

# Desentrañando el conocimiento histórico: validación del contenido de una prueba de evaluación de competencias históricas<sup>1</sup>

Alejandro López-García

Universidad de Murcia 

María Victoria Zaragoza Vidal

Universidad de Murcia 

Jairo Rodríguez-Medina

Universidad de Murcia 

José Monteagudo Fernández

Universidad de Murcia 

<https://dx.doi.org/10.5209/rced.93798>

Recibido: Enero 2024 • Evaluado: Junio 2024 • Aceptado: Julio 2024

**Resumen:** INTRODUCCIÓN. La evaluación del pensamiento histórico implica superar los exámenes tradicionales que buscan valorar la adquisición de contenidos conceptuales mediante el recuerdo de hechos, nombres o fechas. La nueva realidad evaluativa debe avanzar hacia un modelo de prueba emergente que ponga en el centro la demostración de habilidades de pensamiento; esto es, empatizar con el pasado y vehicular puentes hacia situaciones contemporáneas, priorizando el rigor en la selección histórica, el trabajo con fuentes, la argumentación causal en las narrativas históricas, las dialécticas contrafácticas, o el cuestionamiento ético del pasado. MÉTODO. Partiendo de estas aspiraciones, el objetivo de este trabajo es diseñar y validar una prueba, alternativa a la EBAU en España, que permita evaluar las competencias de pensamiento histórico adquiridas por el alumnado que finaliza la etapa de Bachillerato. Con este fin, se ha creado una prueba de evaluación sobre competencias históricas. Se trata de un instrumento que ha sido sometido a un proceso de validación de contenido a través de un proceso inter-jueces en cuatro dimensiones clave, en las que se aplicaron los coeficientes ponderados de concordancia de Bangdiwala ( $B_{WN}$ ) y de Gwet ( $AC_2$ ). Una vez validado, fue optimizado a través de un grupo de discusión con 10 expertos en Didáctica de las Ciencias Sociales. RESULTADOS. Los resultados fueron procesados, analizados y considerados para la construcción de la versión final, a la que se aplicó una revisión por pares con estudiantes de la Facultad de Educación de la Universidad de Murcia, al objeto de terminar de perfilar la prueba y comprobar su aplicabilidad. DISCUSIÓN. Se ha obtenido un instrumento estandarizado y útil para la evaluación de las competencias históricas en la línea de otras propuestas, encaminado a asistir al profesorado de Enseñanza Secundaria hacia un modelo de evaluación por competencias, innovador y transformador.

**Palabras clave:** Bachillerato; evaluación; método histórico; examen escrito.

## ENG Unraveling historical knowledge: validation of the content of a historical skills assessment test

**Abstract:** INTRODUCTION. The assessment of historical thinking implies overcoming traditional tests that seek to assess the acquisition of conceptual content through the recall of facts, names, or relevant dates of the past. The new evaluative reality must move towards an emerging test model that places the demonstration of thinking skills at the center; that is, empathizing with the past and building bridges to contemporary situations, prioritizing rigor in historical selection, working with sources, causal argumentation in historical narratives, counterfactual dialectics, or ethical questioning of the past, among others. METHOD. Based on these aspirations, the aim of this work is to design and validate an alternative test to the EBAU in Spain, which allows the assess of the historical thinking skills acquired by students who complete the Baccalaureate stage. To this end, an assessment test on historical competencies has been developed. It is an instrument that has undergone a content validation process through an inter-judge process in four key dimensions, in which the Bangdiwala ( $B_{WN}$ ) and Gwet ( $AC_2$ ) weighted coefficients of concordance were

<sup>1</sup> El presente trabajo forma parte del proyecto de I+D+i “La enseñanza y el aprendizaje de competencias históricas en bachillerato: un reto para lograr una ciudadanía crítica y democrática” (PID2020-113453RB-I00), financiado por MCIN/AEI/10.13039/501100011033

applied. Once validated, it was optimized through a focus group with 10 experts in Social Science Didactics. RESULTS. The results were processed, analyzed, and considered for the construction of the final version, to which a peer review was applied with students from the Faculty of Education of the University of Murcia, in order to finalize the test and check its applicability. DISCUSSION. A standardized and useful instrument has been obtained for the assessment of historical skills in line with other proposals, aimed at helping Secondary Education teachers towards an innovative and transforming competency-based evaluation model.

**Keywords:** Baccalaureate; assessment; historical method; written examination.

**Sumario:** 1. Introducción. 1.1. Las competencias de pensamiento histórico como motor del cambio. 1.2. Otra forma de evaluar la Historia: adquisición de competencias. 2. Método. 2.1. Objetivos. 2.2. Instrumento. 3. Resultados. 3.1. Objetivo 1. Analizar el grado de acuerdo entre jueces en términos de claridad, coherencia, adecuación y pertinencia. 3.2. Objetivo 2. Calcular las probabilidades acumuladas de pertenencia a los intervalos de referencia. 3.3. Objetivo 3. Calcular las razones y el índice de validez de contenido, en términos de claridad, adecuación, coherencia y pertinencia. 4. Discusión. 5. Referencias bibliográficas.

**Cómo citar:** López-García, A.; Zaragoza Vidal, M. V.; Rodríguez-Medina, J.; Monteagudo Fernández, J. (2025). Desentrañando el conocimiento histórico: validación del contenido de una prueba de evaluación de competencias históricas. *Revista Complutense de Educación* 36(3), 271-288

## 1. Introducción

### 1.1. Las competencias de pensamiento histórico como motor del cambio

El pensamiento histórico y la conciencia histórica suponen dos ejes elementales en la corriente investigadora de las últimas décadas (Seixas, 2017), siendo los contextos inglés, alemán y norteamericano centrales en su reconocimiento (Lévesque, 2023; López Facal & Schugurenky, 2023).

El concepto de pensamiento histórico occidental tiene su máxima expresión en el libro *The Big Six*, de Seixas y Morton (2012). Este enfoque de investigación tiene su origen ideológico en los trabajos de finales de siglo XX, promovidos por Dickinson (Dickinson et al., 1984; Dickinson et al., 1995). En España, destacan los trabajos de la Red 14 de Investigación en Enseñanza de las Ciencias Sociales, en la que se integran los grupos de investigación más consolidados en este campo, que están contribuyendo de forma decisiva al fortalecimiento de esta corriente (Domínguez, 2015; Gómez et al., 2014; Guerrero et al., 2019; López-García, 2023; Sáiz y López-Facal, 2015). Este término aborda la explicación histórica a través del uso de fuentes que, gracias a la argumentación propia del historiador, funcionan como pruebas (Gómez & Sáiz, 2023). También implica la habilidad de identificar, analizar y evaluar los hechos históricos -distinguiendo entre la causalidad cortoplacista y largoplacista-, así como el contexto, los procesos de cambio y continuidad, o las lecciones que se pueden extraer del pasado para comprender mejor el presente (López Facal & Schugurenky, 2023).

La relevancia histórica responde a la habilidad para establecer y comprender qué acontecimientos han sido claves en la historia y por qué, brindándose la oportunidad en las narrativas de seleccionar y diferenciar lo históricamente determinante, de lo curioso, lo raro o lo anecdótico, que en ocasiones puede presentarse de forma conjunta (Egea y Arias, 2018).

Por su parte, la dimensión ética es la habilidad de enjuiciar el pasado y utilizar el pensamiento crítico para valorar el impacto histórico y la repercusión moral de las decisiones que se han tomado en otros tiempos. Un ejemplo son los genocidios históricos (Holocausto judío, genocidio armenio o masacre de Nankín, entre otros) o las políticas expansionistas (Imperio Romano, imperialismo ruso, o germanización de la Alemania nazi, son solo algunas de ellas). En la medida en que el razonamiento ético se encuentra unido con las obligaciones o responsabilidades que tiene la sociedad actual para recordar o responder a las contribuciones, sacrificios e injusticias pasadas, evitando que vuelvan a repetirse (Seixas & Morton, 2012), también se lo ha relacionado con la educación para la ciudadanía (Lévesque & Clark, 2018).

Analizar éticamente el pasado supone también reflexionar sobre los procesos históricos, que implican progresos y decadencias. El concepto de cambio y continuidad es la habilidad de reflexionar sobre qué permanece y qué cambia a lo largo del tiempo. Su complejidad coexiste con los esfuerzos que los profesores de Historia vienen haciendo en el contexto de actividades prácticas que implican el trabajo con líneas de tiempo, gráficos analíticos, metáforas, analogías u otros contextos temporales que buscan la erudición histórica (Counsell, 2017).

Para objetivar los procesos de cambio histórico son claves las evidencias; es decir, el arte de utilizar fuentes primarias como una herramienta fundamental para extraer pruebas que permitan construir el discurso histórico. Es imprescindible saber qué preguntas se deben hacer a la fuente, así como partir del conocimiento que deja el contexto histórico de la misma. Para Ashby (2006) se trata de uno de los conceptos más destacables, pues ofrece la posibilidad al alumnado de obtener la información del pasado, lo que contribuye a generar una comprensión acerca de la construcción del conocimiento histórico, advirtiendo las fuentes como testimonios directos. De esta forma, se pueden obtener el conjunto de inferencias que permitirán generar una explicación sobre el pasado (Cerdeira, 2021).

