

**Cristina Pérez-Guillot**  
**Julia Zabala Delgado**

Prólogo Neus Figueras Casanovas

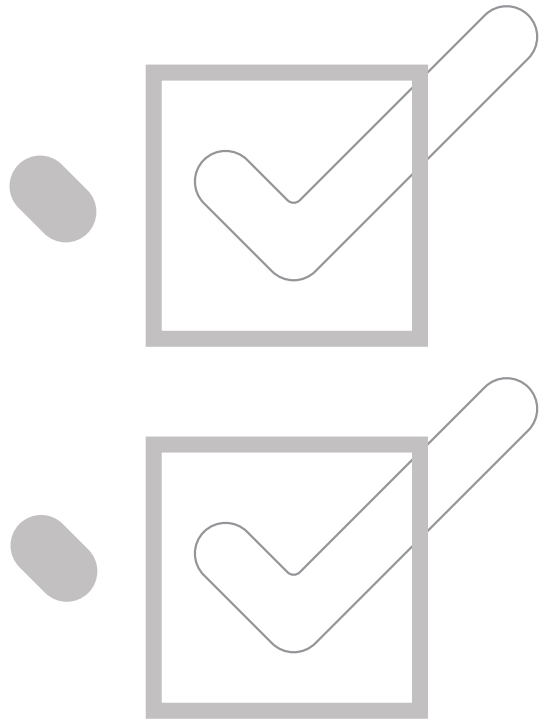


# Guía práctica para la **Elaboración de Pruebas** **Estandarizadas de** **Competencia Lingüística**



**Editorial**

Universitat Politècnica  
de València



**Cristina Pérez-Guillot**  
**Julia Zabala Delgado**

Prólogo Neus Figueras Casanovas

# **Guía práctica para la elaboración de pruebas estandarizadas de competencia lingüística**



**Editorial**  
Universitat Politècnica  
de València

© Cristina Pérez-Guillot  
Julia Zabala Delgado  
© 2020, Editorial Universitat Politècnica de València  
Venta: [www.lalibreria.upv.es](http://www.lalibreria.upv.es) / Ref.: 2028\_01\_01\_01

Diseño y maquetación: Enrique Mateo, Triskelion Diseño Editorial  
Imprime: Byprint Percom, S. L.

ISBN: 978-84-9048-923-9  
Depósito Legal: V-1737-2020  
Impreso bajo demanda

Si el lector detecta algún error en el libro o bien quiere contactar con los autores, puede enviar un correo a [edicion@editorial.upv.es](mailto:edicion@editorial.upv.es)

La Editorial UPV autoriza la reproducción, traducción y difusión parcial de la presente publicación con fines científicos, educativos y de investigación que no sean comerciales ni de lucro, siempre que se identifique y se reconozca debidamente a la Editorial UPV, la publicación y los autores. La autorización para reproducir, difundir o traducir el presente estudio, o compilar o crear obras derivadas del mismo en cualquier forma, con fines comerciales/lucrativos o sin ánimo de lucro, deberá solicitarse por escrito al correo [edición@editorial.upv.es](mailto:edición@editorial.upv.es)

Impreso en España

# Índice

Prólogo .....	3
<b>Parte I. La acreditación de competencia lingüística .....</b>	<b>11</b>
<b>Capítulo 01. Panorama de la acreditación de competencia lingüística y procesos de creación de pruebas estandarizadas. C. Pérez-Guillot y J. Zabala .....</b>	<b>13</b>
1.1. Presente y futuro de la acreditación de competencia lingüística.....	13
1.2. El proceso de creación de una prueba estandarizada.....	17
1.2.1. Planificación de la prueba .....	18
1.2.2. Diseño de la prueba .....	19
1.2.3. Pilotaje de la prueba .....	22
1.2.4. Administración de la prueba .....	23
1.2.5. Análisis de resultados .....	25
1.2.6. Información a los interesados.....	26
<b>Capítulo 02. Principios básicos de la evaluación. J. Zabala.....</b>	<b>27</b>
2.1. Prueba, medición y evaluación .....	27
2.2. Tipos de prueba según su propósito .....	29
2.3. Principios de la evaluación lingüística .....	29
2.3.1. Validez .....	29
2.3.2. Fiabilidad.....	31
2.3.3. Autenticidad.....	33
2.3.4. Practicidad .....	33
2.3.5. Equidad .....	35
2.3.6. Efecto colateral e impacto .....	35
<b>Parte II. La creación de tareas .....</b>	<b>37</b>
<b>Capítulo 03. La comprensión auditiva. C. Pérez-Guillot .....</b>	<b>41</b>
3.1. El proceso de la comprensión auditiva .....	41
3.2. Evaluación de la comprensión auditiva .....	44
3.3. Elección del texto .....	46

3.4. Creación de tareas.....	50
3.5. Tipos de tareas .....	53
3.6. Lista de comprobación de tareas de comprensión auditiva .....	56
<b>Capítulo 04. Expresión e interacción oral. J. Zabala .....</b>	<b>59</b>
4.1. Características de la expresión oral.....	59
4.2. Consideraciones en la evaluación de la expresión oral.....	61
4.2.1. Fiabilidad.....	61
4.2.2. Validez .....	63
4.3. Evaluación de la expresión e interacción oral.....	64
4.4. Lista de comprobación de tareas de expresión oral .....	70
<b>Capítulo 05. La comprensión lectora. C. Pérez-Guillot .....</b>	<b>71</b>
5.1. El proceso de la comprensión lectora .....	71
5.2. Elaboración de tareas.....	72
5.2.1. Elección de textos .....	73
5.2.2. Elección de tareas .....	75
5.3. Los tipos de tareas.....	75
5.4. Lista de comprobación de tareas de comprensión lectora .....	79
<b>Capítulo 06. La expresión escrita. J. Zabala .....</b>	<b>81</b>
6.1. Características de la expresión escrita. ....	81
6.2. Consideraciones en la evaluación de la expresión escrita .....	83
6.2.1. Fiabilidad.....	83
6.2.2. Validez .....	85
6.3. Evaluación de la expresión escrita .....	86
6.3.1. Tareas de escritura imitativa .....	87
6.3.2. Tareas de escritura intensiva.....	87
6.3.3. Tareas de escritura responsiva y extensiva.....	88
6.4. Lista de comprobación de tareas de expresión escrita .....	89
<b>Capítulo 07. Validez en la corrección: escalas de corrección y formación de correctores. J. Zabala .....</b>	<b>91</b>
7.1 Validez en la corrección.....	91
7.2. Escalas de corrección .....	93
7.2.1. Tipos de escalas de corrección.....	93

