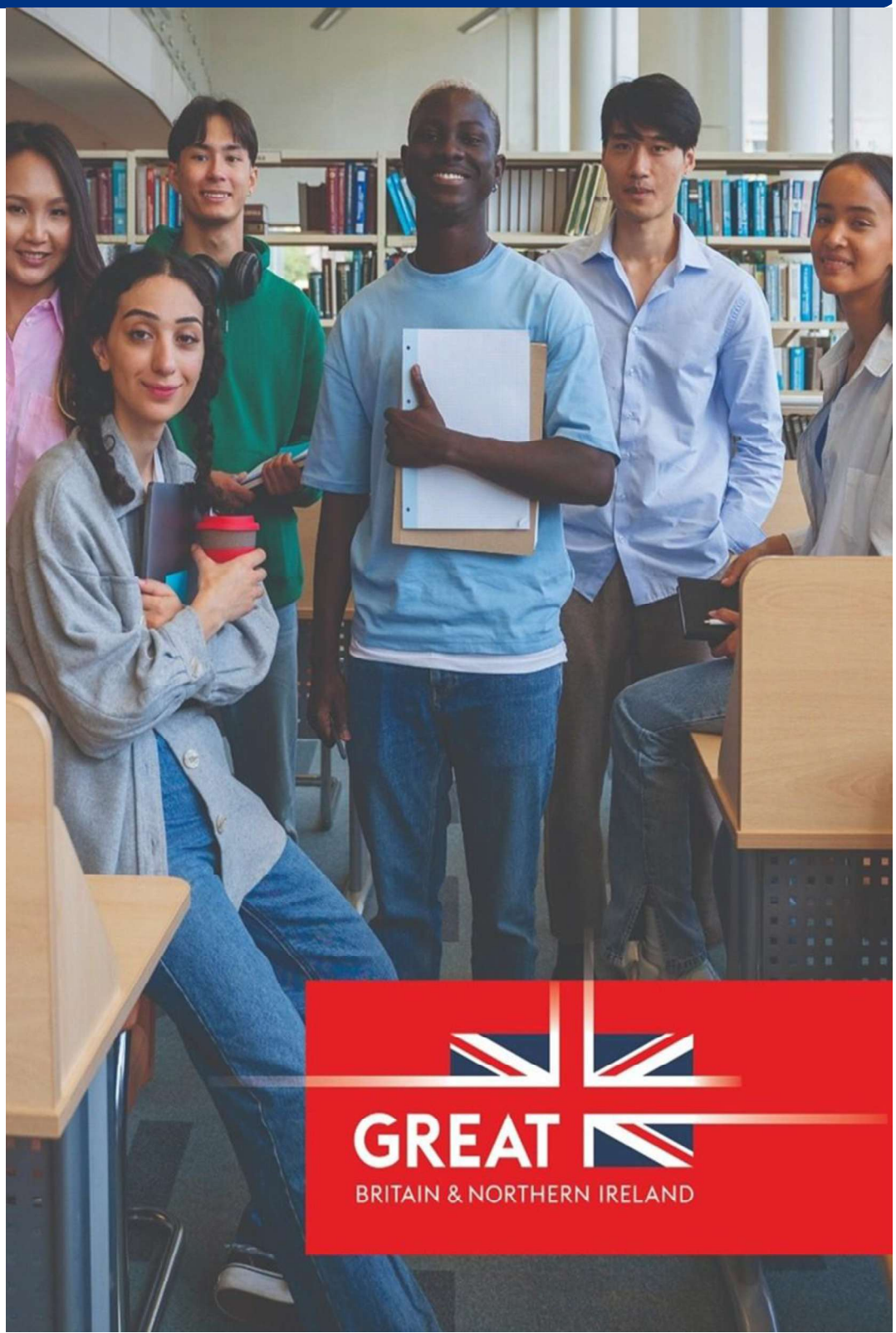
Metodología y herramienta para evaluación y clasificación de exámenes de inglés y lenguas europeas

Cambridge University Press & Assessment

Department for Business and Trade
Education sector
UK Government

Con aportes externos de:
British Council
Oxford University Press
Pearson

Noviembre 2025



Embajada Británica Colombia

UK - COLOMBIA 200 AÑOS

Aviso sobre Independencia Institucional:

Esta metodología y su herramienta complementaria fueron desarrolladas y revisadas por proveedores de evaluación británicos, instituciones y el sector educativo del Departamento de Negocios y Comercio del Reino Unido, con el propósito de apoyar al Ministerio de Educación de Colombia en el fortalecimiento de su política de evaluación de lenguas. Si bien este documento refleja las mejores prácticas internacionales e incorpora aportes de actores británicos, el Ministerio de Educación conserva plena autonomía y autoridad sobre la adopción, adaptación e implementación de cualquier recomendación o herramienta aquí presentada. La inclusión de marcos desarrollados en el Reino Unido no implica respaldo ni obligación; todas las decisiones finales relacionadas con la política de evaluación de lenguas permanecen exclusivamente bajo la jurisdicción del Ministerio de Educación de Colombia.

Créditos

Este documento metodológico y la herramienta fueron desarrollados y revisados como un esfuerzo colaborativo entre expertos en evaluación lingüística y educación internacional, provenientes de proveedores británicos de evaluación y del ámbito académico.

Documento elaborado por:

- **Dr. Szilvia Papp**
  Senior Research Manager, English Research Group, Cambridge University Press & Assessment
- **Dr. Jing Xu**
  Head of Propositions Research, English Research Group, Cambridge University Press & Assessment
- **Safiatu Azevedo Lopes**
  Research Analyst, English Research Group, Cambridge University Press & Assessment
- **Dayana Cortés**
  Education Business Development Manager, British Embassy in Colombia

Con aportes externos de:

- **British Council:**
  Richard Spiby, Head of Assessment Research
  Aidan Holland, Global Assessment Solutions Manager, Wider Europe
- **Oxford University Press:**
  Collin Finnerty, Head of Assessment Research
  Anthony Green, Director of Research
  Martín Caicedo, Regional Business Development Manager - Assessment
- **Pearson:**
  Carmen Quintana, Senior Product Manager - Assessments
  Ezneider Lozano, Assessment Regional Manager

Con la contribución especial del profesor Luke Harding, Universidad de Lancaster

Revisado por:

a) **Kelly Tavares**
   Head of Education LATAC, Department for Business and Trade (DBT)
- **Hazel Mariscal**
  Regional Recognition Manager – Americas, English, Cambridge University Press & Assessment
- **Fabián Eugenio Gastón García**
  Senior Business Development Manager Southern Cone & Andes – English