El concepto de causa y consecuencia sirve para analizar los factores que han provocado un determinado hecho histórico y las múltiples consecuencias que se derivan de este (Seixas & Morton, 2012). Según

Domínguez (2015) el aprendizaje no debe ir enfocado a saber historia, sino a saber hacer historia. En este sentido, la comprensión de la causalidad es fundamental para comprender que esta es una sucesión de hechos interrelacionados y, desarrollar este concepto implica utilizar el razonamiento para construir ese pensamiento histórico. Germinari (2011) señala al respecto que no se trata de la cantidad de información factual que se puede aprender, sino de “aprender a pensar sobre el pasado históricamente” (p. 57), y ello implica conectar evidencias.

Seixas y Morton (2012) definen la perspectiva histórica estableciendo una analogía: debemos mirar al pasado como si fuésemos extranjeros en un país donde el contexto histórico, social y cultural es diferente al nuestro, y este influye en la vida de las personas que habitaron en el pasado. Esta competencia se encuentra especialmente vinculada a la empatía, ya que atiende a los actos y pensamientos de los individuos del pasado, pero a su vez se vincula con la propia experiencia del sujeto actual (Carril-Merino et al., 2023). Por tanto, promueve que los adolescentes puedan entender el pasado sin valerse de sus percepciones del presente, que puedan llevarlos a pronunciarse sobre este de una manera errónea (Lévesque & Clark, 2018).

Una vez descritas estas competencias de pensamiento histórico, es necesario situar en contexto la corriente de conciencia histórica o *historical consciousness*, por sus siglas en inglés. Esta escuela de investigación emerge en Europa central, concretamente en el contexto alemán, a raíz de los postulados propuestos por Rüsen (2004, 2015), y apoyados por Wertsch (2004) o Körber (2017), entre otros. La conciencia histórica europea es el modo en que los ciudadanos ven su pasado colectivo, como plataforma para entender el presente y formar el futuro (Lorentzer, 2000). Este concepto sostiene un papel más social en la enseñanza de la historia, relacionándose con la memoria común, la cultura popular, la identidad, las representaciones mediáticas y la educación moral (Gómez & Sáiz, 2023; Rodríguez-Medina, 2020). También ha sido situado desde una dualidad; como fenómeno colectivo de la sociedad occidental moderna y desde la categoría individual cognitivo-epistemológica (Grever & Adriaansen, 2019). Por otra parte, se integra en entornos informales y experienciales, creando espacios para las perspectivas marginadas, de género, racializadas y aborígenes, con un protagonismo periférico en los métodos tradicionales de enseñar historia (McLean et al., 2017). En definitiva, esta corriente de pensamiento desarrolla una competencia narrativa que permite orientar y comprender la vida práctica del presente gracias al recuerdo, a la representación y a la interpretación de la realidad pasada, cuyo sentido favorece, además de la vivencia actual, la anticipación del futuro (Duquette, 2015; Körber, 2017). Los centros de formación deben ejercer un papel significativo en el desarrollo de la conciencia histórica de los jóvenes, y en el nuevo enfoque multicultural de inclusión, igualdad y apertura ciudadana (Gómez & López Facal, 2023); una educación más cívica y democrática frente al discurso manipulador.

La historia -el pasado- es un recurso fundamental para el devenir del futuro (Arthur et al. 2001; Dewey, 1998). Partiendo de esta máxima, la enseñanza de la historia y la educación ciudadana deben mantener puntos de encuentro, ya que la primera debe ejercer un papel fundamental en el desarrollo de la conciencia cívica, mediante la formación de ciudadanos críticos y comprometidos con la sociedad y el entorno en el que viven (López-Fernández et al., 2023). En este sentido es como, desde las ciencias sociales, se contribuye a generar una memoria común, un pensamiento más crítico, valores como la empatía, la tolerancia y el respeto, o ciertos principios de justicia social (Pagés, 2005).

## 1.2. Otra forma de evaluar la Historia: adquisición de competencias

En el caso de las asignaturas relacionadas con la Historia, el actual Real Decreto 217/2022 de enseñanzas mínimas de Educación Secundaria Obligatoria recoge que, para contribuir a la consecución de los objetivos y competencias que conforman el Perfil de salida del alumnado al término de la enseñanza básica, es necesario trabajar con las herramientas y estrategias propias de los métodos de estudio de esta disciplina, el pensamiento histórico, el cual define como el “proceso por el que se crean narrativas sobre el pasado a través de la reflexión sobre su relevancia, el análisis de fuentes, la discusión sobre las causas y consecuencias de estos acontecimientos, así como el análisis de los cambios y continuidades entre los mismos, desde una perspectiva temporal y contextualizada, y con relación a determinados criterios éticos y cívicos” (p. 96). Este conjunto de intenciones, estrategias y métodos, que orientan el estudio del pasado, se mantiene en la etapa del Bachillerato en las asignaturas de Historia del Mundo Contemporáneo e Historia de España.

Partiendo de que, imperiosamente, las evaluaciones deben estar relacionadas con el proceso de enseñanza-aprendizaje, y habida cuenta de que la enseñanza de las ciencias sociales viene poniendo el foco en las competencias de pensamiento histórico desde hace varios decenios, particularmente en el contexto anglosajón, el sistema de evaluación en esta materia debe priorizar la adquisición de estas competencias -y su demostración fehaciente-, evidenciándose la necesidad de una nueva forma de evaluar la historia.

Moreno et al. (2023) han confirmado que, en el caso de la asignatura de Historia de España, no hay ninguno de los 17 criterios de evaluación que no tenga relación con las competencias del pensamiento histórico, exigiendo a los estudiantes reflexiones más profundas, una mayor comprensión y deducción, capacidades de explicación, argumentación y comprensión de la historia, así como la generación de relaciones multicausales entre los acontecimientos históricos.

Los últimos estudios en la materia han demostrado que las evaluaciones centradas en las respuestas construidas son las que más favorecen la comprobación del nivel de adquisición competencial, seguidas de las preguntas de opción múltiple (Smith et al., 2019; Miguel-Revilla, 2020). Ambas están presentes en este trabajo, particularmente las primeras. En este sentido, Bruner (2005) planteaba a principios de siglo la siguiente pregunta: “¿Qué se gana y qué se pierde cuando los seres humanos recurren al modo narrativo de construir significados?” (p. 23). La propia pregunta evidencia que el pensamiento narrativo se

encuentra omnipresente en la vida de las personas, por lo que se trata de una operación mental imprescindible para representar, organizar y dar sentido a ese pasado histórico, mientras brinda significados y orienta la vida actual mediante la categoría temporal (Carretero y Atorresi, 2014). De hecho, el discurso narrativo promueve la dimensión temporal, el establecimiento de nexos causales y otros medios con los que otorgar importancia histórica y conectar pasado y presente, desde la perspectiva moral (Barton & Levstik, 2004). Así, los hechos históricos y sus protagonistas cumplen con los criterios de significación histórica sólo cuando se demuestre que ocupan un lugar significativo en una narrativa (Lévesque, 2008; Seixas & Morton, 2012), ya sea a través de su sentido de relevancia, su causalidad, su cambios y permanencias o su análisis del contexto histórico.

Este trabajo presenta el diseño y la validación de una prueba de competencias financiada que aspira a poder establecer un marco común de evaluación de competencias en todo el territorio español desde el respeto al marco legislativo, a las corrientes de investigación, y a la implementación práctica en las aulas.

## **2. Método**

Durante el último decenio el grupo DICSO de la Universidad de Murcia (UMU) ha consolidado una amplia producción científica relacionada con el metaconcepto del pensamiento histórico (Álvarez et al., 2020; Arias et al., 2019; Domínguez et al., 2021; Gómez y Miralles, 2015; Gómez, Monteagudo et al., 2018; Gómez, Rodríguez et al. (2018); Miralles y Alfageme 2013; Miralles y Gómez, 2016). A través del actual proyecto de investigación el reto consiste en hacer lo propio en el nivel de Bachillerato. Para ello se propuso el diseño de un instrumento de evaluación capaz de medir el nivel de adquisición de las distintas competencias de pensamiento histórico por parte del alumnado que ha finalizado esta etapa educativa.

Siguiendo el esquema propuesto por Verdú y Gómez (2019), en una primera fase y de cara a diseñar el instrumento de recogida de información, se revisó la literatura sobre las competencias que componen el pensamiento histórico. De dicha revisión surgió un primer borrador de la prueba.

Para testar la validez de contenido, tras la realización de un muestreo deliberado, se realizó una evaluación inter-jueces en la que varios expertos (profesorado de Bachillerato y profesorado universitario de Didáctica de las Ciencias Sociales y de Metodología de la Investigación Educativa, de la UMU, Universidad de Santiago de Compostela, Universidad de Valladolid, Universidad de Valencia) mostraron su parecer sobre esta prueba a través de una escala de valoración sobre el grado de pertinencia, adecuación, coherencia y claridad en la redacción, para cada opción y cada ejercicio. En concreto, la prueba fue sometida a la revisión de 7 jueces, que valoraron cada una de estas variables en una escala que oscila de 1 (nada adecuado) a 5 (totalmente adecuado). Los datos se analizaron a nivel global y para cada una de las versiones (A y B), independientemente.

