

Capítulo 18.

La evaluación estandarizada del aprendizaje en la enseñanza del inglés como lengua extranjera: ¿una medida justa?

Dr. Jorge Gustavo Gutiérrez Benítez
Universidad Autónoma de Baja California

Resumen

Evaluar en el contexto de la enseñanza de lenguas extranjeras nunca ha sido un asunto sencillo. Desde los exámenes más antiguos hasta las más sofisticadas pruebas internacionales estandarizadas como el Test of English as a Foreign Language (TOEFL) o International English Language Testing System (IELTS), la idea de medir lo que un estudiante sabe siempre ha generado debates.

El recorrido histórico muestra cómo la estandarización nace con la intención de garantizar comparabilidad y objetividad. Hoy en día, estos exámenes son considerados instrumentos confiables debido al rigor científico con el que se elaboran, lo que las hace idóneas para tomar decisiones en contextos como la educación y el empleo. Sin embargo, su carácter selectivo y la presión que ejercen sobre los estudiantes han sido criticados por los efectos que produce.

En la enseñanza de lenguas, las tensiones son aún más evidentes: ¿se evalúa realmente la capacidad de comunicarse en la vida diaria o solo el desempeño bajo condiciones de examen? Aquí aparecen fenómenos como el *teaching to the test*, donde lo importante ya no es aprender el idioma, sino aprobar el examen.

Este capítulo explora las bondades y dificultades de la evaluación estandarizada en la enseñanza de lenguas, partiendo de sus orígenes generales hasta los desafíos modernos. Se realiza un recorrido de la historia de la evaluación de una lengua extranjera, se analiza la relación particular entre lengua y prueba, se discuten las tensiones existentes entre la necesidad de evaluar y el vasto universo de implicaciones presentes en la enseñanza de lenguas.

Se invita a la reflexión de la concepción de la evaluación, que más que un fin en sí misma, la evaluación debe ser vista como un medio que acompañe al estudiante en su aprendizaje, lo motive y lo acerque al verdadero propósito de la enseñanza de lenguas que es conectar con otras personas y culturas, potencializar una de las cualidades humanas más importantes: el poder comunicarse.

Palabras clave: evaluación, estándar, lenguas, enseñanza, inglés

Introducción

Pocas situaciones en la vida del estudiante generan tanto nerviosismo como sentarse frente a un examen, y en el caso particular de las evaluaciones de lenguas estos momentos cobran aún más sentido.

No conforme con estar contra reloj, se escucha el audio de una persona conversando en otro idioma, y en la mente del estudiante se empieza a ejecutar un proceso mediante el cual trata de comprender lo que escucha, y al mismo tiempo elegir la respuesta correcta que a su vez también se encuentra en otro idioma.

Para el estudiante de lenguas, pruebas como TOEFL, el Diploma de Español como Lengua Extranjera (DELE) o el IELTS, no sólo son una evidencia de su nivel de dominio en la lengua, sino que también se han convertido en la llave de acceso a las propias instituciones educativas como las universidades o posgrados, empleos y programas de movilidad en todo el mundo. Pero, la pregunta que gira en torno a esto es ¿cómo se puede saber qué tan confiables y justas son realmente? ¿Evalúan lo que se supone deberían evaluar o solo una parte de todas las habilidades deseables en el estudiante?

Son justamente las interrogantes anteriores las que llevan a lo que se conoce como evaluación estandarizada. En el mundo moderno, donde dominar más de un idioma es requisito para la vida profesional, estas pruebas estandarizadas de tipo internacional representan mucho más que solo dar una calificación: son instrumentos de selección, de certificación y, en muchos casos, de exclusión.

Al mismo tiempo, estos tipos de pruebas ofrecen ventajas sobresalientes como la posibilidad de comparar resultados entre países y asegurar

estándares de calidad, pero no solo a escalas globales, sino también en el plano de un aula de clases en las que se puede observar de una manera válida y confiable los niveles de dominio que presentan los estudiantes en determinada área de conocimiento.

