



Instituto Universitario de Ciencias de la Educación (IUCE)
Programa de Doctorado en Formación en la Sociedad del Conocimiento

TESIS DOCTORAL

Diseño de un modelo de adopción tecnológica de dispositivos móviles en evaluación formativa y sumativa

Autor: Alberto Ortiz López

Directores: Susana Olmos Migueláñez y José Carlos Sánchez Prieto

Noviembre de 2023



VNiVERSiDAD D SALAMANCA



TESIS DOCTORAL

Diseño de un modelo de adopción tecnológica de dispositivos móviles en evaluación formativa y sumativa

Doctorando:

Alberto Ortiz López

Directores:

Susana Olmos Migueláñez

José Carlos Sánchez Prieto

VNiVERSiDAD D SALAMANCA

Tesis Doctoral por Compendio de Publicaciones

Autorización de los directores para la presentación de la Tesis Doctoral

La **Dra. Susana Olmos Migueláñez**, Profesora Titular en el Área de Métodos de Investigación y Diagnóstico en Educación, y el **Dr. José Carlos Sánchez Prieto**, Profesor Contratado Doctor en el Área de Métodos de Investigación y Diagnóstico en Educación, ambos pertenecientes al Departamento de Didáctica, Organización y Métodos de Investigación de la Universidad de Salamanca,

Certifican:

Que **D. Alberto Ortiz López** ha realizado la Tesis Doctoral titulada: “**Diseño de un modelo de adopción tecnológica de dispositivos móviles en evaluación formativa y sumativa**” bajo su dirección, y que ésta cumple con los criterios de calidad y originalidad necesarios para optar a la obtención del Grado en Doctor por la Universidad de Salamanca. Además, certifican también que la Tesis se presenta en la modalidad de Compendio de Publicaciones y que el doctorando opta a la mención de Doctor Internacional.

Y para que conste, se firma por los directores de la Tesis Doctoral a todos los efectos oportunos en Salamanca, a 6 de noviembre de 2023.

Susana Olmos Migueláñez

José Carlos Sánchez Prieto

Tesis Doctoral por Compendio de Publicaciones

Publicaciones que avalan el compendio

ANEXO 1

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2021). Calidad en e-learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2). <https://doi.org/10.5944/ried.24.2.29073>

- Revista Q2 – JIF en año de publicación (2021): Factor de impacto 2,689. Posición 111/270. Percentil 59,07.
- Revista Q1 – JCI en año de publicación (2021): Factor de impacto 2,57. Posición 93/743. Percentil 87,55.

ANEXO 2

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-learning en Educación Superior: Una revisión sistemática de la literatura. *Education in the Knowledge Society (EKS)*, 23. <https://doi.org/10.14201/eks.26986>

- Revista Q1 – JCI en año de publicación (2022): Factor de impacto 1,86. Posición 68/742. Percentil 90,90.
- Revista Q2 – SJR último año de índice publicado a fecha de depósito artículo - (2022): Factor de impacto 0,664
- Incluido en índice JIF en 2023, factor de impacto 3,9. Se asignará cuartil en 2024.

ANEXO 3

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12117-5>

- Revista Q1 – JIF en último índice publicado (2022): Factor de impacto 5,5. Posición 19/269. Percentil 93,1
- Revista Q1 – JCI en último índice publicado (2022): Factor de impacto 2,56. Posición 21/742. Percentil 97,24
- Revista Q1 – SJR en último índice publicado (2022): Factor de impacto 1,249.

ANEXO 4

Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (*En prensa*). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. *Journal of New Approaches in Educational Research*.

- Revista Q1 – JCI en último índice publicado (2022): Factor de impacto 2,63. Posición 17/742. Percentil 97,78
- Revista Q1 – SJR en último índice publicado (2022): Factor de impacto 1,034.
- Incluido en índice JIF en 2023, factor de impacto 4,5. Se asignará cuartil en 2024.

ANEXO 5

Ortiz-López, A.; Olmos-Migueláñez, S.; Sánchez-Prieto, J.C. (*En prensa*). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En E. Chiner (Ed.), *Nuevas tendencias interdisciplinarias en Educación y Conocimiento*. Tirant Lo Blanch

- Capítulo de Libro publicado en editorial Q1 Ranking SPI: ICEE

A mis padres, Eloísa y Agustín.

A mi hermana, María.

Y a ti, Sara.

Agradecimientos

A los directores de esta investigación, Susana Olmos Migueláñez y José Carlos Sánchez Prieto, quienes dieron comienzo a este camino y han dirigido una Tesis con cariño, paciencia, dedicación y esfuerzo. Gracias por vuestra confianza.

A Ricardo Pocinho, por hacer de Leiria un lugar donde volver. Gracias por tu implicación, ayuda y compañía en este proceso. Esta Tesis es el comienzo de futuras experiencias compartidas.

Al grupo de investigación GRIAL de la Universidad de Salamanca, por hacerme parte de vuestra (nuestra) familia universitaria. Y en especial a María José, Fran, Lucía, Alicia, Andrea, Sonia... y a todas aquellas personas que han formado, y forman, parte de esta experiencia.

Al Instituto Universitario de Ciencias de la Educación, al Programa de Doctorado Formación en la Sociedad del Conocimiento y a quien a ellos se adscriben: gestores, doctorandos, compañeros y amigos.

A Javier, Miriam y Ana; por vuestra ayuda en los momentos últimos de esta investigación. Y a quienes han sido parte en algún momento durante el desarrollo de esta Tesis. Gracias a los investigadores que han conformado los grupos de expertos. Y gracias también a los alumnos, alumnas y docentes de la Universidad de Salamanca en Ávila, Salamanca y Zamora que han colaborado en el proceso.

A quienes se han sentado en el aula cada día desde el comienzo de mi corta trayectoria docente. Vuestra actitud, confianza y cariño justifica este esfuerzo. Gracias por demostrarme que acompañaros desde el otro lado es una aventura magnífica.

A todos mis compañeros y amigos. En especial, a Jaime, por ser hermano en el camino. Gracias también a Laura, Víctor, Carla, Lau, Iván y Kike; por vuestra paciencia, vuestra compañía y vuestro apoyo continuo.

A mis padres, Eloísa y Agustín, por hacerme ser quien soy. Gracias por vuestras enseñanzas, vuestra educación y vuestros valores. Gracias por hacerme ser constante y por enseñarme a el significado del esfuerzo. Y gracias a mi familia, este camino es también vuestro.

A ti, abuelo, que ya no estás pero que siempre estás. Esta Tesis es también por y para ti.

A mi hermana María. Por enseñarme el sentido del cariño, la amistad y la responsabilidad. Por ser un espejo para recordar el pasado y no olvidar lo recorrido. Por tu acompañamiento en este tiempo. Gracias por estar, siempre.

A Sara. Por dar comienzo y ser impulso en este largo camino. Por todo lo que no se puede reducir a la palabra. Por estar siempre y sin condición. Por tu comprensión infinita. Por hacer de cada éxito una responsabilidad compartida. Qué bien que te pusiste en medio.

Y gracias, por último, a cada persona a quien le debo tiempo, dedicación y momentos compartidos.

En la actualidad, la coexistencia educativo-tecnológica es una realidad en la práctica totalidad de instituciones educativas y escenarios formativos. La transición hacia los entornos virtuales de aprendizaje, impulsada en los últimos años por diferentes políticas educativas e institucionales y acontecimientos como la pandemia, ha derivado en la reubicación de la tecnología como medio principal para el avance de la docencia y como vía de investigación de las innovaciones implementadas en los procesos de enseñanza-aprendizaje.

La Tesis Doctoral que se presenta centra su desarrollo en el papel de los dispositivos móviles en los procesos de evaluación, uno de los factores más rigurosos, objetivos y vinculantes de la calidad educativa. Estos dispositivos abren un nuevo escenario educativo mediante la estimulación de experiencias de aprendizaje y sus posibilidades en lo referente a la disponibilidad de recursos, la interacción ubicua e instantánea y su accesibilidad. La tecnología móvil ha generado unas altas expectativas en la educación digital, replanteándose desde las metodologías de aula hasta las principales políticas legislativas e institucionales; todo respaldado por prácticas formativas y de actualización en materia tecnológica para la adaptación de profesorado y alumnado a los nuevos marcos de enseñanza.

La evaluación no solo se percibe como una herramienta para la medición y rendición de cuentas sobre el progreso del estudiante, sino que adquiere también un carácter procesual en el acompañamiento y la orientación a la mejora en su desarrollo. Por este motivo, adquiere una importancia estratégica abordar la aceptación por parte de docentes en lo referido a sus técnicas, instrumentos y procesos seguidos en los entornos virtuales de aprendizaje. En consecuencia, la tecnología móvil abre una nueva vía en los procesos de evaluación que debe ser abordada desde el estudio previo de la aceptación docente ante su implementación, ya que el conocimiento de los factores condicionantes es un factor fundamental para garantizar una implementación efectiva y eficaz.

La propuesta metodológica que fundamenta la investigación basa su recorrido en el diseño y validación de un nuevo modelo para el estudio de la adopción tecnológica de dispositivos móviles en procesos de evaluación formativa y sumativa. Para definir este modelo, se ha realizado un recorrido conceptual sobre la evaluación del aprendizaje, así como sobre la relación entre evaluación y calidad. En base a esta relación, se ha desarrollado una revisión sistemática de la

literatura sobre su estado actual que ha derivado en la concreción de un nuevo modelo para su estudio en Educación Superior.

Seguidamente, con el objetivo de conocer el estado actual de las investigaciones sobre e-evaluación, la Tesis ha desarrollado un mapeo sistemático de la literatura sobre la temática que ha permitido sintetizar los últimos avances en la materia. Posteriormente, en la combinación de esta variable con los modelos de adopción tecnológica, se ha llevado a cabo una nueva revisión sistemática de la literatura relativa a los modelos de aceptación actuales desarrollados en el campo de la evaluación mediada por dispositivos móviles, la cual ha sentado las bases para el desarrollo de la propuesta realizada y ha mostrado el estado inicial en el que las investigaciones se encuentran.

Una vez finalizada la fase teórica que ha comprendido las revisiones de la literatura referenciadas, se ha desarrollado la propuesta de un nuevo modelo para el estudio de esta aceptación en docentes. Este modelo ha sido sometido a un proceso de validación por jueces expertos, quienes han validado una propuesta que, en fases posteriores, ha sido objeto de aplicación experimental en el estudio intermedio realizado.

El estudio intermedio realizado ha permitido diseñar un modelo de adopción tecnológica con dos versiones diferenciadas: una centrada en la determinación de la intención de uso de dispositivos móviles por parte de docentes en evaluación formativa, y otra orientada a la evaluación sumativa. El resultado final es un modelo compuesto por once constructos y quince hipótesis relacionales que difieren en la finalidad formativa o sumativa y que abren una nueva vía en la investigación y en el desarrollo de los estudios de adopción en el campo de la e-evaluación.

Palabras Clave: Modelos de aceptación tecnológica, evaluación, dispositivos móviles, docentes, calidad, modelos de ecuaciones estructurales.

Abstract

An extended abstract of the Ph.D. Dissertation can be found in **Appendix C (Anexo C)**

The coexistence of education and technology is a reality in almost all educational institutions and learning environments. The transition to virtual learning environments, driven in recent years by different educational and institutional policies and events such as the pandemic, has resulted in the repositioning of technology as the primary means for advancing teaching and as a pathway for investigating innovations implemented in teaching and learning processes.

The presented Doctoral Thesis focuses on the role of mobile devices in assessment, one of the most rigorous, objective, and binding factors for educational quality. These devices have created a new educational scenario by stimulating learning experiences and offering possibilities such as the availability of resources, ubiquitous and instant interaction, and accessibility. Mobile technology has generated high expectations in digital education, prompting a reevaluation of classroom methodologies and key legislative and institutional policies, all supported by training practices and technological updates to adapt teachers and students to new teaching frameworks.

Assessment, therefore, is not only perceived as a tool for measuring and being accountable for student progress but also takes on a procedural character in supporting improvement and guidance. For this reason, it is strategically important to address the technological acceptance by teachers regarding the techniques, tools, and processes used in virtual learning environments. Consequently, mobile technology opens a new avenue in assessment processes that must be approached by studying teachers' acceptance before implementation, as understanding conditioning factors is crucial for ensuring effective and efficient implementation.

The proposed research methodology is grounded in the design and validation of a new model for studying the technological adoption of mobile devices in formative and summative assessment processes. A conceptual journey has been undertaken on learning assessment and its relationship with quality to define this model. Based on this relationship, a systematic literature review has been conducted on its current state, leading to the specification of a new model for its study in higher education.

Subsequently, to understand the current state of research on e-assessment, the thesis has conducted a systematic literature mapping of the topic, synthesizing the latest advances.

Combining this variable with technology adoption models, another systematic literature review has been carried out on current acceptance models developed in the field of mobile-mediated assessment. This review laid the groundwork for the proposed model and revealed the initial state of research in the field.

Upon completing the theoretical phase, which included the referenced literature reviews, a proposal for a new model for studying this acceptance in teachers has been developed. This model has been validated by expert judges, who validated a proposal that, in subsequent phases, underwent experimental application in the intermediate study conducted.

The intermediate study has allowed the design of a technological adoption model with two distinct versions: one focused on determining teachers' intention to use mobile devices in formative assessment, and another oriented towards summative assessment. The result is a model consisting of eleven constructs and fifteen relational hypotheses that differ in formative or summative purposes, opening a new path in research and in the development of adoption studies in e-assessment.

Keywords: Technological acceptance models, assessment, mobile devices, teachers, quality, structural equation models.

EVALUACIÓN DEL APRENDIZAJE, DISPOSITIVOS MÓVILES Y MODELOS DE ADOPCIÓN TECNOLÓGICA	23
1.1. <i>Evaluación del aprendizaje: Contextualización, características y modalidades evaluativas</i>	25
1.1.1. Aproximación al concepto de evaluación	26
1.1.2. Recorrido histórico de la evaluación educativa	28
1.1.3. Funciones de la evaluación	31
1.1.4. Modalidades de evaluación: Evaluación sumativa, formativa, diagnóstica y combinada.....	32
1.2. <i>Uso de la tecnología en procesos de evaluación</i>	36
1.2.1. Aprendizaje mediado por la tecnología	36
1.2.2. Evaluación mediada por la tecnología.....	38
1.2.3. Dispositivos móviles en procesos de evaluación	40
1.3. <i>Modelos de adopción tecnológica</i>	41
1.3.1. Modelo de Aceptación Tecnológica (TAM)	42
1.3.2. Evolución del modelo TAM: TAM2 y TAM3	44
1.3.2. Influencia del modelo TAM en otros modelos relevantes.....	45
1.3.3. El modelo TAM en educación	46
1.4. <i>Justificación</i>	50
DISEÑO DE LA INVESTIGACIÓN: OBJETIVOS, METODOLOGÍA Y MARCO DE TRABAJO	53
2.1. <i>Objetivos</i>	55
2.2. <i>Fases de la investigación</i>	56
2.3. <i>Metodología</i>	57
2.3.1. Mapeos y revisiones sistemáticas de la literatura.....	57
2.3.2. Proceso de validación de modelos.....	58
2.3.3. Análisis estadísticos: Modelos de ecuaciones estructurales (CB-SEM y PLS-SEM).....	59
2.4. <i>Marco de trabajo para el desarrollo de la Tesis Doctoral</i>	60
RESULTADOS: ESTUDIOS DERIVADOS DE LA INVESTIGACIÓN	63
Estudio 1: Calidad y evaluación de la calidad en enseñanzas medidas por la tecnología	64
3.1. <i>Evaluación de la calidad: Dos estudios derivados</i>	65
3.1.1. Evaluación de la calidad: Una revisión sistemática de la literatura en Educación Superior	66
3.1.2. Evaluación de la calidad: Propuesta de un modelo para e-Learning en educación superior	68
3.1.3. Principales aportaciones y prospectiva de los estudios.....	70
Estudio 2: E-evaluación: Un mapeo sistemático de la literatura	73
3.2. <i>Síntesis del mapeo de la literatura sobre e-evaluación</i>	73
3.2.1. Metodología del mapeo	73
3.2.2. Resultados del estudio de mapeo.....	75