Contenidos

Créditos 2

1. Contexto 4

1.1 Contexto histórico del Programa Nacional de Bilingüismo en Colombia 4

1.2 Colaboración Colombo-Británica en evaluación de lenguas y desarrollo de programas ..5

1.3 Apoyo del gobierno británico para una política sostenible de evaluación lingüística 6

2. Metodología para la revisión sistemática y actualización de instrumentos de evaluación lingüística 6

2.1 Propósito, funcionalidad y diseño metodológico 8

3. Definiciones de los tres tipos de evaluación de lenguas 9

Clasificación 9

Proficiencia 10

Certificación 10

Exámenes con propósitos específicos (ESP, EAP) 11

4. Características clave para clasificar exámenes de idiomas 13

5. Procedimiento recomendado para la actualización de la Resolución 018035: 17

6. Limitaciones de la herramienta y próximos pasos 18

Apéndice 1: Herramienta de generación de informes (Excel) 20

Apéndice 2: Definiciones 22

Apéndice 3: Preguntas para los proveedores de exámenes 24

(Incluidas en la herramienta de Excel para subir a herramienta de recolección de datos) 24

Referencias 30

Metodología y herramienta para evaluación y clasificación de exámenes de inglés y lenguas europeas

1. Contexto

1.1 Contexto histórico del Programa Nacional de Bilingüismo en Colombia

El Programa Nacional de Bilingüismo (PNB), lanzado en 2004 por el Ministerio de Educación Nacional (MEN), representa un pilar fundamental de la política educativa de Colombia orientada a fortalecer el aprendizaje de lenguas extranjeras, especialmente el inglés, en los niveles de educación preescolar, básica y media. Basado en la Ley General de Educación (Ley 115 de 1994), el PNB fue concebido como una iniciativa nacional estratégica para mejorar la calidad educativa, promover la equidad y preparar a los estudiantes colombianos para oportunidades globales mediante el fortalecimiento de sus competencias lingüísticas (Mora et al., 2019).

Desde su creación, el PNB ha sido liderado y moldeado por las autoridades colombianas, reflejando una visión del bilingüismo que se ajusta a las diversas realidades socioculturales del país. El Ministerio ha desarrollado e implementado una serie de estándares nacionales, lineamientos curriculares y programas de formación docente, además de promover enfoques pedagógicos innovadores y recursos digitales.

De manera significativa, el PNB ha evolucionado para adoptar un enfoque de plurilingüismo funcional, reconociendo la rica herencia lingüística de Colombia, que incluye más de 65 lenguas nativas, dos lenguas criollas, el romaní y la Lengua de Señas Colombiana, y posicionando el aprendizaje de lenguas extranjeras como una herramienta para el diálogo intercultural, la ciudadanía y el aprendizaje a lo largo de la vida (Ministerio de Educación Nacional de Colombia, s.f.).

Si bien marcos internacionales como el Marco Común Europeo de Referencia para las Lenguas (MCER) han orientado aspectos técnicos del programa, el diseño, la implementación y la adaptación de la política de educación bilingüe permanecen bajo la plena autoridad del gobierno colombiano. El PNB refleja la postura proactiva de Colombia en la construcción de su propio futuro educativo, donde la colaboración internacional actúa como complemento, no como motor, de su política nacional.

1.2 Colaboración Colombo-Británica en evaluación de lenguas y desarrollo de programas

El Ministerio de Educación Nacional (MEN) de Colombia ha liderado la agenda de educación bilingüe del país. Desde el lanzamiento del Programa Nacional de Bilingüismo (PNB), el MEN ha buscado activamente alianzas estratégicas para fortalecer los marcos normativos, los programas de desarrollo docente y los sistemas de evaluación lingüística.

La colaboración internacional, incluida la cooperación con instituciones británicas, ha complementado estos esfuerzos al aportar experiencia técnica y perspectivas globales a la agenda de educación bilingüe. Organizaciones como Cambridge University Press & Assessment, Oxford University Press, Pearson y el British Council han contribuido mediante funciones de asesoría, iniciativas de fortalecimiento de capacidades y el intercambio de mejores prácticas en evaluación y aprendizaje de lenguas. Igualmente importante, los investigadores y universidades colombianas han desempeñado un papel fundamental en el avance de la educación bilingüe a través de investigaciones aplicadas, innovaciones en la formación docente y recomendaciones basadas en evidencia. Esta contribución académica ha fortalecido el desarrollo de políticas del MEN y ha garantizado que las mejoras se fundamenten en realidades locales y en las mejores prácticas pedagógicas.

De manera crucial, todas las decisiones relacionadas con la adopción de herramientas y estándares internacionales, incluido el Marco Común Europeo de Referencia para las Lenguas (MCER), han sido tomadas por las autoridades colombianas. Esto garantiza que las contribuciones extranjeras se alineen con las prioridades nacionales y los objetivos educativos.

La estrategia de educación en lenguas en Colombia se ha configurado principalmente a través de procesos nacionales liderados por el MEN. El reconocimiento formal de los exámenes internacionales mediante la NTC 5580, resoluciones subsiguientes, incluyendo la Resolución 018035 subraya el compromiso del MEN con estándares rigurosos y una alineación transparente con los marcos globales existentes, al tiempo que mantiene plena soberanía sobre su ecosistema de evaluación lingüística.

Sobre la base de esta trayectoria de colaboración, la iniciativa actual tiene como propósito contribuir al perfeccionamiento de la metodología y del marco de clasificación para las evaluaciones lingüísticas utilizadas en Colombia, proporcionando al Ministerio herramientas y evidencia para modernizar la Resolución 018035 y fortalecer los procesos de toma de decisiones.

1.3 Apoyo del gobierno británico para una política sostenible de evaluación lingüística

Con el objetivo de apoyar la estrategia de educación bilingüe de Colombia, el Gobierno Británico, a través del respaldo de su sector educativo y de los principales proveedores de evaluación lingüística, está contribuyendo mediante la entrega de una metodología y una herramienta digital al Ministerio de Educación de Colombia. Esta contribución busca apoyar al Ministerio en sus esfuerzos continuos por actualizar y perfeccionar la Resolución 018035, que regula el reconocimiento de los exámenes internacionales de competencia en inglés y otras lenguas.

El desarrollo de esta herramienta ha estado guiado por las prioridades colombianas en materia de evaluación lingüística y ha sido diseñado para responder a la visión a largo plazo del Ministerio, orientada hacia una política de evaluación transparente, basada en evidencia y sostenible.

Si bien las instituciones británicas han desempeñado un papel asesor, la propiedad, implementación y futura adaptación de la metodología y la herramienta recaen íntegramente en el Ministerio de Educación de Colombia. Esta iniciativa ejemplifica un modelo de cooperación internacional respetuosa, en el que la experiencia extranjera se aprovecha para fortalecer, y no para dirigir, la política nacional.