7.2.1.1. Escalas holísticas.....	94
7.2.1.2. Escalas analíticas.....	95
7.2.2. Proceso de desarrollo de escalas de corrección.....	96
7.3. El rol del corrector.....	98
7.4. Formación de correctores.....	100
<b>Parte III. Controles de calidad.....</b>	<b>103</b>
<b>Capítulo 08. Control de calidad en las pruebas estandarizadas. J. Zabala.....</b>	<b>105</b>
8.1. Juicios expertos.....	105
8.2. Procedimientos de pilotaje.....	108
8.3. Fijación de estándares.....	111
<b>Capítulo 09. Interpretación de análisis estadístico. J. Zabala.....</b>	<b>119</b>
9.1. Análisis a nivel de prueba.....	121
9.1.1. Interpretación de estadística descriptiva.....	121
9.1.2. Fiabilidad.....	124
9.1.3. Análisis de correlaciones.....	126
9.2. Análisis a nivel de ítem: índice de facilidad y discriminación.....	129
<b>Capítulo 10. Certificación de competencia: Modelo CertAcles. C. Pérez-Guillot.....</b>	<b>135</b>
10.1. Introducción.....	135
10.1.1. Principios que rigen la certificación propuesta.....	137
10.1.2. Acreditación de centros.....	137
10.2. Estándares del proceso.....	138
10.2.1. Contenidos, estructura y criterios de evaluación.....	138
10.3. Garantía de calidad.....	139
<b>Referencias bibliográficas.....</b>	<b>141</b>

# Índice de Figuras

Figura 1   Claves del éxito del MCER (Figueras, 2008). .....	14
Figura 2   Equilibrio entre los elementos de la formación y acreditación de lenguas.....	17
Figura 3   Planificación de la prueba. ....	18
Figura 4   Diseño de la prueba. ....	19
Figura 5   Pilotaje de la prueba. ....	22
Figura 6   Administración de la prueba.....	24
Figura 7   Análisis de resultados. ....	25
Figura 8   Información a los interesados. ....	26
Figura 9   Relación entre evaluación, medición y pruebas (Bachman, 1990).....	27
Figura 10   Relación entre la validez y el ciclo de creación de una prueba basado en Bachman (2005). Adaptado de <i>Manual for Language Test Development and Examining</i> . ....	31
Figura 11   Modelo de procesamiento Abajo-Arriba. ....	43
Figura 12   Modelo de procesamiento Arriba-Abajo. ....	43
Figura 13   Modelo de procesamiento interactivo. ....	43
Figura 14   Evaluación de destrezas productivas de McNamara (1996).....	92
Figura 15   Adaptación del marco sociocognitivo de validación de pruebas de Weir (2005).....	93
Figura 16   Diagrama del proceso de corrección adaptado de Cumming, Kantor y Powers (2002). ....	99
Figura 17   Ejemplo gráfico de pilotaje con tareas de anclaje. ....	110
Figura 18   Histograma de los resultados de una prueba. ....	124

## Índice de Tablas

Tabla 1   Textos, tareas y destrezas a evaluar en la comprensión auditiva.....	48
Tabla 2   Mapeo de textos para identificar significado general del texto.....	49
Tabla 3   Mapeo de textos para identificar detalles específicos. ....	49
Tabla 4   Mapeo de textos para identificar ideas principales y accesorias.....	50
Tabla 5   Ejemplo de plantilla de seguimiento de mapeo. ....	50
Tabla 6   Textos, tareas y destrezas a evaluar en la expresión e interacción escrita. ...	88
Tabla 7   Lista de comprobación para juicio de expertos. Adaptada de Bachman y Palmer (1996). ....	106
Tabla 8   Lista de datos cualitativos a recopilar de candidatos y profesores en el pilotaje. ....	110
Tabla 9   Ejemplo de establecimiento de notas de corte de una prueba multinivel.	115
Tabla 10   Ejemplo de juicios según el procedimiento Angoff.....	116
Tabla 11   Estadística descriptiva de una prueba.....	123
Tabla 12   Interpretación de Alfa de Cronbach.....	125
Tabla 13   Interpretación de coeficientes de correlación. ....	127
Tabla 14   Ejemplo de análisis de correlaciones. ....	128
Tabla 15   Interpretación de índices de facilidad. ....	129
Tabla 16   Interpretación de índices de discriminación.....	129
Tabla 17   Ejemplo de análisis de índices de facilidad y discriminación de prueba I.....	130
Tabla 18   Ejemplo de análisis de índices de facilidad y discriminación de prueba II.....	130





# Agradecimientos

Nos gustaría agradecer en primer lugar a Neus Figueras por su gran ayuda y por mostrar su apoyo al aceptar escribir el prólogo de este libro. Con él culminamos una etapa de estudio y trabajo de la que nos sentimos satisfechas. También nos gustaría agradecer a Juana Sanmartin su inestimable ayuda en la revisión del manuscrito final, que sin duda ha facilitado su lectura.

No queremos olvidar, agradecer a los asistentes a nuestros cursos y seminarios en evaluación de estos últimos años su confianza en nosotras. Esta guía surge como respuesta a las necesidades detectadas durante los mismos.

Por último no podemos olvidar a nuestras familias a las que tanto tiempo les hemos robado y sin cuyo apoyo no hubiéramos culminado con éxito esta publicación, en especial a Oscar y Ricardo que tantas veces renuncian a su tiempo para que nosotras podamos seguir nuestro camino y dedicarnos a una de nuestras pasiones: la investigación en búsqueda de nuevos retos.



# Prólogo

Los primeros años del siglo XXI han sido testigos del cambio experimentado en el panorama de la evaluación en general y el de la evaluación de lenguas en concreto, y muy especialmente en el interés que genera. El creciente interés en la evaluación es fruto de la universalización de la educación, que cada vez más necesita contar con indicadores que permitan el análisis y la comparabilidad de los distintos niveles de consecución de objetivos en una economía globalizada (Takala et al., 2014).

El desarrollo de pruebas internacionales como TIMMS, PIRLS o PISA<sup>1</sup>, con la correspondiente difusión de los resultados obtenidos en distintos países, ya procuró una mayor atención a la evaluación por parte de los medios de comunicación y de las instituciones gubernamentales a finales del siglo XX. En el campo de las lenguas –y en especial en el de las lenguas adicionales o extranjeras– el impacto de la publicación del Marco Común Europeo de Referencia para las Lenguas (MCER), con descriptores que permiten la comparabilidad de niveles de dominio de lengua(s) y con recomendaciones para la docencia, el aprendizaje y la evaluación no hizo sino incrementar el interés en la evaluación de lenguas primeras, lenguas segundas y lenguas extranjeras en todos los niveles y contextos educativos.

