

El instrumento LUDEX para el análisis del aprendizaje a través de videojuegos educativos: Una propuesta mediante un estudio Delphi

The LUDEX instrument for analyzing the learning process through educational video games: A proposal based on a Delphi study

Jairo Melo Sánchez

Universidad de Extremadura. Cáceres, España
jairoms@unex.es

Francisco Ignacio Revuelta Domínguez
Universidad de Extremadura. Cáceres, España
fird@unex.es

Jorge Guerra Antequera
Universidad de Extremadura. Cáceres, España
guerra@unex.es

Resumen

Este artículo presenta la metodología de validación de contenido del instrumento LUDEX, un marco innovador y práctico diseñado para evaluar la efectividad de los videojuegos educativos en el contexto de la educación del siglo XXI. El instrumento se estructura en torno a cinco dimensiones clave: Pertinencia Curricular, Usabilidad Orientada al Aprendizaje, Motivación para el Aprendizaje, Evaluación del Progreso, y Trabajo Colaborativo e Interacción Social. A diferencia de otros enfoques, LUDEX ofrece una evaluación integral que incorpora criterios de accesibilidad, igualdad de género y alineación con los Objetivos de Desarrollo Sostenible (ODS), como la educación de calidad y la reducción de desigualdades. Diseñado para ser implementado por docentes sin necesidad de conocimientos técnicos avanzados, el instrumento facilita la formación de ciudadanos globales en entornos digitalizados. El proceso de validación de contenido se llevó a cabo mediante el método Delphi con la participación de 12 expertos del ámbito académico y profesional. Los resultados muestran altos niveles de consenso (medias $> 3.66/4$) y una fiabilidad global aceptable (Alfa de Cronbach = 0.64) para esta fase inicial, brindando evidencia preliminar de validez. Este enfoque asegura que los videojuegos evaluados contribuyan significativamente al desarrollo de competencias clave.

Palabras clave: Videojuegos educativos, Evaluación de la educación, Tecnología educativa, Método Delphi, Accesibilidad a la educación

Abstract

This article presents the content validation methodology for the LUDEX instrument, an innovative and practical framework designed to assess the effectiveness of educational video games in the context of 21st-century education. The instrument is structured around five key dimensions: Curricular Relevance, Learning-Oriented Usability, Motivation for Learning, Progress Assessment, and Collaborative Work and Social Interaction. Unlike other approaches, LUDEX offers a comprehensive assessment that incorporates criteria relating to accessibility, gender equality and alignment with the Sustainable Development Goals (SDGs), such as quality education and the reduction of inequalities. Designed to be implemented by teachers without the need for advanced technical knowledge, the instrument facilitates the development of global citizens in digitalised environments. The content validation process was carried out using the Delphi method with the participation of 12 experts

from academic and professional fields. The results show high levels of consensus (averages > 3.66/4) and acceptable overall reliability (Cronbach's alpha = 0.64) for this initial phase, providing preliminary evidence of validity. This approach ensures that the video games assessed contribute significantly to the development of key competences.

Key words: Educational video games, Educational assessment, Educational technology, Delphi method, Access to education

1. Introducción

La integración de la tecnología en las instituciones educativas ha transformado las aulas en entornos híbridos, planteando el desafío de evaluar la eficacia pedagógica de los recursos digitales. La adopción de videojuegos en el aula ha estado frecuentemente ligada a la asunción de que su carácter lúdico garantiza por sí mismo el aprendizaje significativo. No obstante, aunque la inmersión y la retroalimentación instantánea constituyen ventajas pedagógicas evidenciadas (Granic et al., 2014), la literatura señala una carencia de instrumentos de evaluación validados. Este vacío metodológico limita la capacidad docente para medir la contribución cognitiva y social de estos recursos de forma estructurada (Marín-Suelves et al., 2022).

El presente estudio aborda esta problemática mediante la validación del instrumento LUDEX. Epistemológicamente, la propuesta se fundamenta en la convergencia de la neuroeducación, el Diseño Universal para el Aprendizaje (DUA) y los Objetivos de Desarrollo Sostenible (ODS). A diferencia de modelos previos, LUDEX se articula como una herramienta de contenido que vincula la arquitectura del juego con procesos de aprendizaje profundo, inclusión y ciudadanía digital (Rose, 2024; PNUD, 2023).

2. Marco Teórico

El presente marco teórico analiza los fundamentos científicos que sustentan el uso de los videojuegos como herramientas de aprendizaje, agrupando las perspectivas neuroeducativas, de diseño inclusivo, de evaluación formativa y de impacto social que estructuran el instrumento LUDEX.

2.1. Fundamentos neuroeducativos, constructivismo y aprendizaje basado en juegos

El Aprendizaje Basado en Juegos (ABJ) se consolida en la actualidad no como una mera técnica de animación en el aula, sino como una apuesta por el constructivismo, donde el juego actúa como motor del aprendizaje por descubrimiento activo (Squire, 2011; Gee, 2003). Autores clásicos como Piaget (1976) y Vygotsky (1978) subrayan la importancia de la acción, la interacción social y la resolución de problemas como base para el desarrollo cognitivo. En este sentido, la eficacia de un videojuego educativo no reside en su fidelidad gráfica, sino en su capacidad para otorgar al estudiante un papel activo y autónomo en el aprendizaje (Oliveira et al., 2023) y fomentar la metacognición, entendida por Flavell (1976) como la capacidad de reflexionar sobre los propios procesos de pensamiento mediante un ciclo continuo de hipótesis, experimentación y ajuste.