2. Metodología para la revisión sistemática y actualización de instrumentos de evaluación lingüística

La metodología y la herramienta que la acompaña fueron desarrolladas bajo el liderazgo académico de *Cambridge University Press & Assessment*, en estrecha colaboración con el Departamento de Negocios y Comercio del Reino Unido y la Embajada Británica en Colombia. Esta iniciativa refleja un compromiso compartido para fortalecer la política de evaluación lingüística de Colombia mediante un enfoque transparente, basado en la investigación y sostenible. Especialistas de los principales proveedores británicos de evaluación que operan en Colombia, incluidos *Oxford University Press, Pearson*, el *British Council* y un académico independiente de la Universidad de Lancaster especializado en evaluación lingüística, también contribuyeron proporcionando retroalimentación, garantizando el consenso entre los participantes y la alineación de la metodología y la herramienta con las mejores prácticas internacionales existentes.

Este documento presenta un sólido marco conceptual fundamentado en la lingüística aplicada y en estándares internacionales. Comienza describiendo el propósito, la funcionalidad y el diseño metodológico, e introduce tres categorías clave de evaluación lingüística estrechamente relacionadas con el contexto nacional de Colombia: pruebas de clasificación, pruebas de proficiencia y pruebas de certificación. Cada categoría se define por su propósito, alcance, decisiones comunes basadas en los resultados de las pruebas y su alineación con el Marco Común Europeo de Referencia para las Lenguas (MCER, 2020), de acuerdo con una obra de referencia ampliamente aceptada en evaluación lingüística (Davies et al., 1999). El documento también especifica los criterios mínimos que todos los proveedores de evaluación deben cumplir para clasificar sus pruebas en cada categoría, incluyendo validez, confiabilidad, alineación con el MCER, seguridad de la prueba, transparencia en la calificación, accesibilidad, entre otros.

Para implementar la metodología, este documento se acompaña de una herramienta personalizada en Excel. La herramienta permite al Ministerio de Educación verificar y evaluar sistemáticamente la información presentada por los proveedores de evaluaciones. Facilita la revisión periódica de los productos de evaluación de lenguas europeas, garantizando que solo aquellos que cumplen con los criterios establecidos se incluyan en la versión actualizada de la Resolución 018035. Diseñada para un ciclo de revisión bienal, la herramienta promueve la coherencia, la equidad y la adaptabilidad en el reconocimiento de exámenes de idiomas en todo el sistema educativo colombiano.

Recomendamos que el Ministerio de Educación de Colombia considere adoptar la metodología propuesta y la herramienta en Excel como una solución provisional. Esta recomendación responde a las limitaciones actuales en recursos y experiencia técnica, así como a la necesidad urgente de clasificar, de manera sistemática y práctica, la amplia gama de pruebas de idiomas utilizadas en el país. Reconocemos que este enfoque temporal tiene limitaciones, especialmente porque delega la tarea crítica de revisar la calidad de las pruebas a organismos profesionales externos —como entidades de reconocimiento y acreditación— que pueden tener un conocimiento limitado del contexto local de evaluación en Colombia. En consecuencia, la evidencia sobre la calidad de las pruebas exigida bajo esta solución provisional está en gran medida descontextualizada y no refleja completamente las necesidades nacionales de evaluación ni las prioridades de política. Para cerrar estas brechas, proponemos una solución integral, descrita al final de este documento, que incluye la creación de un panel independiente de expertos en evaluación de idiomas para revisar la evidencia de validez presentada por los proveedores de evaluaciones.

2.1 Propósito, funcionalidad y diseño metodológico

En esta sección, describimos el diseño de la herramienta y cómo puede utilizarse adecuadamente incluso por personas sin experiencia técnica.

2.1.1 Propósito, funcionalidad y diseño metodológico

Se espera que la herramienta se use de las siguientes maneras:

- Verificar y confirmar las afirmaciones que hacen los proveedores de pruebas sobre sus productos.
- Seleccionar evaluaciones específicas para un propósito, contexto o población determinada (ver Anexo 1 – herramienta en Excel y método propuesto de recolección de datos).
- Incluir ayudas visuales para los usuarios finales, como instituciones educativas (colegios, universidades, empresas) y otros actores, que faciliten la interpretación de la clasificación y los resultados de la evaluación.

La herramienta incluye un conjunto de preguntas orientadoras para evaluar la calidad de las pruebas, con énfasis en validez, confiabilidad y relevancia contextual. La metodología promueve un enfoque colaborativo y basado en evidencia, donde el uso previsto de la prueba, su alineación con los objetivos educativos y su pertinencia para la población objetivo son aspectos centrales del proceso de clasificación. La herramienta respalda una evaluación integral que refleja la diversidad de contextos educativos en Colombia.

En segundo lugar, se presenta un conjunto de características clave que distinguen los tres tipos de evaluación (clasificación, proficiencia y certificación). Estas características sirven de base para una serie de preguntas orientadoras fáciles de usar, diseñadas para apoyar la clasificación automática.

En tercer lugar, se ha desarrollado un diagrama de decisión (Figura 2) con la lógica de respuesta incorporada para facilitar el proceso de clasificación automática solicitado por el Ministerio de Educación.

2.1.2 Marcos Complementarios para la Validación Contextual

Además del modelo de validación socio-cognitivo de *Cambridge University Press & Assessment* (Weir, 2005), la metodología reconoce el valor de marcos complementarios como el *Comprehensive Learning System* (O'Sullivan, 2021), *Global Scale of English* (GSE) de Pearson (Mayor, 2024), el *CEFR Alignment Handbook* del British Council (British Council,

UKALTA, EALTA y ALTE, 2022) y el *Assessment Use Argument* (AUA) de Oxford University Press.

Con perspectivas distintas pero compatibles, estos marcos apoyan un enfoque colaborativo y sensible al contexto para la evaluación, priorizando el uso de la prueba, el propósito educativo y la relevancia para los actores involucrados. Esto está alineado con los objetivos de Colombia para una política de evaluación de idiomas inclusiva y centrada en el estudiante. Cada uno contribuye a fortalecer la evaluación de pruebas de idiomas al enfatizar la progresión, la relevancia para los actores, el impacto educativo y la validez del constructo.

3. Definiciones de los tres tipos de evaluación de lenguas

En esta sección describimos las características que definen los tres tipos de evaluación de lenguas según Davies et al. (1999) (ver Anexo 2). Estas características sirven como base para los criterios de clasificación incluidos en la herramienta. El reconocimiento público y la confianza, respaldados por evidencia sobre la calidad de la prueba, son esenciales para las pruebas de certificación.

Cabe destacar que, tras consultar con los numerosos profesionales en evaluación de idiomas involucrados en este proyecto, decidimos ofrecer a los proveedores de pruebas dos opciones para demostrar la calidad de sus exámenes durante este período provisional. Estas dos opciones se explicarán en la siguiente sección.

La razón detrás de esta decisión es que la practicidad y la flexibilidad son consideraciones clave para este método inicial de clasificación, hasta que se pueda desarrollar un enfoque más riguroso. Esto evita el riesgo de excluir pruebas que gozan de buen reconocimiento en otros contextos, al tiempo que se les exige cumplir con un estándar más alto para ser consideradas como pruebas de certificación.