3.2.3. Limitaciones del mapeo.....	77
3.2.4. Implicaciones del mapeo en el desarrollo de la investigación	78
Estudio 3: Aceptación tecnológica de dispositivos móviles.....	79
3.3. <i>Revisión sistemática de la literatura: adopción tecnológica de dispositivos móviles</i>	79
3.3.1. Preguntas de investigación.....	79
3.3.2. Resultados de la revisión.....	81
3.3.3. Limitaciones de la revisión sistemática.....	82
3.3.4. Implicaciones de la revisión sistemática en el desarrollo de la investigación	82
Estudio 4: Diseño y validación de un nuevo modelo de adopción tecnológica	85
3.4. <i>Un nuevo modelo de adopción tecnológica de dispositivos móviles en e-evaluación</i>	85
3.4.1. Propuesta inicial de modelo: Una reflexión sobre aceptación tecnológica.....	85
3.4.2. Proceso de validación de contenido del modelo propuesto.....	87
3.4.3. Descripción del nuevo modelo validado	88
3.4.4. Implicaciones del modelo en el desarrollo de la investigación	89
Estudio 5: Estudio intermedio para la validación del modelo.....	91
3.5. <i>Estudio intermedio: Dos análisis derivados en CB-SEM y PLS-SEM</i>	91
3.5.1. Decisión previa a la aplicación (I): Constructo Ventaja Relativa de la evaluación.....	91
3.5.2. Decisión previa a la aplicación (II): Constructo Autonomía.....	92
3.5.3. Análisis del estudio intermedio.....	92
3.5.4. Instrumento de medida empleado en el estudio intermedio	93
3.5.5. Población y muestra	93
3.5.6. Análisis CB-SEM	94
3.5.7. Análisis PLS-SEM.....	97
3.5.8. Orientaciones y consideraciones finales.....	100
CONCLUSIONES, LIMITACIONES Y LÍNEAS DE INVESTIGACIÓN FUTURAS	101
4.1. <i>Conclusiones</i>	103
Modelo F/S-MAAM: Adopción de dispositivos móviles en evaluaciones formativas y sumativas.....	113
4.2. <i>Implicaciones educativas</i>	115
4.3. <i>Limitaciones</i>	116
4.4. <i>Líneas de investigación futuras</i>	117
4.4.1. Líneas de investigación en exploración actual	117
4.4.2. Líneas de investigación de exploración futura.....	118
REFERENCIAS BIBLIOGRÁFICAS.....	121
ANEXOS DE LA TESIS DOCTORAL	143

Índice de Tablas

Tabla 1. Publicaciones en revista en función de los objetivos específicos de la Tesis Doctoral.....	56
Tabla 2. Publicaciones derivadas en relación con la metodología y técnica de análisis	60
Tabla 3. Dimensiones del modelo propuesto. Adaptado de Ortiz-López et al. (2022)	70
Tabla 4. Hipótesis del modelo validado	88
Tabla 5. Índices de ajuste para los modelos en CB-SEM	96
Tabla 6. Índices de ajuste para los modelos en PLS-SEM.....	99
Tabla 7. Comparativa de dimensiones entre el modelo MBAAM y el modelo propuesto.....	108
Tabla 8. Constructos y varianza explicada del modelo MBAAM y los modelos de la investigación.....	111
Tabla 9. Dimensiones del modelo F/S-MAAM.....	114

Índice de Figuras

Figura 1. Proceso de evaluación. Adaptado de Castillo-Arredondo y Cabrerizo-Diago (2010)	28
Figura 2. Componentes del e-learning. Adaptado de García-Peñalvo y Seoane-Pardo (2010)	37
Figura 3. Modelo TAM. Adaptado de Davis (1989).....	43
Figura 4. Metaanálisis del modelo TAM en educación. Recuperado de Scherer et al. (2019)	46
Figura 5. Modelo GETAMEL. Adaptado de Abduyah y Ward (2016)	48
Figura 6. Modelo MBAAM. Recuperado y adaptado de Nikou y Economides (2017b).....	49
Figura 7. Estructura de la Tesis Doctoral	51
Figura 8. Fases de la investigación	56
Figura 9. Dimensiones identificadas en los modelos de calidad	67
Figura 10. Procesos seguidos en las investigaciones.....	68
Figura 11. Diagrama PRISMA del segundo mapeo de la literatura realizado.....	74
Figura 12. Clasificación de las revistas indexadas en JIF, JCI y SJR	75
Figura 13. Diagrama PRISMA de la revisión sistemática de la literatura.....	81
Figura 14. Propuesta inicial del modelo.....	87
Figura 15. Diseño del modelo validado	89
Figura 16. Esquema del estudio intermedio.....	93
Figura 17. Modelo A del análisis en CB-SEM.....	94
Figura 18. Modelo B del análisis en CB-SEM.....	95
Figura 19. Modelo A del análisis en PLS-SEM	97
Figura 20. Modelo B del análisis en PLS-SEM	98
Figura 21. Modelo C del análisis en PLS-SEM	98
Figura 22. Resultados del análisis de los modelos estructurales de los modelos en PLS-SEM.....	99
Figura 23. Reformulación de la utilidad percibida y la intención conductual de uso.....	112
Figura 24. Modelo F/S-MAAM	115

Introducción:
Evaluación del aprendizaje, dispositivos móviles y
modelos de adopción tecnológica

Introducción:

Evaluación del aprendizaje, dispositivos móviles y modelos de adopción tecnológica

La introducción que se presenta a continuación muestra al lector una aproximación, en líneas generales, al concepto científico de evaluación y su evolución histórica en educación, así como sus principales características y la contextualización de los procesos evaluativos como parte de los entornos virtuales de aprendizaje. Estos nuevos entornos actúan como espacio y sustento de la enseñanza en red y cuentan con funcionalidades que posibilitan una comunicación activa y fluida, que promueve nuevos roles docentes, actuando como espacio moderador para la relación docente-estudiante (Harper, 2018; Li et al., 2019).

Además, se presenta una aproximación a la utilidad de los dispositivos móviles y su relevancia en procesos de evaluación educativa, para continuar con una breve introducción a los modelos de adopción tecnológica en educación. Por último, se presenta la justificación de la Tesis Doctoral.

1.1. Evaluación del aprendizaje: Contextualización, características y modalidades evaluativas

La evaluación del aprendizaje en cualquier contexto educativo formal es, desde una perspectiva histórica, social e ideológica, una de las principales preocupaciones de docentes, alumnos e instituciones educativas. La evaluación puede ser entendida desde infinidad de perspectivas, empleada en relación con multitud de procesos, desarrollándose conceptual y metodológicamente en un proceso de permanente reformulación y adaptación que trascurre de manera paralela al devenir de los cambios que acontecen en los ecosistemas educativos (Black y Wiliam, 1998; Hattie y Timperley, 2007; Ibarra-Sáiz y Rodríguez-Gómez, 2020; Perrenoud, 1991; Scriven, 1967; Shepard, 2000).

El proceso evaluativo es asumido de forma antagónica por aquellos que lo conciben como un fin, y quienes lo asumen como un medio, planteándose desde perspectivas y actuaciones diferenciadas en base a cada premisa. Por tanto, referirse al proceso de evaluación del aprendizaje es describir un proceso sistemático de toma de decisiones, que parte de información válida y fiable que es contrastada mediante unos criterios o referentes, y que deriva en un juicio final, una toma de decisiones y/o una orientación a la mejora.

Las siguientes secciones presentan una aproximación semántica y terminológica al concepto de evaluación del aprendizaje, su recorrido histórico en educación y sus principales elementos caracterizadores.

1.1.1. Aproximación al concepto de evaluación

En el largo recorrido de la evaluación en el sistema educativo, su aproximación terminológica ha sido objeto de evolución y transformación continua. En sus orígenes la evaluación era sinónimo de medición de los resultados de aprendizaje y desempeño de los estudiantes (Rodríguez-Neira et al., 2001). En la actualidad, la noción de evaluación comprende no solo la valoración académica, sino todos los elementos que de manera interrelacionada intervienen en el complejo proceso de enseñanza-aprendizaje (James, 2006). De esta manera, han surgido distintos conceptos relacionados como la evaluación orientada al aprendizaje (Barrientos-Hernán et al., 2020), la evaluación como aprendizaje (Ibarra-Sáiz y Rodríguez-Gómez, 2019), la evaluación sostenible (Boud y Soler, 2016) o la evaluación para el empoderamiento (Fetterman et al., 2015; Fetterman y Wandersman, 2004); entre otras.

La evaluación educativa se concibe como un “proceso de recogida y análisis de la información relevante para descubrir cualquier faceta de la realidad educativa y formular un juicio sobre su adecuación a un patrón o criterio, previamente establecido, como base para la toma de decisiones” (De la Orden, 1982, p. 2). Este proceso se destina tanto a la determinación de la consecución de logros y el cumplimiento de objetivos (Anderson, 2005) como al acompañamiento al estudiante y la retroalimentación en el proceso de enseñanza-aprendizaje (Morris et al., 2021). Entre las distintas perspectivas que clasifican los procesos evaluativos, Tejada (1997) propone concebir la evaluación como consecución de logros (certificación de un mérito), como proceso de ayuda en la toma de decisiones, y como evaluación desde una perspectiva integral.

Entendida como consecución de logros, el término evaluación se ajusta a la definición clásica de Tyler (1949, p. 106), asumida como “the process of determining to what extent the educational objectives are actually being realized” [el proceso que determina hasta qué punto se han conseguido los objetivos educativos]. Esta definición, centrada explícitamente en la consecución de metas,

guarda consonancia con la concepción clásica de evaluación. La aproximación propuesta por Tyler es concebida como uno de los primeros acercamientos de la evaluación a la educación.

Posteriormente, en una línea más orientada a la perspectiva cualitativa, la evaluación se posiciona como un proceso orientado a facilitar la toma de decisiones, que trasciende su capacidad de medición y certificación. En este contexto, Stufflebeam y Shinkfield (1987, p. 92) la presentan como un proceso que posibilita:

identificar, obtener y proporcionar información útil y descriptiva, acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados.

En tercer lugar, surge una corriente que no sólo apela a la consecución de logros, sino que asume la necesidad de la evaluación en la certificación de méritos. En este sentido, Alvira (1991, pp. 10-11) caracteriza la evaluación como el proceso de “emitir juicios de valor o mérito a un programa o intervención basándose en información empírica recogida sistemática y rigurosamente”.

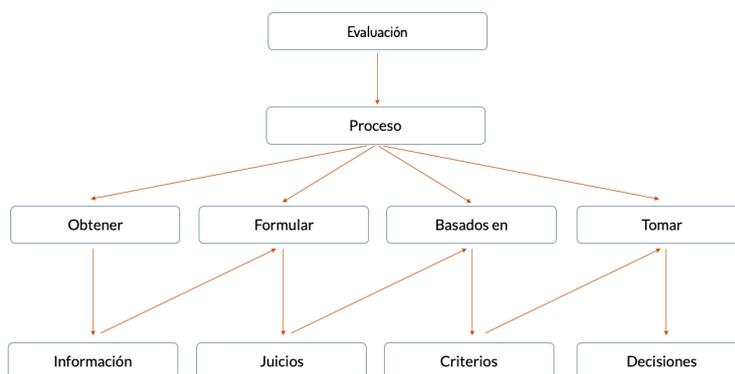
Finalmente, son definiciones más amplias las que contextualizan terminológicamente la evaluación en su conjunto, como la de Pérez-Juste y García-Ramos (1989, p. 23), quienes consideran la evaluación como:

el acto de valorar una realidad, formando parte de un proceso cuyos momentos previos son los de fijación de las características de la realidad a valorar, y de recogida de información sobre las mismas, y cuyas etapas posteriores son la información y la toma de decisiones en función del juicio de valor emitido.

En consecuencia, se observa cómo los autores conciben la evaluación como un proceso continuo, con momentos previos, con fase de recogida de información y con fase de difusión y toma de decisiones. En esta misma línea, Casanova (1995, p. 55) define la evaluación como un:

proceso de recogida de información rigurosa y sistemática para obtener datos válidos y fiables acerca de una situación con objeto de formar y emitir un juicio de valor con respecto a ella. Estas valoraciones permitirán tomar las decisiones consecuentes en orden a corregir o mejorar la situación evaluada.

Las definiciones presentadas conceptualizan la evaluación como un proceso complejo y sistemático en el que destacan cuatro momentos: 1) recogida de información de calidad (válida y fiable), 2) contraste o valoración de dicha información de acuerdo con criterios y/o referentes previamente definidos, 3) formulación de juicios de valor y 4) toma de decisiones orientada a la mejora en un proceso sistemático, organizado y secuenciado (Tejada, 1997) (Figura 1).

FIGURA 1. Proceso de evaluación. Adaptado de Castillo-Arredondo y Cabrerizo-Diago (2010)

Una vez contextualizada la evaluación en educación, se presenta, a continuación, un breve recorrido de su desarrollo histórico en el ámbito educativo, desde comienzos del siglo XX hasta las prácticas, teorías y modelos más recientes (Olmos-Migueláñez y Rodríguez-Conde, 2011).

1.1.2. Recorrido histórico de la evaluación educativa

Desde una perspectiva histórica, Guba y Lincoln (1989) plantean cuatro etapas fundamentales para el estudio del desarrollo de la evaluación: una primera etapa (hasta 1930) que asume el positivismo como base, una segunda etapa (1930-1957) en la que introduce el término evaluación educativa, una tercera etapa (1957-1972) que asocia la evaluación con la toma de decisiones y una cuarta etapa (1973 en adelante) que aúna el gran desarrollo de los modelos de evaluación actuales. Estas fases se pueden categorizar también como la etapa de medición (periodo pre-tyleriano), etapa descriptiva (periodo tyleriano), etapa del juicio (inocencia y realismo) y etapa de la evaluación constructivista (periodos del profesionalismo y la autoevaluación); cuatro periodos que resumen los principales avances conceptuales y epistemológicos del concepto y de las teorías de la evaluación y que sientan las bases de la evaluación educativa.

Profundizando en las consideraciones que estas etapas implican, en relación con el desarrollo temporal del concepto y las prácticas evaluativas, es oportuno subrayar que la primera aproximación sistemática a la evaluación se plantea a mediados del siglo XX, generándose un modelo de evaluación como recurso para la medición (basado en corrientes conductistas) (Serrano, 1990). La evaluación es concebida, en sus estadios iniciales, con el objetivo de la acreditación de instituciones y programas, y con la aparición de las primeras pruebas diagnósticas basadas en la medición precisa de resultados (Escudero, 2003). En esta fase solamente se prioriza la importancia del resultado obtenido de rasgos observables, visibles y evaluables.

En esta mitad del siglo XX, Tyler (1949) (etapa descriptiva) replantea el concepto de evaluación y propone el “Modelo de evaluación por objetivos”, un diseño basado en entender la evaluación como verificación y determinación del grado en el que unos objetivos preestablecidos son

alcanzados, encontrándose ante una evaluación cuantitativa de un producto determinado (Olmos-Migueláñez, 2008). Posteriormente, el concepto de evaluación continúa desarrollándose y deriva en el conocido como “periodo de la inocencia”. En este periodo, la educación comienza su amplio desarrollo en Estados Unidos a través de una oferta exponencial de programas formativos, entroncándose los principios de Tyler con los objetivos educativos y asumiéndose la evaluación educativa como un proceso destinado a la medición de resultados.

Es en los años 60 (periodo del realismo) cuando la evaluación comienza a tomar especial relevancia en educación, surgiendo las primeras preocupaciones por evaluar no sólo los resultados sino también los procesos (Cronbach, 1963). La principal aportación consiste en enfatizar la importancia de la información y su calidad para alcanzar un proceso exitoso y completo, planteando la vinculación de la evaluación con la claridad, la objetividad o la veracidad.

Posteriormente, a partir de los años 70, surge el periodo del profesionalismo de la evaluación, que se relaciona de forma directa tanto con el control y la supervisión de procesos, como con la investigación y el avance de los distintos campos, entre ellos la educación (Stufflebeam y Coryn, 2014).

La elevada rigidez de la corriente de evaluación cuantitativa, basada en la objetividad, la precisión y el establecimiento de objetivos, es causa del surgimiento de una nueva perspectiva que trata de abandonar esta concepción y que termina derivando en un modelo de evaluación sin referencias a objetivos (Scriven, 1991b).

Esta corriente plantea la posibilidad de que el evaluador no conozca los objetivos del proceso educativo de forma previa y refleja la necesidad de establecer criterios y estándares para la determinación de la calidad y la importancia de los análisis externos para la evaluación. Es también en este momento cuando, derivado del desarrollo de la evaluación, Scriven (1967) introduce la distinción entre evaluación formativa y sumativa, otorgando importancia al proceso y no solo a los resultados, como se describirá posteriormente (sección 1.1.5).

En palabras de Fonseca (2007, p. 428), surgen en este momento los modelos cualitativos de evaluación, ante la búsqueda de:

una apertura conceptual que dé cabida a resultados y a acontecimientos no previstos; el inicio de un enfoque que permita la recogida de información tanto de procesos como de productos; el comienzo de una metodología que implique un cambio del monismo al pluralismo metodológico y, una nueva visión ético-política encaminada a estimar los aspectos meritocráticos por encima de los burocráticos.

Esta corriente basa su desarrollo en modelos evaluativos como el “modelo de evaluación orientado por la crítica artística”, propuesto por Eisner (1981) y centrado en asumir la enseñanza

como un arte. En él, el evaluador es un experto en educación que centra su tarea en una observación de carácter descriptivo (explicación detallada de la situación), el carácter interpretativo (forma en la que el evaluador concibe la realidad educativa) y el proceso de realización de juicios de valor (derivado de lo descrito e interpretado en fases previas).

De forma paralela, Parlett y Hamilton (1972) proponen el modelo de evaluación *iluminativa*, un modelo basado en la negociación que se caracteriza por mantener una tendencia holística de la evaluación, la cual debe ser siempre contextual y centrada en la descripción, por encima de la medición, priorizando además la realización de estos procesos en relación con su contexto (Fonseca, 2007). Esta corriente plantea, como fundamentación teórica y metodológica, los sistemas de instrucción (supuestos pedagógicos y modelos que sufren modificaciones después de su aplicación) y los modelos de aprendizaje (contexto psicológico, social y material que rodea a los agentes educativos) (Pérez-Gómez, 1989). Además, esta corriente aboga por la evaluación de los resultados y del propio proceso desde una perspectiva externa; observándose ya dos recorridos diferenciados entre evaluación formativa y sumativa en el campo educativo (Murray, 1984; Scriven, 1991a).