Una vez analizados los resultados y validada la prueba, se llevó a cabo un grupo de discusión con 10 profesores expertos de Bachillerato y de Didáctica de las Ciencias Sociales de la Facultad de Educación de la UMU, al efecto de reflexionar sobre las sugerencias de eliminación y modificación, corregir carencias y optimizar la prueba, perfilando la versión final de la misma. Finalmente, para asegurar que esta estuviese centrada en evaluar el nivel de logro de las competencias históricas del grupo objetivo, la nueva versión fue testada a través de lo que Molina et al. (2016) llaman revisión por pares, en este caso con un grupo de estudiantes de la UMU, que se prestaron a la realización de la prueba, a fin de calcular el tiempo empleado en ello y la evaluación de determinados aspectos técnicos de la misma.

### **2.1. Objetivos**

Con base en lo anterior, el objetivo general de este trabajo es medir la validez de contenido de una prueba de evaluación sobre el nivel de adquisición de competencias de pensamiento histórico del alumnado que ha terminado la etapa de Bachillerato. Este objetivo se concreta en tres objetivos específicos:

- Objetivo 1. Analizar el grado de acuerdo entre jueces en términos de claridad, coherencia, adecuación y pertinencia.
- Objetivo 2. Calcular las probabilidades acumuladas de pertenencia a los intervalos de referencia.
- Objetivo 3. Calcular las razones y el índice de validez de contenido, en términos de claridad, adecuación, coherencia y pertinencia.

### **2.2. Instrumento**

Para este proceso de validación, la escala de valoración aplicada a los jueces-expertos consistió en un cuestionario que, en primer lugar, tiene un apartado de presentación de la prueba. Seguidamente, se propuso una valoración general de la misma de acuerdo con tres variables centrales: el título, las instrucciones para responder y el diseño general de esta. A continuación, para cada ejercicio sobre competencias históricas se incluyó una escala de valoración tipo Likert con cinco opciones (1-nada adecuado; 2-poco adecuado; 3-adecuado; 4-bastante adecuado; 5-muy adecuado), así como una propuesta de mejora, respecto a cada una de las competencias abordadas, de acuerdo con cuatro variables (pertinencia, claridad, coherencia y adecuación).

Respecto a la descripción del primer diseño de la prueba, en su parte inicial, esta incluía un primer apartado con los datos de identificación. Seguidamente, se presentaba un apartado de instrucciones. En su parte principal, la prueba se dividía en seis bloques, cada uno de los cuales se centraba en la evaluación de una competencia de pensamiento histórico. A esta primera propuesta se añadió una batería de ejercicios

alternativa, preparada por profesorado asociado y con experiencia docente en Bachillerato, dejando finalmente este primer diseño inicial de la prueba con dos opciones, A y B; la versión A contaba con 19 ítems y la versión B con 20 ítems, para un total de 39. Cada una de las opciones constaba de varios ejercicios para evaluar cada competencia, de acuerdo con el modelo de competencias que proponen Seixas y Morton (2012). Por tanto, de cada competencia se propusieron dos opciones, siendo labor del validador experto seleccionar o valorar una de ellas, incluso una opción intermedia o una mezcla de las dos.

### 3. Resultados

#### 3.1. Objetivo 1. Analizar el grado de acuerdo entre jueces en términos de claridad, coherencia, adecuación y pertinencia.

Para analizar el acuerdo entre los jueces se calcularon los coeficientes ponderados de concordancia de Bangdiwala ( $B_{WN}$ ) (Bangdiwala, 1987). El índice de acuerdo  $B_{WN}$  de Bangdiwala permite la representación gráfica del grado de acuerdo y proporciona una medida de la fuerza de este. En esta representación, los cuadrados negros muestran el acuerdo observado, en tanto que las áreas grises representan el acuerdo parcial. La zona blanca de cada rectángulo es la representación gráfica del desacuerdo. Los datos se analizaron mediante el software R v. 4.0.4 (R Core Team, 2021).

En la Tabla 1 se muestran los coeficientes de concordancia estrictos (BN) y ponderados o de acuerdo parcial (BWN) de Bangdiwala (Bangdiwala, 1987; Friendly & Meyer, 2016) obtenidos.

Tabla 1. Coeficientes de concordancia de Bangdiwala estrictos (BN) y ponderados (BWN)

|                      | <b>Coef.</b> | <b>Pertinencia</b> | <b>Coherencia</b> | <b>Adecuación</b> | <b>Claridad</b> |
|----------------------|--------------|--------------------|-------------------|-------------------|-----------------|
| <b>Pruebas A y B</b> | $B_N$        | .24                | .44               | .45               | .40             |
|                      | $B_{WN}$     | .49                | .68               | .68               | .62             |
|                      | $AC_2$       | .52                | .69               | .72               | .62             |
| <b>Prueba A</b>      | $B_N$        | .24                | .23               | .25               | .22             |
|                      | $B_{WN}$     | .49                | .60               | .58               | .47             |
|                      | $AC_2$       | .30                | .53               | .54               | .42             |
| <b>Prueba B</b>      | $B_N$        | .46                | .70               | .71               | .62             |
|                      | $B_{WN}$     | .64                | .83               | .87               | .80             |
|                      | $AC_2$       | .64                | .83               | .87               | .76             |

Para completar el análisis se calcularon también los coeficientes ponderados de Gwet (2008) para superar las limitaciones y paradojas propias del coeficiente Kappa de Cohen (Cicchetti & Feinstein, 1990; Feinstein & Cicchetti, 1990). En este caso se empleó el  $AC_2$ , por tratarse de datos ordinales, ya que permite el acuerdo parcial. Los índices de acuerdo  $AC_2$  obtenidos pueden considerarse entre moderados y buenos para las cuatro variables examinadas atendiendo a los criterios propuestos por Muñoz y Bangdiwala (1997).

Los coeficientes BN (sobre los valores de la matriz de coincidencia que se somete al análisis de concordancia) se calculan mediante la fórmula:

$$\frac{\sum n_{ii}^2}{\sum n_i \cdot n_{ii}} = B_N = \frac{\text{área de los cuadrados negros}}{\text{área total de los rectángulos de cada categoría}}$$

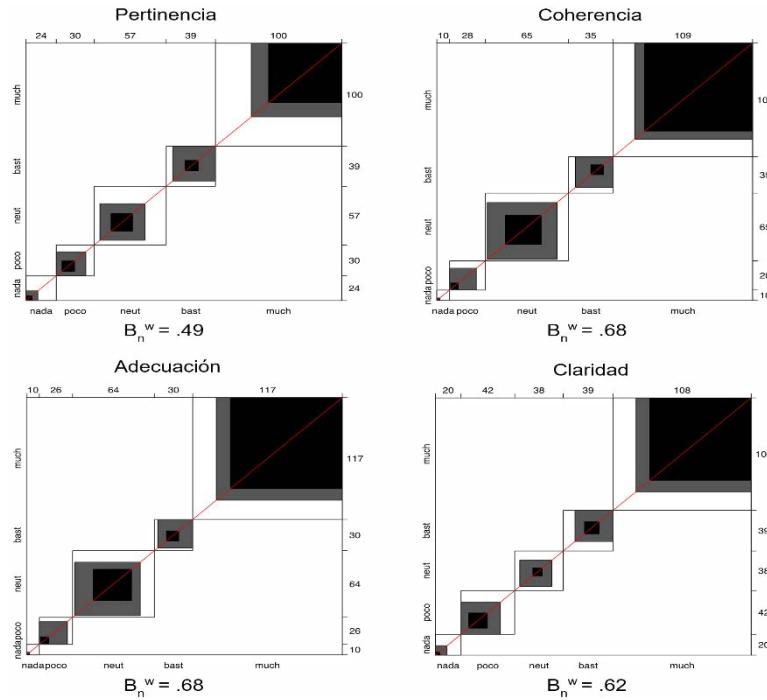
Para contemplar el acuerdo parcial ( $B_{WN}$ ), puesto que se trata de una escala de valoración ordinal, se incluye una contribución ponderada de las celdas fuera de la diagonal en función de los pasos (separación) desde la diagonal principal. Estos acuerdos parciales se incluyen en el gráfico (Figura 1) en forma de cuadrados de un tono más claro (cuadrados grises) que el acuerdo estricto (cuadrados negros). De tal forma que se aplica un patrón de pesos (ponderaciones  $w_1, w_2, \dots, w_b$ ), conforme a la fórmula propuesta por Fleiss y Cohen (1973), a las áreas sombreadas separadas  $b$  pasos de la diagonal. De modo que para calcular el coeficiente de acuerdo parcial ( $B_{WN}$ ) se emplea la siguiente fórmula:

$$B_{WN}^W = 1 - \frac{\sum [n_{ii} + n_{+i} - n_{ii}^2 - \sum w_b A_{bi}]}{\sum n_{ii} + n_{+i}}$$

Donde  $w_b$  representa la ponderación en función de la distancia a la diagonal y  $A_{bi}$  representa el área de las zonas sombreadas con una separación  $b$  de la diagonal. Como se puede observar en la Figura 1, los coeficientes ponderados de concordancia de Bangdiwala (Bangdiwala, 1987) obtenidos fueron:  $B_{WN}^W = .49$  (pertinencia);  $B_{WN}^W = .68$  (coherencia);  $B_{WN}^W = .68$  (adecuación); y  $B_{WN}^W = .62$  (claridad de la redacción).