Este escrito se propone explorar las bondades y dificultades de la evaluación estandarizada en la enseñanza de lenguas, partiendo de sus orígenes generales hasta los desafíos modernos. Para lograr lo anterior, se realiza un recorrido de la historia de la evaluación educativa, se analiza la relación particular entre lengua y prueba, y se discute las tensiones existentes entre la necesidad de evaluar y el vasto universo de implicaciones presentes en la comunicación humana.

La fascinación por medir el aprendizaje

Desde sus inicios, un proceso que nunca puede faltar en la educación es la idea de poder medir lo que se enseña o bien lo que se aprende. A este respecto, retrocediendo varios años es posible señalar las tablillas de arcilla en la antigua Mesopotamia, y en nuestros días se puede dar pauta de los modernos sistemas de evaluación en línea.

Ambos escenarios comparten un punto en común, el profesor de hace siglos y el profesor actual, siempre ha buscado formas para poder comprobar si sus estudiantes aprendieron lo que se les enseñó. En la época reciente, particularmente en el siglo XIX, se dio un auge sustancial sobre esta posibilidad de medir, lo anterior como resultado de la psicometría como la subdisciplina de la psicología experimental dedicada a evaluar y cuantificar los procesos psicológicos y las habilidades cognitivas en las personas (Martínez y Villota, 2022).

El auge anterior, tuvo como resultado la construcción de instrumentos cada vez más sofisticados, diseñados para clasificar y comparar desempeños (Popham, 2017). La lógica detrás de la estandarización era sencilla: si todos los estudiantes respondían al mismo conjunto de preguntas en condiciones similares, entonces sus resultados podían ser comparados objetivamente. Esta forma de concebir la educación dio como resultado el nacimiento de pruebas como el Scholastic Aptitude Test (SAT) en Estados Unidos, misma que sentó las bases en parte para lo que es la evaluación educativa moderna (Villacís et al., 2022).

Claro está, las evaluaciones estandarizadas no están exentas de críticas. Mientras que algunos investigadores las defienden como una herramienta indispensable para la equidad **porque ofrecen un criterio común**, otros señalan que reducen la complejidad del aprendizaje a solo datos numéricos que no reflejan de manera real las múltiples dimensiones del conocimiento (Darling-Hammond, 2017). En el caso de la enseñanza de lenguas, estas diferentes posturas son aún más marcadas, pues lo que está tela de juicio no es solo recordar vocabulario o reglas gramaticales, sino la capacidad propia del estudiante para poder interactuar en contextos diversos.

El auge de los exámenes estandarizados en las lenguas extranjeras

El siglo XX introdujo una necesidad evaluativa profesional muy importante: comprobar el nivel de dominio o bien las habilidades lingüísticas específicas de una manera confiable y comparable. Lo anterior provocado o forzado por los cambios políticos, sociales y tecnológicos que caracterizo ese siglo desde sus primeras décadas, migraciones masivas, la internacionalización de la educación y los intercambios académicos hicieron evidente la importancia de contar con pruebas que certificaran el nivel de dominio de una lengua extranjera (Fulcher, 2015).

Lo anterior inevitablemente provocó la creación de pruebas con las cuales demostrar estas habilidades en una lengua extranjera, fue así como se consolidaron pruebas como el TOEFL, creado en 1964 por la Educational Testing Service, y posteriormente el IELTS, ambos diseñados para evaluar las cuatro habilidades principales: comprensión auditiva, expresión oral, comprensión lectora y producción escrita.

Si bien, un aspecto fundamental de este tipo de evaluaciones es garantizar que los estudiantes no solo alcancen niveles cognitivos de bajo nivel como lo es la memorización o comprensión, sino por el contrario que sean capaces de desenvolverse en contextos reales. Sin embargo, como señalan Bachman y Palmer (2010), cualquier prueba es solo una representación limitada o imperfecta de la realidad, es decir, mide lo que es posible representar en la condición específica de un examen, en el formato que este tenga, lo que a su vez se traduce en la representación parcial y no la totalidad del uso de la lengua.