En este momento se desarrolla también un nuevo modelo que plantea un enfoque tecnocrático de la evaluación. Este modelo trata de correlacionar los resultados obtenidos con el proceso que ha conducido a su obtención, planteando la existencia de relaciones causa-efecto en el proceso de evaluación a través de una metodología cuantitativo-positivista; este posicionamiento se relaciona Rossi et al. (1979) y Rivlin (1971) como sus principales exponentes.

Con la llegada de los años 80, la evaluación continua su progreso mediante autores como Stufflebeam (1987), quien concibe la evaluación como un instrumento de ayuda a la mejora educativa y quien propone también el modelo “Contexto, Entrada, Proceso y Producto” (CIPP). Este modelo integral para la evaluación aborda el contexto, el desarrollo, los recursos empleados en el proceso y los productos finales alcanzados (Stufflebeam, 2000).

Surge también en esta década el concepto de “evaluación respondiente” (Abma y Stake, 2001; Stake, 1982), que asume la evaluación como parte del proceso educativo y promueve la difusión de resultados sistematizada. Este modelo propone tres etapas de evaluación que abordan los antecedentes, las transacciones (interacciones y modificaciones) y los resultados del proceso a través de una matriz de datos con un perfil descriptivo (estudio de intenciones y observaciones). Además, emerge otra corriente evaluativa relacionada con la figura de Kemmis (1982), quien desarrolla la “evaluación participante” y confiere a los sujetos la responsabilidad directa de su evaluación.

Seguidamente, y como uno de los modelos referentes en la evaluación cualitativa, se encuentra el modelo *democrático*. Este modelo asume la evaluación como un proceso basado en metodologías

naturalistas para la toma de decisiones, contemplando las interpretaciones y las opiniones de los evaluados a través de procesos de negociación. El objetivo final es promover cambios reales en el objeto evaluado, siendo el evaluador un orientador del proceso y un facilitador del debate y la negociación (Shohamy, 2001).

La evaluación basada en enfoques cualitativos ha ampliado su alcance para incluir aspectos que no se limitan únicamente al proceso de enseñanza-aprendizaje ni a la evaluación de los procesos y resultados. Desde esta perspectiva, destaca la creciente preocupación por la calidad de la docencia y del proceso evaluativo, surgiendo una corriente literaria fundacional que plantea la necesidad de evaluar la calidad también en educación, tomando una concepción ya instaurada en el ámbito industrial (Fernández-Ramírez et al., 1999). Asimismo, los autores sostienen que:

la calidad se ha convertido en un objetivo crucial perseguido por la mayoría de nuestras universidades. La consecución de ese objetivo implica una adecuada planificación, evaluación y mejora del funcionamiento del conjunto de la institución, en lo relativo a la docencia, la investigación, y la gestión, administración y dispensa de los servicios ofertados (p. 17).

La calidad y su evaluación ha adquirido en los últimos tiempos una especial relevancia en el ámbito educativo, en relación con la evaluación del funcionamiento y desempeño institucional mediante la adopción de diferentes criterios y la creación de agencias destinadas a tal fin (Rodríguez-Espinar, 2001). En virtud de estas consideraciones, la Tesis Doctoral que se presenta dedica una sección específica a este componente del proceso evaluativo (estudio 3.1).

En las últimas décadas la evaluación ha continuado su avance en las líneas anteriormente descritas y ha profundizado también en estudios sobre evaluación de programas (Pérez-Juste, 2000), en la autoevaluación y la coevaluación como partes fundamentales del proceso (Dochy et al., 1999) y en el impulso de la evaluación interna por parte del estudiante y su implicación en procesos de evaluación (Adams y King, 1995; Boud, 1995; McDonald y Boud, 2003; Rodríguez-Gómez et al., 2013; Sullivan y Hall, 1997; Taras, 2010); dirigiendo la evaluación hacia un proceso igualmente riguroso y objetivo, pero más flexible y abierto a la participación de todos los agentes.

1.1.3. Funciones de la evaluación

En el campo educativo, el proceso de evaluación se manifiesta a través de diferentes funciones de carácter performativo que le confieren un carácter consustancial, determinante e inherente al proceso educativo. Su delimitación varía en función de la literatura; así, autores como Hernández (1998) o Posner (1998) especifican cuatro categorías principales: diagnóstica, instructiva, educativa y autoformativa. En base a estos, se definen las funciones de la evaluación como:

- diagnóstica: La evaluación tiene el deber de conocer las características del sujeto objeto de evaluación, para así adaptar el proceso a sus posibilidades y al contexto específico de desarrollo.
- instructiva: La evaluación es, en sí, un proceso de formación y aprendizaje para los evaluadores y los agentes encargados del proceso, ya que supone una experiencia adicional y una cualificación necesaria para garantizar la calidad de este.
- educativa: Los resultados derivados de los distintos procesos de evaluación posibilitan a los evaluadores el diseño de nuevas estrategias que favorezcan la mejora de los resultados obtenidos, ya que este proceso permite el diagnóstico de todas aquellas fortalezas y debilidades, y posibilita la apertura de nuevas vías de avance orientadas a la mejora.
- autoformadora: En el ámbito educativo, la evaluación no solamente ofrece resultados sobre el alumnado y los resultados de aprendizaje, sino que también actúa como un ejercicio de reflexión para el docente o para el evaluador, ya que permite el desarrollo de la conciencia crítica y un aprendizaje de forma previa, durante y posterior al proceso.

A las funciones mencionadas, Pérez-Juste y García-Ramos (1989) añaden la función predictiva, la función orientadora y la función de control, entendidas como:

- predictiva: La evaluación posibilita el conocimiento de los diferentes puntos de partida en educación, conociendo de forma específica el estado inicial del alumnado y pudiendo abordar así sus posibilidades futuras en términos de éxito o fracaso.
- orientadora: Esta función plantea que la evaluación debe promover también situaciones de ayuda a los sujetos en el descubrimiento de sus capacidades, habilidades, intereses, posibilidades y actitudes.
- de control: Entendida como una de las principales acepciones de evaluación, esta función se centra en la clarificación sobre el estado de superación de unos objetivos preestablecidos por parte del sujeto, todo con el objetivo de certificar un avance, un logro o la obtención de un título.

Estas funciones permiten describir la evaluación educativa como un proceso amplio, focalizado en distintos aspectos y fases relativas a la docencia y al ámbito educativo, que son resultado de un proceso continuo de modelado entre los agentes y la evaluación en sí, alcanzándose una serie de modalidades que se describen en la siguiente sección.

1.1.4. Modalidades de evaluación: Evaluación sumativa, formativa, diagnóstica y combinada

La evaluación sumativa y formativa y su distinción corresponden a Scriven (1967), quien delimita ambas modalidades y afirma, de forma genérica, que la primera de ellas (sumativa) se desarrolla tras finalizar el proceso de enseñanza-aprendizaje y se orienta a la toma de decisiones sobre la

certificación o la obtención de títulos; mientras que la evaluación formativa se centra en el proceso en sí con objeto de orientarlo a la mejora en la fase de desarrollo (Olmos-Migueláñez, 2008).

Cabe destacar también como modalidad evaluativa, para algunos autores, la evaluación diagnóstica (Casanova, 1998), centrada exclusivamente en el nivel de conocimiento del alumno de forma previa al comienzo de un curso o materia (Marchesi y Martín, 1998). Esta evaluación es una fase previa que puede condicionar el diseño de las evaluaciones posteriores (sumativas y formativas) y que constituye la base para conocer el punto de partida de los sujetos, con una duración breve y una única función en momentos iniciales o finales (Bejar, 1984; Black, 1983).

1.1.4.1. Evaluación sumativa

La evaluación sumativa guarda estrecha relación con las etapas más iniciales de evaluación (modelos cuantitativos) y con las primeras corrientes generadas en el campo, priorizando la evaluación del producto frente al proceso y orientándose a la determinación del grado de consecución de objetivos o a la certificación del alumno (Harlen y James, 1997; Knight, 2002). Este tipo de evaluación engloba todas las técnicas utilizadas para determinar el logro de objetivos, incluyendo exámenes, pruebas y evaluaciones finales, y representa una parte esencial del proceso global de evaluación (Kibble, 2017).

Entre sus ventajas, la evaluación sumativa destaca por la capacidad de reflejar, de forma precisa, los resultados y avances de un estudiante en un momento específico. Esto permite también poner de relieve, de forma precisa, su nivel de consecución de logros/objetivos para la obtención de una certificación o la determinación de superación de los estándares preestablecidos.

Además, esta evaluación proporciona al docente una instantánea del estado actual del aula en lo que a su rendimiento respecta, pudiendo establecer el punto de partida, y pudiendo determinar también el objetivo o punto final marcado tras el proceso (Biggs, 1998; Olmos-Migueláñez, 2008).

En contraposición, la inflexibilidad inherente a este tipo de evaluación y el empleo de técnicas más limitadas no facilitan la investigación de las causas o de otros elementos relacionados con los resultados, lo que dificulta la capacidad de orientar la mejora de la práctica. Además, su enfoque en la determinación del grado de logro de objetivos impide que sea considerada como una parte activa de un proceso educativo en el que los estudiantes también adquieren valores que no pueden ser medidos mediante evaluaciones de este tipo (Roberts, 2006).

En numerosas ocasiones el concepto de evaluación sumativa ha sido objeto de confusión dado que algunos docentes lo han interpretado equivocadamente como la simple realización de exámenes periódicos, confundiéndolo así con la evaluación formativa (Olmos-Migueláñez, 2008). Este proceso sería, en efecto, una evaluación sumativa a corto plazo que, en palabras de Green

(2004, p. 3), se podría definir como “una evaluación sumativa dividida en momentos y desarrollada a lo largo del curso académico”.

1.1.4.2. Evaluación formativa

Derivado de las crecientes discrepancias de autores como Scriven (1967), Stake (1967) o Stenhouse (1975) con el rígido modelo de evaluación sumativa, y apoyándose en el movimiento progresista americano basado en la reforma educativa y social, la evaluación comienza a ser concebida como sinónimo de valoración orientada a la comparación y a la mejora (González-Monteagudo, 2001).

La evaluación formativa, en contraposición a la sumativa, concibe la evaluación en el proceso completo de enseñanza-aprendizaje y evalúa no solamente resultados, sino también aspectos cognitivos, sociales o contextuales (entre otros), conllevando una toma de decisiones durante el proceso y proporcionando una retroalimentación continua (Marchesi y Martín, 1998). Bell y Cowie (2001, p. 536) resumen esta modalidad como “the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning” [el proceso utilizado por profesores y alumnos para reconocer y responder al aprendizaje de los alumnos con el fin de mejorar dicho aprendizaje, durante el aprendizaje]. Esta evaluación destaca por la relevancia de la retroalimentación, un aspecto abordado desde una perspectiva dual según Miras y Solé (1990) ya que sus impactos son tanto para el estudiante (ayuda en su aprendizaje en un momento dado) como para el profesor (información sobre el estado del proceso de aprendizaje y estímulo de la metacognición) (Morris et al., 2021).

Esta modalidad evaluativa se encuentra ampliamente vinculada a la actividad continua del alumnado, tanto en su consecución de logros como en la valoración de sus fallos de forma individual. Esto implica proporcionar una retroalimentación adecuada y específica para cada estudiante, lo que también representa un esfuerzo adicional por parte del docente (Havnes et al., 2012). En esta modalidad evaluativa destacan las actividades de autoevaluación, la evaluación entre iguales, la coevaluación y la heteroevaluación (Ibarra-Sáiz et al., 2020; Villamar-Coronel, 2022).

Entre sus ventajas, la evaluación formativa es un reflejo del proceso completo de enseñanza-aprendizaje, lo que permite dar respuesta a las necesidades en el momento de su aparición, detectando los posibles problemas derivados y orientando el proceso a la mejora en todas las fases en las que se plantea esta evaluación. En contraposición, la evaluación formativa es un proceso más largo, exigente y duradero que las evaluaciones de carácter sumativo (Green, 2004). Por este motivo, esta modalidad no es aún la más representativa en el sistema educativo actual (Olmos-Migueláñez, 2008).

1.1.4.3. Evaluación formativa y sumativa en la práctica educativa

Examinadas teóricamente las dos modalidades de evaluación (sumativa y formativa), es crucial destacar que en la práctica educativa ambas modalidades no se observan o manifiestan de manera precisa, ni todos los docentes se centran exclusivamente en el planteamiento de una de ellas, habida cuenta de las exigencias del sistema. Esto se debe a que estas modalidades están destinadas a cumplir diferentes objetivos en la evaluación del aprendizaje, ya sea el cumplimiento de estándares o la mejora del proceso de aprendizaje.

En la investigación sobre el desarrollo práctico de las modalidades evaluativas, parecen delimitarse tres corrientes relativas a la función docente y su relación con el desarrollo de estas evaluaciones. La primera asume la no diferenciación entre formativa y sumativa (entendiendo la evaluación como un todo inherente al ejercicio docente), la segunda la coexistencia de ambas, pero su independencia en la práctica, y la tercera la necesaria complementariedad para un proceso evaluativo eficaz (Lau, 2016). Se describen, a continuación, las tres concepciones y su influencia en el proceso de evaluación.

Por un lado, la primera corriente asume la evaluación en líneas genéricas como un conjunto único de procesos, sin diferenciar el proceso del resultado alcanzando y entendiendo su utilidad en términos de eficacia, rendimiento y respuesta a objetivos (Mejía-Pérez, 2012). Esta corriente se corresponde con la concepción clásica de evaluación, sin distinguir entre modalidades formativa y sumativa ni explorar sus posibilidades (Shepard, 2006; Tyler, 1949).

Por otro lado, la segunda corriente de pensamiento descrita asume la independencia de ambas modalidades y trata de disgregar los efectos de una y otra sobre el alumno y sobre la docencia (Harlen y James, 1997; Patton, 1996), planteando evaluaciones formativas por un lado y evaluaciones sumativas por otro (Black y William, 1998). Esta propuesta se enmarca como una de las concepciones más asentadas y clásicas sobre evaluación, siendo aquella que distingue entre sus efectos, y asumiendo que la existencia de un modelo combinado y unitario puede ser un problema para los docentes en la búsqueda de la conjugación de las distintas técnicas, objetivos y finalidades (MacLellan, 2001).

Finalmente, una corriente más reciente asume la dualidad de la evaluación en su carácter formativo y sumativo, lo que le supone al docente la asunción del proceso completo de evaluación como la suma indispensable de ambas (Ahmad y Bhat, 2019; Buchholtz et al., 2018). Bajo esta perspectiva, se asume que la finalidad de la evaluación formativa y la finalidad de la evaluación sumativa guardan una total relación y complementariedad para alcanzar una evaluación integral (Black et al., 2003) en un proceso mediado por la retroalimentación continua y el acompañamiento orientado a la consecución de logros y a su certificación (Dixson y Worrell, 2016; Dolin et al., 2018).

Pese a que las evaluaciones sumativas son un requisito obligatorio en el sistema educativo actual (dada la necesaria certificación del desempeño de los estudiantes), los docentes deben adoptar decisiones sobre el propio proceso de evaluación, desde definir cuándo evaluar, cómo, por qué y para qué, en una decisión condicionada, también, por la introducción de herramientas tecnológicas en el proceso educativo.

El desarrollo de estas herramientas tecnológicas en educación se describe en la próxima sección, que contextualiza el proceso de transición tecnológica de los sistemas educativos y aborda el papel de los dispositivos móviles en los procesos de evaluación.

1.2. Uso de la tecnología en procesos de evaluación

En la actualidad, la integración de la tecnología en el ámbito educativo ha transformado significativamente el proceso de enseñanza-aprendizaje. Esta sección aborda el desarrollo y la presencia de la tecnología, y especialmente de los dispositivos móviles, en los procesos de evaluación, explorando cómo las nuevas tendencias modelan la evaluación en entornos virtuales y proporcionan nuevas oportunidades para la evaluación del aprendizaje de los estudiantes.

1.2.1. Aprendizaje mediado por la tecnología

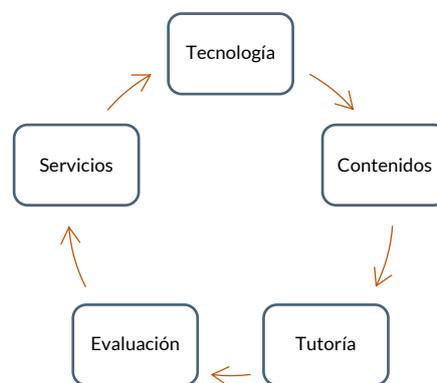
El concepto de aprendizaje mediado por la tecnología (*e-learning*) es un término que surge de los cambios derivados de la implementación de las Tecnologías de la Información y la Comunicación en educación, en las instituciones y en los procesos de enseñanza-aprendizaje, y se define como un:

proceso formativo, de naturaleza intencional o no intencional, orientado a la adquisición de una serie de competencias y destrezas en un contexto social, que se desarrolla en un ecosistema tecnológico en el que interactúan diferentes perfiles de usuarios que comparten contenidos, actividades y experiencias y que, en situaciones de aprendizaje formal, debe ser tutelado por actores docentes cuya actividad contribuya a garantizar la calidad. (García-Peñalvo y Seoane-Pardo, 2015, p. 132).