Para interpretar los coeficientes de concordancia, Muñoz y Bangdiwala (1997), proponen los siguientes criterios: valores por debajo de .09 indican un nivel de acuerdo pobre; entre .09 y .25, acuerdo escaso; entre .25 y .49 moderado; entre .49 y .801 bueno; y superiores a .801, indican un nivel de acuerdo excelente. Los resultados obtenidos, de acuerdo con esta interpretación, sugieren un nivel de acuerdo entre moderado y bueno, de modo que no disponemos de evidencias robustas de que el contenido de la prueba refleje adecuadamente el contenido que se pretende evaluar.

Figura 1. Resultados del análisis de acuerdo entre jueces de Bangdiwala



### 3.2. Objetivo 2. Calcular las probabilidades acumuladas de pertenencia a los intervalos de referencia.

Los índices de acuerdo AC deben considerarse entre moderados y buenos, tomando en consideración la propuesta de Gwet (2021) para interpretar los valores obtenidos. Así pues, el objetivo 2 alude a calcular la probabilidad de pertenencia a cada uno de los intervalos (Interval Membership Probability), cuyos resultados se muestran en la Tabla 2 utilizando la escala de referencia propuesta por Muñoz y Bangdiwala (1997).

Tabla 2. Probabilidades acumuladas de pertenencia a los intervalos de referencia - Instrumento de observación

|                    |           |            | Pertinencia                         | Coherencia                           | Adecuación                           | Claridad                            |
|--------------------|-----------|------------|-------------------------------------|--------------------------------------|--------------------------------------|-------------------------------------|
| Valores de Acuerdo |           | Interpret. | .52 (SE = .197)<br>95%CI [.043 - 1] | .692 (SE = .162)<br>95%CI [.294 - 1] | .718 (SE = .153)<br>95%CI [.344 - 1] | .62 (SE = .198)<br>95%CI [.137 - 1] |
| [.81 - 1]          | Excelente |            | .066                                | .205                                 | .242                                 | .142                                |
| [.49 - .81]        | Bueno     |            | .562                                | .863                                 | .899                                 | .718                                |
| [.25 - .49]        | Moderado  |            | .910                                | .967                                 | .965                                 | .941                                |
| [.09 - .25]        | Escaso    |            | .978                                | .970                                 | .966                                 | .968                                |
| [.01 - .09]        | Pobre     |            | .987                                | .970                                 | .966                                 | .971                                |

Nota. Interpret. = Interpretación.

Atendiendo a los resultados expuestos en la Tabla 2, y asumiendo los valores de la escala de referencia señalada, el nivel de acuerdo obtenido en cuanto a pertinencia y claridad se puede considerar escaso, mientras que, en relación a la coherencia y la adecuación, el grado de acuerdo es moderado, analizándose e interpretándose las cuatro variables con una confianza superior al 95%.

### 3.3. Objetivo 3. Calcular las razones y el índice de validez de contenido, en términos de claridad, adecuación, coherencia y pertinencia.

Para dar respuesta a este objetivo se calcularon las razones de validez de contenido (*Content Validity Ratio*, CVR) (Lawshe, 1975) para cada uno de los ítems en cada una de las variables analizadas (pertinencia, coherencia, adecuación y claridad de la redacción) y el índice de validez de contenido (*Content Validity Index*, CVI) (Lawshe, 1975) para el conjunto de ítems. Se trata de un indicador del acuerdo entre jueces que puede tomar valores entre -1 (desacuerdo total) y +1 (acuerdo total); de tal modo que el valor CVR es negativo si el acuerdo ocurre en menos de la mitad de los jueces; CVR es cero si se tiene exactamente la mitad de los acuerdos entre los jueces expertos; y CVR es positivo si más de la mitad de los jueces están de acuerdo en la valoración del ítem. Para la interpretación de los resultados con 7 jueces, Ayre y Scally (2014) proponen el valor crítico de CVR = 1 ( $p = .008$ ), lo que supone que los 7 jueces estén de acuerdo en la valoración del ítem y supera la probabilidad de acuerdo por efecto del azar a un nivel de confianza del 95% ( $\alpha = .05$ ).

Para el cálculo de los valores CVR, se agruparon las valoraciones de los jueces que seleccionaron las opciones 4 (bastante) y 5 (mucho) y se calcularon las razones aplicando la fórmula propuesta por Lawshe (1975):

$$RVC = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}$$

Donde  $n_e$  es igual al número de jueces que consideran el ítem bastante o muy pertinente, adecuado, coherente, o claro y  $N$  es el número total de jueces. Finalmente, el Índice de Validez de Contenido (CVI) se obtuvo mediante el cálculo del promedio de la validez de contenido (RVC) de cada uno de los ítems en cada una de las variables contempladas y globalmente.

Los valores RVC oscilaron entre -.71 y 1 en las cuatro variables consideradas (pertinencia, coherencia, adecuación y claridad de la redacción). Solo un ítem (B.1.3) alcanzó el valor RVC crítico propuesto por Ayre y Scally (2014) de .1 para 7 jueces, en las variables pertinencia y coherencia; del mismo modo, solo un ítem (B.1.1) alcanzó este valor en la variable claridad. Ningún ítem alcanzó este valor en la variable adecuación. El CVI global para el conjunto de ítems fue .077 (pertinencia), .055 (coherencia), .077 (adecuación) y .077 (claridad).

Estos resultados indican que, a juicio de los expertos, la prueba carece de evidencias de validez referidas al contenido, es decir, el contenido de la prueba no refleja adecuadamente el contenido que se pretende evaluar. En cualquier caso, aparentemente la prueba B ha mostrado un nivel ligeramente superior de acuerdo entre los jueces respecto a su pertinencia, coherencia, adecuación y claridad que la prueba A.

#### 4. Discusión

Dentro del diseño inicial del instrumento para la evaluación de las competencias históricas que pretende medir el nivel de adquisición de dichas competencias por el alumnado que ha cursado en 2º de Bachillerato la asignatura de Historia de España, el modelo B resultó cumplir mejor con las propiedades psicométricas aceptables para ser empleado como herramienta de evaluación.

En este sentido, tanto la prueba del coeficiente ponderado de concordancia como el coeficiente de acuerdo arrojaron unos resultados muy positivos, siendo buenos en el caso de la claridad (.80 y .76, respectivamente) y pertinencia (.64 en ambos casos), y con un nivel excelente en la adecuación (.87) y la coherencia (.83).

Por otro lado, el Índice de Validez de Contenido (CVI), para un grupo de 7 jueces recogía un global para el conjunto de ítems de .077 (pertinencia), .055 (coherencia), .077 (adecuación) y .077 (claridad). De este modo, siguiendo lo expuesto por Pedrosa et al. (2013), podría decirse que se trata de un instrumento útil pero prescindible. Con otras palabras, nos encontrábamos con un instrumento que servía para los fines propuestos pero que podía ser sustituible, de modo que precisaba de un proceso de optimización a fin de que resultase esencial y mida lo que se propone.

A partir de estos resultados, el grupo de discusión, como medio cualitativo, permitió reflexionar sobre este proceso y acometer las mejoras oportunas, siguiendo las sugerencias de cambio, y en consonancia con los resultados de la investigación, de manera que se eligió el modelo B y fue perfeccionado, modificando o suprimiendo aquellos ítems con resultados mejorables. Esto redundó en una nueva versión de la prueba que prioriza la evaluación del alcance en el nivel de competencias históricas, acercándolo también al público al que va iba dirigido.

Por su parte, la revisión por pares que efectuó el grupo de alumnos de la UMU, realizando la prueba, permitió ajustar el tiempo necesario para su desarrollo y depuró algunos detalles formales y técnicos. En este proceso de optimización de la prueba, a nivel general los resultados arrojaron que los ítems no precisaban de grandes cambios para poder ser entendidos por una población con características similares. No obstante, algunas ligeras observaciones realizadas por los estudiantes relacionadas con la redacción de los ejercicios fueron incorporadas a la prueba para su perfeccionamiento final. Tras esta evaluación por pares, el instrumento volvió a ser presentado al resto de investigadores de las distintas universidades involucrados en este proceso de mejora. En una segunda revisión por este grupo de discusión se pulieron todavía más los ejercicios para evaluar competencias históricas, a fin de que el alumnado no tuviese problemas relacionados con la comprensión de lo que se le demandaba ni con el tiempo de realización.