La mayor parte de países europeos de habla no inglesa (España se cuenta entre ellos), y también países hispanoamericanos como Chile, Argentina o Brasil habían apostado con fuerza en la última década del siglo XX por el dominio de la lengua inglesa, pero también por el dominio de otras lenguas extranjeras en la enseñanza

---

<sup>1</sup> TIMMS: Trends in International Mathematics and Science Study (Estudio de las Tendencias en Matemáticas y Ciencias).  
PIRLS: Progress in International Reading Literacy Study (Progreso en el Estudio Internacional de Competencia en Lectura).  
Las pruebas de TIMMS y PIRLS son elaboradas y gestionadas por la International Association for the Evaluation of Education Achievement (Asociación Internacional para la Evaluación del Aprovechamiento Escolar) <https://www.iea.nl/>  
PISA: Programme for International Student Assessment (Programa Internacional para la Evaluación de Estudiantes). Las pruebas PISA son elaboradas y gestionadas por la OCDE (Organización para la Cooperación y el Desarrollo Económicos) <http://www.oecd.org/pisa/test/>

obligatoria, post obligatoria y universitaria. Estas iniciativas o nuevas políticas lingüísticas se habían avanzado a las propuestas del proceso de Bolonia iniciado el año 1999 para mejorar la calidad de los estudios universitarios, y también a las conclusiones de la reunión del Consejo Europeo en Barcelona de marzo de 2002. En esta reunión se acordó recomendar a los estados miembros del Consejo Europeo la docencia de dos lenguas adicionales a la lengua de escolarización a partir de una edad temprana para facilitar la movilidad y la competitividad. Dado que estas propuestas requerían un aumento de recursos de todo tipo con sus correspondientes inversiones, los gobiernos vieron la necesidad de poder demostrar su rentabilidad o éxito, lo que se tradujo en el uso de pruebas o exámenes que certificaran los logros conseguidos.

Este escenario marcó dos caminos diferenciados. Por una parte, el que llevaba al incremento en la demanda de pruebas internacionales estandarizadas, que creció exponencialmente (pensemos en los exámenes de Cambridge o Trinity College London en lengua inglesa, en los DELF y DALF en lengua francesa, en las pruebas del Goethe Institut en lengua alemana o en las de la Università per Stranieri de Siena en lengua italiana, por citar los más populares). Por otra parte, se inició también el camino que priorizaba la “localización” de las pruebas y que facilitó la creación y desarrollo de sistemas de evaluación de ámbito nacional o regional adaptados a su propio contexto para evaluar – y certificar – los conocimientos adquiridos y demostrar la utilidad de la inversión de recursos en los procesos de docencia-aprendizaje (las escuelas oficiales de idiomas o los centros de idiomas de las universidades son ejemplos de estas iniciativas).

En ambos casos, la publicación del MCER en 2001 dio una respuesta a la necesidad de contar con un referente transparente y comprensible y que a la vez facilitara la comparabilidad (¿Es mi B1 como tu B1?). También en ambos casos se generó un aumento de las demandas de responsabilidad por parte de los autores y gestores de pruebas o exámenes, y de claridad por parte de los usuarios de tales pruebas, lo que forzó al Consejo de Europa a elaborar un Manual (2003, 2009, 2005) para facilitar el uso adecuado de los contenidos del MCER en evaluación. El Manual facilitó a todos aquellos involucrados en la elaboración, administración y corrección de pruebas, el rendir cuentas sobre la calidad de sus actuaciones de forma transparente y veraz, contribuyendo a una mayor profesionalización.

A partir de ahí, los cambios que se han observado en el campo de la evaluación de lenguas desde los inicios del siglo XXI cobran gran relevancia, ya que permiten incrementar la calidad en las distintas actuaciones evaluativas y facilitan la transparencia. El hecho más señero al respecto en el siglo XXI y que ha influido en los cambios observados que se repasan brevemente a continuación, es que la evaluación de lenguas se ha ganado un espacio como disciplina en el campo de la lingüística aplicada y cuenta en la actualidad con numerosas conferencias monográficas, un nutrido número de publicaciones especializadas, periódicas y no periódicas, multitud de manuales y libros de recursos específicos, enciclopedias y sitios web especializados. Asimismo, la evaluación de lenguas está presente de forma sostenida en encuentros y/o publicaciones genéricas. Hacemos referencia, por ejemplo, a las reuniones anuales de ILTA, EALTA, ALTE o SICELE (ver lista de referencias) y publicaciones periódicas

especializadas en evaluación de lenguas como *Language Testing* o *Language Assessment Quarterly*, Manuales como el de Fulcher y Davidson (2007) o Tsagari y Banerjee (2016), Enciclopedias completas como la editada por Kunnan (2013) en 4 volúmenes o la reedición del volumen editado por Shohamy, Or y May (2n, 2017) en la *Encyclopedia of Language and Education*.

El haber conseguido erigirse en disciplina específica ha generado más interés por parte de la comunidad científica, que ha obtenido también más recursos, y que junto con el incremento en la demanda de pruebas o certificaciones de lenguas ha contribuido a aumentar la calidad y la cantidad de investigación en evaluación de lenguas, que hoy en día cuenta con un amplio abanico de recursos a su disposición y una amplia red de comunicaciones para difundir sus resultados. Entre los avances positivos y más significativos se encuentran cuestiones como una mayor precisión en la definición y operacionalización de aquello que se evalúa, el constructo (Bachman y Palmer, 2010), y de los conceptos básicos en evaluación de lenguas, como la validez y la fiabilidad (Weir, 2004; Kane, 2013; Spolsky, 2017). Los últimos años del siglo XX y los primeros del siglo XXI han visto también un mayor interés en los aspectos sociales de la evaluación, como su impacto y sus consecuencias (Bachman, 2010; Tsagari y Cheng, 2017) y en cuestiones relativas a la profesionalización de todos aquellos involucrados en actuaciones evaluativas (Davies, 2017) o en la importancia de la localización (O’Sullivan, 2017).

Los avances en evaluación de lenguas –y sobre todo la preocupación por una mayor profesionalización de aquellos que la ejercen– han impulsado dos vías de desarrollo en la evaluación de lenguas que se complementan y retroalimentan. Por una parte, existe en la actualidad, tal y como ya se ha mencionado, una mayor actividad de las iniciativas asociativas en el ámbito internacional (en especial con el desarrollo de directrices y códigos de buenas prácticas, como las de asociaciones muy implantadas como ALTE, EALTA, ILTA, SICELE, CERCLES). Y también se observa el interés en la creación de asociaciones más locales y alejadas del mundo anglosajón como por ejemplo ACLES en el ámbito universitario español, y las recientes LAALTA en hispanoamérica y GIELE en España. Por otra parte, y a partir del artículo de Taylor (2013) en *Language Testing*, se ha generado gran interés en identificar y definir qué quiere decir saber de evaluación de lenguas (Kremmel y Harding 2015)<sup>2</sup> y se han llevado a cabo multitud de proyectos que tienen como principal objetivo aumentar los conocimientos sobre evaluación de lenguas (Cambridge Assessment Webinars, British Council Animations, o TALE ).