Clasificación

- Ubicar a los estudiantes en grupos, clases, cursos, programas o niveles adecuados.
- Tradicionalmente, consiste en una serie de tareas o preguntas que corresponden a distintos niveles de idioma, organizadas por dificultad; más recientemente, se usan pruebas adaptativas.

- Decisiones de bajo impacto, fáciles de corregir si se cometen errores basados en los resultados.
- Los aspectos de confiabilidad y validez son menos críticos.
- A veces se utiliza para diagnosticar fortalezas y debilidades del estudiante.

Proficiencia

- No está basada en un curso de instrucción.
- Estandarizada para uso a nivel mundial.
- Los resultados indican la habilidad en una amplia variedad de contextos generales (por ejemplo, social, turístico, personal, público).
- A veces se utiliza para propósitos de ingreso o egreso.
- Puede evaluar una o más destrezas.
- Puede ser directa (implica uso real del idioma), indirecta (evalúa conocimientos subyacentes) o semidirecta (usa estímulos y respuestas grabadas).
- La vigencia de los resultados suele ser de 2 años.
- El proveedor de la prueba debe presentar evidencia de validez y confiabilidad.
- Sin embargo, la prueba puede no estar acreditada por autoridades o entes reguladores.

Certificación

- Reconocimiento público de un nivel de dominio del idioma.
- El reconocimiento puede otorgarlo gobiernos, instituciones, empleadores o evaluadores independientes.
- La acreditación por organismos externos (p. ej., Ofqual, ALTE, otros) es una vía para el reconocimiento, pero no es un requisito.
- Generalmente alineada con los niveles del MCER (CEFR).
- Se utiliza para decisiones de alto impacto (p. ej., estudio, trabajo, migración).
- Referenciada a criterios, basada en estándares de desempeño claramente definidos.
- Puede adoptar diversas formas (directa, indirecta, semidirecta), sin estar ligada a un formato específico.
- Debe presentarse evidencia de validez y confiabilidad para el uso previsto y la población objetivo.

- El período de vigencia de los resultados lo determina la institución que otorga el reconocimiento.
- Los resultados pueden reportarse como aprobado/reprobado o mediante puntajes escalados, con servicios de verificación disponibles.

El dominio de uso del idioma (TLU, por su acrónimo en inglés, *Target Language Use*) o las situaciones en las que los examinados utilizarán el idioma evaluado es una consideración fundamental para la validez de la prueba. En algunos contextos de evaluación, el TLU se refiere al uso del idioma en dominios específicos, como las prácticas comunicativas en las profesiones de salud o en la aviación. Las pruebas diseñadas para evaluar el uso del idioma en dominios específicos se denominan pruebas para propósitos específicos (ESP, por su acrónimo en inglés, *English for Specific Purposes*). La elaboración de estas pruebas suele involucrar a expertos en el área.

Aunque el TLU de una prueba no contribuye directamente a su clasificación, esta información es crucial para determinar la validez del examen y, por lo tanto, debe ser claramente presentada por los proveedores de pruebas.

Exámenes con propósitos específicos (ESP, EAP)[1]

Las pruebas para propósitos específicos se incluyen dentro de las categorías generales de evaluación de clasificación, proficiencia y certificación, pero se distinguen por su enfoque dirigido al uso del idioma en ámbitos ocupacionales o académicos definidos. Estas pruebas están diseñadas para evaluar si el estudiante posee la competencia comunicativa necesaria para desenvolverse eficazmente en contextos reales específicos.

Características principales:

• **Diseño basado en necesidades:** Se desarrolla a partir de un análisis exhaustivo de las necesidades del grupo objetivo, identificando tareas comunicativas, géneros y funciones lingüísticas relevantes para el dominio (p. ej., salud, aviación, ingeniería, academia).

[1] ESP: English for Specific Purposes; EAP: English for Academic Purposes

• **Contenido específico del dominio:** Incluye vocabulario especializado, patrones discursivos y tipos de tareas alineados con las demandas comunicativas del campo (p. ej., para profesionales de la salud, controladores aéreos, pruebas académicas para educación superior).

• **Propósito real:** Evalúa si el examinado puede realizar tareas relevantes para el dominio, como comprender conferencias, redactar informes, participar en interacciones profesionales o traducir material técnico.

• **Funcionalidad flexible:** Puede cumplir múltiples propósitos de evaluación: clasificación, logro, dominio o certificación, según el contexto y el nivel de importancia.

• **Alineación con el MCER:** Aunque se adapta a dominios específicos, estas pruebas deben demostrar una alineación transparente con los descriptores del MCER, especialmente los relacionados con mediación, interacción y comunicación profesional.

• **Validez y confiabilidad:** Deben presentar evidencia de validez del constructo y confiabilidad, especialmente cuando se usan para decisiones de alto impacto.

4. Características clave para clasificar exámenes de idiomas

Como se describe en la sección 3, las evaluaciones de idiomas pueden clasificarse en tres categorías principales: clasificación, proficiencia y certificación. Aunque estas categorías comparten algunas características, se diferencian en su propósito, alcance y en las decisiones que respaldan.

El *CEFR Companion Volume* (2020) señala que las distinciones más significativas surgen del uso previsto de la prueba, el tipo de decisiones que permiten los resultados y la evidencia de validez y confiabilidad que respalda esas decisiones.

Los resultados de las pruebas de clasificación suelen utilizarse para orientar la enseñanza y el aprendizaje, donde las implicaciones de las decisiones para los usuarios son relativamente de bajo impacto. La cuestión de si una prueba se clasifica como de proficiencia o certificación surge cuando las decisiones son de alto impacto (high stakes) (ver Figura 1).

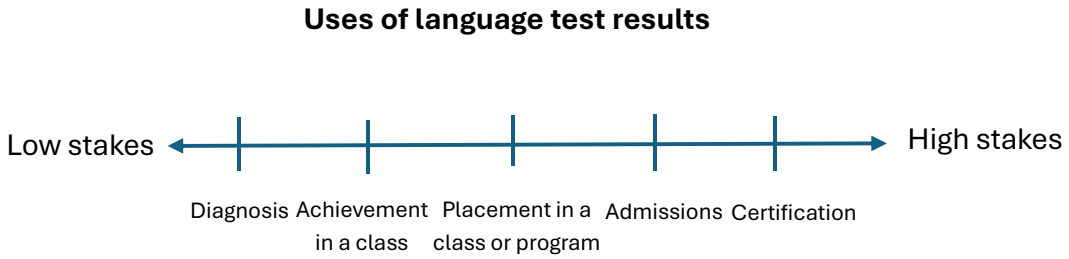


Figura 1. Usos de los resultados de los exámenes de lemguas (adaptado de Stoynoff & Chapelle, 2005, p. 11)

Es importante destacar que la certificación se define por el reconocimiento público de la calidad de la prueba, a veces respaldado por una revisión formal realizada por un tercero independiente mediante procesos de regulación o acreditación que implican examinar las afirmaciones del proveedor y la evidencia presentada. Aunque no es un requisito esencial para obtener el estatus de certificación, la validación por una autoridad externa refuerza la confianza y el reconocimiento público.