Existen en el caso español iniciativas anteriores enfocadas a la evaluación del pensamiento histórico (González et al., 2011), pero más dilatadas en el tiempo. Así, los intentos más recientes y similares de evaluar las competencias históricas mediante una única prueba, en este caso parecidas a las de PISA (Domínguez et al., 2017), contaron con un proceso de diseño y validación similar al descrito a través de una validación cualitativa realizada por expertos en distintas especialidades y otra más cuantitativa y experimental con estudiantes de 4.º de ESO.

En contraste, la novedad de nuestro intento por confeccionar un único instrumento de verificación capaz de evaluar el nivel de adquisición de competencias históricas es que se centra en los exámenes de acceso a la universidad que debe realizar el alumnado que ha terminado la etapa del Bachillerato, con la intención de ofrecer una alternativa a este tipo de pruebas que apunte directa y claramente a las capacidades cognitivas superiores desarrolladas por los historiadores en el desempeño de su labor. El resultado ha sido la Prueba titulada *Evaluación de Competencias Históricas* (ECOHI), presentada en el Anexo 1 de este trabajo, la cual, teniendo presente el método seguido, anclado en la validación por jueces-expertos, y el posterior pulido en los grupos de discusión y evaluación por pares, cumple con dicho cometido, ya que, en base a los objetivos

de nuestra investigación, los resultados muestran que se fundamenta en un buen grado de claridad, es decir, la redacción de los ejercicios es comprensible formalmente y sin ambigüedades; en unos excelentes niveles de adecuación y coherencia, esto es, que los ejercicios incorporados se ajustan al nivel cognitivo de los participantes, y que se relacionan con la competencia histórica que abordan; y en un buen grado de pertinencia, en otras palabras, que los ejercicios favorecen una demostración del conocimiento competencial.

La relevancia y utilidad educativas de propuestas como la presentada aquí y otras similares (Seixas et al., 2015; VanSledright, 2014), esto es, la confección de una prueba de evaluación sobre el pensamiento histórico centrada en el trabajo con fuentes en base a preguntas de opción múltiple y ejercicios de tipo ensayo que posibilitan el tratamiento complejo de temas históricos, quedan de manifiesto en el reciente diseño de propuestas didácticas en entornos digitales dirigidas a estudiantes de Bachillerato (Pérez de Iriarte y Andía, 2024).

Tras el proceso seguido, se concluye que se requiere de un gran conocimiento didáctico del contenido (Shulman, 2005) a fin de encapsular, en una única prueba, una amplia gama de ejercicios conducentes a valorar el grado de adquisición de cada una de las citadas competencias que conforman el pensamiento histórico.

En prospectiva, las limitaciones podrán venir de la aplicación de la prueba en un estudio piloto, dirigido a comprobar la validez y fiabilidad de la prueba en sí misma, como instrumento para evaluar el pensamiento histórico del alumnado. De él se derivarán posibles necesidades de mejora, antes de ser administrada a la muestra final de destino. A pesar de ello, resulta perentorio poner en valor la necesidad de un cambio en los modelos evaluativos del conocimiento histórico, por lo que esta prueba aspira a establecer una alternativa innovadora y transformadora en los procesos de evaluación del pensamiento histórico.

## 5. Referencias bibliográficas

- Álvarez, J. M., Trigueros, F. J., Miralles, P., & Molina, J. (2020). Assessment by competences in social sciences: secondary students perception based on the EPECOCISO scale. *Sustainability*, 12(23), 10056, 1-15. <https://www.mdpi.com/2071-1050/12/23/10056>
- Arias, L., Egea, A., Sánchez, R., Domínguez, J., García, J. y Miralles, P. (2019). ¿Historia olvidada e historia no enseñanza? El alumnado de Secundaria español y su conocimiento sobre la Guerra Civil. *Revista Complutense de Educación*, 30(2), 461-478. <https://doi.org/10.5209/RCED.57625>
- Arthur, J., Davies, I., Wrenn, A., Haydn, T., & Kerr, D. (2001). *Citizenship through Secondary History*. Routledge.
- Ashby, R. (2006). Desenvolvendo um conceito de evidencia histórica: as idéias dos estudantes sobre testar afirmações factuais singulares (E.M. Schmidt, L. Braga, M.A. Schmidt y T. Braga, trad.). *Educar, Curitiba, Especial* (pp. 151-170). Editora UFPR (Texto original publicado en 2005). <https://doi.org/10.1590/0104-4060.404>
- Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79–86. <https://doi.org/10.1177/0748175613513808>
- Bangdiwala, S. (1987). Using SAS software graphical procedures for the observer agreement chart. *Proceedings of the SAS Users Group International Conference*, 12, 1083-1088.
- Barton, K. C., & Levstik, L. S. (2004). *Teaching History for the Common Good*. Routledge.
- Bruner, J. S. (2005). Past and Present as Narrative Constructions. In J. Straub (Ed.), *Narration, Identity and Historical Consciousness* (pp. 23-43). Berghahn Books.
- Carretero, M. y Atorresi, A. (2014). Pensamiento narrativo. En M. Carretero y M. Asensio (Eds.), *Psicología del pensamiento* (2<sup>a</sup> Ed.) (pp. 269-289). Alianza.
- Carril-Merino, M. T., Sánchez-Agustí, M. y López-Torres, E. (2023). Empatía y empatía histórica. Su correlación a través de un dilema ético. *Profesorado. Revista de currículum y formación del profesorado*, 27(2), 331-315. <https://doi.org/10.30827/profesorado.v27i2.22645>
- Cerdeira, A. S. (2021). Operações do pensamento histórico de jovens estudantes: um estudo sobre a concepção de evidência. *Revista Brasileira de História*, 41(86), 67-86. <https://doi.org/10.1590/1806-93472021v41n86-03>
- Cicchetti, D. V., & Feinstein, A. R. (1990). High agreement but low Kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology*, 43(6), 551-558. [https://doi.org/10.1016/0895-4356\(90\)90158-L](https://doi.org/10.1016/0895-4356(90)90158-L)
- Counsell, C. (2017). Historical change and continuity: how history teachers are advancing the field. In I. Davies (Ed.), *Debates in History Teaching* (2<sup>a</sup> Ed.) (pp. 113-129). Routledge. <https://doi.org/10.4324/9781315642864>
- Dewey, J. (1998). *Democracia y Educación*. Ediciones Morata.
- Dickinson, A., Gordon, P., Lee, P., & Slater, J. (Eds.). (1995). *International Yearbook of History Education*. Routledge. <https://doi.org/10.4324/9781315030357>
- Dickinson, A. K., Lee, P. J., & Rogers, P. J. (Eds.). (1984). *Learning History*. Heinemann.
- Domínguez, J. (2015). *Pensamiento histórico y evaluación de competencias*. Graó.
- Domínguez, J., Arias, L., Sánchez, R., Egea, A. y García, F. (2017). Primeros resultados de una prueba piloto para evaluar el pensamiento histórico de los estudiantes. *Clio & Asociados. La historia enseñada*, 24, 38-50. <https://doi.org/10.14409/cya.v0i24.6672>
- Domínguez, J., Arias, L., Sánchez, R., Egea, A., García, J., & Miralles, P. (2021). A competence-based test to assess historical thinking in secondary education: design, application and validation. *Historical Encounters* 8(1), 30-45. <https://doi.org/10.52289/hej8.103>
- Duquette, C. (2015). Relating Historical Consciousness to historical thinking through assessment. In K. Ercikan, & P. Seixas (Eds.), *New Directions in Assessing Historical Thinking* (pp. 51-63). Routledge.