En lo que respecta a sus dimensiones, García-Peñalvo y Seoane-Pardo (2010) describen los cinco aspectos que contextualizan al e-learning, de cuyas diferentes combinaciones derivan los distintos modelos de enseñanza basados en la tecnología. Estas dimensiones son la dimensión tecnológica (soporte empleado), la dimensión relativa al contenido (calidad, organización y estructuración de los mismos), la dimensión relativa a la tutorización (forma de relación entre alumno y docente), la dimensión propia de la evaluación (nueva concepción y desarrollo en entornos virtuales) y aquella relativa a los nuevos servicios derivados de esta modalidad (cargos de gestión, de supervisión, relativos a la acción docente, evaluación de la calidad, etc.) (Figura 2).

En este momento, el e-learning es un marco sobre el que se amparan diferentes modalidades específicas de docencia mediada por la tecnología. Entre ellas destacan, desde el punto de vista de la presencialidad: *electronic learning* o *e-learning* (presencial), *blended learning* o *b-Learning* (combinación de sesiones presenciales con formación en red) y el *online learning* u *o-learning* (formación completamente online). Por otro lado, desde el punto de vista del soporte empleado, la literatura distingue entre *e-learning* (dispositivos electrónicos, generalmente ordenadores), *m-learning* (con dispositivo móvil) o *u-learning* (aprendizaje ubicuo en el que el acceso no depende del dispositivo) (Parlakkiliç, 2020).

FIGURA 2. Componentes del e-learning. Adaptado de García-Peñalvo y Seoane-Pardo (2010)



La tecnología y su progresiva implementación en el ámbito educativo ha supuesto una revolución en la concepción clásica de la didáctica desde los niveles obligatorios hasta la educación superior (Gros-Salvat, 2018). Esta inclusión ha modificado los procesos de enseñanza-aprendizaje y supuesto para el docente una posibilidad de ampliar los recursos didácticos y metodológicos disponibles en el aula (Colás-Bravo et al., 2018). El desarrollo tecnológico ha supuesto también una posibilidad para las instituciones en lo referido al trabajo en red institucional (Souabi et al., 2021) y la apertura de un nuevo escenario enfocado en las enseñanzas no presenciales basadas en plataformas virtuales de enseñanza (García-Peñalvo, 2020) mediante el asentamiento de dispositivos como ordenadores, teléfonos o *tablets* como soporte principal.

La presencia de la tecnología en educación se remonta a la década de los 50 con la introducción en las aulas de los medios audiovisuales para apoyar los procesos formativos y con el estudio de la relación tecnológica con materiales educativos como primera línea de avance (Vidal, 2006). En los estadios iniciales, el desarrollo se encontraba limitado a la observación del avance de la tecnología en los medios de comunicación en unos cambios que serán progresivamente adaptados en años posteriores a la tecnología educativa. A partir de los años setenta, gracias al desarrollo de la informática, comienza la implementación de ordenadores con finalidades educativas y con la enseñanza individualizada como objetivo principal (Cabero-Almenara, 1999). En los años ochenta,

el desarrollo comienza a centrarse en las posibles integraciones tecnológicas en educación, generándose también numerosas críticas a este proceso (Area-Moreira, 2000).

Es en la década de los años noventa cuando, impulsado por la llegada de internet, se empieza a asumir que la tecnología en educación debe ser un medio fundamental, comenzando a partir del año 2000 el desarrollo sobre las metodologías, los recursos didácticos y los materiales curriculares para docentes. En el año 2009, el Ministerio de Educación implanta el programa Escuela 2.0 para la integración de las TIC y la modernización de los centros (Area-Moreira et al., 2014). Este programa basa sus actuaciones en la creación de aulas digitales, la mejora de la conectividad en el aula, la formación para el profesorado (técnica, metodológica y socialmente), la creación de materiales digitales, la garantía de acceso digital para profesorado, alumnado y familias y la formación en materia de responsabilidad a toda la comunidad educativa (Pérez-Sanz, 2011).

La inclusión tecnológica en educación es resultado de numerosas experiencias innovadoras (que no investigaciones educativas) realizadas a título individual o institucional en las aulas (Aguilar et al., 2019). Unas experiencias que tratan de determinar el efecto de la tecnología en el rendimiento académico, en los resultados de los estudiantes y en la docencia (Wentworth y Middleton, 2014). La mayoría de las intervenciones realizadas en el aula, ocasionalmente también investigaciones, concluyen con un alto porcentaje de satisfacción y de mejora de resultados. Además, estas investigaciones abogan por una continuidad en el desarrollo tecnológico en educación y por una mayor investigación y desarrollo en lo referido a programas educativos innovadores (Chandra et al., 2020).

1.2.2. Evaluación mediada por la tecnología

Ante el reto digital descrito, los sistemas educativos y las instituciones involucradas deben garantizar la calidad y la correcta transición de todas y cada una de las partes del proceso de enseñanza-aprendizaje, siendo la evaluación del aprendizaje uno de los procesos más importantes que demanda un mayor rigor en su quehacer (Guskey, 2009). Es en esta transición donde surgen las primeras cuestiones referidas a la forma en la que la digitalización de la evaluación se desarrolla: ¿es cuestión de digitalizar instrumentos físicos al medio digital, o es necesaria una reformulación completa de la evaluación? Estudios comparativos realizados en este sentido (Ruiz-Morales et al., 2018) revelan la diferencia en las competencias evaluables y la forma de hacerlo según el contexto. Entre sus conclusiones, destaca la necesidad de replantear la evaluación desde la fase inicial de la construcción de un escenario e-learning o m-Learning, abandonando la fórmula del traspaso de instrumentos y del mantenimiento de los mismos estándares de calidad.

Los rápidos cambios descritos en materia de tecnología en educación han derivado en una rápida adaptación de la evaluación a nuevos entornos (Rodríguez-Conde, 2005). Sin embargo, la

transformación efectiva sigue siendo un desafío para investigadores y docentes (Tomas et al., 2015) en la búsqueda de un enfoque holístico que valore la evaluación de aprendizajes significativos, fomente la participación y la colaboración entre los estudiantes, y promueva la autoevaluación, la evaluación entre iguales y la coevaluación (Cubero-Ibáñez et al., 2018; Ibarra-Sáiz y Rodríguez-Gómez, 2020) beneficiándose de las ventajas de la tecnología (Hernández-Sellés et al., 2018)

Estas ventajas no hacen más que enfatizar el potencial de la tecnología (y de los dispositivos móviles en particular) en educación en una sociedad que debe asumir que la relación alumno-teléfono móvil es una unión ya indisociable (Brazuelo-Grund et al., 2017). Todo ello supone una oportunidad para la introducción de innovaciones y la inclusión efectiva de las Tecnologías de la Información y la Comunicación (TIC) en la mayoría de los niveles educativos (Bogdanovic et al., 2014).

La evaluación en entornos virtuales trata de contraponerse hoy con la rigidez que caracteriza a los métodos clásicos de evaluación, focalizados muchos de ellos en la rendición de cuentas y en la calificación del alumnado y denostando la creatividad y la originalidad, dos aspectos esenciales para la motivación del alumnado (Pearse-Romera y Ruiz-Cecilia, 2019). Por tanto, las instituciones afrontan un desafío completo en la superación de los retos que esta adaptación supone (García-Peñalvo et al., 2020), mejorando la articulación formación-evaluación en contextos mediados por la tecnología. Todo ello en un proceso coordinado por nuevas figuras que garanticen los criterios necesarios para el cumplimiento de los estándares de la evaluación ya descritos (Vergés, 2018).

El reto que implica este cambio para la acción docente cuenta con una serie de limitaciones que lo dificultan, entre las que destacan la accesibilidad a los recursos y la necesaria formación y actualización docente en competencia digital y en materia de calidad (Domingo-Coscollola et al., 2020; García-Aretio, 2021; Lee y Fanguy, 2022). Pese a ello, son numerosas las investigaciones que presentan experiencias exitosas de implementación de tecnología en procesos educativos y formativos, demostrando que las ventajas superan ampliamente tanto a las desventajas como a los posibles riesgos derivados del proceso (Alqahtani y Rajkhan, 2020; Rangel-Baca, 2018).

El avance en materia tecnológica y educativa es, por tanto, una tendencia que trasciende la práctica docente y que se respalda en prácticas de investigación educativa. Estas experiencias de desarrollo evaluativo en contextos virtuales se encuentran lideradas, entre otros, por grupos de investigación como el grupo *Evalfor* (Evaluación en Contextos Formativos) de la Universidad de Cádiz, a través del establecimiento de marcos innovadores como el Marco de Acción FLOASS. Este marco se basa en la tecnología para la mejora de la evaluación, el desarrollo de analíticas de aprendizaje y la orientación, implementación y seguimiento de procesos de evaluación de resultados de aprendizaje actualizados (Ibarra-Sáiz y Rodríguez-Gómez, 2019b)

1.2.3. Dispositivos móviles en procesos de evaluación

En este contexto, la tecnología móvil emerge como un soporte sólido para el futuro de la enseñanza, las metodologías y la evaluación (Fombona-Cadavieco y Rodil-Pérez, 2018). Esta modalidad de aprendizaje (*mobile learning* o *m-learning*) es definida como la implementación de un dispositivo móvil en sistemas e-learning para el desarrollo de experiencias de aprendizaje (Elkheir y Abdul-Mutalib, 2015). Este proceso permite pensar en una docencia innovadora basada en la ubicuidad, la inmediatez, la movilidad, el acceso permanente y abierto y la contextualidad (Gabriel-Pedro et al., 2018).

En la actualidad, esta tecnología se encuentra ampliamente extendida entre los estudiantes que promocionan en el sistema educativo, siendo la edad media de tenencia del primer dispositivo móvil en España los 12 años (Caldeiro-Pedreira et al., 2021), y teniendo a su vez una representatividad muy superior a la tenencia de ordenadores portátiles en nuestro país (Sánchez-Prieto et al., 2019). Este nuevo contexto generado por la introducción tecnológica en el aula ha sido también objeto de debate en su implementación y desarrollo en educación (Grande-de-Prado et al., 2021).

El uso de estos dispositivos en procesos evaluativos presenta unos beneficios que pasan por la estimulación y el desarrollo de nuevas experiencias y situaciones de aprendizaje, la posibilidad de ampliar la variedad de técnicas e instrumentos, la facilitación de la interacción instantánea, el menor coste de acceso que un ordenador convencional, la posibilidad de geolocalización, el empleo de la cámara o el uso de la realidad aumentada y códigos QR para la evaluación (Zamora-Delgado, 2019). Estos avances derivan en el surgimiento de nuevos contextos que permiten el desarrollo en línea de modalidades como la autoevaluación, la evaluación entre iguales o la heteroevaluación (en función del agente implicado); y de evaluaciones formativas, basadas en el rendimiento, sumativas o competenciales (en función de la finalidad) (Rodríguez-Gómez et al., 2013).

Por otro lado, la tecnología ha sido un factor determinante en situaciones como la pandemia de la COVID-19 (2020) (García-Peñalvo et al., 2020). En este escenario, la tecnología ha generado altas expectativas que parten de diseños metodológicos y evaluativos propuestos durante la crisis sanitaria, así como de las distintas políticas institucionales generadas, que han posibilitado la presencia de los dispositivos en las aulas (Kuklinski y Cobo, 2021). Esta tarea ha requerido la necesaria formación del profesorado y alumnado en materia digital (Bashkireva et al., 2020; Williamson et al., 2020). Por otro lado, es fundamental considerar el rápido avance tanto en desarrollo como en accesibilidad al usuario de la inteligencia artificial y la realidad aumentada (Samala et al., 2023), lo que puede suponer un cambio en la educación actual reforzando el papel de la tecnología, y de los dispositivos móviles, como parte central del proceso de enseñanza-aprendizaje (Jia et al., 2022).

Sin embargo, a pesar de los beneficios descritos, estos dispositivos cuentan con unas limitaciones que comprometen su presencia en el proceso de enseñanza-aprendizaje. La brecha digital persistente y la disparidad en el acceso los mismos, pueden suponer desigualdades que condicionen su introducción (Warf, 2019). Además, la distracción potencial que pueden suponer y la falta de supervisión durante las evaluaciones en línea pueden plantear desafíos en términos de integridad académica (Kaisara y Bwalya, 2023).

Por otro lado, la legislación y las regulaciones en el ámbito educativo son una limitación adicional en su implementación efectiva. La legislación educativa no ha evolucionado de forma paralela con el desarrollo de la tecnología móvil y sus potencialidades en las aulas. Esta falta de directrices respecto al uso de dispositivos móviles en evaluaciones puede generar ambigüedades y desafíos legales. Además, la protección de la privacidad y la seguridad de los datos del estudiante se convierte en una preocupación importante para centros educativos e instituciones, especialmente cuando se utilizan aplicaciones y plataformas en línea (Adejo et al., 2018).

Por último, las diferencias en las políticas de los centros educativos y las diferentes estrategias de implementación pueden generar desigualdades entre las instituciones, reflejando una falta de uniformidad en la aplicación de diseños y programas relacionados con el uso de dispositivos móviles, situación que puede dar lugar a disparidades en la calidad de la educación y en la equidad de acceso entre los estudiantes.

En virtud de lo anteriormente señalado, se puede concluir que los dispositivos móviles generan hoy unas altas expectativas de cambio educativo (Ozturk, 2019) respaldado por prácticas de formación a docentes para la adaptación a los nuevos entornos de aprendizaje. En este proceso, se espera que los docentes adopten los dispositivos móviles en su actividad (tanto metodológica como evaluativa) para facilitar y apoyar el proceso de enseñanza-aprendizaje y para una mejora en la calidad de su experiencia docente (Aldunate y Nussbaum, 2013).

La siguiente sección presenta una aproximación a los modelos de adopción tecnológica y a su desarrollo en el ámbito de la educación y la evaluación mediada por la tecnología.

1.3. Modelos de adopción tecnológica

La comprensión del panorama actual de las investigaciones sobre la adopción tecnológica en educación es esencial para contextualizar y fundamentar las decisiones que los docentes toman respecto a su uso e implementación. Los modelos de adopción tecnológica, sustento principal de los estudios de aceptación, posibilitan el análisis de los factores que conducen al docente al empleo de estos dispositivos en procesos de evaluación.

La aceptación tecnológica ha sido descrita en numerosas investigaciones y ha derivado en diferentes teorías como la Teoría de Difusión de Innovaciones (IDT) (Rogers, 1962), la Teoría del

Ajuste Tarea-Tecnología (TTF) (Goodhue y Thompson, 1995) o la Teoría Unificada de la Aceptación y Uso de la Tecnología (UTAUT) (Venkatesh et al., 2003); entre otras.

1.3.1. Modelo de Aceptación Tecnológica (TAM)

Esta investigación centra su recorrido en el Modelo de Aceptación Tecnológica (TAM) (Davis, 1989), el más abordado en la literatura sobre la adopción de la tecnología (García-Peñalvo et al., 2015) dada su solidez teórica y la simplicidad de su instrumento, siendo capaz de explicar un alto porcentaje de la varianza en estudios de aceptación (Hernández-García, 2012). Este modelo es resultado de los principios de la Teoría de la Acción Razonada (TRA) (Fishbein y Ajzen, 1975) y la Teoría del Comportamiento Planeado (TPB) (Ajzen, 1985).

La primera de ellas, la Teoría de la Acción Razonada, establece que las actitudes hacia la realización de una determinada conducta se encuentran asociadas positivamente con la intención de llevar a cabo dicho comportamiento (Fishbein y Ajzen, 1975).

En estudios de adopción tecnológica el análisis se focaliza en los aspectos actitudinales del comportamiento para la determinación de la intención de uso, entendida como la probabilidad subjetiva de un individuo de realizar un comportamiento determinado, a través de dos factores. Estos factores son la actitud hacia el uso, definida como “the person’s general feeling of favorableness or unfavorableness toward some stimulus object” [sentimiento hacia un estímulo] (Fishbein y Ajzen, 1975, p. 216), y la norma subjetiva, definida como “the person’s perception that most people who are important to him think he should or should not perform the behavior in question” [percepción del sujeto sobre el pensamiento de las personas importantes para él acerca de la realización de una conducta] (Fishbein y Ajzen, 1975, p. 302).

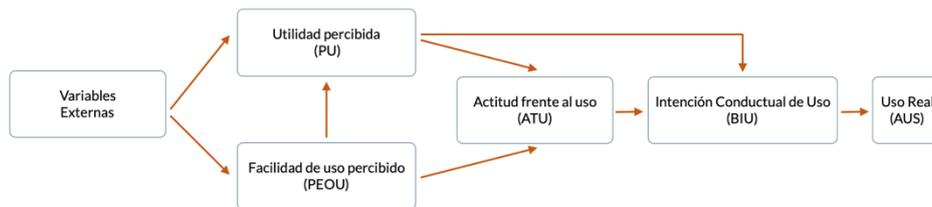
A pesar de que esta teoría supuso un avance en los modelos de adopción, asumir que un determinado comportamiento es realizado voluntaria y conscientemente por el individuo limitaba el valor predictivo de los modelos cuando esta situación no se cumplía (Sánchez-Prieto et al., 2016).

