTENIDOS Y DOMINIOS DE LA TEORÍA SUSTANTIVA DE LA HABILIDAD DE LA EXPRESIÓN ESCRITA, EL MCER Y EL PNIEB DE EDUCACIÓN SECUNDARIA.
	SELECCIONAR LOS COMPONENTES PARA EL DESARROLLO DE LA PRUEBA.
	DEFINIR EL MODELO DE LA TEORÍA SUSTANTIVA (NOMOLÓGICA-COGNITIVA) PARA EL DISEÑO DE LA PRUEBA.
II. DESARROLLO DE LOS ÍTEMS.	ELABORAR LAS ESPECIFICACIONES DE LA PRUEBA
	CONFORMAR EL COMITÉ DE ELABORACIÓN DE ESPECIFICACIONES (CEE) DE LA PRUEBA.
	ELABORAR INSUMOS PARA LA INTEGRACIÓN Y ACTIVIDADES DEL COMITÉ DE DESARROLLO DE ÍTEMS.
	DESARROLLAR LAS ESPECIFICACIONES PARA LOS ÍTEMS.
III. VALIDACIÓN DEL ASPECTO DE CONTENIDO Y DEL PROCESO DE RESPUESTA DE LOS ÍTEMS.	DESARROLLAR LAS RÚBRICAS PARA CALIFICAR LAS RESPUESTAS A LOS ÍTEMS DE ENSAYO DE LA PRUEBA.
	ELABORAR LAS VERSIONES DE LOS ÍTEMS DE LA PRUEBA.
	DEFINIR EL SISTEMA DE CATEGORÍAS DEL MODELO COGNITIVO DE LA TEORÍA SUSTANTIVA Y DE LOS PROCESOS DE VARIANZA IRRELEVANTE PROBABLE DE LOS ÍTEMS.
	DISEÑAR Y APLICAR LOS PROTOCOLOS VERBALES A ESTUDIANTES PARA EL ANÁLISIS COGNITIVO DE LOS ÍTEMS.
	ANALIZAR LAS VERBALIZACIONES RESULTANTES DE LAS TÉCNICAS DE PENSAMIENTO EN VOZ ALTA CON BASE EN EL SISTEMA DE CATEGORÍAS DEL MODELO DE LA TEORÍA SUSTANTIVA Y EL MODELO UNIVERSAL DE DISEÑO (MUD).
	AJUSTAR LOS ÍTEMS CON BASE EN LOS RESULTADOS DEL ESTUDIO COGNITIVO DE LOS ÍTEMS.
	DISEÑAR Y APLICAR LOS PROTOCOLOS VERBALES A INTEGRANTES DEL CDE PARA EL ANÁLISIS COGNITIVO Y VALIDACIÓN DE LOS ÍTEMS.
V. ANÁLISIS DE LAS PROPIEDADES PSICOMÉTRICAS DE LA PRUEBA.	VALIDAR EN COMITÉ LA ALINEACIÓN DE LOS ÍTEMS CON LAS ESPECIFICACIONES DEL CONTENIDO DE LA PRUEBA Y EL CURRÍCULO.
	AJUSTAR LOS ÍTEMS CON BASE EN LOS RESULTADOS DE LA VALIDACIÓN DE LOS ÍTEMS.
	INTEGRAR LAS VERSIONES DE LA PRUEBA PARA SU APLICACIÓN.
	DEFINIR LA MUESTRA DE PARTICIPANTES PARA EL PILOTEO DE LA PRUEBA.
	APLICAR EL ESTUDIO PILOTO DE LA PRUEBA.
	RECOLECTAR Y CODIFICAR RESPUESTAS DE LOS EVALUADOS.
	IDENTIFICAR A DOCENTES DE LA ASIGNATURA DE INGLÉS DE SECUNDARIA PARA INTEGRAR AL EQUIPO DE JUECES CALIFICADORES DE LAS RESPUESTAS DE LOS ESTUDIANTES ANTE LOS ÍTEMS DE ENSAYO DE LA PRUEBA.
	CONTACTAR PARA AGENDAR ENTREVISTA Y SOLICITAR CURRÍCULO A LOS DOCENTES QUE PARTICIPARÁN COMO JUECES CALIFICADORES.
	ENTREVISTAR Y EVALUAR A ASPIRANTES A JUECES CALIFICADORES.
	ELABORAR INSUMOS PARA LA INTEGRACIÓN Y ACTIVIDADES DE LOS JUECES CALIFICADORES.
SELECCIONAR Y CAPACITAR A TRES JUECES PARA CALIFICAR LAS RESPUESTAS DE LOS EXAMINADOS CON BASE EN LAS RÚBRICAS DE LA EVALUACIÓN DE LA EXPRESIÓN ESCRITA.	
EVALUAR LOS RESULTADOS DE LOS EXAMINADOS CON BASE EN LAS RÚBRICAS DE LA EVALUACIÓN DE LA EXPRESIÓN ESCRITA.	
EVALUAR LA CONCORDANCIA ENTRE JUECES CALIFICADORES.	
ANALIZAR LAS PROPIEDADES PSICOMÉTRICAS DE CADA UNO DE LOS ÍTEMS DE LA PRUEBA.	
ELABORAR LOS REPORTES DE RESULTADOS PRELIMINARES DE LA PRUEBA PARA ESTUDIANTES, DOCENTES Y AUTORIDADES EDUCATIVAS.	