Sin embargo, y a pesar del evidente progreso en todo lo que se refiere a la evaluación de lenguas a nivel internacional, los países no anglosajones, con muy pocas excepciones, no han seguido la evolución ni adoptado las propuestas surgidas en el mundo anglosajón a ambos lados del Atlántico ni han generado recursos o publicaciones alternativas. Esto es debido a diferencias en los enfoques docentes y en las tradiciones educativas pero sobre todo a la dificultad que experimentan muchos profesionales de la educación a la hora de acceder y utilizar textos y materiales en inglés.

2 [https://lancasteruni.eu.qualtrics.com/SE/?SID=SV\\_dgUnkyGldhQGhNz&Q\\_JFE=0&Q\\_SE=ZGdV](https://lancasteruni.eu.qualtrics.com/SE/?SID=SV_dgUnkyGldhQGhNz&Q_JFE=0&Q_SE=ZGdV)

En el mundo hispanohablante, la escasez de traducciones de libros o artículos fundamentales en evaluación es palmaria (con algunas excepciones como la de la colección en español de Cambridge University Press a finales de los años 90 que publica la traducción al español del libro de Alderson, Clapham y Wall [1998], entre otros libros de referencia obligada).

Un ejemplo de la importancia de tener acceso directo a la literatura especializada es la muy distinta difusión que han tenido los contenidos del MCER, traducido a la lengua española en 2002 y que ha generado multitud de artículos, publicaciones y proyectos en lengua española. En el ámbito de las publicaciones propias, el número de publicaciones periódicas especializadas en evaluación en español (exceptuando el monográfico sobre evaluación de MARCOELE, 2008) y la muy breve lista de libros o manuales de evaluación publicados originalmente en español (Bordón, 2006; Prati, 2007; Martínez Baztán, 2011; Figueras y Puig, 2013; entre los más citados) es escaso, algo que también contrasta con las publicaciones acerca y sobre el MCER (Cassany, 2006; Llorián, 2007; Soler Canela, 2006; Pikabea et al., 2008; entre otros). Esta escasez en lo que se refiere a la literatura específica sobre evaluación en español hace muy difícil el desarrollo profesional en evaluación de aquellos que no dominan la lengua inglesa o que trabajan en un contexto no anglosajón.

Este libro quiere dar respuesta al creciente interés en la evaluación de lenguas en España y en Sudamérica y contribuir a una mayor toma de conciencia sobre lo que representa una evaluación estandarizada de calidad de acuerdo con las propuestas y pautas acordadas por la comunidad especializada internacional, propuestas y pautas con las que las autoras están muy familiarizadas gracias al estudio de la teoría, su activa participación en la evolución y profesionalización de ACLES (Asociación de Centros de Lenguas en la Enseñanza Superior en España), y gracias sobre todo a la experiencia ganada con el trabajo de campo en el diseño, desarrollo y puesta en práctica de las pruebas de certificación del centro de lenguas en la Universidad Politécnica de Valencia.

El libro quiere ser una guía práctica, dirigida a aquellos que – en distintos contextos, con distintas responsabilidades y con distinta formación – están involucrados en pruebas de evaluación de la competencia lingüística. En cada uno de sus ocho capítulos el lector puede familiarizarse con los distintos aspectos y pasos en el proceso de diseño, elaboración, administración, corrección o análisis de pruebas de dominio estandarizadas desde una perspectiva teórico práctica. En cada capítulo no sólo se hace referencia a lo que dice la teoría sino que ésta también se ejemplifica, mediante propuestas concretas y los gráficos y las ilustraciones facilitan, a modo de resumen, la comprensión de lo tratado.

Aunque los distintos capítulos forman un todo, el lector puede seleccionar o priorizar aquellos que le resultan más interesantes o relevantes para su trabajo. Un administrador o gerente, por ejemplo, puede decidir centrarse en los primeros dos capítulos, que tratan sobre el proceso de creación y desarrollo de una prueba y sobre los principios básicos que deben sustentarlo, sólo en los últimos, que se centran en el análisis, únicamente en los dos primeros y en el capítulo 9 centrado en la estadística, o

también en los capítulos que tratan la evaluación de la comprensión oral, la expresión interacción oral, la comprensión escrita y la expresión e interacción escrita. Deberá decidir según sus responsabilidades y según el tiempo de que disponga. Un redactor de pruebas, en cambio necesitará centrarse en la lectura del capítulo sobre la destreza en la que trabaja y consultar cuando lo considere necesario el resto de capítulos. Para facilitar este tipo de lectura, cada capítulo, aunque siga una misma estructura, es independiente.

El primer capítulo presenta la estructura que debe seguir todo proceso de creación de una prueba estandarizada y describe de forma pormenorizada los pasos a seguir. Además, y a modo de ilustración o ejemplo, facilita un relato detallado del contexto en el que operan los centros de lenguas de las universidades españolas, describe las razones socioeconómicas y políticas que han impulsado el proceso de desarrollo de sus pruebas estandarizadas y repasa los éxitos logrados y los retos pendientes a nivel local, nacional e internacional. Lo descrito en este relato resultará familiar al lector, aunque su contexto sea distinto, y le puede resultar útil a la hora de poner en práctica su proyecto.

El segundo capítulo trata los conceptos básicos en evaluación que deben conocerse y aplicarse para obtener pruebas de calidad. Se hace referencia a la teoría y a las investigaciones más recientes y se presentan ejemplos prácticos para facilitar la comprensión e internalización de lo descrito.

Los capítulos 3, 4, 5 y 6 se centran en la evaluación de lo que se ha venido describiendo de forma tradicional como las cuatro destrezas (aunque el MCER las denomina actividades y ha ampliado su ámbito para incluir en la expresión el término interacción). Estos capítulos están estructurados con una introducción y un listado de pasos y recomendaciones a seguir, junto con ejemplos para su consideración y análisis.

El capítulo 7 trata sobre la elaboración y uso de escalas (criterios o rúbricas) para la evaluación y la corrección de pruebas, En cuanto a la elaboración se describe el proceso a seguir y los aspectos a considerar, siempre de acuerdo con el propósito de cada prueba y su contexto de aplicación que garantice su validez. En cuanto a su uso se analiza el rol del corrector (y también del interlocutor en la expresión e interacción oral) y la importancia de contar con un programa de formación y control de calidad que garantice su fiabilidad.

El capítulo 8 se centra en el control de calidad, la importancia de contar con la monitorización por parte de jueces a lo largo de todo el proceso de elaboración y administración, la utilidad de llevar a cabo experimentaciones o pilotajes y la importancia de tener un sistema claro de fijación de estándares.