- Egea, A. y Arias, L. (2018). ¿Qué es relevante históricamente? Pensamiento histórico a través de las narrativas de los estudiantes universitarios. *Educação e Pesquisa*, 44, 168641. <https://doi.org/10.1590/S1678-4634201709168641>.
- Feinstien, A. R., & Cicchetti, D. V. (1990). High agreement but low Kappa: I. The Problems of two paradoxes. *Journal of Clinical Epidemiology*, 43(6), 543–549. [https://doi.org/10.1016/0895-4356\(90\)90158-I](https://doi.org/10.1016/0895-4356(90)90158-I)
- Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33(3), 613–619. <https://doi.org/10.1177/001316447303300309>
- Friendly, M., & Meyer, D. (2016). *Discrete data analysis with R: Visualization and modeling techniques for categorical and count data*. Chapman & Hall.
- Germinari, G. D. (2011). Educação histórica: a constituição de um campo de pesquisa. *Revista HISTEDBR online*, 11(42), 54-70. <https://doi.org/10.20396/rho.v11i42.8639866>
- Gómez, C. J., & López Facal, R. (2023). Introduction. Re-imagining the teaching of European history from historical thinking and civic engagement. In C. J. Gómez (Ed.), *Reimagining the Teaching of European History. Promoting Civic Education and Historical Consciousness* (pp. 1-10). Routledge. <https://doi.org/10.4324/9781003289470>.
- Gómez, C. J. y Miralles, P. (2015). ¿Pensar históricamente o memorizar el pasado? La evolución de contenidos históricos en educación obligatoria en España. *Revista Estudios Sociales*, 52, 52-68.
- Gómez, C. J., Monteagudo, J. y Miralles, P. (2018). Conocimiento histórico y evaluación de competencias en los exámenes de Educación Secundaria. Un análisis comparativo España-Inglatera. *Educatio Siglo XXI*, 36(1), 85-106.
- Gómez, C. J., Ortúñoz, J. y Molina, S. (2014). Aprender a pensar históricamente. Retos para la historia en el siglo XXI. *Tempo e Argumento*, 6(11), 5-27. <https://doi.org/10.5965/2175180306112014005>.
- Gómez, C. J., Rodríguez, R. A. y Mirete, A. B. (2018). Metodología didáctica y concepciones epistemológicas sobre la enseñanza de la historia. Una investigación con futuros maestros. *Revista Complutense de Educación*, 24(1), 237-250. <https://doi.org/10.5209/RCED.52233>.
- Gómez, C. J., & Sáiz, J. (2023). The origin and development of research into historical thinking: a key concept in the renewal of history education. In C. J. Gómez (Ed.), *Reimagining the Teaching of European History. Promoting Civic Education and Historical Consciousness* (pp. 25-41). Routledge. <https://doi.org/10.4324/9781003289470>.
- González, N., Pagès, J. y Santisteban, A. (2011). ¿Cómo evaluar el pensamiento histórico del alumnado? En P. Miralles, S. Molina y A. Santisteban (Eds.), *La evaluación en el proceso de enseñanza y aprendizaje de las Ciencias Sociales* (Vol. 1, pp. 221-232). AUPDCS.
- Grever, M., & Adriaansen, R.-J. (2019). Historical consciousness: the enigma of different paradigms. *Journal of Curriculum Studies*, 51(6), 814-830. <https://doi.org/10.1080/00220272.2019.1652937>
- Guerrero, C., López-García, A. y Monteagudo, J. (2019). Desarrollo del pensamiento histórico en las aulas a través de un programa formativo para Enseñanza Secundaria. *REIFOP: Revista Electrónica Interuniversitaria de Formación del Profesorado*, 22(2), 81-93. <https://doi.org/10.6018/reifop.22.2.363911>.
- Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61(1), 29–48. <https://doi.org/10.1348/000711006X126600>
- Gwet, K. L. (2021). *Handbook of inter-rater reliability*. AgreeStat Analytics.
- Körber, A. (2017). Historical consciousness and the moral dimension. *Historical Encounters: A Journal of historical consciousness, historical cultures, and history education*, 4(1), 81-89. <https://doi.org/10.52289/hej4.100>.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563–575.
- Lévesque, S. (2008). *Thinking historically: Educating students for the 21st century*. University of Toronto Press.
- Lévesque, S. (2023). Narratives of the past. A tool to understand history. In C. J. Gómez (Ed.), *Reimagining the Teaching of European History. Promoting Civic Education and Historical Consciousness* (pp. 53-67). Routledge. <https://doi.org/10.4324/9781003289470>.
- Lévesque, S., & Clark, P. (2018). Historical thinking: Definitions and educational applications. En S. Alan & L. McArthur (Eds.), *The Wiley International Handbook of History Teaching and Learning* (pp. 119-148). Wiley.
- López Facal, R., & Schugurensky, D. (2023). History education and democracy. In C. J. Gómez (Ed.), *Reimagining the Teaching of European History. Promoting Civic Education and Historical Consciousness* (pp. 13-24). Routledge. <https://doi.org/10.4324/9781003289470>
- López-Fernández, C., Tirado-Olivares, S., Minguez-Pardo, R., & Cázar-Gutiérrez, R. (2023). Putting critical thinking at the center of history lessons in primary education through error- and historical thinking-based instruction. *Thinking Skills and Creativity*, 49, 101316 <https://doi.org/10.1016/j.tsc.2023.101316>.
- López-García, A. (2023). Effectiveness of a teaching methodology based on the theory of historical thinking through active methods and digital resources in Spanish adolescents. *Frontiers in Education*, 8, 1175123. <https://doi.org/10.3389/feduc.2023.1175123>.
- Lorentzer, S. (2000). Key Aspects of European Historical Consciousness. In S. Macdonald (Ed.), *Approaches to European Historical Consciousness. Reflections and Provocations* (pp. 32-40). Körber-Stiftung.
- McLean, L. R, Cook, S. A., Lévesque, S., Stanley, T. J., Rogers, P., & Baroud, J. (2017). Introduction: Historical thinking, historical consciousness. *Canadian Journal of Education/Revue Canadienne De l'éducation*, 40(1), 1-5. <https://journals.sfu.ca/cje/index.php/cje-rce/article/view/3122>

- Miguel-Revilla, D. (2020). ¿Cómo evaluar las competencias históricas en la escuela? Desafíos, nuevos marcos teóricos y avances en la evaluación del pensamiento histórico. *Revista San Gregorio*, (40), 132-144. [http://scielo.senescyt.gob.ec/scielo.php?script=sci\\_arttext&pid=S2528-79072020000300132](http://scielo.senescyt.gob.ec/scielo.php?script=sci_arttext&pid=S2528-79072020000300132)
- Miralles, P. y Alfageme, M. B. (2013). Educación, identidad y ciudadanía en un mundo globalizado y posmoderno. *Educatio Siglo XXI*, 31(1), 11-24.
- Miralles, P. y Gómez, C. J. (2016). Sin "Cronos" ni "Kairós". El tiempo histórico en los exámenes de 1.º y 2.º de Educación Secundaria Obligatoria. *Enseñanza de las Ciencias Sociales*, 15, 15-26.
- Molina, S., Felices, M. D. M. y Chaparro, Á. (2016). Diseño y validación de un cuestionario para evaluar la concepción y uso del patrimonio del profesorado de Ciencias Sociales de Educación Secundaria. *Her&Mus. Heritage & Museography*, 17, 149-166, <https://raco.cat/index.php/Hermus/article/view/315307>.
- Moreno, J. R., Rodríguez, R. A. y Monteagudo, J. (2023). Competencias de pensamiento histórico en Bachillerato: análisis de los niveles cognitivos en los criterios de evaluación. *Áreas. Revista Internacional de Ciencias Sociales*, (45), 89-107. <https://doi.org/10.6018/areas.528181>
- Muñoz, S. R., & Bangdiwala, S. (1997). Interpretation of kappa and B statistics measures of agreement. *Journal of Applied Statistics*, 24(1), 105-112. <https://doi.org/10.1080/02664769723918>
- Pagés, J. (2005). Educación cívica, formación política y enseñanza de las ciencias sociales, de la geografía y de la historia. *Iber. Didáctica de las Ciencias Sociales, Geografía e Historia*, (44), 45-55.
- Pedrosa, I., Suárez-Álvarez, & García-Cueto, E. (2013). Content Validity Evidences: Theoretical Advances and Estimation Methods. *Acción Psicológica*, 10(2), 13-18. <http://dx.doi.org/10.5944/ap.10.2.11820>.
- Pérez de Iriarte, J. y Andía, L. (2024). El desarrollo de las competencias de pensamiento histórico en la escuela digital: un caso práctico con estudiantes de Bachillerato. *Didáctica de las ciencias experimentales y sociales*, 46, 37-56. <https://doi.org/10.7203/DCES.46.28470>
- R Core Team. (2021). *R: A language and environment for statistical computing* (Versión, 4.0.4) [Programa informático]. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria. (BOE núm. 76, de 30 de marzo de 2022). <https://www.boe.es/eli/es/rd/2022/03/29/217/con>
- Rodríguez-Medina, J., Gómez-Carrasco, C. J., López-Facal, R. y Miralles-Martínez, P. (2020). Tendencias emergentes en la producción académica de educación histórica. *Revista de Educación*, (389), 211-242. <https://doi.org/10.4438/1988-592X-RE-2020-389-460>
- Rüsen, J. (2004). Historical consciousness: Narrative structure, moral function and ontogenetic development. In P. Seixas, P. (Ed.), *Theorizing historical consciousness* (pp. 63-85). University of Toronto Press.
- Rüsen, J. (2015). *Teoria da História. Uma teoria da história como ciência* (E. C. de Rezende, Trad.). UFPR.
- Sáiz, J. y López-Facal, R. (2015). Competencias y narrativas históricas: el pensamiento histórico de estudiantes y futuros profesores españoles de educación secundaria. *Revista de Estudios Sociales*, (52), 87-101. <http://dx.doi.org/10.7440/res52.2015.06>.
- Seixas, P. (2017). Historical consciousness and historical thinking. In M. Carretero, M. Grever & S. Berger (Eds), *Palgrave handbook of research in historical culture and education* (pp. 59-72). Palgrave McMillan.
- Seixas, P., Gibson, L., & Ercikan, K. (2015). A design process for assessing historical thinking: The case of a one-hour test. In K. Ercikan & P. Seixas (Eds.), *New directions in assessing historical thinking* (pp. 102-116). Routledge.
- Seixas, P., & Morton, T. (2012). *The big six historical thinking concepts*. Nelson Education.
- Shulman, L. S. (2005). Conocimiento y enseñanza: fundamentos de la nueva reforma. Profesorado. *Revista de currículum y formación del profesorado*, 9(2), 1-30.
- Smith, M., Breakstone, J., & Wineburg, S. (2019). History Assessments of Thinking: A Validity Study. *Cognition and Instruction*, 37(1), 118-144. <https://doi.org/10.1080/07370008.2018.1499646>
- Vansledright, B. (2014). *Assessing historical thinking and understanding. Innovation design for new standards*. Routledge.
- Verdú, D. y Gómez, C. J. (2019). Enseñanza de la historia y concepciones identitarias de los maestros en formación. Diseño y validación del cuestionario CIAMEH. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 22(2), 41-65. <http://dx.doi.org/10.6018/reifop.22.2.370011>
- Wertsch, J. V. (2004). Specific narratives and schematic narrative templates. En P. Seixas (Ed.), *Theorizing historical consciousness* (pp. 49-62). University of Toronto Press.