Desde la perspectiva neuroeducativa, los videojuegos tienen el potencial de generar un "estado de flujo" que implica plenamente al alumnado, optimizando sus recursos

cognitivos y potenciando funciones ejecutivas clave como la planificación, la toma de decisiones y el control de impulsos (Luna, 2023; León et al., 2025). Sin embargo, este proceso exige gestionar adecuadamente la arquitectura del juego. Siguiendo la Teoría de la Carga Cognitiva (Sweller, 2011), si el estudiante debe invertir demasiada energía mental en comprender mecánicas complejas o interfaces recargadas (carga cognitiva extraña), se reduce drásticamente su capacidad para asimilar el contenido educativo. Por ello, es imperativo superar la paradoja de una gamificación superficial —basada únicamente en puntos o rankings efímeros— y diseñar experiencias donde la tecnología se vuelva "transparente", permitiendo que el esfuerzo cerebral se enfoque en la carga germinal y el aprendizaje profundo.

2.2. Diseño inclusivo, motivación y evaluación formativa en entornos híbridos

Para que el aprendizaje mediado por videojuegos sea verdaderamente efectivo y equitativo, debe sustentarse en la inclusión y la motivación. En este ámbito, la aplicación del Diseño Universal para el Aprendizaje (DUA) resulta innegociable. El diseño de los recursos educativos digitales debe concebirse de manera universal desde su origen, garantizando múltiples formas de representación y asegurando que la interfaz no suponga una barrera cognitiva, motora o sensorial (Rose, 2024; Del Puerto, 2025; Hernán y Vásquez, 2022).

A la par de la accesibilidad, el compromiso del estudiante (engagement) actúa como motor del aprendizaje. La Teoría de la Autodeterminación (Ryan y Deci, 2017) postula que el compromiso real es multidimensional e incluye componentes cognitivos (concentración), emocionales (disfrute) y conductuales (persistencia ante el fracaso). La capacidad narrativa de los videojuegos permite a los estudiantes conectar emocionalmente con la experiencia y desarrollar empatía (Isbister, 2016), lo cual, sumado a una evaluación formativa y retroalimentación constante (feedback en tiempo real), es crucial para la autorregulación y la mejora del rendimiento académico (Hattie y Timperley, 2007; Black et al., 2003; Hernández et al., 2021). Finalmente, los videojuegos transforman el aula creando un entorno de aprendizaje híbrido (Potter y McDougall, 2017), donde la experiencia trasciende la pantalla para fomentar la interacción social, la resolución pacífica de conflictos y el debate colaborativo.

2.3. Trazabilidad teórica: Competencias clave, ODS y el instrumento LUDEX

A pesar de la sólida base teórica existente, el profesorado carece frecuentemente de rúbricas operativas y fundamentadas para la realidad del aula. LUDEX se concibe como una respuesta a esta brecha metodológica: un instrumento que conecta la teoría educativa con la práctica, alineándose explícitamente con el desarrollo de competencias clave para el aprendizaje permanente (Unión Europea, 2006) y con la Agenda 2030 (UNDP, 2023). A diferencia de otros marcos de evaluación (Connolly et al., 2012), LUDEX evalúa de forma directa cómo los videojuegos apoyan la educación de calidad (ODS 4), la igualdad de género a través de narrativas no estereotipadas (ODS 5), la reducción de desigualdades mediante la accesibilidad (ODS 10) y la promoción de sociedades pacíficas e inclusivas a través de la cooperación social (ODS 16).

Para garantizar la trazabilidad exigida entre estos marcos epistemológicos complejos y la herramienta propuesta, la Tabla 1 sintetiza cómo la fundamentación teórica se operacionaliza en las cinco dimensiones de LUDEX y en sus indicadores de evaluación directa.

Tabla 1.

Contenido teórico, dimensión LUDEX e indicadores

Fundamentación teórica	Dimensión LUDEX	Indicador	Evidencia observable
Constructivismo y Metacognición	Pertinencia Curricular	Reflexión sobre el aprendizaje	Existencia de diarios, preguntas de reflexión o mecanismos de autoevaluación
DUA	Usabilidad Orientada al Aprendizaje	Accesibilidad	Opciones de personalización, subtítulos, ayudas visuales o auditivas
Autodeterminación	Motivación para el Aprendizaje	Agencia del estudiante	Capacidad de elegir rutas, estrategias o decisiones relevantes
Evaluación Formativa	Evaluación del Progreso	Calidad del feedback	Feedback específico, comprensible y orientado a la mejora
Aprendizaje Sociocultural	Trabajo Colaborativo e Interacción Social	Empatía y cooperación	Mecánicas que exigen coordinación y toma de decisiones conjunta

3. Método

El método Delphi se emplea en esta investigación como estrategia de validación de contenido del instrumento LUDEX, a través del juicio estructurado de expertos. Su objetivo principal es alcanzar un alto nivel de consenso sobre la pertinencia, claridad y relevancia de los indicadores que conforman el instrumento. De este modo, se busca asegurar que LUDEX mida de forma coherente las dimensiones propuestas y que su aplicación resulte válida en contextos educativos reales.

3.1. Un estudio descriptivo y de validación de contenido

Esta investigación adopta un diseño descriptivo orientado a la validación de contenido y contenido del instrumento LUDEX. Dado que el objetivo es establecer la pertinencia y relevancia de las dimensiones que componen el marco de evaluación, se optó por el método Delphi como estrategia de validación. Esta técnica es especialmente adecuada en ámbitos donde no se busca una validación psicométrica basada en el rendimiento de sujetos, sino la validación de un contenido teórico a través del consenso de expertos (Cabero-Almenara y Llorente-Cejudo, 2013). El hilo conductor de la validación reside en

la convergencia de juicios expertos sobre la representatividad de los ítems respecto a las dimensiones teóricas propuestas.

3.2. El método Delphi como herramienta de validación

La validación de contenido se llevó a cabo mediante el método Delphi. Este procedimiento constituye una técnica estructurada de consulta a expertos que busca alcanzar consenso a través de varias rondas de evaluación sucesiva (Reguant-Álvarez y Torrado-Fonseca, 2016). Su principal fortaleza radica en el anonimato de las respuestas, lo que reduce la influencia de figuras dominantes y favorece la independencia de los juicios emitidos por los participantes. El proceso se desarrolló como un “diálogo a ciegas” en tres fases:

1. **Preparación:** Se definió el objeto de la consulta y se elaboró la primera versión de LUDEX.
2. **Rondas de consulta:** Los expertos analizaron la propuesta, puntuando cada ítem y, lo más importante, proporcionando retroalimentación cualitativa.
3. **Consenso:** Se ajustaron los componentes del instrumento hasta alcanzar un acuerdo generalizado sobre su solidez.