Nota: Todos los integrantes de los comités, así como los jueces calificadores y equipo técnico del proyecto deben firmar la carta de aceptación y confidencialidad

Los resultados en general fueron satisfactorios mostrando el logro de los criterios y estándares de las evidencias de validez de constructo de los aspectos del contenido, del proceso de respuesta y de la estructura interna de la prueba diseñada.

Conclusiones

La enseñanza y evaluación del inglés con fines comunicativos es relevante y debe ser tomado en cuenta por las políticas educativas. La promoción de lenguas extranjeras en los sistemas educativos responde en gran medida a los discursos de globalización y a la conformación de bloques político-económicos regionales y continentales y evidencia su valor en la proyección de un modelo determinado de ciudadanía. Así los conceptos de lengua, sociedad y política están estrechamente vinculados (Reyes, 2011). No son sólo individuos lo que cruza la frontera; bienes y servicios, formas de vida, lenguas, cultura, mano de obra, hábitos de consumo y divisas se encuentran en constante movimiento entre ambos países. Ambos gobiernos reaccionan ante este fenómeno con tratados de comercio y políticas de migración, económicas y educativas, entre las últimas interesándonos la planificación de la enseñanza y aprendizaje de lenguas desde el punto de vista político-lingüístico que toma en cuenta el escenario en el que estos procesos ocurren (Canale, 2010).

Las pruebas aplicadas de manera estandarizada permiten conocer el nivel en el que los estudiantes están aprendiendo los contenidos de un currículo al final de un período y evaluar el efecto de un programa en el logro educativo (Ravela, 2008). La importancia de la elaboración de una prueba que mida la habilidad de expresión escrita en el idioma inglés en los alumnos de educación básica radica en la necesidad de conocer estos datos para aportar evidencia para la toma de decisiones sobre cambios y mejoras a realizar en el programa y en la manera de implementarlo.

Cabe señalar que las pruebas disponibles en el mercado miden el nivel de dominio del idioma inglés de los estudiantes de acuerdo a una norma. Su resultado aporta información sobre el nivel en el que un individuo se encuentra con respecto a un grupo tomado como referente. Estos resultados son útiles en cuanto a la necesidad de probar o no que el aspirante posee un nivel adecuado en términos de certificación o de admisión a alguna institución educativa pero no dan información sobre el nivel de logro de aprendizaje con respecto al currículo determinado. Considerando que no existe una prueba alineada al currículo de inglés del nivel secundaria, esta prueba aportar información no solo a los tomadores de decisiones del sistema educativo estatal sino también a directivos de escuelas, maestros y padres de familia.