## Anexo 1. Prueba de Evaluación de Competencias Históricas (ECOHI)

### EVALUACIÓN DE COMPETENCIAS HISTÓRICAS

La presente prueba es una actividad voluntaria una vez finalizado el Bachillerato. Los datos extraídos de la misma son confidenciales. Si marcas la casilla con una "X" significa que das tu consentimiento a cumplimentar la prueba. De lo contrario, no la rellenes.

#### DATOS IDENTIFICATIVOS

- Código de la prueba:** \_\_\_\_\_ (No llenar)

#### DATOS SITUACIÓN ACADÉMICA ACTUAL

- Comunidad Autónoma:
- Nombre de la Universidad:
- Titulación:
- Grupo bilingüe (Sí o No):

#### DATOS SITUACIÓN ACADÉMICA ANTERIOR

- Qué estudios cursaste (Bachillerato, Ciclo Formativo de Grado Superior, Prueba de acceso a la universidad para mayores de 25 años, Prueba de acceso a la universidad para mayores de 45 años):
- En caso de Bachillerato, qué tipo realizaste (Investigación, Internacional, Bilingüe, Ordinario, otro):
- En caso de Bachillerato, qué modalidad realizaste (Ciencias y Tecnología, Humanidades y Ciencias Sociales, Artes, General):
- Nota cualitativa en la asignatura de Historia al finalizar el Bachillerato (sobresaliente, notable, aprobado, insuficiente):
- Nota de la prueba EBAU de Historia (sobresaliente, notable, aprobado, insuficiente):
- Sexo (rodea): Mujer / Hombre / Prefiero no responder
- Nivel de estudios de la madre (rodea): 1 sin estudios / 2 estudios básicos (Graduado Escolar, ESO) / 3 estudios secundarios (Bachillerato/Formación Profesional Grado Medio) / 4 estudios superiores (universitarios, Formación Profesional Grado Superior, Danza, Música, otros)
- Nivel de estudios del padre (rodea): 1 sin estudios / 2 estudios básicos (Graduado Escolar, ESO) / 3 estudios secundarios (Bachillerato/Formación Profesional Grado Medio) / 4 estudios superiores (universitarios, Formación Profesional Grado Superior, Danza, Música)



#### FINANCIAS:

Este trabajo forma parte del proyecto de I+D+i *La enseñanza y el aprendizaje de competencias históricas en bachillerato: un reto para lograr una ciudadanía crítica y democrática* (PID2020-113453RB-I00), financiado por MCIN/AEI/10.13039/501100011033.

#### COLABORAN:



#### INSTRUCCIONES:

##### LEE CON ATENCIÓN LAS SIGUIENTES PAUTAS PARA CUMPLIMENTAR CON ÉXITO LA PRUEBA

- Esta prueba sobre competencias históricas forma parte de un estudio de investigación de la Universidad de Murcia orientado a hacer la enseñanza y la evaluación de la asignatura de Historia durante el Bachillerato más interesante y atractiva, favoreciendo así el análisis y el razonamiento del contenido cursado.
- Tus respuestas serán de gran ayuda para los futuros estudiantes, por lo que es muy importante las aportaciones que realices en cada una de las cuestiones que se plantean. Concretamente, la prueba consta de cinco bloques diferentes.
- Encontrarás preguntas con distintos documentos (mapas, imágenes, textos, etc.). Para responderlas solo tienes que fijarte con atención y comprender los documentos que acompañan a estas preguntas.

Tienes una hora para completar las seis partes del cuestionario. ¡GRACIAS POR TU COLABORACIÓN!1.  
**COMPETENCIA HISTÓRICA: RELEVANCIA HISTÓRICA**

- A partir de los siguientes carteles de propaganda responde a las siguientes preguntas.

**Cartel 1**



- 1.1. Identifica a qué bando de la guerra pertenece esta imagen.**

---

---

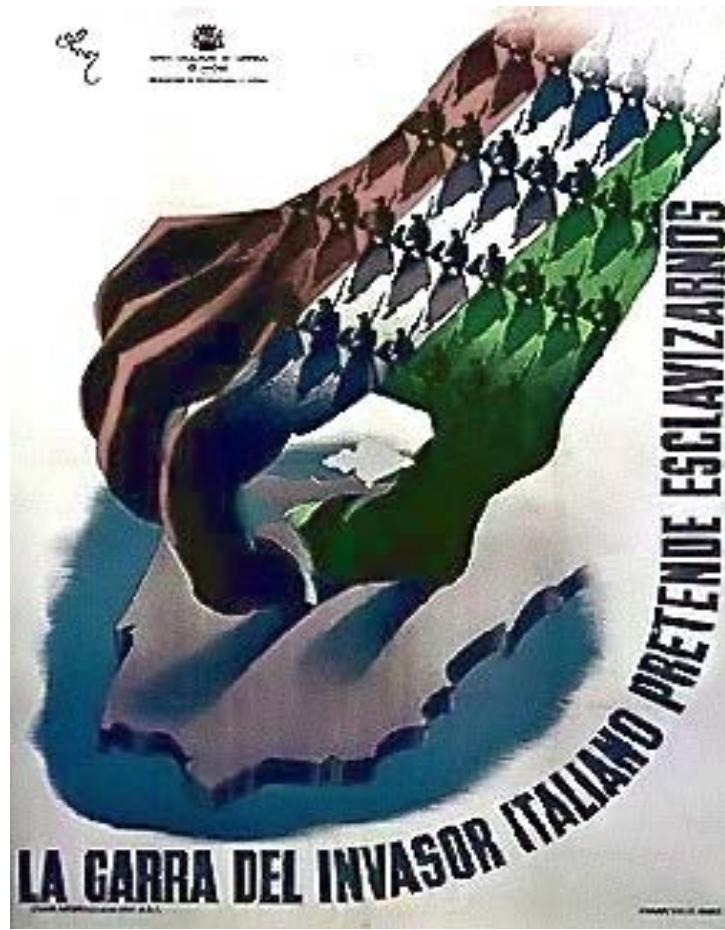
---

- 1.2. Explica el contenido y la ideología que se deducen de este cartel.**

---

---

**Cartel 2**



**1.3. Identifica a qué bando de la guerra pertenece esta imagen.**

---

---

**1.4. Explica el contenido y la ideología que se deducen de este cartel.**

---

---

**1.5. Relaciona la importancia de la cartelería de la época con la concienciación y movilización de cada bando.**

---

---

**2. COMPETENCIA HISTÓRICA: CAUSAS Y CONSECUENCIAS**

- Lee el escrito de Manuel Azaña al presidente de las Cortes de la República (véase DOCUMENTO 3) y expresa sus ideas dando respuesta a las preguntas que se te plantean:

*“Excelentísimo señor: Desde que el general jefe del Estado Mayor Central, director responsable de las operaciones militares, me hizo saber que la guerra estaba perdida para la República, sin remedio alguno, y antes de que, a consecuencia de la derrota, el Gobierno aconsejara y organizara mi salida de España, he cumplido el deber de recomendar y proponer al Gobierno, en la persona de su jefe, el inmediato ajuste de una paz en condiciones humanitarias, para ahorrar a los defensores del régimen y al país entero nuevos y estériles sacrificios [...]”*

Dimisión de Azaña de la presidencia de la República.