Si bien la acreditación y la regulación pueden fortalecer y contribuir a la confianza pública, no son prerrequisitos. Este enfoque provisional de clasificación adopta la postura de que

las pruebas alineadas con el MCER y utilizadas para fines oficiales —como aquellas reconocidas por ministerios de educación, empleadores o evaluadores independientes— no deben perder su estatus de certificación por carecer de acreditación formal.

Este enfoque respalda, por ahora, un marco de clasificación más inclusivo, flexible y sensible al contexto, ofreciendo dos rutas alternativas para demostrar la calidad requerida en el nivel más alto de certificación.

Para garantizar una clasificación coherente y sensible al contexto de las pruebas de idiomas, se proponen los siguientes criterios, ordenados según su relevancia para la toma de decisiones y su alineación con el MCER:

1. **Uso de los resultados de la prueba / Tipo de decisión que respalda**
    - ¿Qué decisiones permiten tomar los resultados de la prueba??
    - Implicaciones de bajo (*low-stakes*) o alto impacto (*high stakes*), por ejemplo, relacionadas con empleo, ingreso académico, inmigración, rendición de cuentas.
2. **Prueba de calidad / Evidencia de validez**
    - ¿Existe evidencia documentada que demuestre que la prueba mide lo que afirma medir y que es adecuada para su propósito??
    - Incluye evidencia en documentación pública sobre el propósito de la prueba, el diseño de las tareas y los posibles usos (por ejemplo, un manual técnico o un informe de validación que describa las especificaciones de la prueba y haga referencia a conceptos clave en los marcos de validación, como afirmaciones de validez, justificaciones y refutaciones, confiabilidad, equidad, alineación con el MCER y el impacto de la prueba).
3. **Seguridad de la prueba**
    - ¿Existe una infraestructura dedicada que garantice la seguridad de la prueba o medidas antifraude?
    - Incluye la infraestructura, procedimientos, procesos, departamentos y equipos que garantizan una administración segura de las pruebas (incluyendo la entrega segura del examen, verificación de identidad del candidato, supervisión, controles contra malas prácticas, etc.).

4. **Que la prueba tenga:**

    a. **Reconocimiento institucional internacional**

    - ¿Está el examen reconocido por universidades, ministerios o empresas en más de 10 países??, O

    b. **Aprobación/verificación por terceros (Ej. Acreditación y regulación)**

    - ¿Existe aprobación o validación por parte de un organismo profesional calificado que también haya aprobado o validado pruebas de otros proveedores? (por ejemplo, ¿el examen ha sido auditado externamente o aprobado por organismos reguladores o de acreditación?)
    - Aunque no es esencial para el estatus de certificación, contribuye a la confianza y reconocimiento público.

5. **Verificación de resultados**

    - ¿Existen mecanismos seguros y transparentes que permitan a las partes interesadas —como instituciones educativas, empleadores, organismos gubernamentales y otros— confirmar los resultados de los titulares de certificados y verificar la autenticidad y validez del certificado, reduciendo el riesgo de fraude y aumentando la confianza en el proceso de certificación?

    - Incluye una descripción detallada del sistema de verificación de resultados, incluyendo los medios tecnológicos o procedimentales mediante los cuales terceros pueden acceder o confirmar la información del certificado (por ejemplo, portales de verificación en línea, códigos QR, bases de datos seguras). La descripción debe incluir los protocolos de acceso, medidas de protección de datos y ejemplos de cómo se realiza habitualmente la verificación.

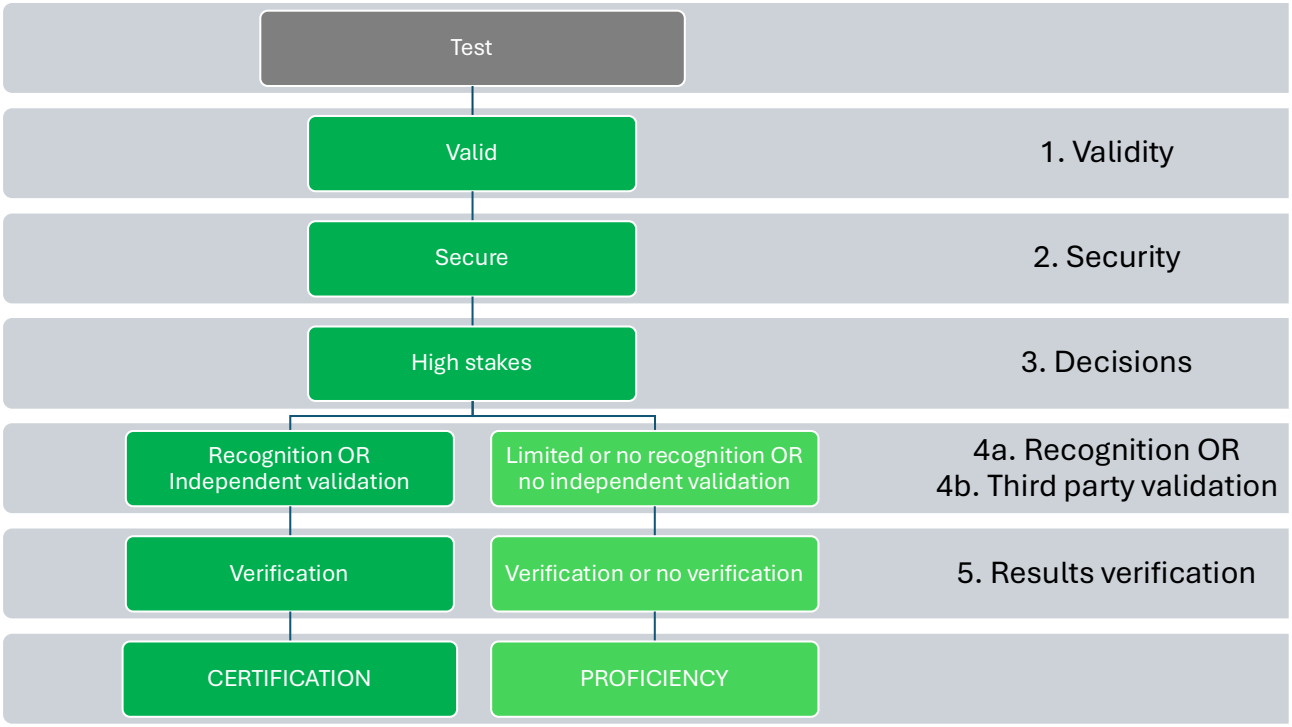


Figura 2. Diagrama de decisión que ilustra el proceso de clasificación en CERTIFICACIÓN o PROFICIENCIA

4.1 Preguntas específicas para los proveedores de pruebas

Las características de los criterios anteriores se han utilizado para crear preguntas fáciles de usar que informan la clasificación automatizada. Un conjunto de características adicionales y preguntas relacionadas permite la selección de evaluaciones específicas para un propósito, contexto o población objetivo de examinados (ver Apéndice 3). Estas preguntas también sirven como base para un enfoque más riguroso de clasificación que se desarrollará tras el período provisional.