Agregado a esto, la mayoría de estas pruebas adquirieron una etiqueta de carácter socioeconómico y político. Los puntajes en exámenes como TOEFL o IELTS se convirtieron en requisito indispensable para acceder a universidades de prestigio, la obtención de becas o incluso para ser acreedor de una visa (McNamara, 2019). Fue de esta manera como el examen o la evaluación estandarizada dejó de ser una herramienta para el diagnóstico del aprendizaje para transformarse en una poderosa herramienta de selección y movilidad social (Shohamy, Or & May, 2017).

De manera natural junto con el prestigio que adquirieron las pruebas estandarizadas también conlleva una parte negativa. Mientras que, para algunos (Reyes et. al, 2023) los exámenes estandarizados están entrelazados con ideales meritocráticos que se manifiestan globalmente, aunque se adaptan a contextos locales.

Para otros, constituyen una barrera que privilegia a quienes tienen acceso a cursos de preparación costosos, materiales de práctica y tecnología adecuada, generando desigualdades profundas (Shohamy, Or & May, 2017).

Enseñanza del inglés como lengua extranjera y evaluaciones: una combinación compleja

En el contexto específico de la evaluación del aprendizaje, las lenguas son un reto particular. Lo anterior cobra mayor significado al compararse con materias como matemáticas o física, donde las respuestas comúnmente son únicas y verificables con procedimientos específicos.

Sin embargo, el lenguaje se caracteriza por su variabilidad, creatividad y contexto. Lo que permite que más de un estudiante pueda expresar la misma idea, pero de formas distintas e igual de válidas. Por este motivo autores como Brown & Abeywickrama (2019) expresan la dificultad existente de querer reducir o evidenciar el desempeño lingüístico a una calificación numérica.

En el caso de las diferentes dimensiones implicadas en el aprendizaje de una lengua, la competencia comunicativa, concepto introducido por Hymes (1972), señala que saber una lengua o idioma no es solo dominar su gramática, sino también saber usarla de una manera adecuada en situaciones reales.

Lo anterior implica por lo tanto aspectos pragmáticos, socioculturales y discursivos. Sin embargo, una situación observable en la mayoría de las pruebas estandarizadas, es el hecho de que, aunque intentan aproximarse a esa variedad de aspectos, terminan privilegiando actividades o evidencias de desempeño que pueden calificarse con relativa objetividad y rapidez, como opciones múltiples o ensayos con rúbricas predefinidas.

Por ello, algunos autores como Simone et al. (2023) señalan que este tipo de pruebas contribuyen el fenómeno conocido como *teaching to the test* o *backwash effect*, describiendo la influencia que las evaluaciones ejercen en lo que se enseña y en la forma se enseña, lo que de igual forma repercute en el comportamiento de los estudiantes y docentes. Esto se observa en los cursos de idiomas cuando estos se enfocan en preparar a los estudiantes para superar el examen más que en desarrollar una competencia comunicativa auténtica (Cheng, 2008).

Esto produce por lo tanto un aprendizaje parcial, es decir, por un lado, el alumno puede saber o dominar muy bien una prueba o examen, pero no serlo en una situación real o por ejemplo una conversación cotidiana. El alumno se prepara o estudia para la prueba, pero no para el aprendizaje significativo.

Aun con estas situaciones presentes, no es posible ignorar ni reconocer que las pruebas estandarizadas han contribuido a establecer parámetros comunes que permiten comparar niveles entre distintos países e instituciones. Un ejemplo claro de este beneficio es el Marco Común Europeo de Referencia para las Lenguas (MCER) publicado oficialmente en el año 2001, mismo que ofrece una escala que facilita ubicar a los estudiantes en niveles de competencia, lo que se ha convertido en un lenguaje común en la enseñanza de lenguas a nivel mundial (Council of Europe, 2020).