Con el objetivo de superar estas limitaciones de la TRA surge la Teoría del Comportamiento Planeado (TPB) (Ajzen, 1985). Esta teoría tiene como objetivo predecir el comportamiento volitivo y comprender los determinantes psicológicos añadiendo, a los constructos anteriormente descritos, un nuevo constructo que refiere a factores exógenos: el control conductual percibido, entendido como la dificultad para llevar a cabo una conducta determinada (Ajzen, 1985).

Tomando como referencia las teorías descritas, Davis (1989) propone el Modelo de Aceptación Tecnológica (TAM) (Figura 3). Este modelo mantiene la actitud frente al uso como antecedente de la intención conductual de uso y elimina la norma subjetiva por su difícil medición. Además, incorpora como antecedentes la utilidad percibida, entendida como “the degree to which a person believes that using a particular system would enhance his or her job performance”) [el grado en que

un sistema mejorará el rendimiento] (Davis, 1989, p. 320; y la facilidad de uso percibida, concebida como “the degree to which a person believes that using a particular system would be free of effort” [grado en que un sistema reducirá el esfuerzo] (Davis, 1989, p. 320). La actitud frente el uso y la intención conductual de uso han sido definidas en la aproximación a la TRA.

FIGURA 3. Modelo TAM. Adaptado de Davis (1989)



El modelo TAM establece cinco hipótesis que relacionan los constructos descritos y contribuyen a la determinación de la intención de uso. Estas hipótesis son:

- La facilidad de uso percibida se encuentra directamente relacionada con la utilidad percibida.
- La facilidad de uso percibida se encuentra directamente relacionada con la actitud hacia el uso.
- La utilidad percibida se encuentra directamente relacionada con la intención conductual de uso.
- La utilidad percibida se encuentra directamente relacionada con la actitud hacia el uso.
- La actitud hacia el uso se encuentra directamente relacionada con la intención de uso conductual.
- La intención conductual de uso se encuentra directamente relacionada con el uso real del sistema de información.

Por último, Davis (1989) introduce la existencia de variables externas antecedentes de la utilidad percibida y de la facilidad de uso percibido, sin especificar unos constructos determinados. Por este motivo, y dado que el modelo no abarca todos los antecedentes, TAM es un modelo ampliable mediante la búsqueda de estas variables condicionantes de la adopción tecnológica.

Pese a que TAM es el modelo de adopción tecnológica más utilizado por su fiabilidad, su solidez, y su fácil adaptación a diferentes ámbitos que contemplan el uso tecnológico (King y He, 2006), cuenta con una serie de limitaciones entre las que destacan la falta de inclusión de variables externas o los bajos niveles de varianza alcanzados en estudios exploratorios; entre otras (Hernández-García, 2012; Legris et al., 2003).

Derivado de estas limitaciones, el modelo TAM ha sufrido modificaciones estructurales y funcionales en su desarrollo posterior. Estos cambios han sido categorizados por King y He (2006)

en: la investigación de las variables externas que afectan a la utilidad percibida y a la facilidad de uso percibida (1), la referencia a la inclusión de elementos de otras teorías con el objetivo de aumentar la capacidad predictiva (2), la introducción de elementos referidos al contexto para estudiar sus efectos como factores moderadores (3) o los cambios en la medición de elementos finales del modelo (4).

1.3.2. Evolución del modelo TAM: TAM2 y TAM3

Con la necesidad de superar de las limitaciones del modelo TAM surgen años después de la propuesta de Davis (1989) dos evoluciones del modelo: el modelo TAM2 (Venkatesh y Davis, 2000) y el modelo TAM3 (Venkatesh y Bala, 2008).

La segunda versión del modelo (TAM2) mantiene los constructos del modelo TAM a excepción de la actitud hacia el uso, eliminada por su limitado efecto moderador constatado en estudios previos (Davis y Venkatesh, 1996) y posteriores (Hu et al., 2003) al modelo TAM2, relacionando de forma directa la utilidad percibida y la facilidad de uso percibida con la intención conductual de uso. Este modelo añade la norma subjetiva como factor que condiciona la utilidad percibida y la intención de uso y añade también la experiencia y la voluntariedad de uso como elementos moderadores. Además, Venkatesh y Davis (2000) introducen como antecedentes de la utilidad percibida, la imagen, la relevancia laboral, la demostrabilidad del resultado y la calidad del resultado.

El resultado es un modelo más complejo que pierde la sencillez de la propuesta original, razón por la que apenas se utiliza en la actualidad (Sánchez-Prieto et al., 2016). Por este motivo, y con el objetivo de superar las discrepancias al no abordar un mayor número de factores explicativos, Venkatesh y Bala (2008) proponen el modelo TAM3 que combina el modelo TAM2 con el Modelo de los Determinantes de la Facilidad de Uso Percibida (Venkatesh, 2000), desarrollando un nuevo modelo integrado de aceptación de la tecnología.

La tercera versión del modelo (TAM3) mantiene todos los constructos propuestos en su predecesor (TAM2), y agrupa los antecedentes de la facilidad de uso percibida en dos categorías: factores de anclaje y factores de adaptación. Venkatesh y Bala (2008) definen estos factores como:

- **Factores de anclaje:** Estos factores impulsan los juicios iniciales sobre la facilidad de uso percibida de forma previa a la implementación de un sistema tecnológico (Venkatesh y Bala, 2008). Los factores de anclaje sugeridos por Venkatesh (2000) y presentes en TAM3 son: *computer self-efficacy*, *computer anxiety*, *computer playfulness and the perceptions of external control (or facilitating conditions)* [autoeficacia, ansiedad, disfrute, y percepción del control externo (o condiciones facilitantes)]. Los tres primeros representan diferencias individuales, es decir, creencias generales asociadas al ordenador y a su uso.

- **Factores de ajuste:** Los factores de ajuste se definen como la respuesta del individuo a los juicios iniciales una vez adquirida la experiencia con el nuevo sistema empleado (Venkatesh, 2000). El modelo TAM3 presenta dos factores de ajuste relacionados con las características del sistema: “the perceived enjoyment and the objective usability” (Venkatesh y Bala, 2008, p. 278) [entretenimiento percibido y usabilidad objetiva].

Con relación a estos factores, Venkatesh (2000) propuso que, incluso en sujetos con una nivel de experiencia con el sistema elevado, el papel de dos anclajes (autoeficacia y percepciones de control externo) continuaban siendo relevantes una vez introducido el sistema. Sin embargo, afirmó que los efectos de los otros dos factores de anclaje disminuían con el tiempo.

1.3.2. Influencia del modelo TAM en otros modelos relevantes

El modelo TAM no es solo el modelo más empleado para el estudio de la adopción tecnológica, sino que es también base para el desarrollo de nuevos modelos de aceptación y de modelos relevantes ya asentados, como el modelo C-TAM-TPB y la Teoría Unificada de Aceptación y Uso de tecnología (UTAUT).

El primer modelo, el C-TAM-TPB, es resultado de la reflexión realizada por Taylor y Todd (1995) al comparar el Modelo de Aceptación Tecnológica y la Teoría del Comportamiento Planeado. Este modelo refuerza el estudio de la adopción mediante la inclusión de factores de control (control conductual percibido, TPB) y factores sacionormativos (norma subjetiva, TAM).

Por otro lado, un segundo modelo relevante derivado de TAM es el modelo UTAUT (Venkatesh et al., 2003), uno de los modelos más populares en estudios de adopción (Williams et al., 2015). Este modelo basa su propuesta en el estudio de la intención de uso de la tecnología mediante cuatro antecedentes de la intención conductual de uso: las condiciones facilitantes, la influencia social, las expectativas de rendimiento y las expectativas de esfuerzo. El modelo UTAUT incluye como variables moderadoras entre las relaciones propuestas el género, la edad, la experiencia y la formación.

Posteriormente, el modelo experimenta modificaciones hasta alcanzar su segunda versión, UTAUT2 (Venkatesh et al., 2012), que añade al modelo original los constructos de la motivación hedonista, la experiencia y el hábito, y el valor del precio/beneficio en un modelo vinculado con el uso de la tecnología en el ámbito del consumo.

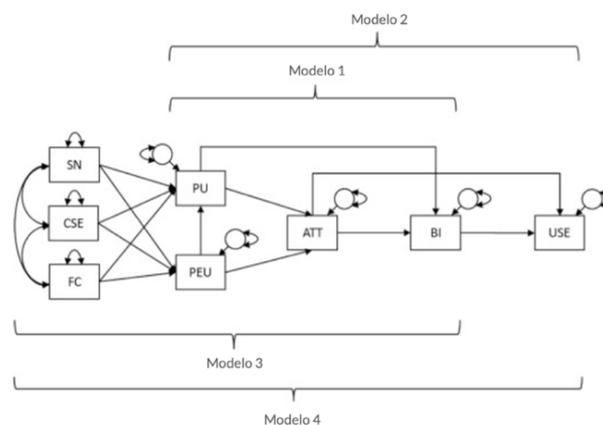
En base a este recorrido, se confirma que la literatura presenta hoy un amplio marco de modelos para el estudio de la adopción tecnológica. Una vez terminada su contextualización, se presenta a continuación un recorrido por el desarrollo del modelo TAM en el ámbito educativo por ser el modelo más empleado (Lai, 2020).

1.3.3. El modelo TAM en educación

El modelo TAM ha sido desarrollado de manera extensa y específica en educación. Esta sección busca profundizar en las investigaciones realizadas en este área para concluir con una aproximación a la evolución del modelo TAM en estudios de adopción de evaluaciones mediadas por dispositivos móviles. Este modelo ha sido objeto de numerosas investigaciones y aplicaciones en educación.

En este sentido, el metaanálisis llevado a cabo por Scherer et al. (2019) sintetizó 114 modelos referidos a aplicaciones del modelo TAM en educación para comprobar el ajuste del modelo en sus diferentes versiones (Figura 4).

FIGURA 4. Metaanálisis del modelo TAM en educación. Recuperado de Scherer et al. (2019)



Nota: ATT (Actitud frente al uso), BI (Intención conductual de uso), CSE (autoeficacia), FC (Condiciones facilitantes), PEU (Facilidad de uso percibida), PU (Utilidad percibida), SN (Norma subjetiva), USE (Uso real).

Entre las conclusiones de los autores, el modelo destaca por su buen funcionamiento y su capacidad explicativa sobre la aceptación de la tecnología en estudios de adopción en educación. Sin embargo, los autores consideran determinante el papel de ciertos constructos clave, la importancia de las variables externas y los antecedentes en el funcionamiento del modelo, derivando en la recomendación de establecer modelos específicos para cada área del conocimiento y para cada campo de investigación.

Por este motivo, se presenta una aproximación al modelo en ámbitos específicos del área educativa. En primer lugar, se realiza un recorrido por la vinculación entre TAM y las enseñanzas mediadas por la tecnología.

1.3.3.1. TAM en enseñanzas e-Learning

El modelo TAM y su adaptación a modelos de adopción en el campo de la educación mediada por la tecnología es una línea de investigación ampliamente desarrollada en la actualidad. En este

sentido, destacan investigaciones relevantes como la de Park (2009), quien propone un modelo para conocer el proceso de adopción del e-learning en estudiantes universitarios. El modelo, con base en el modelo TAM, incluía la autoeficacia, la norma subjetiva y la accesibilidad del sistema para su estudio. El estudio concluyó que el Modelo de Aceptación de Tecnología es eficaz para comprender la aceptación del e-learning y que la autoeficacia fue el factor más significativo, seguido de la norma subjetiva.

En lo referido al colectivo docente, resulta relevante y significativa la investigación llevada a cabo por Mailizar et al. (2021) con el objetivo de examinar los factores que influían en la intención de utilizar sistemas e-learning en su docencia. Este estudio también tomó como base el Modelo de Aceptación de la Tecnología (TAM) e incluyó la experiencia en e-learning como constructo adicional. Los resultados mostraron que la actitud hacia el uso y la experiencia fueron los constructos más significativos para predecir el uso.

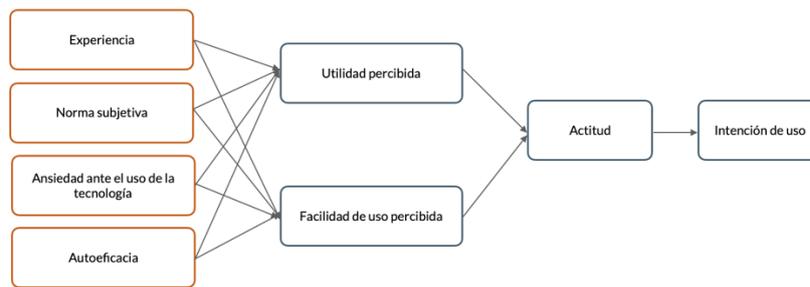
Los modelos de adopción y su relación con la educación mediada por la tecnología también han sido objeto de estudio durante la pandemia. En este sentido, es interesante la investigación de Sukendro et al. (2020) para la validación de un modelo ampliado con las condiciones facilitantes como factor externo. Este estudio mostró relaciones significativas entre las condiciones y la facilidad de uso percibida y entre estas y la utilidad percibida.

En la búsqueda de revisiones que sinteticen la relación entre el modelo de Davis (1989) y la enseñanza mediada por la tecnología, Abdullah y Ward (2016) realizaron un metaanálisis de 107 modelos de adopción con el objetivo de identificar los factores externos más utilizados en el contexto de la adopción del e-learning. Los resultados mostraron que la autoeficacia, la norma subjetiva, la ansiedad ante la tecnología, el disfrute percibido y la experiencia son los factores externos al modelo TAM más utilizados.

Con respecto al funcionamiento de estos constructos sobre las variables del modelo TAM, Abduyah y Ward (2016) concluyen que el mejor predictor de la facilidad de uso percibida de los sistemas de e-learning es la autoeficacia, seguida del disfrute y la experiencia. Respecto a la utilidad percibida, los constructos más determinantes fueron el disfrute percibido, la norma subjetiva y la autoeficacia.

Utilizando estos factores externos y los tamaños de efecto analizados, este estudio propuso el Modelo General Extendido de Aceptación de Tecnología para E-Learning (GETAMEL), uno de los modelos más significativos del campo dada su fundamentación teórica (Figura 5).

FIGURA 5. Modelo GETAMEL. Adaptado de Abduyah y Ward (2016)



1.3.3.2. TAM en la adopción de dispositivos móviles en educación

Entre los modelos de adopción tecnológica basados en TAM vinculados con el uso de dispositivos móviles en educación, destacan investigaciones relevantes como la desarrollada por Liu et al. (2010), quien modifica el modelo TAM mediante la introducción de factores relativos a la utilidad percibida a corto y largo plazo en el diseño. Los resultados mostraron que esta utilidad y la capacidad de innovación personal influyen significativamente en la intención de uso, mientras que la utilidad percibida a largo plazo afecta significativamente a la utilidad percibida a corto plazo.

Una segunda investigación destacada en esta relación entre dispositivos móviles y educación es la llevada a cabo por Sánchez-Prieto et al. (2017a). Este estudio presenta un modelo ampliado que introduce la autoeficacia y la ansiedad en el uso de dispositivos móviles, unos constructos ya referenciados en la anterior sección para sistemas e-learning. Los resultados alcanzados respaldaron las hipótesis relacionales propuestas y reflejaron que las relaciones más fuertes se encontraban entre la utilidad percibida y la intención conductual, entre la facilidad de uso y la utilidad percibida, y entre la autoeficacia y la facilidad de uso percibida.

En tercer lugar, es interesante referenciar el estudio desarrollado por Huang et al. (2007). Esta investigación amplió el modelo TAM con las dimensiones “disfrute” y “valor percibido de la movilidad” para tratar de aumentar su poder explicativo. Los resultados mostraron que las diferencias individuales poseen una influencia significativa en la intención de uso y que los dos nuevos constructos contribuyeron positivamente a predecir las intenciones de utilizar dispositivos móviles.

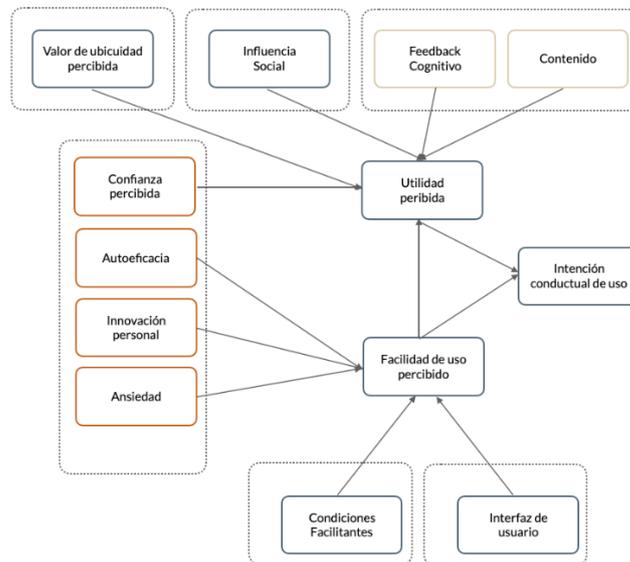
Los modelos de adopción específicos para este campo han sido sintetizados por Alsharida (2021) mediante una revisión sistemática de 64 modelos relevantes. Los principales hallazgos han confirmado que la autoeficacia es el factor más frecuente empleado en modelos TAM en el campo del m-learning, seguido por la norma subjetiva, el disfrute percibido, la ansiedad en el uso de dispositivos móviles, las condiciones facilitantes, la influencia social, la innovación y la satisfacción; respectivamente. Este estudio confirma también la tendencia creciente de publicación de modelos

TAM ampliados, lo que refleja una elevada confianza en la propuesta de Davis como base para el estudio de las intenciones de uso de dispositivos móviles en educación (Alsharida, 2021).