5. Procedimiento recomendado para la actualización de la Resolución 018035:

En esta sección describimos detalles sobre cómo llevar a cabo la clasificación de la evaluación lingüística utilizando la herramienta, así como las directrices que los proveedores de pruebas deben seguir.

- **Recolección de datos**

  Se recomienda que el Ministerio de Educación recopile información sobre cada examen utilizando una herramienta de recopilación de datos en línea, como Google Forms, Microsoft Forms, Microsoft Lists, SurveyMonkey o una plataforma similar. Ver Apéndice 1.

- **Establecer instrucciones claras para el registro de la información.**

  Comunicar claramente en la convocatoria pública la fecha límite de presentación, la documentación requerida y los criterios de evaluación. Enfatizar que todas las evaluaciones deben demostrar alineación con el Marco Común Europeo de Referencia para las Lenguas (MCER), en el caso de las lenguas europeas.

- **Establecer un cronograma de respuesta**

  Los procedimientos y el plazo para responder a los proveedores de evaluaciones deben ser determinados por el Ministerio de Educación de Colombia, de acuerdo con sus procesos internos, capacidad y prioridades de política. Esta flexibilidad permite al Ministerio adaptar el mecanismo de revisión a las necesidades cambiantes y garantiza que el proceso siga siendo transparente y eficiente. Se recomienda que el Ministerio comunique los procedimientos elegidos, por ejemplo, mediante una nota informativa en Colombia Aprende, para fomentar la confianza y la claridad entre las partes interesadas.

- **Establecer un ciclo de revisión bienal**
  Actualizar la Resolución correspondiente cada dos años, asegurando que la lista de exámenes reconocidos se mantenga actualizada y alineada con los estándares internacionales.

- **Compartir una comunicación pública en la plataforma Colombia Aprende explicando:**
    - Las tres categorías de evaluaciones lingüísticas: clasificación, proficiencia y certificación.
    - El propósito de la resolución actualizada.
    - La importancia de utilizar la lista oficial como referencia para universidades, escuelas y el sector corporativo al seleccionar evaluaciones lingüísticas.

- **Organizar seminarios web informativos para el público en general**

Organizar seminarios web de acceso abierto dirigidos a informar al público en general —en particular a educadores, estudiantes, padres y actores institucionales— sobre las diferentes categorías de evaluaciones lingüísticas (clasificación, proficiencia y certificación). Estas sesiones también deben explicar el propósito de la Resolución, cómo interpretar la lista oficial de exámenes reconocidos y cómo esta información puede orientar la toma de decisiones en contextos educativos y profesionales.

6. Limitaciones de la herramienta y próximos pasos

Es importante reconocer que la metodología y la herramienta propuestas forman parte de una solución recomendada de forma provisional hasta que se puedan elaborar lineamientos más completos y matizados. Reconocemos que la clasificación de los tres tipos de pruebas de idioma solo puede ser sólida y rigurosa mediante una revisión independiente de expertos, basada en la información y evidencia presentada por los proveedores de pruebas.

Los participantes en la discusión para la finalización de este documento sugirieron de manera unánime que el Ministerio de Educación constituya a futuro un panel independiente integrado por profesionales locales y globales en evaluación lingüística, con el apoyo de un organismo profesional independiente, como la *International Language Testing Association (ILTA) o la Latin American Association for Language Testing and Assessment (LAALTA),* para

adaptar la herramienta y la metodología subyacente al contexto y propósitos de evaluación locales.

También se recomienda que los proveedores de pruebas paguen una tarifa razonable para apoyar financieramente este proceso de revisión independiente.

Apéndice 1: Herramienta de generación de informes (Excel)

La clasificación final en la que se basa el código de Excel se fundamenta en las siguientes características de criterio:

La categoría de examen "CERTIFICACIÓN" debe cumplir con estos cinco criterios:

1. Evidencia de calidad / Evidencia publicada de validez
2. Seguridad de la prueba
3. Decisiones que respalda: *high stakes*
4. Que cumpla con:
    4a. Reconocimiento en más de 10 países, **o bien**,
    4b. Aprobación / validación / verificación por terceros
5. Verificación

La categoría de examen de "PROFICIENCIA" debe cumplir con estos tres criterios:

1. Evidencia de calidad / Evidencia publicada de validez
2. Seguridad de la prueba
3. Decisiones que respalda: *high stakes*

La categoría de examen de "CLASIFICACIÓN" cumple con los siguientes criterios:

1. Decisiones que respalda: *low stakes*

Sugerencias para la recolección de los datos:

Se recomienda que el Ministerio de Educación recopile la información de cada prueba utilizando una herramienta en línea, como Google Forms, Microsoft Forms, Microsoft Lists, SurveyMonkey u otra plataforma similar. Esto debe permitir enviar el enlace directamente a los proveedores de exámenes y garantizar que todos los datos permanezcan bajo la custodia exclusiva del Ministerio.

La herramienta también puede incluir datos de contacto de la organización evaluadora, como una dirección de correo electrónico o página de soporte, para que las instituciones que utilizan los exámenes o los candidatos puedan resolver cualquier duda.

Instrucciones para usar la hoja de cálculo en Excel.

(ver Apéndice 3 y la herramienta en Excel para la lista de preguntas incluidas)

1. Crear la encuesta utilizando las preguntas propuestas en la herramienta de Excel proporcionada con este documento.
2. Envíe el enlace a los proveedores de exámenes junto con las instrucciones para su diligenciamiento. Se debe completar un formulario por cada examen que se desee incluir en

la Resolución. Establezca una fecha límite para las respuestas; se recomienda un máximo de quince días calendario.

3. Transfiera las respuestas correspondientes a la hoja “Respuestas de proveedores de exámenes” en el archivo de Excel.
4. La pestaña “Classification” devolverá automáticamente CERTIFICATION, PROFICIENCY o PLACEMENT para cada examen.
5. En aquellos casos en los que no sea posible clasificar según la lógica de decisión, aparecerá la categoría “Not able to classify”.
6. Si faltan respuestas en las características clave, el sistema de clasificación automática devolverá la categoría “Missing info”.
7. Después de revisar y verificar la evidencia suministrada por los proveedores de exámenes, el Ministerio podrá cambiar manualmente las respuestas (por ejemplo: si el proveedor indicó “Yes” en Verificación de resultados, pero no proporcionó evidencia de cómo se verifican los resultados, el MEN puede cambiar el “Yes” a “No” y la clasificación cambiará automáticamente de CERTIFICATION a PROFICIENCY).