3.3. Participantes y criterios de selección

El panel de validación estuvo compuesto por 12 expertos seleccionados mediante un procedimiento de muestreo intencional no probabilístico. Para garantizar la idoneidad del panel, se establecieron criterios de inclusión estrictos: poseer un mínimo de 10 años de experiencia profesional comprobable y mantener una actividad laboral vigente en áreas de especialización vinculadas a la tecnología educativa, el diseño de software o la pedagogía aplicada.

La composición final del panel integró tres perfiles complementarios: 6 doctores e investigadores universitarios (especializados en tecnología educativa y neuroeducación), 3 expertos profesionales en el diseño y desarrollo de videojuegos, y 3 docentes en activo (en las etapas de educación primaria y secundaria). Esta triangulación de perfiles se justifica en la necesidad de evaluar el instrumento desde una perspectiva integral: el rigor científico-pedagógico, la viabilidad técnica de los indicadores y su aplicabilidad real y directa en el entorno de aula.

3.4. Procedimiento Delphi y evolución del instrumento

La validación no se limitó a una consulta aislada, sino que se estructuró a través del método Delphi mediante dos rondas reales y sucesivas de evaluación. El umbral estadístico de consenso se definió a priori estableciendo que un ítem sería aceptado si alcanzaba un nivel de acuerdo igual o superior al 75% entre todos los expertos.

En la primera ronda, los jueces evaluaron la versión inicial de cinco dimensiones. Se les solicitó puntuar cuantitativamente la claridad, pertinencia y relevancia de los ítems

mediante una escala Likert (1 a 4) y aportar retroalimentación cualitativa abierta. El análisis de esta fase evidenció que, aunque el contenido teórico era pertinente, los enunciados poseían un nivel de abstracción que dificultaba su aplicación directa por parte del docente.

A partir de esta retroalimentación, se aplicaron criterios de reformulación orientados a transformar conceptos teóricos en indicadores operativos observables, lo que derivó en las siguientes modificaciones sustanciales retenidas para el instrumento: 1) Se incorporó una "fase cero" de contextualización obligatoria para definir el nivel madurativo (código PEGI) y los requisitos técnicos previos; 2) Se renombraron las dimensiones para reflejar un enfoque más neuroeducativo y menos técnico; 3) Se reformuló la dimensión de retroalimentación para exigir que el feedback del juego fuera explicativo y orientador, no solo correctivo; y 4) Se integraron criterios de accesibilidad para la interacción social.

En la segunda ronda, el panel revisó la rúbrica reestructurada con los nuevos ítems operativos. Tras esta iteración, todas las dimensiones superaron el umbral de consenso del 75% previamente definido, consolidándose así la versión definitiva del instrumento LUDEX sin necesidad de rondas adicionales.

3.5. Criterios de reformulación y versión final del instrumento

Durante la primera ronda de evaluación, el análisis cualitativo evidenció que los enunciados originales poseían un nivel de abstracción que dificultaba su aplicación directa en el aula. Con base en la retroalimentación del panel, se aplicaron criterios de reformulación orientados a transformar conceptos teóricos en indicadores operativos observables.

Las principales modificaciones estructurales retenidas fueron: 1) Se incorporó una "fase cero" obligatoria de contextualización para definir el nivel madurativo (código PEGI) y los requisitos técnicos previos; 2) Se unificó la nomenclatura de las dimensiones hacia un enfoque neuroeducativo (por ejemplo, pasando a denominarse "Pertinencia Curricular" o "Motivación para el Aprendizaje"); 3) Se desglosó el criterio de evaluación para exigir que la retroalimentación del juego fuera explicativa y orientadora; y 4) Se incluyeron criterios de accesibilidad social para proteger a los estudiantes con perfiles introvertidos. Estas modificaciones permitieron que, en la segunda ronda, todas las dimensiones superaran el umbral de consenso preestablecido.

3.6. Análisis de datos

Para el tratamiento cuantitativo de las valoraciones emitidas por el panel Delphi, se diseñó una estrategia de análisis estructurada en tres niveles metodológicos: estadística descriptiva, análisis de consistencia interna e inferencia de magnitud.

En el primer nivel, la tendencia central y la dispersión de los datos se evaluaron mediante el cálculo de la media, la mediana, la desviación estándar y el rango intercuartílico (Q1-Q3). El objetivo de estas medidas fue determinar la concentración de las puntuaciones en la zona superior de la escala Likert (1-4). Paralelamente, el nivel de consenso absoluto se definió a priori como un porcentaje de acuerdo interjueces igual o superior al 75% en

todas las dimensiones; este umbral operó como criterio de corte para la retención estructural de los ítems en la versión definitiva del instrumento.

En el segundo nivel, relativo a la fiabilidad y concordancia, se aplicó el estadístico Alfa de Cronbach (tanto a nivel global como desglosado por subdimensiones) para estimar la consistencia interna inicial del contenido. Para analizar el grado de acuerdo más allá del azar, se calculó el Kappa de Fleiss (indicado para múltiples evaluadores) y el Kappa de Cohen promedio. La literatura metodológica advierte que, en procesos Delphi donde se espera una alta convergencia de expertos hacia la máxima puntuación, estos estadísticos pueden verse afectados por la reducción drástica de la varianza marginal, un fenómeno conocido metodológicamente como "techo de puntuación".