El capítulo 9 se centra en la interpretación de los datos resultantes de los análisis estadísticos llevados a cabo durante la fase de experimentación o pilotaje de las pruebas y también de los datos recogidos una vez se ha administrado la prueba en su versión final. El análisis estadístico cierra el proceso de elaboración de toda prueba y permite la revisión y modificación de las decisiones tomadas a partir de su inicio.



El capítulo prioriza los conceptos y métodos de análisis básicos y más útiles para los profesionales no especialistas en estadística y los describe en relación con ejemplos prácticos que permiten al lector el poder utilizarlos de forma inmediata.

El capítulo 10 tiene como objetivo la descripción exhaustiva del sistema de acreditación de CertACLES, privativo de las universidades españolas pero anclado en el sistema de centros de idiomas en universidades europeas. Aunque muy centrado en el contexto universitario, el sistema CertACLES, por su carácter pionero en España, resultará de interés para aquellos interesados en una organización clara de un sistema de acreditación.

En todos los capítulos del libro-guía, la importancia de los contenidos y propuestas en el MCER tienen un espacio muy importante. Esto es así dado el impacto del documento a nivel internacional, y en especial en Europa, ya que se viene utilizando como referente obligado en la definición del constructo a evaluar y en su operacionalización de contenidos en la comprensión y en la expresión. Dada la novedad de los contenidos del Volumen de Acompañamiento del MCER (*Companion Volume*) publicado en la web del Consejo de Europa en 2018, y de próxima publicación en formato libro, en el que se presentan algunos cambios sobre el documento original y nuevos descriptores, las autoras se limitan a apuntar al posible impacto de algunos cambios (como por ejemplo el de la escala de Control Fonológico en la evaluación de la expresión e interacción oral y en la elaboración de escalas) y de algunos nuevos descriptores (como por ejemplo los descriptores para la mediación, tanto en la elaboración y diseño de tareas como en la elaboración de escalas). En cualquier caso, el lector encontrará a lo largo de los ocho capítulos suficiente información práctica que le permita acceder y utilizar los contenidos y las recomendaciones del nuevo Volumen de Acompañamiento que le resulten relevantes en su contexto.

Aunque el libro se centra en el proceso de elaboración de pruebas estandarizadas, lo tratado en los distintos capítulos es de utilidad también para aquellos con responsabilidades en un centro escolar o universitario, tanto a nivel administrativo como académico. Por una parte, es importante que los receptores de los efectos o consecuencias de toda prueba estandarizada tengan suficiente información para comprender su funcionamiento e interpretar sus resultados. Por otra parte, el control de calidad a lo largo de todo el proceso de elaboración de una prueba o dispositivo de evaluación, la importancia de la planificación, del rigor en los pasos a seguir en la selección y elaboración de tareas, en los criterios para la elaboración en el uso de escalas de corrección y en el análisis de resultados debe formar parte ya de cualquier iniciativa en evaluación y considerarse imprescindible en la formación de profesionales en la docencia, aprendizaje y evaluación de lenguas.

## Referencias

- Alderson, C., Clapham, C., Wall, D. (1998). *Exámenes de Idiomas*. Madrid: Cambridge University Press.
- Bordón, T. (2006). *La evaluación de las lenguas en el marco de E/LE*. Madrid: Arco Libros.
- Bachman, L., Palmer, A. (2010). *Language Assessment in Practice*. Oxford: Oxford University Press.
- British Council. *How Language Assessment works*. Accessed 14 March 2020 <https://www.britishcouncil.org/exam/aptis/research/assessment-literacy>
- Cambridge Assessment Webinars. Understanding Assessment. What Every teacher should know. Accessed March 2020. [https://www.youtube.com/watch?v=WeC\\_j3AhbZU](https://www.youtube.com/watch?v=WeC_j3AhbZU)
- Cassany, D. (ed) (2006). *El Portfolio Europeo de Las Lenguas y sus aplicaciones en el aula*. Ministerio de Educación y Ciencia. Secretaría General de Educación
- Council of Europe. (2003). *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Pilot Manual*. Language Policy Division, Strasbourg.
- Council of Europe. (2009). *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual*. Language Policy Division, Strasbourg. Accessed 14 March 2020. <https://www.coe.int/en/web/common-european-framework-reference-languages/relating-examinations-to-the-cefr>
- Council of Europe. (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with New Descriptors*. Strasbourg. Accessed 14 March 2020 <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989>
- Davies, A. (2017). Ethics, Professionalism, Rights and Codes. En Shohamy et al. *Language Testing and Assessment*. Springer. [https://doi.org/10.1007/978-3-319-02261-1\\_27](https://doi.org/10.1007/978-3-319-02261-1_27)
- Figueras, N., North, B., Takala, S., Verhelst, B., van Avermaet, P. (2005). Relating examinations to the Common European Framework: a manual. *Language Testing*, 22(3), 261-279
- Figueras, N., Puig, F. (2013). *Pautas para la Evaluación del Español como Lengua Extranjera*. Madrid: Edinumen.
- Figueras, N., Mingarro, P., Puig, F. (2011). *Docencia, aprendizaje y evaluación de segundas lenguas en las escuelas oficiales de idiomas*. Barcelona: ICE. Horsori.
- Kane, M.T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement*. 50(1), 1-73. <https://doi.org/10.1111/jedm.12000>

- Kremmel, B., Harding, L. (2015). *Assessment Literacy Survey* Accessed March 2020 [https://lancasteruni.eu.qualtrics.com/SE/?SID=SV\\_dgUnkyGIDhQGtNz&Q\\_JFE=0&Q\\_SE=ZGdV](https://lancasteruni.eu.qualtrics.com/SE/?SID=SV_dgUnkyGIDhQGtNz&Q_JFE=0&Q_SE=ZGdV)
- Llorián, S. (2007). *Entender y utilizar el Marco común europeo de referencia desde el punto de vista del profesor de lenguas*. Salamanca: Santillana.
- Martinez Baztán, A. (2011). *La Evaluación de las lenguas: Garantías y Limitaciones*. Barcelona: Editorial Octaedro.
- O'Sullivan, B. (2017) *Localisation of Assessment: Theory & Practice. New Directions Shanghai*. Accessed 20 March 2020. [https://www.britishcouncil.cn/sites/default/files/plenary\\_speaker\\_barry\\_osullivan\\_0.pdf](https://www.britishcouncil.cn/sites/default/files/plenary_speaker_barry_osullivan_0.pdf)
- Pikabea, I., Lukas, F., Figueras, N. (2009). Proceso de comparabilidad de titulaciones de euskara y adecuación al marco común europeo de referencia (MCER). *Revista de Investigación Educativa*, 27(2), 469-486
- Prati, S. (2007). *La evaluación del español como lengua extranjera*. Buenos Aires. Libros de La Araucaria.
- Soler Canela, O. (2006). La evaluación de ELE y el Marco Común Europeo de Referencia. ANPE. *I Congreso Español en Noruega*. <https://www.mecd.gob.es/dam/jcr:b34c18cb-de1d-4182-bb52-6206e331e628/2006-esp-09-21soler-pdf.pdf>
- Spolsky, B. (2017). History of Language Testing, en Shohamy et al. *Language Testing and Assessment*. Springer. [https://doi.org/10.1007/978-3-319-02261-1\\_32](https://doi.org/10.1007/978-3-319-02261-1_32)
- Takala, S. (2014). International Assessments. En Kunnan, A. (ed) *The Companion to Language Assessment*. Volume I. Abilities, Contexts, and Learners. Part 3. Assessment Contexts. <https://doi.org/10.1002/9781118411360.wbcla052>
- Takala, S., Erickson, G., Figueras, H., Gustafson, J.E. (2016). Future Prospects and Challenges in Language Assessment, en Banerjee, J. y Tsagari, D. *Contemporary Language Assessment*. London. Bloomsbury.
- Teachers Assessment Literacy Enhancement (TALE project) Accessed March 2020 <http://taleproject.eu/>
- Tsagari, D., Cheng, L. (2017). Washback, impact, and consequences revisited. E. Shohamy, I. Or, S. May (3rd Ed.), *Language Testing and Assessment. Encyclopedia of Language and Education*. Springer Cham. [https://doi.org/10.1007/978-3-319-02261-1\\_24](https://doi.org/10.1007/978-3-319-02261-1_24)
- Weir, C. (2004). *Language Testing and Validation. An evidence-based approach*. Palgrave.