1.3.3.3. TAM en la adopción de dispositivos móviles en procesos de evaluación

En tercer lugar, en relación con la asociación entre dispositivos móviles y evaluación, son todavía muy reducidas las investigaciones centradas en el estudio de la adopción tecnológica (Alrofouh et al., 2019). En este ámbito, destacan como modelos TAM ampliados más relevantes para el estudio de la adopción el Modelo de Aceptación de la Evaluación basada en Dispositivos Móviles (MBAAM) (Nikou y Economides, 2017b) (Figura 6) y el Modelo de Aceptación de la Evaluación Móvil en Profesores (TAMBA) (Nikou y Economides, 2019).

FIGURA 6. Modelo MBAAM. Recuperado y adaptado de Nikou y Economides (2017b)



Los estudios realizados hasta el momento muestran que el Modelo de Aceptación Tecnológica es un soporte sólido para el estudio de la adopción en este campo específico, y tratan de confirmar el efecto de factores ya descritos en otros ámbitos y de nuevos factores como antecedentes de la intención de uso de dispositivos móviles en procesos de evaluación.

Por tanto, en base a la actual relevancia de estos modelos y considerando que investigaciones previas han profundizado en los factores vinculados con la evaluación, pero no con las características específicas del proceso; se puede afirmar que el desarrollo de los estudios de adopción para este campo es una línea futura de investigación.

En este epígrafe se ha descrito las características de la evaluación del aprendizaje, el desarrollo tecnológico en educación y evaluación, la implementación de dispositivos móviles en estos procesos y la introducción a los modelos de adopción tecnológica. Se presenta, a continuación, la justificación de la investigación propuesta.

1.4. Justificación

La Tesis Doctoral que se propone defiende, como punto de partida, la importancia de la evaluación mediada por dispositivos móviles, las diferencias específicas entre las modalidades formativa y sumativa de la evaluación, y los modelos de adopción tecnológica.

Investigaciones previas han explorado distintos aspectos del proceso evaluativo en entornos virtuales. Heil e Ifenthaler (2023) enfocaron su revisión sistemática en las características y modalidades evaluativas en la educación superior, destacando el potencial de la evaluación en el proceso educativo. Por su parte, Topuz et al. (2022) analizaron las características de los sistemas de evaluación en línea y concluyeron la falta de soporte de determinados sistemas para su uso en dispositivos móviles, los beneficios de los sistemas de gestión del aprendizaje (LMS) en la evaluación continua y la retroalimentación y el amplio catálogo de recursos disponible para evaluar en línea.

Otras revisiones han abordado la incorporación de la inteligencia artificial (Chiu et al., 2023; Samala et al., 2023) y la realidad virtual en procesos educativos y de evaluación (Ifanov et al., 2023); lo que muestra una tendencia creciente en la investigación sobre evaluación y tecnología (Noah y Das, 2021).

Por tanto, a pesar de la extensa literatura disponible sobre evaluación y uso tecnológico, la inclusión de dispositivos móviles con fines evaluativos en las aulas se enfrenta hoy a resistencias multifactoriales que comprenden desde limitaciones legislativas (Álvarez-Herrero, 2018) hasta barreras asociadas con la adopción docente (Hu et al., 2003). Esta reticencia de los docentes ha sido identificada como una de las principales limitaciones para una transformación efectiva de la docencia (Córica y García-Aretio, 2018) y de la evaluación (Nikou y Economides, 2019), siendo estos algunos de los actores clave que predeterminan la introducción de tecnologías en el aula. Esta resistencia conduce hoy a implementaciones esporádicas (Sánchez-Prieto et al., 2017b; Taherdoost, 2019) y convierte la tecnología educativa en un debate social (Grande-de-Prado et al., 2021).

En última instancia, la integración de la evaluación con dispositivos móviles y el estudio de la adopción docente de estos avances se erige como una línea de investigación crucial en el panorama educativo actual, con el objetivo de promover no solo la optimización del proceso de enseñanza-aprendizaje, sino también una inclusión tecnológica más efectiva.

Por este motivo, y dado que los modelos actuales de adopción tecnológica en este ámbito no refieren a características específicas del proceso de evaluación, se pretende en esta investigación diseñar un **modelo de adopción tecnológica para el estudio de los factores que condicionan el uso de dispositivos móviles en evaluación por parte del profesorado.**

En la figura 7 se presenta la organización y estructura de la Tesis Doctoral que se resume en introducción (1), objetivos, metodología y marco de trabajo (2), resultados de la investigación (3, cinco estudios derivados) y las conclusiones, las limitaciones y las líneas de investigación futuras (4).

FIGURA 7. Estructura de la Tesis Doctoral



Finalizada la introducción, que ha descrito la evaluación del aprendizaje, el papel de los dispositivos móviles en evaluación y los modelos de adopción tecnológica; la próxima sección presenta los objetivos, la metodología y el marco de trabajo de la investigación.

Diseño de la investigación:
Objetivos, metodología y marco de trabajo

Diseño de la investigación: **Objetivos, metodología y marco de trabajo**

La segunda sección presenta los objetivos establecidos en el marco de la investigación, la metodología utilizada y las etapas y desarrollos procedimentales llevados a cabo. En primer lugar, se describen los objetivos.

2.1. Objetivos

La propuesta que aquí se realiza pretende contribuir al desarrollo de la literatura académica sobre el objeto de estudio y trata de resolver un problema de investigación derivado de la falta de conocimiento sobre el proceso de adopción de dispositivos móviles en procesos de evaluación por parte del profesorado. En consecuencia, se propone como objetivo general:

Diseñar un modelo de adopción tecnológica para el estudio de los factores que condicionan el uso de dispositivos móviles en evaluación por parte del profesorado.

Este objetivo general se desglosa en cinco objetivos específicos que guardan relación los estudios realizados en el contexto de esta tesis:

- OE1. Analizar la relación entre evaluación y calidad, y su relevancia en el proceso de enseñanza-aprendizaje mediado por la tecnología.
- OE2. Describir el estado actual de las investigaciones en materia de e-evaluación en los últimos años, sus tendencias y las líneas futuras de investigación.
- OE3. Identificar los principales modelos de adopción tecnológica de dispositivos móviles en procesos de evaluación.

OE4. Validar una propuesta inicial de modelo de adopción tecnológica de dispositivos móviles en evaluación.

OE5. Determinar, de manera empírica, el funcionamiento de los factores incluidos en el modelo inicial.

El documento presentado sintetiza la elaboración de una de Tesis Doctoral en formato compendio de publicaciones. En la tabla 1 se muestra la relación entre los objetivos específicos y las publicaciones derivadas de la Tesis (en formato artículo), dos de las cuales se encuentran en fase de revisión a fecha de depósito.

TABLA 1. Publicaciones en revista en función de los objetivos específicos de la Tesis Doctoral

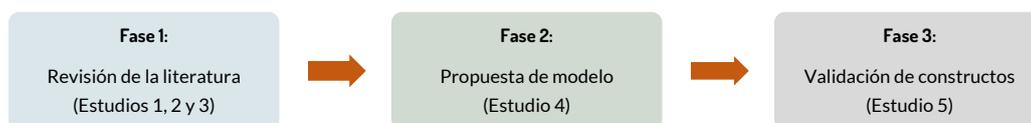
O.E.	Publicación derivada
OE1	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-learning en Educación Superior: Una revisión sistemática de la literatura. <i>Education in the Knowledge Society (EKS)</i> , 23. https://doi.org/10.14201/eks.26986 Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2021). Calidad en e-learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. <i>RIED. Revista Iberoamericana de Educación a Distancia</i> , 24(2). https://doi.org/10.5944/ried.24.2.29073
OE2	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. <i>Education and Information Technologies</i> . https://doi.org/10.1007/s10639-023-12117-5
OE3	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (En revisión). Mobile-based assessment acceptance: A systematic literature review in the educational context
OE4	Ortiz-López, A.; Olmos-Migueláñez, S.; Sánchez-Prieto, J.C. (En prensa). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En E. Chiner (Ed), <i>Nuevas tendencias interdisciplinarias en Educación y Conocimiento</i> . Tirant Lo Blanch
OE5	Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En prensa). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. <i>Journal of New Approaches in Educational Research</i> . Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En revisión). Mobile technology in assessment processes: Comparing its usefulness and antecedents in CB-SEM

Elaboración propia

2.2. Fases de la investigación

Los procesos realizados dan respuesta a los objetivos establecidos a través de tres fases que comprenden la revisión de la literatura (Fase 1), la propuesta inicial de un modelo (Fase 2) y la validación empírica de los constructos del modelo diseñado (Fase 3) (Figura 8).

FIGURA 8. Fases de la investigación



Estas fases se desarrollan a través de cinco estudios que formalizan la investigación y que dan respuesta a los objetivos establecidos:

1. El **primer estudio** aborda la evaluación de la calidad como componente inherente a la evaluación mediante una revisión sistemática de la literatura y la propuesta de un nuevo modelo para su determinación.
2. El **segundo estudio** presenta un mapeo sistemático de la literatura sobre e-evaluación y las principales tendencias de investigación.
3. El **tercer estudio** muestra una revisión sistemática de la literatura sobre los principales modelos de adopción tecnológica de dispositivos móviles en procesos de evaluación.
4. El **cuarto estudio** centra su desarrollo en la propuesta y validación por expertos de un nuevo modelo para la determinación de la adopción tecnológica en evaluación que será aplicado.
5. El **quinto estudio**, por último, valida los constructos incluidos en el modelo inicial en una muestra de futuros docentes para mejorar el modelo planteado en la fase anterior.

A continuación, se presenta la metodología seguida en los estudios derivados de la investigación y el marco de trabajo al que se adscribe la tesis doctoral.

2.3. Metodología

Durante el desarrollo de la tesis se ha seguido un aparato metodológico que integra distintos métodos, procedimientos y técnicas de análisis: mapeos y revisiones de la literatura, validaciones de modelos y análisis estadísticos (modelados de ecuaciones estructurales).

2.3.1. Mapeos y revisiones sistemáticas de la literatura

Los mapeos y las revisiones sistemáticas de la literatura son dos procesos para la búsqueda y selección crítica de estudios centrados en la recopilación de estos para su análisis a través de procedimientos objetivos y estructurados (García-Peñalvo, 2017). Estos métodos posibilitan la identificación, evaluación e interpretación del trabajo de investigadores, académicos y profesionales en un campo elegido (Fink, 1998).

El mapeo y la revisión sistemática son procesos similares pero que se conforman en torno a objetivos claramente diferenciados. El primer proceso tiene como objetivo la categorización de las publicaciones en función de sus hallazgos clave y sus datos bibliométricos (revistas, año de publicación autores, etc.); mientras que las revisiones sistemáticas promueven, como proceso esencial, un análisis y una síntesis más profunda de los trabajos seleccionados (Sáenz, 2001).

Ambos procesos de revisión constituyen, *per se*, procesos de indagación de carácter deductivo que incorporan todas las exigencias formales y metodológicas del paradigma científico dados sus

métodos preestablecidos y la replicabilidad que posibilitan (Gisbert y Bonfill, 2004), permitiendo también la reflexión profunda sobre el estado del objeto de estudio en un momento específico y facilitando la detección de nuevas líneas de investigación (Cooper, 2016). Las revisiones sistemáticas de la literatura siguen el esquema procedimental propuesto por Kitchenham y Charters (2007) y establecen una secuencia de pasos, que son:

1. Establecimiento de preguntas de investigación.
2. Diseño de criterios de inclusión y exclusión.
3. Selección de bases de datos.
4. Diseño de la cadena de búsqueda.
5. Selección de publicaciones en base a los criterios establecidos.
6. Extracción de la información para dar respuesta a las preguntas de investigación.

Este proceso de revisión y selección de publicaciones se ha ajustado también a las fases propuestas por Manterola et al. (2013), secuenciadas en: extracción de resultados (1), eliminación de duplicados (2), aplicación de criterios de inclusión (3) y la aplicación de criterios de calidad (4). Además, las revisiones presentadas en esta Tesis Doctoral se ajustan al diagrama de flujo PRISMA (Page et al., 2021) para este tipo de procesos, una lista para la comprobación de las fases y criterios cuyo cumplimiento debe ser garantizado por una revisión sistemática de la literatura.

2.3.2. Proceso de validación de modelos

Los dos procesos de validación de contenido por jueces expertos realizados han seguido las líneas propuestas por García-Llorente et al. (2019) para su desarrollo. Las validaciones de contenido han sido llevadas a cabo por evaluadores expertos nacionales e internacionales en las temáticas de los dos modelos propuestos (calidad, adopción tecnológica y evaluación).

Las validaciones han constado de tres fases. La primera fase ha consistido en la evaluación individual de los ítems. Para ello, se ha propuesto que los expertos evalúen, partiendo de las recomendaciones de Escobar-Pérez y Cuervo (2008) en una escala tipo Likert, su coherencia, relevancia y claridad.

En segundo término, se ha solicitado a los expertos una validación de los constructos que componen el modelo. Para esto, se ha incluido otra escala en la que medir su validez y un espacio para las consideraciones cualitativas oportunas. Finalmente, se concluye con una valoración global del modelo en la que evaluar la adecuación, la validez y la relevancia de los constructos propuestos. Además, se vuelve a incluir un nuevo apartado final de comentarios y observaciones para el evaluador.

2.3.3. Análisis estadísticos: Modelos de ecuaciones estructurales (CB-SEM y PLS-SEM)

La investigación ha empleado modelos de ecuaciones estructurales (SEM) para la validación de la propuesta. Estos modelos son un método de análisis de datos que permite analizar relaciones complejas entre constructos e indicadores. Este proceso metodológico puede partir de dos enfoques diferenciados para su cálculo, uno basado en el análisis de la matriz de covarianzas (*Covariance-Based – Structural Equation Modeling*, CB-SEM), y un segundo basado en el análisis de mínimos cuadrados parciales, denominado PLS-SEM (*Partial Least Squares – Structural Equation Modeling*) (Hair et al., 2021).

La primera técnica, el análisis de ecuaciones estructurales basado en la covarianza (CB-SEM), es una técnica centrada en estimar los parámetros del modelo que minimizan las diferencias entre la matriz de covarianza de la muestra observada (calculada antes del análisis) y la matriz de covarianza estimada una vez confirmado el modelo teórico revisado (Hair et al., 2017). Esta técnica sigue un proceso de dos fases en las que se determina la validez del modelo de medida (fiabilidad y validez) y la evaluación del modelo estructural atendiendo a las relaciones entre los constructos propuestos (Hair et al., 2012) Los análisis correspondientes a esta modalidad han sido desarrollados a través del software estadístico JASP (v. 0.17.1) (Love et al., 2019).

La segunda técnica empleada ha sido el análisis de ecuaciones estructurales basado en mínimos cuadrados parciales (PLS-SEM). Esta técnica aplica la regresión de mínimos cuadrados con el objetivo de minimizar los errores de medida de los constructos endógenos a través de la estimación de coeficientes (relaciones del modelo) para predecir (dentro de la muestra) y maximizar los valores de la varianza de los constructos (Hair et al., 2021). La técnica PLS-SEM consta de un análisis en varias etapas: la evaluación del modelo de medida (análisis de validez convergente, de fiabilidad, de colinealidad y comprobación de cargas) y la evaluación del modelo estructural (Hair et al., 2021).

PLS-SEM posibilita, además, el modelado de constructos tanto formativos como reflectivos (Ajms, 2015; Bollen, 2011). Por un lado, en la teoría de la medición formativa, las medidas causan el dominio del constructo, “indicating a predictive (causal) relationship” [indicando una relación predictiva (causal)] (Hair et al., 2021, p. 8). Por otro lado, los modelos con constructos reflectivos indican “the assumption that the construct “causes” the measurement (more precisely, the covariation) of the indicator variables” [la asunción de que el constructo “causa” la medición (más exactamente, la covariación) de las variables indicadoras] (Hair et al., 2021, p. 8). Los constructos reflectivos llevan asociado un error con cada indicador, caso que no sucede con constructos formativos (Hair et al., 2021).

En resumen, en modelados formativos son los indicadores los que causan el constructo, mientras que desde una perspectiva reflectiva los indicadores son causados por el constructo (Hair et al.,

2014). Los análisis centrados en esta técnica se han llevado a cabo a través del software estadístico *SmartPLS* (v. 3.2.9) (Hair et al., 2019; Sarstedt et al., 2016).

En la tabla 2 se muestra la relación entre las metodologías y técnicas de análisis empleadas y las publicaciones derivadas.