En el tercer nivel, para compensar dicha limitación matemática y estimar la verdadera direccionalidad de las valoraciones, se calculó el tamaño del efecto mediante la d de Cohen. Es imperativo precisar que este estadístico no se empleó en ningún caso como un índice de concordancia interevaluador. Su aplicación se justificó estrictamente como una medida de magnitud para contrastar las puntuaciones medias otorgadas por el panel frente al valor neutro teórico de la escala (2.5), permitiendo así determinar si las altas valoraciones se alejan significativamente de una distribución aleatoria.

3.7. Ética

La validación de un instrumento que aborda valores y ODS requirió un enfoque ético riguroso. Todo el proceso Delphi se llevó a cabo garantizando el anonimato de los expertos tal y como señalan autores como Israel y Hay (2006) y Santi (2016) se trata de garantizar la protección de las personas participantes, el uso honesto de los datos, la transparencia metodológica y la prevención de prácticas fraudulentas, así como las orientaciones del Comité de Bioética y Bioseguridad de la Universidad de Extremadura para los estudios en el ámbito de las Ciencias Sociales. Se utilizó una plataforma cifrada y todos los expertos firmaron un consentimiento informado. Este procedimiento aseguró que las opiniones fueran libres y sinceras, y la integridad de la investigación confiere valor a los resultados presentados a continuación.

4. Resultados

El análisis de los estadísticos descriptivos refleja una convergencia favorable en las valoraciones del panel de expertos. En una escala de 1 a 4, la media global obtenida fue de 3.66, un valor que apunta hacia una pertinencia inicial adecuada del instrumento. Este resultado se contextualiza con desviaciones estándar que oscilan entre 0.50 y 0.66; esta variabilidad sugiere que, si bien existen matices y diversidad de criterio propios de un panel multidisciplinar, se alcanza un nivel de acuerdo razonable para esta fase de estudio.

El nivel de consenso se corrobora mediante los porcentajes de acuerdo interjueces, que superan el 80% en las dimensiones de "Evaluación del Progreso" y "Trabajo Colaborativo e Interacción Social". Por su parte, la fiabilidad global (α de Cronbach = 0.64) muestra una consistencia interna moderada, considerada asumible en aproximaciones exploratorias. En conjunto, estas métricas aportan evidencia preliminar

de validez de contenido para el instrumento LUDEX, sugiriendo su viabilidad teórica para la evaluación de videojuegos educativos a la espera de futuras comprobaciones empíricas.

Tabla 2.

Estadísticos Descriptivos y Consenso por Dimensión

Dimensión	Media	Mediana	Desv. Est.	Q1	Q3	Consenso (%)
Pertinencia Curricular	3.75	4.0	0.50	3.67	4.0	77.78%
Usabilidad Orientada al Aprendizaje	3.67	4.0	0.66	3.67	4.0	77.78%
Motivación para el Aprendizaje	3.69	4.0	0.60	3.67	4.0	77.78%
Evaluación del Progreso	3.75	4.0	0.56	3.92	4.0	80.56%
Trabajo Colaborativo e Interacción Social	3.72	4.0	0.59	3.67	4.0	80.56%

Nota: Los cuartiles Q1 y Q3 oscilaron mayoritariamente entre 3.67 y 4.0 para todas las dimensiones, concentrándose las valoraciones en la parte adecuada de la escala.

4.2. Análisis de Fiabilidad, Concordancia y Tamaño del Efecto

El análisis psicométrico preliminar de las valoraciones expertas revela un escenario característico de los instrumentos multidimensionales en fase exploratoria. El análisis de fiabilidad arrojó un Alfa de Cronbach global de 0.64, un valor que representa una consistencia interna inicial aceptable y sugiere que, en su conjunto, el contenido teórico LUDEX es coherente.

Sin embargo, el análisis desglosado evidenció una marcada variabilidad interna, arrojando coeficientes bajos en dimensiones específicas: 0.11 en Usabilidad Orientada al Aprendizaje, 0.37 en Evaluación del Progreso y 0.45 en Trabajo Colaborativo e Interacción Social. Estos índices denotan que los jueces, dada su naturaleza interdisciplinar (pedagogos, diseñadores y neuroeducadores), interpretaron los ítems de estas subdimensiones con un rango de dispersión mayor. Esta limitación estadística obliga a interpretar la fiabilidad del cuestionario con extrema cautela y certifica la necesidad ineludible de someter el instrumento a futuros análisis factoriales (exploratorios y confirmatorios) utilizando muestras empíricas amplias de usuarios finales para depurar su estructura interna.

En cuanto al análisis de acuerdo interevaluador, el coeficiente Kappa de Fleiss arrojó valores próximos a cero (oscilando entre -0.080 y 0.020). Tal como se anticipó en el diseño metodológico, este resultado no refleja una falta de consenso real, sino que es consecuencia directa de la "paradoja del Kappa" ante una distribución altamente homogénea. Al concentrarse la inmensa mayoría de las valoraciones en las puntuaciones máximas (3 y 4), se genera un techo de puntuación que anula la variabilidad marginal;

sin esta varianza, el estadístico es incapaz de reflejar matemáticamente una concordancia que supere el azar, subestimando artificialmente el nivel de acuerdo.

Precisamente para complementar este vacío interpretativo, la magnitud de las valoraciones se contrastó a través del tamaño del efecto (*d* de Cohen). Los resultados obtenidos fueron marcadamente elevados en todas las dimensiones, con valores que oscilaron entre 1.76 y 2.49 (véase Tabla 3). Esta magnitud tan pronunciada confirma que las puntuaciones emitidas por el panel de expertos se distancian estadísticamente del punto neutro de la escala. En consecuencia, se evidencia una direccionalidad sólida y consistente hacia la validación positiva de los indicadores, demostrando que el alto consenso descriptivo no es producto de una distribución azarosa.