# Parte I

## La acreditación de competencia lingüística

En la actualidad, la acreditación de competencia lingüística se ha convertido en un elemento esencial en el entorno universitario, puesto que es un requisito para graduarse en muchas universidades del territorio nacional y una de las exigencias para acceso a la mayoría de los másteres.

Esta necesidad de acreditación se plantea desde el ámbito universitario, no como un trámite burocrático o administrativo, sino como un mecanismo que permita implementar mejoras en la enseñanza y el aprendizaje de lenguas mediante su inclusión en los sistemas de calidad en educación terciaria. Cuando hablamos de acreditación de competencia lingüística en este ámbito, hablamos de acreditación mediante exámenes estandarizados y diseñados a través de procesos de desarrollo que aseguren la generalización de sus resultados y la extrapolación de los mismos a la capacidad del candidato en el mundo real. Este tipo de pruebas, como veremos, cumplen con una serie de principios entre los que se encuentra la consideración del impacto de las mismas, que es una de las consideraciones fundamentales cuando se habla de educación.

La evaluación así entendida, debe tener un impacto positivo en la adquisición de competencia por parte de los alumnos, y ese impacto, no debería limitarse al ámbito universitario, sino actuar como impulsor de mecanismos de evaluación y mejora de la calidad en todos los niveles educativos. No debemos de perder de vista que el requisito de acreditación lingüística en la universidad supondrá la necesidad de mejorar la formación en lenguas desde etapas iniciales, para que a su llegada a la universidad la ruta sea más sencilla y las posibilidades de éxito mayores.

Somos conscientes de que la evaluación de la competencia lingüística se ha visto sometida, por la gran demanda, a una gran presión política, institucional y social. Todos los organismos relacionados con la enseñanza de idiomas han tenido que considerar las opciones y se han buscado diversas vías de homogeneización y comparabilidad.

Las entidades certificadoras de competencia lingüística, entre las que se encuentran las universidades, responden a las demandas de los diversos actores sociales e institucionales y trabajan para desarrollar certificaciones equiparables que cumplan los principios básicos de evaluación, que se encuentren bien anclados a un nivel del MCER, y que estén reconocidos en ámbitos educativos y laborales. En el caso de instituciones educativas, la clave está en conocer las exigencias y los problemas que los alumnos tendrán en sus diferentes etapas de formación, de manera que todos los que nos dedicamos a la enseñanza, sea esta primaria, secundaria o terciaria, tengamos un poco más claro el camino a seguir y los objetivos a conseguir en lo que a evaluación de lenguas se refiere, para de este modo ayudar a la mejora en la calidad de la enseñanza y el aprendizaje.

# 01

## Panorama de la acreditación de competencia lingüística y procesos de creación de pruebas estandarizadas

*C. Pérez-Guillot y J. Zabala*

### 1.1. Presente y futuro de la acreditación de competencia lingüística

La Declaración de Bolonia, firmada en junio de 1999, sienta las bases para la creación del Espacio Europeo de Educación Superior, cuyo principal objetivo es favorecer, en materia de educación, la convergencia europea. La integración de la universidad española en EEES ha supuesto una creciente movilidad de estudiantes y un avance hacia la internacionalización, así como una mejora del empleo. Esta integración ha propiciado que la acreditación de competencia lingüística se haya convertido en un elemento clave en la carrera, no solo del docente, sino también del egresado universitario.

Un elemento clave sin duda es en el año 2001, cuando el Consejo de Europa publica el Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación (MCER). Su principal propósito era establecer una serie de referentes comunes que facilitasen el intercambio de información sobre conocimientos de la lengua y, por tanto, permitiesen responder a esa necesidad de homologación de las distintas certificaciones existentes. Actualmente, se ha convertido en un instrumento clave en el mundo de la enseñanza y el aprendizaje de lenguas extranjeras.

Los años de debate sobre la vinculación de las pruebas y los niveles de competencia en lenguas con los niveles del MECR han sido indudablemente útiles. Conferencias, congresos, talleres y mesas redondas en toda Europa han advertido sobre los obstáculos y peligros en el uso y abuso de los niveles del MCER. La organización de una gran diversidad de actividades ha sido fundamental en la concienciación de los profesionales y de las autoridades sobre la necesidad de contar con una enseñanza y evaluación de idiomas de calidad, con la creación y desarrollo de redes de profesionales concienciados en su trabajo y dispuestos a compartir conocimientos y experiencias.

Una primera justificación del éxito del MCER, siguiendo a Figueras (2008), es la solvencia de la institución que lo publica. Las primeras páginas del propio documento lo identifican como la culminación del amplio programa del Consejo de Europa, y lo caracterizan como “la última etapa de un proceso activo desde 1971 y que debe mucho a la colaboración de muchos miembros de la profesión docente en toda Europa y más allá” (2001).



Figura 1 | Claves del éxito del MCER (Figueras, 2008).

Siguiendo esta misma línea, una segunda justificación es la que podemos denominar de tipo estratégico, y que es la que se relaciona más directamente con la naturaleza del MCER, donde hay una combinación de política educativa y didáctica. Ya en los primeros capítulos del documento se destaca la necesidad de contar en Europa con un marco en el que se puedan reconocer los distintos esfuerzos que se llevan a cabo en los diferentes países en el campo de la docencia, del aprendizaje y de la evaluación de lenguas.