Apéndice 2: Definiciones

Adaptado de: Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of language testing*. Cambridge University Press.

Examen de clasificación	Una prueba diseñada para proporcionar información que ayude a ubicar a los estudiantes en clases apropiadas. Para que sea más efectiva, las pruebas de clasificación deben reflejar las características del contexto de enseñanza (como el nivel de competencia de las clases, la metodología y el tipo de programa). Por ejemplo, una prueba de clasificación basada en gramática puede no ser la más adecuada cuando el programa está orientado a tareas. En contextos donde las admisiones masivas de estudiantes son la norma, la eficiencia en la administración y corrección suele ser una consideración clave en el desarrollo de los procedimientos de ubicación. Dado que es relativamente fácil corregir errores derivados de los resultados de la prueba (por ejemplo, estudiantes ubicados incorrectamente), los aspectos de fiabilidad y validez pueden considerarse menos cruciales en el desarrollo o selección de una prueba de clasificación que en pruebas de alto impacto, donde los resultados son serios e irreversibles (como las pruebas de selección o cribado). Las pruebas de clasificación suelen cumplir una doble función: ubicación y diagnóstico. Cuando este es el caso, los docentes pueden utilizar la información proporcionada por la prueba sobre cada estudiante para orientar la planificación del curso.
Examen de proficiencia general vs. examen de proficiencia para propósitos específicos	Una prueba en la que el comportamiento que se examina es la competencia lingüística, generalizable a una amplia gama de contextos no especificados ni especificables. Existen numerosas pruebas de idioma que buscan realizar una evaluación general del conocimiento o la habilidad lingüística, aunque no siempre se explicita con claridad qué incluye esa competencia general. Estas pruebas pueden evaluar la capacidad para usar una o más macrodestrezas, o centrarse en el conocimiento de los sistemas lingüísticos; pueden ser pruebas directas o indirectas. Las pruebas de competencia general se contrastan con las pruebas de propósito específico, donde el dominio de uso del lenguaje está más claramente definido.
Examen de proficiencia para un contexto específico de uso de la lengua	Una prueba que mide cuánto de un idioma ha aprendido una persona. A diferencia de una prueba de logro, una prueba de competencia no se basa en un curso de instrucción específico. Una prueba de competencia suele medir lo que el candidato ha aprendido en relación con un propósito real, por ejemplo, si sabe lo suficiente del idioma objetivo para seguir una conferencia, formarse como ingeniero o trabajar como instructor de esquí en ese idioma, o para traducir con el nivel requerido desde ese idioma.

Certificación

El reconocimiento público del desempeño en una prueba. En la evaluación de lenguas, la certificación puede estar vinculada a la capacidad para realizar tareas específicas basadas en el idioma (por ejemplo, presentar un informe oral o completar formularios adecuadamente), o a la capacidad del candidato para cumplir con los requisitos de un trabajo determinado (por ejemplo, guía turístico o docente). También puede reflejar el progreso del estudiante en un curso o su finalización. Aunque la certificación suele asociarse con la evaluación basada en competencias, las pruebas de lengua utilizadas para la certificación pueden adoptar muchas formas y reflejar diferentes concepciones de la competencia lingüística.
Cabe destacar que algunas pruebas de proficiencia han sido estandarizadas y reconocidas para uso mundial. El alto nivel de rigor y calidad de estas pruebas estandarizadas se establece mediante investigaciones exhaustivas sobre validez, confiabilidad, equidad e impacto en diversos contextos de uso. Estas pruebas de competencia reconocidas internacionalmente suelen estar avaladas por organismos gubernamentales (por ejemplo, UKVI) o entidades de acreditación (por ejemplo, ALTE). Sirven como herramientas de certificación que demuestran que los candidatos cumplen con los requisitos mínimos para estudios académicos o procesos migratorios.

Apéndice 3: Preguntas para los proveedores de exámenes

(Incluidas en la herramienta de Excel para subir a herramienta de recolección de datos)

Type of test: How do you describe the test?

a. Placement
b. Proficiency
c. Certification

Proof of quality/ Validity evidence: Is there documented evidence that the test measures what it claims to measure and is fit for purpose, including information on the validity and reliability of the test?

- No
- Yes

1. **Test Security:** Is there a dedicated infrastructure that ensures test security or anti-fraud measures?
   d. No
   e. Yes
2. **Decisions supported**: What kind of decisions are allowed by the results?
   a. Low-stakes decisions (e.g. teaching and learning, diagnosis, self-assessment)
   b. High-stakes decisions (e.g. employment, academic entry, immigration, accountability)
3. **EITHER / OR**
   a. **International institutional recognition**: Is the test recognized by universities, ministries, or corporations in 10+ countries?
      - No
      - Yes
   b. **Third party verification**: Is there third-party approval/validation from a qualified body that has also approved/validated tests from other providers?
      - No
      - Yes
4. **Results verification:** Do you offer a results verification service?
   - No

- Yes

5. **Test purpose**: What is the testing purpose?
    - To diagnose learners' strengths and weaknesses
    - To place learners at a specific level / into groups
    - To check learning at the end of a unit/semester/year/course
    - To measure general language proficiency
    - To certify language proficiency levels
6. **CEFR level and descriptor alignment**: Are the test tasks and scoring aligned with CEFR levels and descriptors, including mediation and interaction?
    a. No
    b. Yes
7. **Domain specificity**: Is the test designed for general or specific-purpose use?
    - General proficiency
    - Specific purpose (EAP, ESP)
8. **External framework**: Is the test related to an external framework?
    - CEFR
    - National curriculum guidelines
    - Content / domain specifications based on needs analysis
9. **Context / Purpose of language use:** What communicative contexts are prioritized?
    - Personal/social/public/daily communications
    - Study/academic/educational
    - Work/occupational/professional/business and commerce
    - Migration
10. **Use of test results:** For what purpose are the results used?
    - General purposes (e.g. self-assessment)
    - For decisions on teaching and learning in schools
    - For admissions into higher education degree in Colombia
    - For university admission abroad
    - For selection in the work environment
    - For (re)-qualification in the work environment
    - For the qualification of English language teachers
    - For migration: as part of a procedure for immigration
    - For social cohesion: as part of an application for citizenship