Ventajas de la evaluación estandarizada

Aunque se han señalado algunos aspectos desfavorables u objeto de crítica en las evaluaciones estandarizadas, existen también una variedad de atributos positivos que ofrecen ventajas significativas, especialmente en el ámbito de la enseñanza de lenguas.

En un primer beneficio de las pruebas estandarizadas se encuentra la comparabilidad global. Este atributo es relevante, ya que permite evaluar estudiantes de distintos contextos bajo un mismo criterio, lo que resulta especialmente útil para universidades que, por ejemplo, reciben postulaciones internacionales. De esta forma un puntaje en TOEFL o IELTS se convierte en un insumo directo para que las instituciones puedan tomar decisiones de admisión basadas en estándares relativamente consistentes, lo que de cierta manera contribuye a garantizar un grado de equidad en la selección de candidatos (Kunnan, 2018).

Otro beneficio importante es la estandarización de la calidad de los instrumentos de evaluación. Este tipo de exámenes se caracterizan por rigurosos procedimientos metodológicos en su construcción y diseño, así como de validación psicométrica, lo que asegura que las preguntas sean confiables, que realmente midan las competencias que pretenden evaluar y que los resultados se mantengan estables en el tiempo (Gutiérrez y Acuña, 2023).

Aunque se acepta la realidad de que en toda medida existe un error, y por lo tanto ningún instrumento o prueba pueda considerarse perfecta, este nivel de control difícilmente podría alcanzarse mediante evaluaciones que no implementen alguna especie de control o procedimiento establecido para su construcción.

La objetividad alcanzada constituye un tercer punto a favor. A diferencia de la evaluación subjetiva habitualmente dada por el docente, que inherentemente está cargada de sesgos, algunos de manera inconsciente, los exámenes estandarizados tienen como premisa una igualdad de probabilidades.

Lo anterior significa que todos los estudiantes que desempeñen la prueba cuentan con igual probabilidad de logro, es decir, todos los sustentantes se enfrentan a la misma dificultad y la misma posibilidad de evidenciar el nivel de dominio sobre el universo de conocimientos comprendido por la prueba (Gutiérrez y Acuña, 2022).

Finalmente, las pruebas estandarizadas poseen un reconocimiento internacional significativo. Una evaluación que provea de un certificado que tenga aceptación global se traduce de manera directa en otro tipo

de oportunidades profesionales y académicas. En el mercado laboral actual con un alto sentido de competitividad, un desempeño destacado en una prueba estandarizada de aceptación internacional puede marcar significativamente la diferencia.

Aunque pueden señalarse más atributos favorables de las pruebas estandarizadas, los antes mencionados demuestran que este tipo de evaluaciones no solo funcionan como instrumentos de medición válidos y confiables, sino que también actúan como verdaderos diferenciadores en el mercado laboral, o como una especie de moneda en la economía del conocimiento, explicando así su popularidad, amplia difusión y permanencia en contextos educativos de todo el mundo.

Limitantes a considerar

Si bien los beneficios son claramente señalados, también lo son las limitaciones. La evaluación estandarizada en el contexto de la enseñanza de lenguas se ve expuesta a cuestionamientos profundos en torno a ciertas formas de validez, equidad y consecuencias pedagógicas.

Una primera limitante de las pruebas estandarizadas tiene que ver con los sesgos culturales y lingüísticos. En este sentido, los ítems suelen reflejar perspectivas particulares, muchas veces anglocéntricas, lo que puede poner en desventaja a estudiantes de contextos culturales diferentes (Cheng, 2008).

Por otro lado, uno de los desafíos más señalado es la reducción del lenguaje a un atributo numérico o cuantificable. Lo anterior cobra sentido al expresar los diferentes atributos presentes en proceso propio del lenguaje, como lo son la ironía, humor, gestos y variaciones dialectales. Todos estos constructos psicológicos, reducidos a una calificación numérica facilitan la administración de las pruebas, pero el principal problema es que casi en forma proporcional reduce la totalidad de la competencia comunicativa que es recuperada en las pruebas (Gozalo, 2014).