Un último punto a destacar como clave es de tipo académico-pedagógico, el MCER aparece en un momento clave en el que hay una clara necesidad de considerar a la lengua como elemento de comunicación y que todo el mundo involucrado en el área de enseñanza aprendizaje reconozca este papel.

A todo esto, podemos añadir que surge en un momento en que la Europa multilingüe necesita un referente para la enseñanza / aprendizaje de lenguas, así como su evaluación. Conviene recordar la diferencia de significado entre multilingüismo y plurilingüismo. El multilingüismo implica conocimiento de varias lenguas o la coexistencia de distintas lenguas en una sociedad determinada, mientras que plurilingüismo enfatiza el hecho de que conforme se expande la experiencia lingüística de un individuo en los entornos culturales de una lengua, se desarrolla una competencia comunicativa en otras lenguas.

A este respecto el MCER nos indica: “*Competencia plurilingüe y pluricultural: La capacidad de utilizar las lenguas para fines comunicativos y de participar en una relación intercultural en que una persona, en cuanto agente social, domina -con distinto grado- varias lenguas y posee experiencia de varias culturas*” (MCER 202\_167) y en la literatura podemos encontrar numerosas referencias a la necesidad de promover el conocimiento de las lenguas en los estudiantes:

- La globalización del empleo y la internacionalización de los nuevos puestos de trabajo, en las que se considerará de forma especial la adquisición de conocimientos, habilidades y competencias de internacionalización con el fin de preparar a nuestros futuros egresados para un mundo cada vez más globalizado a nivel de empleo (Estrategia de Internacionalización de las Universidades EIU, 2011).
- La promoción de competencias internacionales como la lingüística y la intercultural (Haug y Vilalta, 2011).
- La importancia de dominar la lengua en la que se aprenden contenidos para acceder al conocimiento (Schleppegrell, 2008).
- El denominado proceso de Bolonia ha contribuido a plantearnos una serie de prioridades, siendo la internacionalización una de ellas. Unido a la internacionalización encontramos la necesidad de mejorar en idiomas y, de manera especial, en el uso del inglés como instrumento de comunicación (Alcon, 2011).

Llegamos así al punto de ver como en la actualidad, la acreditación de competencia lingüística se ha convertido en elemento esencial en el entorno universitario, donde es requisito para graduarse en muchas universidades del territorio nacional y una de las exigencias para acceso a la mayoría de los másteres. No debemos de perder de vista el impacto positivo que esto ha tenido en las aulas por la clara necesidad de mejorar la formación en lenguas desde etapas iniciales de su formación para que a su llegada a la universidad la ruta sea más sencilla y las posibilidades de éxito mucho mayores.

Somos conscientes de que la evaluación de la competencia lingüística se ha visto sometida, por la gran demanda, a una gran presión política, institucional y social. Todos los organismos relacionados con la enseñanza de idiomas han tenido que considerar las opciones y se han buscado diversas vías de homogeneización y comparabilidad. Desde un principio se ha intentado establecer certificaciones equiparables que cumplan los principios básicos de evaluación y que se encuentren bien ancladas a un nivel del MCER para poder ser reconocidas en ámbitos educativos y laborales. En este sentido, el MCER es una herramienta esencial para la elaboración de pruebas lingüísticas comparables entre instituciones y su manual para el desarrollo de pruebas lingüísticas asegura el establecimiento de buenas prácticas en la evaluación. Pero el MCER no es estático, el paso del tiempo ha llevado a la necesidad de revisión y la última, publicada, en su versión definitiva, en mayo de 2020, presenta muchas novedades:

- Actualización de los descriptores del MCER de mediación y competencia plurilingüe y pluricultural.



- Definición de los niveles “plus” y Pre-A1.
- Descriptores más elaborados para comprensión oral y lectora.
- Mejora de las descripciones de A1 y C2.
- Nuevas escalas para discusión y conversación online.

El desarrollo de los descriptores de mediación ha sido la parte más larga y compleja del proyecto del *CEFR companion volume*. En el MCER se define la mediación como la actividad del lenguaje consistente en reformular, de manera oral o escrita, un texto con el propósito de posibilitar la comprensión mutua de otras personas. La mediación, como la propia palabra sugiere, implica que el hablante está “en el medio”. No le concierne su propia expresión lingüística, sino que su papel consiste en crear el espacio y las condiciones adecuadas que permitan la comunicación o el aprendizaje, colaborando con sus interlocutores para construir significados

Se establecen las diferencias entre mediación social, cultural, pedagógica, lingüística o profesional. Pero sea cual sea la forma que adopte, implica siempre ser un intermediario en la comunicación. Por tanto, el mediador colabora con otro usuario que tiene problemas de comprensión del mensaje, debido a factores diversos como registro determinado, parafraseando el mensaje y facilitando la comunicación.

Asimismo, esta versión actualizada del Marco añade nuevos descriptores que complementan el conjunto original contenido en el texto de 2001 y profundiza en la importancia de los hablantes como agentes sociales. Es evidente, atendiendo a la experiencia anterior, que la implantación de las modificaciones propuestas, su difusión y puesta en marcha, tendrán un impacto directo en las aulas de enseñanza/aprendizaje de lenguas.

Y precisamente hablando de mediación y para centrarnos un poco más en el panorama de la certificación, hablaremos de la Asociación de Centros de Lenguas en Educación Superior (ACLES) y su papel mediador para intentar dotar de estructura al sistema de certificación lingüística en España. Esta labor está siendo realizada, no solo a través de la creación de su propio certificado de competencia, CertAcles, sino a través de la colaboración con los distintos colectivos y a la elaboración de una tabla de certificados, en sintonía con la mesa lingüística de la CRUE, de las certificaciones lingüísticas de los distintos idiomas, reconocidas nacional e internacionalmente. Este papel, lo realiza a su vez junto a su colaborador a nivel europeo, la Confederación Europea de Centros de Lenguas en Educación Superior (*European Confederation of Language Centres in Higher Education*, CercleS) que agrupa 290 centros en 22 países. ACLES trabaja en CERCLES dentro de NULTE (*Network of University Language Testers in Europe*), su red de examinadores europeos en pro del reconocimiento mutuo de los certificados a nivel de instituciones de educación superior europeas, lo que sin duda colabora al plan de internacionalización de las universidades españolas.

Podemos afirmar, que la acreditación de la competencia lingüística en un idioma extranjero se ha convertido en una necesidad para todos nosotros por cuestiones de trabajo o de estudio. Asimismo, la evaluación certificativa ha ido adquiriendo una mayor importancia y ocupando de forma progresiva un espacio cada vez más amplio en la investigación sobre la enseñanza y el aprendizaje de lenguas extranjeras y segundas lenguas.