Tabla 3.
Análisis de Concordancia, Fiabilidad y Tamaño del Efecto

Dimensión	Kappa de Fleiss	Kappa de Cohen (Avg)	Alfa de Cronbach	d de Cohen
Pertinencia Curricular	-0.034	0.341	0.60	2.49
Usabilidad Orientada al Aprendizaje	0.005	0.185	0.11	1.76
Motivación para el Aprendizaje	0.020	0.366	0.65	1.98
Evaluación del Progreso	-0.080	0.292	0.37	2.23
Trabajo Colaborativo e Interacción Social	-0.020	0.330	0.45	2.08

Nota: El Alfa de Cronbach Global del instrumento es de 0.64, lo que se considera aceptable para escalas en desarrollo temprano. Los valores de *d* de Cohen (Lambda) indican un tamaño del efecto muy alto, sugiriendo que las puntuaciones se alejan significativamente del valor neutro hacia la pertinencia.

4.3. Retroalimentación cualitativa y refinamiento del instrumento LUDEX

Más allá de los índices de concordancia cuantitativa, el análisis de contenido de las observaciones abiertas formuladas por el panel de expertos en la primera ronda Delphi resultó determinante para la validez de contenido del instrumento. Esta retroalimentación evidenció que la versión preliminar presentaba un grado de abstracción teórica que limitaba su fiabilidad empírica. En consecuencia, se procedió a la operacionalización de las dimensiones, transformando contenidos latentes en indicadores de evaluación directos.

El primer ajuste estructural de calado fue la incorporación de una fase preliminar de contextualización. Los expertos advirtieron sobre el sesgo de evaluar la herramienta desvinculada de su ecosistema de implementación. Por ello, se estableció como prerrequisito que el evaluador defina la madurez evolutiva del público objetivo (mediante el sistema PEGI) y la alineación directa con los estándares de aprendizaje de la asignatura.

A nivel dimensional, el refinamiento cualitativo introdujo las siguientes modificaciones sustanciales:

- **Pertinencia Curricular:** Se sustituyó la nomenclatura original ("Lógica de Aprendizaje") para adoptar un enfoque estrictamente neuroeducativo. Los jueces subrayaron que la estimulación cognitiva no debía medirse de forma binaria, sino a través de la progresión de la dificultad, exigiendo que los retos escalen de forma coherente con la zona de desarrollo próximo del estudiante. Asimismo, se incluyó la exigencia de medir mecanismos explícitos de transferencia del conocimiento a contextos reales.
- **Usabilidad Orientada al Aprendizaje:** Para evitar el solapamiento conceptual entre la arquitectura del software y el proceso cognitivo del sujeto, se redefinió la dimensión. Se incorporó la "retroalimentación del sistema" como indicador clave de usabilidad, evaluando la claridad con la que la interfaz comunica la progresión y el error. Adicionalmente, se exigió incluir criterios de adaptación dinámica al ritmo de aprendizaje como mecanismo de diseño inclusivo.
- **Motivación para el Aprendizaje:** El panel advirtió sobre el riesgo reduccionista de equiparar el engagement (compromiso) con el tiempo de exposición al juego. Atendiendo a la Teoría de la Autodeterminación, la validación del contenido se amplió hacia una taxonomía multidimensional que evalúa el compromiso cognitivo (concentración), emocional (disfrute) y conductual (persistencia ante el reto). Todo ello, mediado por el fomento de la agencia del estudiante en la toma de decisiones.
- **Evaluación del Progreso:** Esta dimensión experimentó una reestructuración crítica. Se introdujo el concepto de pertinencia temporal del feedback, estipulando que la evaluación formativa debe integrarse orgánicamente en la interfaz sin interrumpir el "estado de flujo" (inmersión) del jugador. Asimismo, se refinó el indicador de retroalimentación para exigir que esta trascienda el nivel correctivo (señalar el error) y alcance un nivel explicativo y orientador.
- **Trabajo Colaborativo e Interacción Social:** Las observaciones del panel obligaron a desligar esta dimensión de las mecánicas puramente multijugador. Se reenfocó la evaluación hacia la capacidad del recurso para generar una experiencia de aprendizaje sociocultural que trascienda la pantalla, fomentando el debate cívico en el ecosistema del aula. Finalmente, se introdujo un criterio de accesibilidad social para garantizar que las dinámicas de competición no resulten excluyentes o generen ansiedad en estudiantes con perfiles de interacción introvertidos.

En síntesis, la fase cualitativa del método Delphi garantizó la evolución de LUDEX hacia una rúbrica sólidamente estructurada. La rigurosa operacionalización de sus indicadores asegura que el instrumento pueda ser aplicado por el profesorado con plenas garantías metodológicas, minimizando la variabilidad en la interpretación y asegurando la medición precisa del potencial educativo del software.

5. Discusión

La validación del instrumento LUDEX mediante el método Delphi permite situar esta propuesta dentro del panorama actual de la evaluación de tecnología educativa. El análisis de los resultados, y especialmente la retroalimentación cualitativa aportada por el panel de expertos, proporciona claves fundamentales para comprender cómo debe estructurarse la evaluación de un videojuego para que trascienda lo puramente lúdico y se integre de forma efectiva en el currículo.

5.1. Contextualización y operacionalización frente a modelos previos

Al contrastar LUDEX con marcos de evaluación preexistentes, se observa una evolución significativa en la concepción del recurso. Modelos consolidados como RETA (Marín-Suelves et al., 2022) o E-GAMES (Oliveira et al., 2023) ponen el acento mayoritariamente en la calidad del software, la arquitectura multimedia y la jugabilidad intrínseca. Sin embargo, las aportaciones del panel Delphi revelaron una premisa pedagógica ineludible: la evaluación de un videojuego no puede realizarse en el vacío.

Como señalaron los expertos en la primera ronda de consulta, asumir que un videojuego es intrínsecamente educativo constituye un error de base; su validez depende estrictamente de su contextualización. Esta observación justificó la inclusión de una "fase cero" en el instrumento LUDEX, exigiendo al docente la definición previa del nivel madurativo del alumnado (mediante el código PEGI) y la alineación directa con los estándares de la asignatura. Así, el instrumento transforma conceptos abstractos en indicadores operativos, permitiendo evaluar el videojuego no solo como un producto de consumo, sino como un recurso pedagógico situado.