En este mismo tenor, es innegable abordar las implicaciones psicológicas y emocionales que las pruebas estandarizadas sobre todo las de alto impacto tienen en los estudiantes. La presión que hay detrás por alcanzar

un puntaje que bien puede definir el futuro académico o profesional puede afectar tanto su rendimiento como su bienestar emocional (Horwitz, 2010).

Por otro lado, otra limitante es la necesidad de experticia para la elaboración de pruebas estandarizadas, dada la naturaleza de este tipo de pruebas, es necesario aplicar procedimientos metodológicos rigurosos y con ello el uso de tratamientos estadísticos que requieren de una experticia tanto en el proceso de construcción de la prueba como en las evaluaciones psicométricas con las que se da pauta de la calidad técnica de la misma. Lo anterior también incluye el uso de software especializado para análisis psicométrico, la interpretación de diversos atributos observables de validez y confiabilidad entre algunas otras implicaciones.

En definitiva, las pruebas estandarizadas son herramientas poderosas y ampliamente utilizadas, pero no son neutrales. Mientras ofrecen comparabilidad, objetividad y reconocimiento global, también reflejan sesgos, alteran o condicionan lo que se enseña y como se enseña, lo que a su vez puede ampliar desigualdades. También se hace latente una marcada accesibilidad a la construcción y aplicación de las mismas, por los costos que implican y la necesidad de personal especializado para el desarrollo de las mismas.

Alternativas a la evaluación estandarizada

Considerando las limitaciones antes mencionadas, diversas investigaciones en evaluación de enseñanza de lenguas han explorado distintas alternativas que puedan reducir los aspectos no observables con las pruebas estandarizadas o sustituir incluso algunas actividades evaluativas representadas por exámenes estandarizados tradicionales.

Una de ellas es el fortalecimiento y robustecer los esquemas la evaluación formativa, que es la que se efectúa de manera continua en las clases, enfatizando el estímulo que el estudiante recibe para fomentar su desarrollo y concreción de conocimientos, priorizando estos procesos por encima de ver como un fin la asignación de un puntaje.

A este respecto un aspecto fundamental de todo proceso evaluativo, es la retroalimentación, y en el caso particular de la evaluación formativa,

a través de la retroalimentación constante, los estudiantes pueden identificar sus fortalezas y determinar si se han alcanzado los objetivos de aprendizaje establecidos sin estar limitados a una nota o calificación (Defas et al., 2025).

Un ejemplo claro de instrumentos formativos son los portafolios digitales, en los cuales se recopilan trabajos escritos, grabaciones de audio y proyectos colaborativos. Este tipo de herramienta provee al docente de diversas ventajas para ampliar la visión del progreso lingüístico de cada estudiante (Luelmo del Castillo, 2021).

Otras herramientas **clave para valorar** el nivel de dominio de una lengua son las rúbricas de desempeño y la evaluación basada en tareas. A diferencia de las pruebas tradicionales que miden respuestas correctas o incorrectas, las rúbricas detalladas permiten valorar criterios como claridad, coherencia, adecuación sociocultural y creatividad en la expresión (Brown & Abeywickrama, 2019). Este tipo de herramientas permiten observar de manera más integral la competencia comunicativa, al centrarse en habilidades de comunicación auténtica y contextualizada.

Por otro lado, una innovación tecnológica que está cambiando a la humanidad en muchas esferas es la inteligencia artificial (IA), y en este sentido, las evaluaciones no han sido la excepción en relación a sufrir cambios y mejoras al emplear IA. Los sistemas inteligentes o de adaptación automática ajustan la dificultad de las preguntas según el desempeño del estudiante, lo que permite obtener diagnósticos más precisos y proporcionar a su vez interacciones personalizadas (Kalidindi, Das & Dandu, 2025). Además, la IA puede analizar patrones de error, identificar áreas de oportunidad o mejora, así como generar retroalimentación inmediata, lo que ahora convierte a la evaluación en una herramienta para el aprendizaje de manera continua.