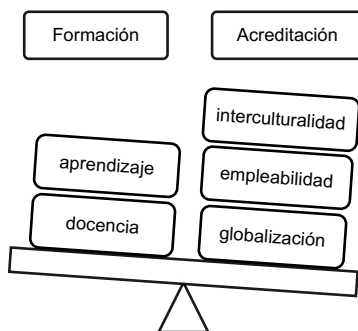


Figura 2 | Equilibrio entre los elementos de la formación y acreditación de lenguas.

Para concluir, podemos ver representado gráficamente en la Figura 2, el equilibrio que debe existir entre todos los elementos que participan en la formación y acreditación de lenguas. Todos ellos bien combinados contribuyen al buen funcionamiento de la enseñanza y/o aprendizaje de lenguas en un mundo global orientado a la interculturalidad y la empleabilidad.

## 1.2. El proceso de creación de una prueba estandarizada

El proceso de elaboración de una prueba de competencia lingüística es un proceso iterativo, es un ciclo continuo en el que cada fase se alimenta de la anterior y a su vez enlaza con la siguiente. Del mismo modo, cada ciclo de prueba proporciona datos valiosos para la mejora continua y por tanto enlaza con el siguiente ciclo de manera natural. Esta iteración hace que en ocasiones el proceso parezca complejo y ajeno a la práctica de enseñanza y aprendizaje, pero nada más lejos de la realidad.

La existencia de una serie de procesos más o menos estructurados a la hora de crear una prueba, no dificulta su comprensión ni pretende crear una apariencia de campo desconocido para los no profesionales de la evaluación; lo que realmente pretende, es dotar de sistematicidad a actividades que probablemente ya se realizan en el aula, para de este modo asegurar el cumplimiento de buenas prácticas en la evaluación que son de sobra conocidas para aquellos involucrados en la educación, tanto del lado de los enseñantes como en el de los alumnos.

A continuación, vamos a analizar una propuesta de sistematización de proceso que lo divide en cinco grandes pasos, que a su vez incluyen cuatro sub-pasos cada uno. Esta propuesta (Zabala, 2014) pretende servir de lista de comprobación del proceso de elaboración de la prueba con un nivel de detalle suficiente como para ser usada de esqueleto, pero sin una exhaustividad restrictiva. De este modo, los pasos principales son estándares mínimos que cualquier proceso debería cumplir, y los sub-pasos son una extensión de los principales que a su vez pueden ser más o menos desarrollados según las posibilidades de cada institución. Todos los pasos, como corresponden en un proceso iterativo, están interrelacionados y como veremos, justifican y son justificados por el resto.

### 1.2.1. Planificación de la prueba

El primer paso para la creación de una prueba estandarizada es la planificación previa. Puesto que la validez de una prueba está intrínsecamente ligada a su utilidad para el propósito para el que fue creada, una planificación cuidadosa es el primer paso para asegurar su validez, puesto que necesitaremos conocer cuál es el objetivo de la misma previamente a comenzar su desarrollo.



Figura 3 | Planificación de la prueba.

Dentro de la planificación, cualquier prueba de competencia debe de comenzar teniendo en cuenta su contexto, que podrá ir desde contextos locales, hasta nacionales o internacionales. En cada caso, es necesario ser conscientes del mismo, así como realizar un estudio de las necesidades que en él surgen. Mediante cuestionarios, análisis de la institución, de los posibles candidatos, etc. podremos averiguar qué se requiere de la prueba y para qué va a ser usada.

Una vez analizado el contexto y detectadas las necesidades, estaremos en la posición de determinar el **propósito** de nuestra prueba, diseñado para cubrir las necesidades detectadas y las consecuencias de la misma. La determinación de las consecuencias y de su carácter “decisivo” es necesaria para establecer no sólo sus contenidos sino también nuestros márgenes de error aceptables.

Las necesidades y el consecuente propósito de la prueba dependerán a su vez de los **recursos disponibles**. En un mundo ideal, la prueba diseñada lo estaría de tal manera que cubriera todas las necesidades de nuestro entorno de manera óptima. Sin embargo, en el mundo real, los recursos no son ilimitados y en ocasiones, a pesar de su abundancia, hay limitaciones externas que resultan insalvables. A la hora de diseñar nuestra prueba, resulta útil considerar estas limitaciones en la fase inicial del proceso para poder subsanarlas en la medida de lo posible.

Por último, necesitaremos formar a un equipo de especialistas encargados de desarrollar la prueba. El **equipo de desarrollo** de la prueba debería conocer bien el contexto, los candidatos y sus peculiaridades, y combinar formación lingüística, docente y de evaluación. La formación proporcionada debería incluir fundamentos teóricos de evaluación para sustentar sus decisiones y práctica tutelada en el desarrollo de ítems, que les permita dar forma a los instrumentos que darán pie a los candidatos a demostrar su capacidad.

### 1.2.2. Diseño de la prueba

Una vez planificada la prueba, es hora de comenzar con el diseño. El diseño incluirá, no sólo el diseño de los ítems y las tareas, sino también de la estructura de la prueba (número de ítems, de tareas, partes de la prueba, duración de la prueba, etc.), de las escalas de corrección y la formación de correctores.



Figura 4 | Diseño de la prueba.

El primer paso en el diseño de la prueba es el **desarrollo de las especificaciones**. Las especificaciones se basan en el propósito de la prueba y el contexto definidos en el paso anterior y son el verdadero esqueleto de la prueba, son la hoja de ruta de los candidatos, profesores, desarrolladores de ítems, correctores, etc. Puesto que no todas las especificaciones de la prueba son de interés para todos los actores interesados, es habitual desarrollar especificaciones para los diferentes colectivos: profesionales de la evaluación o validadores de la prueba, desarrolladores de ítems, administradores, correctores, candidatos e incluso agentes externos tales como empleadores o instituciones. En ocasiones, para hacer las especificaciones más accesibles, se les puede dar formato de guía. Alderson, Clapham y Wall (1995, p.38) proporcionan una lista de comprobación para el desarrollo de especificaciones que incluye:

Propósito
Descripción del candidato
Nivel de la prueba
Constructo (marco teórico de la prueba)
Descripción de material preparatorio
Numero de partes/secciones
Tiempo para cada sección
Puntuación asignada a cada sección (peso de cada sección)
Uso de lenguaje objetivo
Tipos de texto
Longitud de los textos
Destrezas a evaluar
Elementos de la lengua a evaluar
Tareas
Métodos de evaluación
Instrucciones de las tareas
Criterios de evaluación
Respuestas modelo satisfactorias para cada nivel
Relación de la tarea con la capacidad lingüística del candidato en el mundo real
Pruebas modelo o de ejemplo
Ejemplos de respuestas de candidatos a diferentes niveles

**Para seguir leyendo, inicie el proceso de compra, [click aquí](#)**