Comunicación dirigida al presidente de las Cortes

27 de febrero de 1939

**2.1. Señala las principales razones por las que dimitió Manuel Azaña y las consecuencias que tuvo.**

---



---



---

**2.2. ¿Qué trascendencia tuvo la dimisión de Manuel Azaña para el desarrollo posterior de la sociedad española?**

---



---



---

**3. COMPETENCIA HISTÓRICA: EVIDENCIAS HISTÓRICAS  
DOCUMENTO 4. NIÑOS DE LA GUERRA EXILIADOS 1937**



- 3.1. Identifica el tipo de fuente en los documentos y describe qué representa (personajes, situación política, conflicto...).**
- 3.2. ¿Qué contexto histórico se da, que justifica la salida de los niños/as? Sitúa en el espacio y en el tiempo los hechos representados en la fuente (documento 4).**
- 3.3. ¿Conoces otros escenarios históricos donde los niños/as han sufrido/sufren las consecuencias de conflictos, persecuciones...? Valora y justifica tu respuesta contextualizando situaciones.**

**4. COMPETENCIA HISTÓRICA: PERSPECTIVA HISTÓRICA**

- Lee los tres documentos (véase DOCUMENTOS 5, 6 y 7) sobre la villa de Sepúlveda y marca si las afirmaciones que se encuentran después de dichos documentos son verdaderas (V) o falsas (F):

1. Otorgo a los vecinos de Sepúlveda, presentes y futuros, todo su término municipal [...].
3. Si un forastero hiere o matare en riña a un vecino, pagará el doble de la indemnización determinada por el Fero. Pero si el vecino, defendiéndose en la pelea, matare el forastero, no tendrá que pagar nada [...].
11. Los vecinos de Sepúlveda no pagarán impuesto por sus ganados.
12. Cristianos, moros, judíos, libres o siervos [...] no tendrán que responder por haber incurrido en enemistad, ni por deudas [...], ni por ningunos otros actos llevados a cabo antes de haberse poblado Sepúlveda [...].
22. Os concedo dos ferias. Y el que vaya a estas ferias, ya sea cristiano, moro o judío, irá seguro. Al que matare le enterrarán vivo debajo del muerto, y al que causare lesiones le cortarán la mano. El que robare pagará mil maravedíes de multa al Rey, y si fuere insolvente, le tirarán por las peñas [...].

Alfonso VI a Sepúlveda, 1076 (Adaptación)

La vigésimo segunda campaña de Almanzor fue también contra Sepúlveda, otra vez; emplazó las catapultas contra ella y la combatió noche y día hasta conquistarla por la fuerza. En ella consiguió cautivos y botín en una cantidad incalculable. La destruyó y regresó por Barcelona matando y destruyendo.

*El Qikr bilad al-Andalus, ss. XIV-XV. (Adaptación)*

Sepúlveda [...] es sinónimo de Patrimonio, tanto histórico, artístico, cultural como natural, ya que conserva gran parte de sus edificios construidos desde época medieval [...] Hecho que produjo en 1951 que fuera declarada Conjunto Histórico-Artístico, con la intención de proteger sus valores arquitectónicos y culturales. Entre el caserío sepulvedano sobresalen diferentes casas blasonadas como la Casa de los Proaño, la de los González de Sepúlveda, la de las Conchas o la del Conde. Además, desde el año 2016 este municipio forma parte de la Asociación de los Pueblos más Bonitos de España.

*CyLTV.es (Extracto)*

La situación de la villa de Sepúlveda fue una excepción entre las localidades cristianas del norte peninsular en el siglo XI.

- Verdadero.
- Falso.
- No sabe, no contesta.

Un joven cristiano de 18 años tendría pocos alicientes para ir a vivir a Sepúlveda a fines del siglo XI.

- Verdadero.
- Falso.
- No sabe, no contesta.

La concesión del fuero a Sepúlveda supuso un antes y un después en la historia de la villa.

- Verdadero.
- Falso.
- No sabe, no contesta.

Los beneficios derivados del fuero de Sepúlveda son visibles hoy en día en la villa.

- Verdadero.
- Falso.
- No sabe, no contesta.

## 5. COMPETENCIAS HISTÓRICAS: DIMENSIÓN ÉTICA DE LA HISTORIA Y CAMBIO Y CONTINUIDAD

➤ Lee los discursos de Victoria Kent y Clara Campoamor y responde a las preguntas.

[...] se discute, en este momento, el voto femenino [...] creo que no es el momento de otorgar el voto a la mujer española [...] Y es necesario, señores diputados aplazar el voto femenino, porque yo necesitaría ver, para variar de criterio, a las madres en la calle pidiendo escuelas para sus hijos; yo necesitaría haber visto en la calle a las madres prohibiendo que sus hijos fueran a Marruecos; yo necesitaría ver a las mujeres españolas unidas todas pidiendo lo que es indispensable para la salud y la cultura de sus hijos [...] cuando la mujer española se dé cuenta de que sólo en la República están garantizados los derechos de ciudadanía de sus hijos, de que sólo la República ha traído a su hogar el pan que la monarquía no les había dejado, entonces, señores diputados, la mujer será la más ferviente, la más ardiente defensora de la República [...] Por hoy, señores diputados, es peligroso conceder el voto a la mujer.

*Discurso en el Congreso, 1 de octubre de 1931 (Extracto)*

Respecto a la serie de afirmaciones que se han hecho esta tarde contra el voto de la mujer, he de decir, con toda la cordialidad necesaria, con toda la consideración necesaria, que no están apoyadas en la realidad [...] ¿Cómo puede decirse que cuando las mujeres den señales de vida por la República se les concederá como premio el derecho a votar? ¿Es que no han luchado las mujeres por la República? [...] Y desde el punto de vista práctico, utilitario, ¿de qué acusáis a la mujer? ¿Es de ignorancia? [...] la mujer, hoy día, es menos analfabeta que el varón. [...] he visto que a los actos públicos acudía una concurrencia femenina muy superior a la masculina, y he visto en los ojos de esas mujeres la esperanza de redención, he visto el deseo de ayudar a la República, he visto la pasión y la emoción que ponen en sus ideales. La mujer española espera hoy de la República la redención suya y la redención del hijo.

*Discurso en el Congreso, 1 de octubre de 1931 (Extracto)*

**5.1. ¿Qué posición adopta cada autora con relación al voto femenino? Justifica tu respuesta.**

---



---



---

**5.2. Teniendo en cuenta tus conocimientos sobre el periodo histórico en el que se pronunciaron estos discursos, ¿consideras que el reconocimiento del derecho al voto para las mujeres fue una decisión positiva o negativa para el régimen republicano? Justifica tu respuesta.**

---



---



---

**5.3. Realiza una valoración de las dificultades que tuvieron que superar las personas defensoras del sufragio femenino a inicios del siglo XX y si consideras que suponen un ejemplo a seguir para las reivindicaciones del siglo XXI.**

---



---



---

**5.4. En las siguientes afirmaciones se habla de los cambios en la historia, teniendo en cuenta las dimensiones temporales y la existencia de diferentes ritmos y direcciones en dichos cambios. Marca la opción que más se ajusta al proceso de cambio y continuidad.**

En la España de principios del siglo XX, el debate en el Congreso de los Diputados sobre el sufragio femenino supuso un gran cambio, para alcanzar la igualdad política entre hombres y mujeres.

En la España de principios del siglo XX, el debate en el Congreso de los Diputados sobre el sufragio femenino supuso un gran cambio, fruto de años de reivindicaciones de muchas mujeres (pese a las resistencias).

En la España de principios del siglo XX, el debate en el Congreso de los Diputados sobre el sufragio femenino supuso un gran cambio, pero existieron posturas diferentes por miedo a que peligrase el sistema republicano.

- **A partir de las imágenes, responde a los ejercicios propuestos. La vida diaria en los guetos. En el cartel puede leerse en alemán: Prohibido entrar en la zona residencial de los judíos.**

**Fuente:** <https://atravesdelasescrituras.com/2019/11/20/la-historia-del-holocausto-shoa-%D7%94%D7%A9%D7%95%D7%90%D7%94/>



(DOCUMENTO 11). Sistema de discriminación y control a mujeres palestinas. Fuente: <https://palestinalacrimamea.blogspot.com/2016/08/de-ce-este-israel-un-stat-de-apartheid.html>



(DOCUMENTO 12). Fotografía de Elliot Erwitt en la que se muestra el sistema de segregación racial en Estados Unidos conocido como “separados pero iguales”. Fuente: National Science and Media Museum.



(DOCUMENTO 13). El jugador de fútbol Vinicius Jr. encarándose con un aficionado por insultos racistas. Fuente: <https://contilnetnoticias.com.br>



**6.1. ¿Consideras que la persecución, represión y faltas de respeto efectuadas en el pasado, por las diferencias en el color de piel, de religión, sexual, género... siguen presentes en la sociedad de hoy?**

Siempre o casi siempre.

Nunca o casi nunca.

Algunas veces.

Muchas veces.

**Justifica tu respuesta**

---

---

---

**6.2. Marca cuál de las siguientes acciones consideras la más importante para la convivencia en una sociedad democrática, cívica y crítica:**

Analizar comportamientos discriminatorios en el pasado teniendo en cuenta valores democráticos que se asumen en la sociedad actual: respeto, convivencia, multiculturalidad.

Analizar y valorar problemas actuales de convivencia y/o discriminación de personas o colectivos relacionándoles con antecedentes o situaciones similares producidas en el pasado.

Identificar en el pasado la génesis u orígenes de problemas actuales.

**6.3. Marca cuál de las siguientes acciones crees que ayuda mejor a conectar el pasado y el presente:**

Imaginar posibles futuros sabiendo que el pasado ha dejado su huella en el presente en la sociedad actual.

Tener una perspectiva histórica que ayude a la comprensión de las circunstancias actuales.

Identificar los aspectos del pasado que influyen en las circunstancias del mundo actual.

**ACABAS DE LLEGAR AL FINAL DE LA PRUEBA ¡MUCHAS GRACIAS POR TU COLABORACIÓN!**