Adicionalmente, atendiendo a que el aprendizaje se puede entender como el cambio observado en conocimientos y capacidades a lo largo de un período y que su medición depende de la aplicación de pruebas (Ravela, 2008); los resultados de la prueba aplicada durante el ciclo básico podrían ser utilizados como diagnóstico del nivel de conocimientos de los estudiantes de los tres grados de secundaria. Se puede concluir que se lograron cabalmente los objetivos del diseño y validación de la prueba, sin embargo, es importante aplicar un piloteo con una muestra lo suficientemente grande como para realizar análisis más sofisticados que aporten a las evidencias de la calidad técnica de la prueba.

Referencias

- Contreras, L. (2000). *Desarrollo y Pilotaje de un Examen de Español para la Educación Primaria en Baja California*. (tesis de maestría). Instituto de Investigación y Desarrollo Educativo, Ensenada, México.
- Cushing, S. (2002). *Assessing Writing*. New York USA: Cambridge University Press.
- Embretson, S. & Gorin, A. (2001). Improving Construct Validity With Cognitive Psychology Principles. *Journal of Educational Measurement*, 2(4),343-368.
- Hussain, S., Tadesse, T., & Sajid, S. (2015). Norm-Referenced and Criterion-Referenced Test in EFL Classroom. En *International Journal of Humanities and Social Science Invention* (págs. 24-30)
- Mac Arthur S. (2017). *Handbook of Writing Research*. New York, United States: Guilford
- Mendoza, A. (2015). La validez en los exámenes de alto impacto. Un enfoque desde la lógica argumentativa. *Perfiles Educativos* xxxvii (149) pp 169 – 186
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*. 50, 741-749.
- Ministerio de Educacion Cultura y Deporte (2002). Recuperado de: http://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf
- Mislevy (2007). Comments on Validity by Design. *Educational Researcher*, 36 (8). pp463-469.
- Nitko, A. (1994). A model for developing curriculum-driven criterion-referenced and norm-referenced national examinations for certification and selection of students. Ponencia presentada en la conferencia internacional sobre evaluación y medición educativas, de la Asociación para el Estudio de la Evaluación Educativa en Sudáfrica (ASSESA).
- Nichols, P., Kobrin, J., Lai, E., Koepfler, J. (2017). The Role of Theories of Learning and Cognition in Assessment Design and Development. In: *The Handbook of Cognition and Assessment: Frameworks, Methodologies, and Applications*, Editor Andre Rupp, first edition. (pp 16-40). USA John Wiley and Sons, Inc.
- OCDE (2014). En *Revisiones de la OCDE sobre la Evaluación en Educación México*.
- Pedrosa, I., Suárez-Alvarez y García-Cueto, E. (2014). Evidencias sobre la Validez de Contenido: Avances Teóricos y Métodos para su Estimación. *Acción Psicológica*, 10(2), 3-18. ISSN: 1578-908X.
- Ramírez, J., Pamplón, E. & Cota, S. (2012). Poblémática de la enseñanza de inglés en las primarias públicas de México: una primera lectura cualitativa. En *Revista iberoamericana de Educación* 60/2 (págs. 02-12).
- Ravela, P. (2000). En *Para comprender las evaluaciones educativas*. Santiago de Chile. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=2232546>.
- Subsecretaria de Educación Media Superior (2014). *dgb.sep.org.mx*. Obtenido de [dgb.sep.org.mx](http://www.dgb.sep.org.mx/02-ml/03-iacademica/01-programasdeestudio/ler_SEMESTRE/Lengua_Adicional_Espanol_I_biblio2014.pdf): http://www.dgb.sep.org.mx/02-ml/03-iacademica/01-programasdeestudio/ler_SEMESTRE/Lengua_Adicional_Espanol_I_biblio2014.pdf
- Subsecretaria de Educación Media Superior (2014). En *Programas de Estudio/Lengua Adicional al Español*. Obtenido de http://www.dgb.sep.org.mx/02-ml/03-iacademica/01-programasdeestudio/ler_SEMESTRE/Lengua_Adicional_Espanol_I_biblio2014.pdf