TABLA 2. Publicaciones derivadas en relación con la metodología y técnica de análisis

Metodología / Técnica de análisis	Publicación derivada
1. Mapeos y revisiones sistemáticas de la literatura	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-learning en Educación Superior: Una revisión sistemática de la literatura. <i>Education in the Knowledge Society (EKS)</i> , 23. https://doi.org/10.14201/eks.26986
	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. <i>Education and Information Technologies</i> . https://doi.org/10.1007/s10639-023-12117-5
	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (En revisión). Mobile-based assessment acceptance: A systematic literature review in the educational context
2. Validación de contenido	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2021). Calidad en e-learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. <i>RIED. Revista Iberoamericana de Educación a Distancia</i> , 24(2). https://doi.org/10.5944/ried.24.2.29073
	Ortiz-López, A.; Olmos-Migueláñez, S.; Sánchez-Prieto, J.C. (En prensa). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En E. Chiner (Ed), <i>Nuevas tendencias interdisciplinarias en Educación y Conocimiento</i> . Tirant Lo Blanch
3. Análisis de ecuaciones estructurales basado en mínimos cuadrados parciales (PLS-SEM)	Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En prensa). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. <i>Journal of New Approaches in Educational Research</i> .
4. Análisis de ecuaciones estructurales basado en la covarianza (CB-SEM)	Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En revisión). Mobile technology in assessment processes: Comparing its usefulness and antecedents in CB-SEM

Elaboración propia

2.4. Marco de trabajo para el desarrollo de la Tesis Doctoral

Esta Tesis se enmarca en el Programa de Doctorado Formación en la Sociedad del Conocimiento (García-Peñalvo, 2014), un programa doctoral interdisciplinar de la Universidad de Salamanca que tiene como objetivo analizar los procesos de enseñanza-aprendizaje que se desarrollan en el paradigma de la Sociedad del Conocimiento a través de un total de nueve líneas de investigación abiertas.

En segundo término, la Tesis Doctoral que se presenta se sitúa dentro del contexto del Instituto Universitario de Ciencias de la Educación (IUCE) de la Universidad de Salamanca y en el Grupo de Investigación en InterAcción y e-learning (GRIAL) (García-Peñalvo et al., 2019).

Además, han sido soporte para el desarrollo de la Tesis todos aquellos recursos facilitados por la Universidad de Salamanca, así como el acceso a los distintos repositorios científicos y a diferentes herramientas y software educativo necesario para su desarrollo.

Resultados:

Estudios derivados de la investigación

PUBLICACIONES DERIVADAS:

Anexo 1

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2021). Calidad en e-Learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2). <https://doi.org/10.5944/ried.24.2.29073>

Anexo 2

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-Learning en Educación Superior: Una revisión sistemática de la literatura. *Education in the Knowledge Society (EKS)*, 23. <https://doi.org/10.14201/eks.26986>

Anexo 3

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12117-5>

Anexo 4

Ortiz-López, A., Olmos-Migueláñez, S., Sánchez-Prieto, J.C. (En prensa). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En E. Chiner (Ed.), *Nuevas tendencias interdisciplinares en Educación y Conocimiento*. Tirant Lo Blanch

Anexo 5

Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En prensa). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. *Journal of New Approaches in Educational Research*.

Anexo 6

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (En revisión). Mobile-based assessment acceptance: A systematic literature review in the educational context.

Anexo 7

Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En revisión). Mobile technology in assessment processes: Comparing its usefulness and antecedents in CB-SEM.

Estudio 1:**Calidad y evaluación de la calidad en enseñanzas medidas por la tecnología****Anexo 1**

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2021). Calidad en e-learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2). <https://doi.org/10.5944/ried.24.2.29073>

Anexo 2

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-learning en Educación Superior: Una revisión sistemática de la literatura. *Education in the Knowledge Society (EKS)*, 23. <https://doi.org/10.14201/eks.26986>

3.1. Evaluación de la calidad: Dos estudios derivados

Este primer apartado de los resultados presenta una revisión sistemática de la literatura sobre la evaluación de la calidad en enseñanzas virtuales y una propuesta de modelo para su evaluación en educación superior, que pueden ser consultados en los anexos 1 y 2.

La evaluación y la calidad son dos términos presentes en el léxico educativo y en la práctica regulatoria de los sistemas e instituciones de enseñanza, dos conceptos que deben ser entendidos de forma dual y permanentemente conectada para una adecuada delimitación. En palabras de Municio (2004):

se dice evaluación del aprendizaje, entendiendo que se quiere decir evaluación de la “calidad del aprendizaje”. La ausencia del término calidad implica que evaluación y calidad son dos conceptos tan estrechamente relacionados que no es posible la existencia de uno sin el otro. En el caso de no incluir uno de ellos se da el otro por implícito en el utilizado” (p. 102).

Esta relación bidireccional garantiza el adecuado desarrollo del proceso educativo, ya que la evaluación es el componente que contribuye, en mayor medida, a la mejora de la calidad educativa (Kellaghan y Greaney, 2001; Koslowski, 2006; Olmos-Migueláñez, 2008); pero también la calidad es un componente fundamental sujeto a evaluación en educación (García-Soto et al., 2020).

En la nueva realidad educativa hacia la que se dirigen hoy los sistemas, inmersos en procesos continuos de cambio y adaptación a nuevos contextos sociales, institucionales y tecnológicos, la evaluación de la calidad con la que se desarrollan los procesos de enseñanza-aprendizaje concita un interés creciente por parte de la comunidad académica. Esto se refleja, por ejemplo, en el elevado número de agencias destinadas a tal fin que operan en la actualidad o el elevado número de modelos disponibles para su evaluación (Brown et al., 2018; Fernández, 2005).

En este escenario, la incorporación de la tecnología como herramienta clave en la educación ha añadido una dimensión adicional a la evaluación de la calidad. La tecnología no solo ha transformado el proceso de enseñanza-aprendizaje, sino que también ha generado nuevas demandas y expectativas en cuanto a la calidad de estos procesos. Por este motivo, es fundamental abordar la intersección entre calidad, evaluación y tecnología con la finalidad de responder a los nuevos desafíos de la educación.

La investigación realizada no solo busca comprender la dinámica de la evaluación de calidad en este nuevo contexto, sino que busca alcanzar un marco sólido y consensuado para su evaluación, reconociendo la influencia que la tecnología tiene en la calidad. En última instancia, la interrelación entre calidad, evaluación y tecnología se posiciona como un estudio fundamental para entender los desafíos, comprender las limitaciones docentes y superar las barreras actuales que condicionan el alcance de una educación de calidad mediada por la tecnología.

3.1.1. Evaluación de la calidad: Una revisión sistemática de la literatura en Educación Superior

Se presenta una síntesis de un artículo derivado de la investigación (anexo 2), que tiene como objetivo conocer los modelos existentes sobre evaluación de la calidad en los últimos años mediante la realización de una revisión sistemática de la literatura, con el objetivo de sintetizar la producción existente y ofrecer los últimos avances sobre evaluación de la calidad en la educación mediada por la tecnología.

3.1.1.1. Metodología de la revisión sistemática

Para el desarrollo de la revisión se han seguido las recomendaciones propuestas por Kitchenham y Charters (2007) para revisiones sistemáticas de la literatura y se ha tomado el método PRISMA (Page et al., 2021) como base para el desarrollo de las fases. Las preguntas de investigación propuestas en la revisión se han centrado en conocer las dimensiones relativas a la evaluación de la calidad en e-learning, así como los procesos mediante los que se evalúa dicha calidad. Por otro lado,

se han descrito las principales aportaciones de las investigaciones y las vías de investigación que proponen.

Para la revisión se han propuesto seis criterios de inclusión y exclusión. En base a estos, y tomando los repositorios Web of Science y Scopus para la búsqueda, identificación y selección de publicaciones, se ha empleado la siguiente cadena de búsqueda: (“Evaluación Calidad” OR “Quality Assessment” OR “Quality Evaluation”) AND (“elearning OR “e-learning”) AND (“Educación Superior” OR “Universidad” OR “Higher Education” OR “University”).

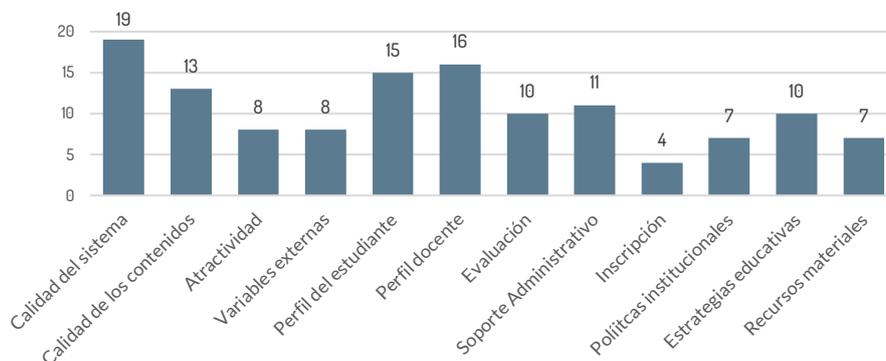
Con respecto a la limitación temporal de las publicaciones, se han recogido las comprendidas entre 2015 y 2021. El proceso completo de revisión con la extracción de publicaciones, la aplicación de criterios y la respuesta a cada pregunta puede ser consultada en: <https://bit.ly/3kFz7IM>. La búsqueda inicial permitió recuperar un total de 697 publicaciones en Scopus y 56 en Web of Science, con 38 duplicados eliminados entre ambas para alcanzar un total de 715 investigaciones. Después de la aplicación de criterios, se eliminaron 690 investigaciones y se alcanzaron un número final de 25 publicaciones que informan sobre los modelos de evaluación de la calidad disponibles en la literatura.

3.1.1.2. Resultados de la revisión sistemática

En la revisión se identificaron varias de las dimensiones clave abordadas en los estudios (Figura 9). Estas dimensiones incluyen, entre otras las características de los estudiantes, la función docente, la infraestructura tecnológica o el entorno y feedback proporcionado. En cuanto al proceso de evaluación de calidad, se categorizaron los estudios en cuatro grupos principales:

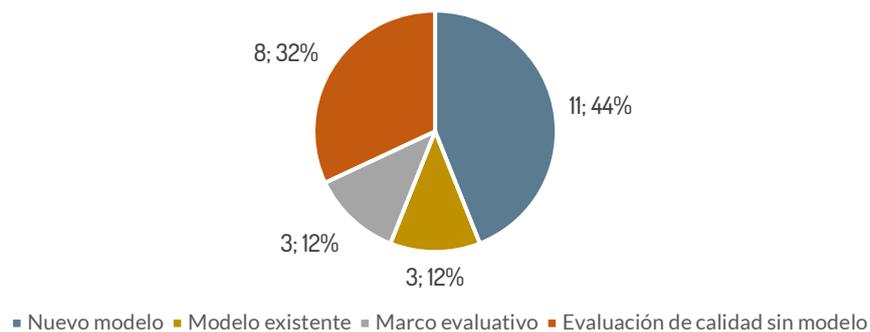
- 1) aquellos que generan un nuevo modelo,
- 2) los que aplican un modelo existente,
- 3) los que proponen un marco evaluativo,
- 4) los que evalúan la calidad sin desarrollar un nuevo modelo.

FIGURA 9. Dimensiones identificadas en los modelos de calidad



La segunda pregunta de investigación describe el proceso seguido en las investigaciones, categorizados en cuatro grupos principales: aquellas investigaciones que generan un modelo, aquellas que aplican un modelo ya existente, aquellas que proponen un marco evaluativo y aquellas que evalúan la calidad sin la generación de un nuevo modelo explícito y complejo; de forma que todas las publicaciones quedan categorizadas en una u otra categoría (Figura 10).

FIGURA 10. Procesos seguidos en las investigaciones



En tercer lugar, la revisión ha abordado las principales aportaciones de los estudios. Para ello, se han descrito cinco grupos de aportaciones que categorizan el total de resultados. De ellos se extrae como línea principal, y por tanto tendencia en la investigación, la búsqueda constante de nuevas dimensiones que puedan afectar y condicionar la calidad del e-Learning, estando presente en prácticamente la mitad de las investigaciones. En contraposición, son líneas incipientes la búsqueda de protocolos y mecanismos que posibiliten la comparación objetiva entre distintos sistemas e-Learning, y la propuesta de un modelo que aúne y categorice los principales avances.

En conclusión, la revisión muestra un interés creciente en la evaluación de calidad en e-learning, pero aún no existe un modelo único y ampliamente aceptado. La variedad de dimensiones y enfoques hace que la determinación de la calidad dependa del modelo de evaluación elegido. En virtud de esta coyuntura, se sugiere la creación de un nuevo modelo que integre las propuestas existentes y permita un marco más amplio y consensado.

3.1.2. Evaluación de la calidad: Propuesta de un modelo para e-Learning en educación superior

Con el objetivo de consolidar y sintetizar las diversas perspectivas y enfoques relacionados con la evaluación de la calidad en el contexto del e-learning, se ha llevado a cabo la propuesta de un nuevo modelo de evaluación. Este modelo se encuentra disponible como publicación derivada de la Tesis Doctoral en el anexo 1.

El nuevo modelo de evaluación de calidad en e-learning, basado en los modelos existentes, se estructura en dos secciones diferenciadas. La primera sección está destinada a la evaluación de

cuatro dimensiones fundamentales (calidad de la institución, calidad de la función docente, calidad del sistema de aprendizaje y calidad de la evaluación del programa). La segunda parte del modelo se dedica a la evaluación cualitativa de la calidad por parte del evaluador.

Finalizado el diseño, se llevó a cabo un proceso de validación de contenido basado en las recomendaciones de García-Llorente et al. (2019) y Galicia et al. (2017). Este proceso incluyó la evaluación de la suficiencia y de la relevancia de ítems para las dimensiones, así como la valoración de la coherencia, relevancia y claridad de los propios ítems, evaluándose también la suficiencia y representatividad de las subdimensiones. Para el proceso de validación de contenido se seleccionó a un total de cuatro evaluadores expertos en el ámbito educativo y tecnológico, así como en el campo de la calidad en la evaluación.

En el análisis de la validación, se eliminaron ítems con puntuaciones inferiores a 2, se realizaron modificaciones profundas en los ítems comprendidos entre 2 y 3 y se modificó la formulación de aquellos con valores inferiores a 3,5.

En una escala 1-4 para el análisis de las dimensiones, los resultados indican una media de suficiencia elevada (3,84) y una representatividad positiva de los constructos (3,92), lo que sugiere una formulación sólida de las dimensiones. En cuanto a las subdimensiones, los resultados también son favorables, con una medida media de suficiencia de 3,63 y una representatividad de 3,74. En lo que respecta a la evaluación individual de los ítems, los resultados globales fueron altamente positivos. El valor medio de coherencia para los ítems fue de 3,90, con una relevancia de 3,85 y un valor de claridad de 3,68.

Desde un enfoque cualitativo, los evaluadores proporcionaron observaciones tanto sobre las dimensiones específicas como sobre el instrumento general. Las observaciones principales se centran en la propuesta de reemplazo de términos en determinados ítems. También se abordó la sugerencia de incorporar subdimensiones e ítems adicionales para enriquecer la medida del constructo, junto con la recomendación de eliminar o reubicar algunos de ellos en el modelo final.

La validez de contenido del instrumento, de las dimensiones, y de los ítems, se determinó con el coeficiente de validez de contenido (CVC) (Hernández-Nieto, 2002), admitiéndose aquellos con puntuaciones superiores a 0,7 (Pedrosa et al., 2013). Los estadísticos pueden ser consultados en: <http://bit.ly/3PAgfUr>.

Después del proceso de validación, el modelo cuenta con las cuatro dimensiones, 14 subdimensiones y un total de 80 indicadores (Tabla 3). Este modelo se encuentra disponible al completo en: <http://bit.ly/3EyVzWv>

TABLA 3. Dimensiones del modelo propuesto. Adaptado de Ortiz-López et al. (2022)

Dimensión	N Ítems
Dimensión 1: Calidad de la institución	
Política Institucional	6
Contexto Institucional	4
Respuesta institucional	5
Coordinación pedagógica	4
Dimensión 2: Calidad docente	
Función docente	8
Estrategias docentes	5
Materiales y recursos docentes	8
Tutoría	5
Dimensión 3: Calidad del sistema de aprendizaje	
Funcionamiento	10
Accesibilidad	5
Dimensión 4: Calidad de la evaluación	
Evaluación de la preparación previa	3
Evaluación procesual	4
Evaluación final	9
Evaluación de la calidad del programa	4

3.1.3. Principales aportaciones y prospectiva de los estudios

En resumen, tras explorar en profundidad los modelos actuales centrados en la evaluación de la calidad en e-learning, se confirma la importancia de su evaluación en el ámbito educativo. Esta evaluación se erige como uno de los factores críticos para impulsar una mejora en los procesos educativos y el desarrollo efectivo de la evaluación en nuevos entornos de aprendizaje.

Si se garantiza un proceso de evaluación de alta calidad en entornos virtuales, y si los docentes adoptan las adaptaciones necesarias en estos procesos, la educación se encontrará más próxima al avance hacia los nuevos ecosistemas de aprendizaje (García-Holgado et al., 2019; García-Holgado y García-Peñalvo, 2017; García-Peñalvo et al., 2015).