5.2. Neuroeducación, estado de flujo y el rediseño de la evaluación formativa

Uno de los debates más enriquecedores suscitados por el panel de validación giró en torno a la gestión cognitiva y emocional del aprendizaje. Los expertos advirtieron sobre el riesgo de mantener una visión superficial de la gamificación basada únicamente en la retención temporal del jugador o en recompensas extrínsecas. En respuesta a ello, la dimensión de *Motivación para el Aprendizaje* se reestructuró para abordar el compromiso (*engagement*) desde una taxonomía multidimensional, tal como propone la Teoría de la Autodeterminación (Ryan y Deci, 2017), midiendo la concentración (cognitiva), el disfrute (emocional) y la persistencia ante el fracaso (conductual).

A nivel neuroeducativo, la discusión cualitativa del panel obligó a replantear drásticamente la dimensión de *Evaluación del Progreso*. Los jueces enfatizaron que un exceso de correcciones puede resultar contraproducente. Si el sistema interrumpe constantemente al jugador para señalar errores, se rompe el "estado de flujo" o inmersión (Luna, 2023). Por tanto, LUDEX incorporó el criterio de "pertinencia temporal", exigiendo que la retroalimentación se entregue en momentos oportunos. Además, los expertos señalaron que el *feedback* no debe limitarse a ser correctivo, sino explicativo y orientador, fomentando la autorregulación y la metacognición sin penalizar la motivación intrínseca del estudiante.

5.3. El ecosistema sociocultural y la inclusión como eje transversal (DUA y ODS)

El panel Delphi también evidenció que la inclusión digital se interpreta frecuentemente de forma parcial, limitándose a la usabilidad del software. Las observaciones cualitativas permitieron llevar el Diseño Universal para el Aprendizaje (DUA) un paso más allá. En la dimensión de *Trabajo Colaborativo e Interacción Social*, los jueces alertaron de que las mecánicas multijugador competitivas pueden generar ansiedad o resultar contraproducentes para estudiantes con perfiles sociales introvertidos. Atendiendo a esta crítica, LUDEX integró indicadores específicos para garantizar que el ecosistema social del juego sea seguro y ofrezca alternativas de participación inclusivas.

Asimismo, se concluyó que el impacto de un videojuego educativo debe "trascender la pantalla". El consenso experto determinó que la herramienta alcance su máxima valoración solo cuando fomente dinámicas sociológicas en el ecosistema del aula, induciendo debates cívicos, el desarrollo de la empatía y la resolución pacífica de conflictos. Esta aproximación confirma la alineación de LUDEX con la Agenda 2030, dotando al profesorado de una métrica estructurada para evaluar cómo los recursos digitales contribuyen a la educación de calidad (ODS 4) y a la reducción de las desigualdades (ODS 10).

5.4. Limitaciones del estudio y prospectiva de investigación

A pesar de la solidez teórica alcanzada tras el refinamiento cualitativo y los altos índices de consenso, los hallazgos de este estudio deben interpretarse reconociendo sus limitaciones metodológicas. La validación se ha sustentado en el juicio de 12 expertos, lo cual constituye un marco adecuado para una fase exploratoria de diseño, pero resulta insuficiente para confirmar estadísticamente la dimensionalidad de la validación del contenido.

La presencia de un "techo de puntuación" en las valoraciones redujo la variabilidad de los datos (afectando a estadísticos como el Kappa de Fleiss) y reveló valores bajos en el Alfa de Cronbach para subdimensiones específicas. La ausencia de un pilotaje empírico con usuarios finales, la falta de un análisis factorial confirmatorio y la no aplicación del instrumento en un entorno real de aula constituyen restricciones objetivas. Estas limitaciones impiden aportar evidencias definitivas sobre la estructura interna del cuestionario o su validez externa. En consecuencia, futuras líneas de investigación deberán abordar estas carencias mediante diseños cuasiexperimentales y validaciones psicométricas aplicadas a muestras amplias de docentes en contextos de enseñanza auténticos.

6. Conclusiones

La validación del instrumento LUDEX mediante el método Delphi aporta evidencia preliminar de validez de contenido para la evaluación integral de videojuegos educativos. El proceso iterativo y la consulta cualitativa a expertos han sido determinantes para transformar un marco teórico abstracto en una herramienta operativa, neuroeducativa e inclusiva, articulada en cinco dimensiones: Pertinencia Curricular, Usabilidad Orientada

al Aprendizaje, Motivación para el Aprendizaje, Evaluación del Progreso, y Trabajo Colaborativo e Interacción Social.

Las recomendaciones del panel de jueces permitieron superar enfoques de evaluación tradicionalmente tecnocéntricos. La incorporación de la fase de contextualización previa, la exigencia de un feedback explicativo que respete la inmersión del estudiante y la consideración de la accesibilidad para perfiles sociales introvertidos dotan a LUDEX de una profunda pertinencia pedagógica. Esta estructura proporciona al profesorado un andamiaje analítico claro para decidir, de manera fundamentada, qué recursos digitales fomentan verdaderamente competencias clave, autorregulación y valores ciudadanos.

Si bien los índices de acuerdo y el tamaño del efecto (d de Cohen) respaldan la coherencia teórica de los indicadores propuestos, la naturaleza exploratoria del estudio obliga a mantener una interpretación rigurosamente cautelosa. Afirmar en esta etapa que la herramienta posee una estructura metodológicamente inamovible excedería el alcance de los datos, dado que el diseño metodológico actual carece de análisis factorial y de implementación inmersiva.