El futuro de evaluar el inglés como lengua extranjera

En la actualidad el panorama que se observa en el campo de la educación sugiere que emplear evaluaciones estandarizadas para determinar el nivel de dominio en las lenguas no va a dejar de existir, sin embargo, se hace evidente su evolución. Las tendencias apuntan hacia modelos más flexibles, integradores y conscientes de la diversidad de los estudiantes.

La globalización y la demanda de movilidad internacional siguen representando factores de influencia que impulsan el uso de exámenes estandarizados para poder clasificar sustentantes, sobre todo en aspectos laborales profesionales. Pero, por otro lado, existen otras corrientes que están empujando fuertemente a la concientización sobre equidad, inclusión y diversidad cultural, lo que se traduce en la necesidad de innovar para humanizar la evaluación y reflejar mejor las diferentes habilidades del dominio de una lengua, como lo es la competencia comunicativa real (Darling-Hammond, 2017).

Como se ha dicho, la evaluación en la enseñanza de lenguas está experimentando una transformación significativa. Existe una presión constante por mejorar las formas en que se pueda observar la competencia lingüística de una manera significativa, por ello se está optando por complementar las evaluaciones tradicionales, como las pruebas estandarizadas con otros tipos de evidencias de desempeño. Investigaciones recientes sugieren que las evaluaciones empleadas de manera tradicional, que son las que mayormente se emplean en las aulas en los diferentes niveles educativos no logran capturar completamente la diversidad de habilidades lingüísticas de los estudiantes (Chuang & Yan, 2025).

Herramientas que están basadas e incluso construidas con IA, como lo son los chatbots y sistemas de retroalimentación automatizada, están siendo utilizadas para proporcionar evaluaciones cada vez más individuales llevando la personalización a niveles no alcanzados anteriormente, y con ellos ambientes más dinámicos (Alshehri, 2025). Agregado a esto, este tipo de tecnologías permiten dar al estudiante una retroalimentación automática e inmediata, lo que contribuye a mejorar los niveles de motivación y la autonomía del aprendizaje en los estudiantes.

Sin embargo, la implementación de IA en la evaluación sigue planteado desafíos éticos y prácticos. Por un lado, la precisión de estas herramientas en diferentes contextos lingüísticos y culturales aún está perfeccionándose, ya que herramientas accesibles que en su mayoría son gratuitas manejan niveles de error muy considerables. Por otro lado, las mejores herramientas se exceden en costos, mismos que escapan a la mayoría de los usuarios, y agregado a esto existe el riesgo de que se sigan repitiendo sesgos existentes en las evaluaciones si estas no se diseñan cuidadosamente (Yan et al., 2023).

La pandemia de la COVID-19 aceleró la adopción de evaluaciones en línea, lo que ha llevado a una reconsideración de los métodos de evaluación tradicionales. Esto naturalmente debido a la flexibilidad y accesibilidad que las evaluaciones en línea ofrecen y que a su vez promueven oportunidades para una mayor inclusividad (ECML, 2023).

El futuro de la evaluación en lenguas parece dirigirse hacia un enfoque híbrido que combine la tecnología con la comprensión humana. Es indispensable que las instituciones educativas diseñen y desarrollen sistemas de evaluación que sean justos, inclusivos y que reflejen con precisión las habilidades lingüísticas de los estudiantes haciendo un uso responsable de las nuevas tecnologías que dotan de ventajas significativas al quehacer educativo.

Conclusión: reflexionando sobre evaluar el lenguaje

Al mirar todo lo que se ha explorado, queda claro que evaluar el aprendizaje de lenguas no es solo un asunto de exámenes y puntajes. Las pruebas estandarizadas, tales como TOEFL e IELTS, han transformado la forma en que se miden las competencias, generando oportunidades de estudio, trabajo y movilidad internacional. Pero también muestran las limitaciones existentes al querer reducir el lenguaje a números y porcentajes.