11. **Proficiency coverage**: What proficiency range is tested?
    - Single level (level-based) (e.g. B2)
    - Multilevel (e.g. A1-C2)
12. **CEFR level**: What CEFR levels are assessed?
    - Pre-A1
    - A1
    - A2
    - B1
    - B2
    - C1
    - C2
13. **Skills coverage:** Does the test assess one or multiple skills?
    - One skill
    - Multiple skills
14. **Format** (i.e. four skills, modular, combined, or integrated formats)
    - Four skills (all four skills, L,R,S,W)
    - Modular (i.e. skills taken in separate modules)
    - Combined skills (e.g. listening and speaking together, L&S)
    - Integrated skills (e.g. listening into speaking, L->S)
15. **Skills assessed**: What skills are assessed?
    - L,R,S,W
    - L
    - R
    - S
    - W
    - L&S
    - R&W
    - L->S
    - L->W
    - R->S
    - R->W
16. **Target candidature**: Who is the test designed for?
    - General public (young people and adults)

- Very young children (age 4 to 6)
- Children 6 to 9 years old
- Children 9 to 12 years old
- Teens (13 to 17 years old)
- Prospective higher education students
- Students (18-24 years old)
- Prospective employees
- Workers
- Professionals
- Teachers
- Asylum seekers or immigrants

17. **Accessibility and Fairness**: Does the test offer accommodations or design features for test takers with special needs and disabilities?
    - No
    - Yes – limited accommodations
    - Yes – comprehensive accessibility features
18. **Recognising organisations**: Who recognises the test?
    - Not applicable
    - HE institutions
    - Employers
    - Governments (e.g. Ministry of Education, UKVI, Home Office)
    - Accreditation body (e.g. ALTE)
    - Regulator (e.g. Ofqual)
    - Other, please specify
19. **Test information available**: Do you provide information about the test?
    - not applicable
    - handbook / test manual
    - sample or practice papers
    - assessment criteria
    - guidance on how to interpret results
    - special requirements
    - all of the above
20. **Test information provided for**: Who do you provide test information materials for?

- Not applicable
- Test takers
- Teachers
- Parents
- HE institutions
- Employers

21. **AI and technology**: Is AI or technology used in the following?
    - Not applicable
    - Test development
    - Test delivery
    - Scoring/marking
22. **Test administration**: How is the test administered?
    - Face-to-face, paper and pencil
    - Computer-based (PC, laptop, tablet/iPad)
    - Mobile
23. **Test delivery**: How is the test delivered?
    - in person: administered one on one
    - in person: in pairs
    - in person: in groups
    - online: with a human examiner
    - online: with a human examiner and another test taker
    - online: with the computer / avatar as partner
    - online: in groups
24. **Test security method**: How do you check test takers' identity?
    - None (self-administered at home or in school)
    - Human invigilation / on-site supervision
    - Remote monitoring by humans
    - Remote proctoring by machine technology/AI
25. **Results processing time**: How long before results are provided?
    - Immediate
    - 1-3 days
    - 1 week
    - More than 1 week

26. **Results verification method**: How do you make sure stakeholders can confirm certificate holders' results and verify the authenticity and validity of the certificate?
    - not applicable
    - secure and transparent mechanisms through exam centre
    - secure and transparent mechanisms by a registered user on the platform
27. **Results interpretation**: Do you provide an explanation to users of the appropriate uses of test results?
    - not applicable
    - how to interpret the results for the intended purpose
    - how to avoid specific misuses of results
28. **Results enquiries**: Are test takers able to query their results?
    - not applicable
    - have their responses checked
    - have their answers remarked
    - re-take the test
29. **Results expiry**: Is there a 'results valid for' period?
    - no 'results valid for' period
    - 2 years (e.g. SELT tests for visa and citizenship)
30. **Test impact**: Do you collect evidence on how the test benefits stakeholders and society?
    - not applicable
    - better learning outcomes
    - progress in English
    - better awareness of strengths and weaknesses
    - higher learner motivation
    - learning beyond the test
    - academic preparedness
    - better (higher standard) of teaching
    - improved classroom practice
    - increased teacher motivation
    - improved assessment culture and practice
    - economic growth through educational policy

Referencias

British Council. (2015). *English in Colombia: An examination of policy, perceptions and influencing factors*. https://www.britishcouncil.co/sites/default/files/colombia_version_final_-_ingles.pdf

British Council, UKALTA, EALTA and ALTE. (2022). *Aligning language education with the CEFR: A handbook*. cefr_alignment_handbook_layout.pdf

Council of Europe. (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing. https://book.coe.int/en/education-and-modern-languages/8152-common-european-framework-of-reference-for-languages-learning-teaching-assessment-companion-volume.html

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of language testing*. Cambridge University Press.Ministerio de Educación Nacional de Colombia. (2007). *Norma Técnica Colombiana NTC 5580: Programas de formación para el trabajo en el área de idiomas. Requisitos*. https://www.mineducacion.gov.co/1621/articles-157089_archivo_pdf_NTC_5580.pdf

Mayor, M. (2024) The Global Scale of English Reflections on 10 years of innovation. pearson-gse-10-years-of-innovation.pdf

Ministerio de Educación Nacional de Colombia. (2021). *Resolución 018035 del 21 de septiembre de 2021*. https://www.mineducacion.gov.co/1780/articles-406979_pdf.pdf

Ministerio de Educación Nacional de Colombia. (n.d.). *Programa Nacional de Bilingüismo*. https://educacionrindecuentas.mineducacion.gov.co/pilar-1-educacion-de-calidad/programa-nacional-de-bilinguismo/

Ministerio de Educación Nacional de Colombia. (n.d.). *About National Bilingual Program*. https://eco.colombiaaprende.edu.co/about-bilingualism/

Ministerio de Educación Nacional de Colombia. (n.d.). *Bases para una nación bilingüe y competitiva*. https://www.mineducacion.gov.co/1621/article-97498.html

Mora, R. A., Chiquito, T., & Zapata, J. D. (2019). Bilingual education policies in Colombia: Seeking relevant and sustainable frameworks for meaningful minority inclusion. In *Bilingualism and bilingual education: Politics, policies and practices in a globalized society* (pp. 55-77). Cham: Springer International Publishing.

O'Sullivan, B. (2021). The comprehensive learning system. British Council. https://www.britishcouncil.org/sites/default/files/the_comprehensive_learning_system_new_layout.pdf

Stoynoff, S., & Chapelle, C. A. (Eds.). (2005). *ESOL tests and testing: A resource for teachers and administrators*. TESOL.

Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan.

Department for Business and Trade



Department for Business and Trade

The UK's Department for Business and Trade is an economic growth department. We ensure fair, competitive markets at home, secure access to new markets abroad and support businesses to invest, export and grow. Our priorities are the Industrial Strategy, Make Work Pay, the Trade Strategy and the Plan for Small Business.

Legal disclaimer
Whereas every effort has been made to ensure that the information in this document is accurate, the Department for Business and Trade does not accept liability for any errors, omissions or misleading statements, and no warranty is given or responsibility accepted as to the standing of any individual, firm, company or other organisation mentioned.

© Crown copyright 2025