En conclusión, los resultados posicionan a LUDEX como un marco de referencia sólidamente estructurado desde la perspectiva teórica y experta, pero se subraya la estricta necesidad de validación empírica posterior. Su futura aplicación práctica en aulas reales permitirá contrastar empíricamente su fiabilidad psicométrica y su eficacia definitiva como instrumento de evaluación en el marco de la educación del siglo XXI.

Presentación del artículo: 22 de abril de 2026
Fecha de aprobación: 16 de junio de 2026
Fecha de publicación: 1 de julio de 2026

Melo-Sánchez, J., Revuelta-Domínguez, F. I., & Guerra-Antequera, J. (2026). El instrumento LUDEX para el análisis del aprendizaje a través de videojuegos educativos: Una propuesta mediante un estudio Delphi. <i>Revista de Educación a Distancia</i> , 26(84). http://dx.doi.org/10.6018/red.710781

Financiación

Ayudas para la contratación de personal investigador predoctoral en formación”, con número de expediente 5939110060220230000651 y PD23038, financiado por la Junta de Extremadura, la UE, el Ministerio de Trabajo y Economía Social y los Fondos Europeos.

Declaración de las personas autoras sobre el uso de LLM

Este artículo no ha utilizado textos provenientes (o generados) de un LLM (ChatGPT u otros) para su redacción.

Declaración de las contribuciones de las personas autoras

Conceptualización, J.M.S. y F.I.R.D.; curación de datos, J.M.S.; software, J.M.S.; análisis formal, J.M.S., F.I.R.D. y J.G.A.; obtención de fondos, J.M.S.; investigación, J.M.S.; metodología, J.M.S.; administración del proyecto, J.M.S.; recursos, J.M.S.; supervisión, J.M.S., F.I.R.D. y J.G.A.; validación, J.M.S.; visualización, J.M.S.; redacción – borrador original, F.I.R.D. y J.M.S.; redacción – revisión y edición, J.G.A. y J.M.S.

Referencias

- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for Learning: Putting it into practice*. Open University Press. <http://www.mcgraw-hill.co.uk/html/0335212972.html>
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79-122. <https://doi.org/10.3102/0034654315582065>
- Cobos, J. C. (2022). El uso de la gamificación para aumentar la participación y el compromiso estudiantil. *Nexus Research Journal*, 1(1), 34-42. <https://doi.org/10.62943/nrj.v1n1.2022.5>
- Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2), 661-686. <https://doi.org/10.1016/j.compedu.2012.03.004>
- Del Puerto, D. A. (2025). El Diseño Universal para el Aprendizaje como marco para analizar la accesibilidad de los Recursos Educativos Abiertos. *EduTec*. <https://edutec.es/revista/index.php/edutec-e/article/view/3909>
- Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American Psychologist*, 69(1), 66-78. <https://doi.org/10.1037/a0034857>
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences*, 3025-3034. <https://doi.org/10.1109/HICSS.2014.377>
- Hernán, C., & Vásquez, M. (2022). Análisis de la accesibilidad de los recursos educativos digitales: Una revisión sistemática. *Revista Científica Kosmos*, 1(1), 27-38. <https://doi.org/10.62943/rck.v1n1.2022.37>
- Hernández, V. M., Santana, P. J., & Sosa, J. J. (2021). Feedback y autorregulación del aprendizaje en educación superior. *Revista de Investigación Educativa*, 39(1), 227-248. <https://doi.org/10.6018/rie.423341>
- Israel, M., & Hay, I. (2006). *Research ethics for social scientists: Between ethical conduct and regulatory compliance*. SAGE Publications. <https://doi.org/10.4135/9781849209779>

- Isbister, K. (2016). *How games move us: Emotion by design*. The MIT Press.
<https://doi.org/10.7551/mitpress/9267.001.0001>
- L León, M. M. C., Morales, M. E. C., Jácome, B. E. V., et al. (2025). Estrategias didácticas basadas en la neuroeducación para mejorar la atención y memoria en estudiantes de educación básica. *SAGA: Revista Científica*, 102.
<https://revistasaga.org/index.php/saga/article/view/102>
- Luna, M. A. (2023). Análisis sobre funciones ejecutivas desarrolladas en los videojuegos. *Psicomotricidad, Movimiento y Emoción*, 9(1), 1-15.
- Marín-Suelves, D., Esnaola-Horacek, G., & Cassany, D. (2022). Videojuegos y educación: Análisis de tendencias en investigación. *Revista Colombiana de Educación*, (84), 209-228. <https://doi.org/10.17227/rce.num84-12125>
- Medina, P. A., Pilatasig, M. R., Ibáñez, J. E., et al. (2024). Evaluación formativa digital: Herramientas y técnicas para mejorar el feedback y el aprendizaje continuo. *Ciencia Latina Revista Científica Multidisciplinar*, 8(4), 10052-10073.
https://doi.org/10.37811/cl_rcm.v8i4.13115
- Mendoza, O., Cedeño, F., Verduga, A., & Castillo, L. J. (2024). Evaluación de la enseñanza universitaria mediante el análisis crítico del feedback estudiantil. *Revista de Ciencias Sociales*, 30(Especial 10), 115-129.
- Mishra, P., & Mehta, R. (2017). What we educators get wrong about 21st-century learning: Results of a survey. *Journal of Digital Learning in Teacher Education*, 33(1), 6-19. <https://doi.org/10.1080/21532974.2016.1242392>
- Oliveira, W., Hamari, J., Shi, L., et al. (2023). Tailored gamification in education: A literature review and future agenda. *Education and Information Technologies*, 28(1), 373-406. <https://doi.org/10.1007/s10639-022-11122-4>
- Potter, J., & McDougall, J. (2017). *Digital Media, Culture and Education: Theorising Third Space Literacies*. Palgrave Macmillan.
- Programa de las Naciones Unidas para el Desarrollo [PNUD]. (2023). *Estrategia digital del PNUD 2022-2025*. <https://www.undp.org/es/digital-strategy>
- Reguant-Álvarez, M., & Torrado-Fonseca, M. (2016). El método Delphi. *REIRE Revista d'Innovació i Recerca en Educació*, 9(1), 87-102.
<https://doi.org/10.1344/reire2016.9.1916>
- Rose, D. H. (2024). *Universal design for learning in the classroom: Practical applications for K-12 and beyond* (T. E. Hall, K. H. Robinson, & D. Gordon, Eds.; 2.^a ed.). The Guilford Press.
- Rose, P. (2024). Global citizenship education and teachers missing in action? *International Journal of Educational Development*, 104, Artículo 102941.
<https://doi.org/10.1016/j.ijedudev.2023.102941>
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press.
<https://doi.org/10.1521/978.14625/28806>