Como se ha dicho, el lenguaje es vivo, diverso y es una característica única en el ser humano. Cada conversación, cada mensaje escrito o cada interpretación de un texto lleva consigo rasgos únicos, culturales, sociales y personales que ninguna prueba puede capturar por completo. Por eso, aunque un examen pueda decir que “el estudiante alcanzó el nivel C1”, no puede decir cómo se expresa con naturalidad en una pizzería en Italia, o cómo convence a un colega en Londres de escuchar cierto tipo de música, o cómo tiene que improvisar al platicar con un niño que le pide un dulce en una tienda en Francia.

Aun así, no se pueden desestimar los beneficios de la evaluación estandarizada. Ya que posibilita un lenguaje común para comunicar, identificar o determinar niveles de competencia, facilita además la movilidad internacional y, en muchos casos, brinda motivación a los estudiantes para alcanzar metas claras. El reto, sin embargo, está en no permitir que estas pruebas sean el único indicativo del nivel de dominio, logro alcanzado o aprendizaje que evidencien los alumnos.

Cuando la enseñanza se enfoca únicamente en superar un examen, se corre el riesgo de perder lo esencial: que el estudiante aprenda a comunicarse, a explorar, a equivocarse y a desarrollarse con el idioma.

El futuro de la evaluación parece inclinarse hacia un equilibrio entre estandarización y diversidad, donde idealmente se creen escenarios en los que las pruebas no sea un obstáculo, sino una guía: donde cada estudiante pueda recibir retroalimentación personalizada, reconocer sus fortalezas, trabajar en sus debilidades y, sobre todo, aprender a usar la lengua como un puente hacia otras personas y culturas.

Esta reflexión transmite algo fundamental: evaluar no es sólo medir, y mucho menos es un proceso neutral, la evaluación no es un fin, sino un medio. Las pruebas son herramientas, y como cualquier herramienta, su valor depende de cómo se utilicen. Si se integran con conciencia, sensibilidad y creatividad, pueden ayudar a construir un aprendizaje más inclusivo, motivador y auténtico. Si, por el contrario, se les trata como finales absolutos, se corre el riesgo de confundir calificación con competencia y de reducir la riqueza del lenguaje a un simple número.

La conclusión de este **capítulo nos deja una reflexión fundamental y de profundas implicaciones**: aprender un idioma va mucho más allá de sólo aprobar un examen o de obtener una calificación alta. Cada palabra que se pronuncia, cada texto que se escribe y cada conversación que se inicia es un paso hacia comprender y conectar con el entorno, es un proceso mucho más complejo que no puede ser observado o determinado por una sola tarea. Y si una evaluación significativa acompaña ese viaje, en lugar de limitarlo, entonces se habrá logrado su verdadero propósito: apoyar al aprendizaje y la enseñanza del inglés como lengua extranjera, no solo medirlo.

Referencias

- Alshehri, A. (2025). AI's effectiveness in language testing and feedback provision. *Social Sciences & Humanities Open*, 12, 101892. <https://doi.org/10.1016/j.ssaho.2025.101892>
- Brown, H., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson.
- Cheng, L. (2008). *Adapting language tests for use in new contexts: Challenges and strategies*. Cambridge University Press.
- Chuang, P.-L., & Yan, X. (2025). Language assessment in the era of generative artificial intelligence: Opportunities, challenges, and future directions. *System*, 134, 103846. <https://doi.org/10.1016/j.system.2025.103846>
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2020/16809ea0d4>
- Darling-Hammond, L. (2017). *The flat world and education: How America's commitment to equity will determine our future*. Teachers College Press.
- Defas, E., Toro, A., Victoria, Díaz, M., Rivera, D., & Gavilánez, M. (2025). Evaluación formativa versus sumativa: efectos en el aprendizaje significativo. *Sinergia Académica*, 8(5), 833-844.
- ECML. (2023). *The future of language education in the light of Covid*. European Centre for Modern Languages. <https://www.ecml.at/en/ECML-Programme/Programme-2020-2023/The-future-of-language-education>
- Reyes, V., Phillips, L., Hamid, M. O., & Hardy, I. (2024). Navigating datascares: mapping testing practices within and across national and global contexts. *Learning, Media and Technology*, 49(1), 109–121. <https://doi.org/10.1080/17439884.2023.2218645>
- Fulcher, G. (2015). *Practical language testing* (3rd ed.). Routledge.

Gozalo, A. (2014). Hacia el impacto positivo de la evaluación: pruebas estandarizadas de idiomas. *Avances En Supervisión Educativa*, 22. <https://doi.org/10.23824/ase.v0i22.63>

Gutiérrez, J., & Acuña, L. (2023). Evaluación estandarizada de los aprendizajes en el tronco común de lenguas: Un estudio en la Universidad Autónoma de Baja California, México. *Revista Electrónica Educare*, 27(3), 1-22. <https://doi.org/10.15359/ree.27-3.17218>

Gutiérrez, J., & Acuña, L. (2022). Evaluación estandarizada de los aprendizajes: una revisión sistemática de la literatura. *Revista CPU-e*, (34), 321-351.

Horwitz, E. (2010). *Becoming a language teacher: A practical guide to second language learning and teaching*. Pearson Longman.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269-293). Penguin.

Kunnan, A. J. (2018). *Evaluating Language Assessments*. Routledge.

Luelmo del Castillo, M. (2021). Portfolios electrónicos de lenguas: un estudio comparativo. *Revista Espacios*, 42(1), 201-214.

Martínez, R. y Villota, H. (2022). La psicometría. *Revista SIGMA*, 18(1), 23-29.

McNamara, T. (2019). *Language testing*. Oxford University Press.

Popham, W. (2017). *Classroom assessment: What teachers need to know* (8th ed.). Pearson.

Shohamy, E., Or, I. G., & May, S. (Eds.). (2017). *Language testing and assessment* (3.^a ed.). Springer.

Simone, P. M., Whitfield, L. C., Bell, M. C., Kher, P., & Tamashiro, T. (2023). Shifting students toward testing: impact of instruction and context on self-regulated learning. *Cognitive Research: Principles and Implications*, 8(14). <https://doi.org/10.1186/s41235-023-00470-5>

Villacís, J., Morales, C., Santander, L., & Estrella, A. (2022). El SAT y su importancia en la evaluación de la apercepción en adultos mayores. *Ciencia Latina Revista Científica Multidisciplinar*, 6(6), 8694-8715. https://doi.org/10.37811/cl_rcm.v6i6.4027

Kalidindi, M., Das, D., & Dandu, A. (2025). Artificial intelligence for language learning: *Opportunities and challenges*. *International Research Journal on Advanced Engineering and Management (IRJAEM)*, 3(7), 2560–2566. <https://doi.org/10.47392/IRJAEM.2025.0404>

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90-112. <https://doi.org/10.48550/arXiv.2303.13379>



Fondo
Editorial para la
Investigación
Académica

Este libro reúne experiencias de docentes que han situado el error en el centro de la reflexión pedagógica, no solo para identificarlo, sino también para interpretarlo, analizarlo y transformarlo en una oportunidad formativa. Desde distintos contextos educativos, los proyectos que aquí se presentan exploran prácticas de corrección que van más allá de la señalización de fallas y se orientan a la construcción de sentido.

A través de estas experiencias, la obra pone en diálogo el error con las estrategias de retroalimentación, entendidas como procesos intencionales que acompañan al estudiante, favorecen la toma de conciencia y promueven la autorregulación del aprendizaje. Cada capítulo ofrece no solo propuestas didácticas sino también decisiones pedagógicas fundamentadas que permiten repensar la intervención docente. Más que corregir para sancionar, este libro invita a corregir para comprender, a retroalimentar para orientar y a enseñar para transformar.



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE QUINTANA ROO