- Santi, M. F. (2016). Controversias éticas en torno a la privacidad, la confidencialidad y el anonimato en investigación social. *Revista de Bioética y Derecho*, 37, 5–21. https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1886-58872016000200002
- Squire, K. (2011). *Video games and learning: Teaching and participatory culture in the digital age*. Teachers College Press.
- Sweller, J. (2011). Cognitive Load Theory. En J. P. Mestre & B. H. Ross (Eds.), *The Psychology of Learning and Motivation: Cognition in Education* (Vol. 55, pp. 37-76). Academic Press.
- UNDP. (s. f.). Sustainable Development Goals. Recuperado el 3 de diciembre de 2024 de <https://www.undp.org/sustainable-development-goals>

ANEXOS

Rúbrica LUDEX Orientada a la Validación de Videojuegos

DIMENSIÓN	EXCELENTE (4)	BUENO (3)	SATISFACTORIO (2)	NECESITA MEJORA (1)
1. PERTINENCIA CURRICULAR (L)	Los desafíos y la dificultad escalan de manera progresiva y coherente con el desarrollo cognitivo. Existe una alineación directa con los objetivos curriculares. Exige alta reflexión y ofrece oportunidades claras de transferir el conocimiento a la vida real.	La progresión es adecuada, aunque con saltos bruscos menores. La alineación con el currículo es parcial o implícita. Exige reflexión, pero los mecanismos para transferir el aprendizaje fuera del juego son limitados.	Dificultad poco escalonada (muy fácil o frustrante). La relación con el currículo es débil. Ofrece estímulos cognitivos básicos, basados más en la memorización o mecánicas repetitivas que en la reflexión.	La dificultad es plana o desajustada. No hay alineación curricular. El juego se basa en tareas mecánicas sin reflexión ni transferencia de conocimientos.
2. USABILIDAD ORIENTADA AL APRENDIZAJE (U)	Interfaz totalmente accesible e intuitiva (menús claros, botones visibles). El sistema comunica logros, progreso y errores de forma óptima y accesible . Se adapta dinámicamente a diversas necesidades, ritmos de aprendizaje y perfiles inclusivos.	Interfaz mayormente accesible, con algún elemento confuso. Buena comunicación de progreso/errores. La adaptación al ritmo del estudiante es adecuada, pero no cubre todas las diversidades o necesidades de inclusión.	Interfaz algo confusa que interrumpe la fluidez del aprendizaje. La comunicación de progreso o errores es básica o poco clara. Escasa o nula adaptación a diferentes ritmos de aprendizaje.	Interfaz inaccesible que bloquea la interacción con el juego. No hay retroalimentación del sistema sobre errores o progreso. Cero adaptación dinámica a las necesidades del estudiante.
3. MOTIVACIÓN PARA EL APRENDIZAJE (M)	Mantiene engagement multidimensional (alta concentración, disfrute y persistencia) a lo largo de todas	Mantiene buena motivación e interés, aunque puede decaer en algunos niveles o etapas.	Logra captar la atención solo en momentos puntuales. La agencia del estudiante es muy limitada (el juego es	Pierde rápidamente el interés de los estudiantes. Nula agencia del jugador , lo

	las etapas del juego. Fomenta fuertemente la agencia del estudiante (toma de decisiones), generando gran motivación intrínseca.	Permite cierta agencia al estudiante, pero el avance podría ser un poco más libre o flexible.	demasiado dirigido), lo que reduce el componente emocional y conductual de la motivación.	que genera aburrimiento o reduce el aprendizaje a una obligación sin motivación.
4. EVALUACIÓN DEL PROGRESO (E)	Proporciona feedback explicativo y orientador (por qué se falló y cómo mejorar), entregado en el momento oportuno sin romper la inmersión (estado de flujo). Ofrece tableros de progreso altamente claros, comprensibles y exportables para docentes y estudiantes (metacognición).	El feedback es útil y mayormente claro, pero a veces es solo correctivo (señala el error) o interrumpe levemente el flujo del juego. Los paneles para el docente son adecuados, pero podrían ser más detallados.	Retroalimentación básica, tardía o que interrumpe bruscamente la inmersión del jugador. Los informes de progreso son muy simples, difíciles de interpretar o poco accionables para el docente.	No ofrece retroalimentación útil para mejorar. Los informes de progreso o paneles docentes son inexistentes o los datos extraídos son completamente irrelevantes para el aula.
5. TRABAJO COLABORATIVO E INTERACCIÓN SOCIAL (X)	Fomenta el desarrollo de empatía, toma de perspectiva y negociación. La experiencia trasciende la pantalla , generando debates en el aula. El entorno es seguro e inclusivo para perfiles introvertidos o con dificultades sociales.	Fomenta la interacción social o competencia sana, pero de forma algo superficial. Genera algunas reflexiones en el aula. Es mayormente accesible, aunque ciertas mecánicas pueden incomodar a perfiles más introvertidos.	Interacción social forzada, centrada solo en ganar (colaboración obligada). Rara vez genera debate fuera de la pantalla. Puede resultar contraproducente para estudiantes con dificultades sociales.	Fomenta el individualismo excesivo o una competitividad tóxica. No hay reflexión social en el aula. Mecánicas excluyentes y frustrantes para perfiles introvertidos.