En este sentido, y tras los estudios realizados, se pueden apuntar como posibles futuras líneas de investigación:

- **Ampliación de los modelos de calidad:** En primer lugar, la investigación debe diseñar y validar nuevos modelos para la evaluación de la calidad de enseñanzas mediadas por la tecnología que aborden el mayor número posible de dimensiones vinculadas con este proceso.
- **Desarrollo del modelo propuesto:** El nuevo modelo diseñado ha generado un marco para la evaluación de la calidad que sintetiza los principales modelos actuales. Investigaciones futuras deberán realizar la aplicación real del modelo para comprobar su funcionamiento en procesos formativos mediados por la tecnología.

- **Simplificación del proceso evaluativo:** En este sentido, los modelos futuros de evaluación de la calidad deben simplificar sus procedimientos para posibilitar la realización de este proceso con un menor número de recursos empleados y un menor número de agentes implicados.

La siguiente sección describe el segundo estudio derivado, que da respuesta al estado actual del campo de la e-evaluación mediante un mapeo sistemático de la literatura.

Estudio 2:**E-evaluación: Un mapeo sistemático de la literatura****Anexo 3**

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12117-5>

3.2. Síntesis del mapeo de la literatura sobre e-evaluación

En este epígrafe se presenta un mapeo de la literatura sobre las publicaciones relativas a la evaluación mediada por la tecnología comprendidas entre 2017 y 2021. Esta revisión puede consultarse en el anexo 3.

El mapeo realizado ha permitido ampliar el campo de los estudios sobre e-evaluación y agrupar las publicaciones disponibles en la literatura en el marco temporal establecido. De esta forma, se posibilita la identificación de las publicaciones actuales sobre evaluación mediada por la tecnología y la detección de nuevas vías de investigación una vez analizados los últimos avances (Paul y Barari, 2022). Se describe, por tanto, la e-evaluación sin restringir la búsqueda a un nivel educativo específico, a una modalidad evaluativa o a un proceso concreto.

3.2.1. Metodología del mapeo

La metodología seguida para esta revisión ha consistido en un mapeo sistemático de la literatura. Una explicación más exhaustiva de la metodología puede consultarse en la sección correspondiente de este documento, así como en la publicación derivada.

3.2.1.1. Preguntas de investigación y proceso seguido

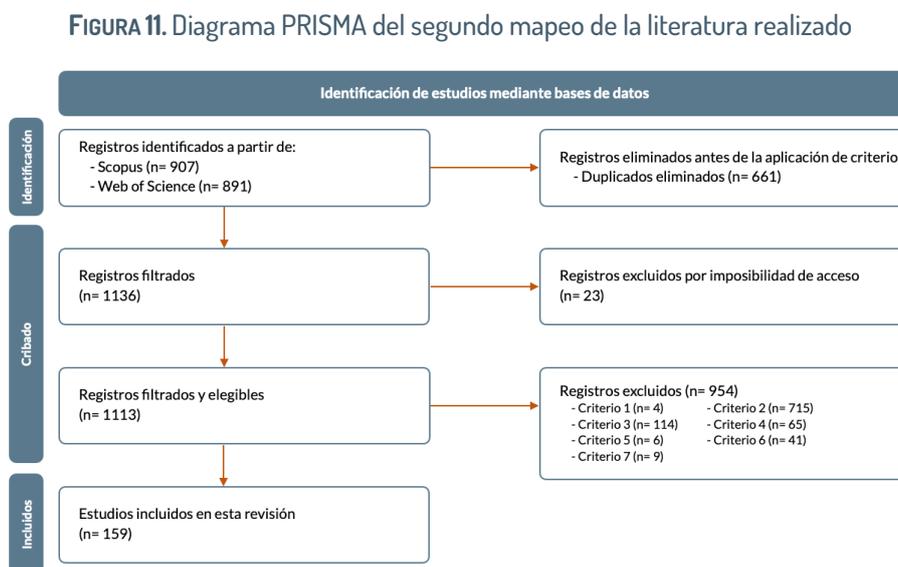
En la definición de las preguntas de investigación se plantearon cuatro cuestiones: la relevancia de las investigaciones (1), el estado actual del campo en los últimos cinco años (2), la concepción y delimitación de los estudios sobre la e-evaluación (3) y las tendencias futuras de investigación (4). Para el proceso de selección se establecieron los siguientes criterios de exclusión:

- 1) No se aceptó ninguna publicación que no estuviera recogida en una revista científica indexada en Web of Science o Scopus.
- 2) Se excluyeron los artículos no centrados en la evaluación en entornos virtuales.
- 3) Se excluyeron los artículos que no presentaban metodologías empíricas.
- 4) Se excluyeron los artículos no publicados en inglés.
- 5) Se excluyeron los artículos que no describían la población y la muestra de la investigación.

Para la búsqueda y selección de publicaciones se emplearon las bases de datos Web of Science y Scopus. La cadena de búsqueda empleada incluyó términos clave relacionados con la e-evaluación: (“e-assessment” OR “eassessment” OR “Computer Based Assessment” OR “Mobile Based Assessment”). La búsqueda realizada se limitó a los últimos cinco años desde la fecha de búsqueda (2017-2021).

El proceso de selección de publicaciones se realizó en noviembre de 2021. Inicialmente se obtuvieron 1798 resultados, una cifra que se redujo a 1137 publicaciones tras la eliminación de duplicados. Finalmente, se incluyeron 159 artículos tras la aplicación de los criterios de exclusión (Figura 11).

A continuación, y concluida la fase metodológica, se presentan los resultados de la investigación. El proceso completo de la revisión realizada puede ser consultado en: <http://bit.ly/3N8xBXf>.



3.2.2. Resultados del estudio de mapeo

Los resultados más relevantes del mapeo realizado se presentan de forma sintética en cuatro categorías de análisis planteadas a raíz de las preguntas de investigación iniciales.

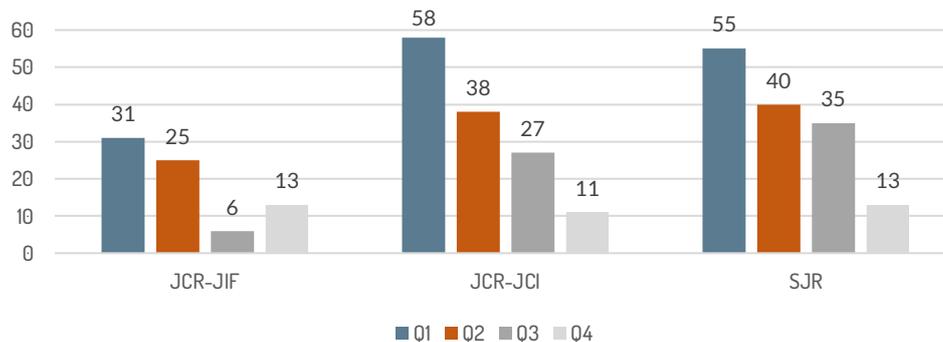
3.2.2.1. Relevancia de los estudios actuales sobre e-evaluación

La relevancia de los estudios incluidos en esta revisión se evaluó a través de dos criterios: el número de citas recibidas y los indicadores bibliométricos de las publicaciones.

Para calcular el número de citas recibidas se empleó Scopus como base de datos de búsqueda en enero de 2023. En promedio, se determinó una citación media de 14,21 citas por documento (para un total de 145 publicaciones incluidas, 14 no disponibles en Scopus). Las publicaciones más citadas abordan la adopción tecnológica de la e-evaluación y fueron desarrolladas por Nikou y Economides (2017a, 2017b), acumulando un total de 354 citas recibidas entre las dos publicaciones (181 y 173, respectivamente).

El análisis bibliométrico concluyó con la extracción de las métricas de cada revista correspondientes al año de publicación del artículo. Los resultados mostraron 75 revistas indexadas en *Journal Impact Factor* (JIF) (31 en el primer cuartil, 25 en el segundo, 6 en el tercero y 13 en el cuarto) y 134 en presentes en la clasificación *Journal Citation Index* (JCI) (58 en primer cuartil, 38 en el segundo, 27 en el tercero y 11 en el cuarto). En lo relativo al índice *Scimago Journal & Country Rank* (SJR), se encontraron 143 publicaciones indexadas (Figura 12).

FIGURA 12. Clasificación de las revistas indexadas en JIF, JCI y SJR



3.2.2.2. Relevancia de los estudios actuales sobre e-evaluación

Para comprender la evolución del campo de la e-evaluación, se tomaron el año de publicación de las investigaciones, el país de realización del estudio, los autores, las revistas y los enfoques metodológicos seguidos. En términos de temporalidad, se observó un crecimiento constante en el número de investigaciones pasando de 28 estudios en 2017 a 38 en 2021.

En cuanto al país de origen de las investigaciones, Estados Unidos lidera el campo con 24 estudios publicados. Los autores más representativos, de un total de 486, son Nikou y Economides. La revista *International Journal of Emerging Technologies in Learning* fue la más destacada con un total de siete artículos incluidos en esta revisión. En lo referido a los niveles educativos, destaca el ámbito de la educación superior, con 133 de las 159 investigaciones incluidas en la revisión.

3.2.2.3. Contextualización de las investigaciones en e-evaluación

En tercer lugar, se abordaron las principales características de las investigaciones. Para la delimitación del objeto de estudio se establecieron ocho categorías excluyentes. La categoría más abordada se centró en el análisis de rendimiento con un total de 48 investigaciones que exploraron este área. Estas categorías se corresponden con diversos estudios:

- centrados en actitudes: 39 investigaciones,
- de aceptación tecnológica. 15 investigaciones,
- relativos a la eficacia: 8 investigaciones,
- referentes a procesos de innovación: 17 investigaciones,
- focalizados en la supervisión de procesos evaluativos: 8 investigaciones,
- orientados al rendimiento: 48 investigaciones,
- orientados a la calidad: 4 investigaciones,
- de satisfacción: 20 investigaciones.

Metodológicamente, 104 investigaciones han desarrollado estudios de corte cuantitativo, 29 estudios cualitativos y 26 investigaciones mixtas. Con respecto a las técnicas e instrumentos aplicados en las investigaciones bajo dichos enfoques metodológicos, se han identificado 78 estudios que aplican cuestionarios, 68 centrados en pruebas objetivas, 19 entrevistas, 9 procesos de validación de cuestionarios, 8 grupos focales y 4 procesos de observación.

En términos de herramientas tecnológicas se pone de relieve el predominio de los LMS como principal recurso abordado (55% del total de investigaciones). Entre estos sistemas destaca la presencia de Moodle (42 investigaciones), Blackboard (14) y Canvas (3), junto con otras herramientas propias de las instituciones educativas.

Respecto a las investigaciones basadas en el empleo de tecnología móvil, los estudios se han centrado en aplicaciones destinadas a la evaluación (Examsoft, tExam), para la organización de la docencia (Snappet, Tronclass, Socrative), aplicaciones para el trabajo en equipo y coordinado (Slack, Viber, WeChat) o aplicaciones para el aprendizaje de campos específicos del conocimiento (GeoGebra, GotItLanguage), entre otras.

El resto de las herramientas representan el 31% del total, algunas orientadas a e-evaluación (como Kahoot, Dewis o JCQuest), y otras destinadas a la supervisión, autenticación y seguridad (Proctorio, Tinfollec). Por otro lado, se identifican también estudios basados en las herramientas de Google para la docencia o en software destinado a la comunicación (Microsoft Teams, Yammer o WebEX; entre otras).

3.2.2.4. Líneas de investigación futuras en e-evaluación

Finalmente, se identificaron las vías futuras de investigación en e-evaluación identificadas en las investigaciones analizadas. En este sentido, destaca la línea orientada al desarrollo metodológico de la e-evaluación, incluyendo la creación de nuevo software de calidad y la mejora de la fiabilidad de las herramientas utilizadas. La segunda línea más compartida, entre los estudios abordados, se relacionó con la replicación de estudios en diferentes poblaciones y contextos para su validación y la comparación de resultados.

Como tercera línea de investigación propuesta se encuentra la ampliación de los estudios de aceptación tecnológica en el campo de la e-evaluación (19,9% de las investigaciones). Autores como Bacca-Acosta y Ávila-Garzón (2020) o Nikou y Economides (2018) abogan por encontrar nuevas dimensiones que, de forma específica, sean condicionantes intrínsecos de la aceptación de la e-evaluación por parte de docentes, mientras que autores como Pu y Xu (2021) proponen explorar la influencia de los factores socioculturales en esta.

Por otro lado, un grupo de autores centra su propuesta en el fortalecimiento de la colaboración institucional en el ámbito de la evaluación digital, resaltando la necesidad de desarrollar sinergias efectivas entre entidades educativas (Cheng et al., 2021; DeMara et al., 2017; Slade et al., 2022). Asimismo, se enfatiza la importancia de la formación docente como un elemento fundamental para mejorar la calidad de la e-evaluación (Abdou, 2020; Cutumisu, 2018; Wu y Wang, 2021; Zhang et al., 2021).

Por último, algunas investigaciones proponen la necesidad de una reflexión futura sobre el efecto de estas evaluaciones en el rendimiento (Acosta-Gonzaga y Walet, 2018; Boote et al., 2021; Eurboonyanun et al., 2021; Kühbeck et al., 2019; Máñez-Sáez et al., 2019) y la ampliación del marco teórico que sustenta esta modalidad evaluativa (Guerrero-Roldán et al., 2020; Sirianni et al., 2017).

3.2.3. Limitaciones del mapeo

En primer lugar, el número de bases de datos seleccionadas ha sido una limitación en la investigación dado que se excluyen otros repositorios que pueden contener publicaciones relevantes. Una segunda limitación se corresponde con el alcance temporal, reducido a cinco años por el elevado número de publicaciones. Finalmente, es también una limitación del mapeo el posible sesgo de publicación de las investigaciones seleccionadas.

3.2.4. Implicaciones del mapeo en el desarrollo de la investigación

El mapeo realizado ha sintetizado el estado de la e-evaluación en el momento actual y ha derivado en la detección de las áreas susceptibles de exploración futura que guiarán los avances en esta línea. Estas áreas se corresponden con investigaciones orientadas a los siguientes ámbitos:

- **Profesorado universitario:** La mayoría de las publicaciones recogidas centran su foco en estudiantes universitarios. El profesorado es una vía de investigación todavía poco explorada pese a ser un agente clave en el proceso de enseñanza-aprendizaje.
- **El empleo de tecnologías móviles para procesos de evaluación:** El estudio de la e-evaluación se encuentra altamente desarrollado en lo referente a plataformas de enseñanza, software educativo y herramientas disponibles en la red. Por tanto, se abre una vía de investigación menos explorada referida al uso de tecnologías móviles en e-evaluación.
- **La aceptación tecnológica y calidad de la e-evaluación:** La evaluación en entornos virtuales se encuentra ampliamente abordada en lo relativo a la satisfacción de los usuarios y el rendimiento. El mapeo constata que la aceptación tecnológica y la evaluación de la calidad son áreas susceptibles de desarrollo.
- **Niveles educativos alternativos a la educación superior:** La educación superior ha destacado como el nivel educativo más representado en los estudios. Se propone, por tanto, la exploración de la e-evaluación, sus características y su desarrollo en otros niveles.

Esta revisión ha tratado de retratar la situación de la investigación en e-evaluación en el momento actual. Entre ellas, la aceptación tecnológica de dispositivos móviles en procesos de evaluación relacionados con colectivos menos abordados (docentes) se posiciona como uno de los principales ámbitos de exploración, constituyendo el tema central de la tesis.

Por consiguiente, el próximo estudio presenta una revisión sistemática de la literatura más específica que describe las publicaciones relativas a la aceptación tecnológica de dispositivos móviles en procesos de evaluación.

Estudio 3:**Aceptación tecnológica de dispositivos móviles en evaluación****Anexo 6**

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (En revisión). Mobile-based assessment acceptance: A systematic literature review in the educational context

3.3. Revisión sistemática de la literatura: adopción tecnológica de dispositivos móviles

El tercer estudio plantea una revisión sistemática de la literatura sobre la aceptación tecnológica de los dispositivos móviles en evaluación, una de las líneas de desarrollo identificadas en la literatura extraídas en el mapeo anterior (anexo 6).

Por lo tanto, el objetivo de esta revisión consistió en proyectar una panorámica sobre el estado de la investigación, como un análisis más en profundidad sobre los resultados e implicaciones de las investigaciones realizadas que permitiera identificar los constructos más relevantes, la efectividad de los modelos de adopción ya desarrollados y las limitaciones de los estudios. Por este motivo, se optó por presentar una combinación de preguntas de mapeo y de revisión sistemática.

3.3.1. Preguntas de investigación

Esta revisión sistemática planteó once preguntas de investigación (ocho preguntas de mapeo y tres preguntas de revisión sistemática) a las que dar respuesta con las investigaciones seleccionadas. En primer lugar, las preguntas de mapeo han sido:

- M1) ¿Cómo ha evolucionado el número de investigaciones sobre adopción tecnológica de la evaluación basada en móviles a lo largo de los años?

- M2) ¿Quiénes son los autores con un mayor número de investigaciones en el campo?
- M3) ¿Qué medios centran la difusión de las investigaciones?
- M4) ¿Cuáles son las publicaciones más relevantes?
- M5) ¿En qué países se han llevado a cabo las investigaciones?
- M6) ¿En qué colectivo se centran los estudios para abordar la adopción de la evaluación basada en tecnologías móviles?
- M7) ¿Qué metodología se emplea para el análisis de los datos?
- M8) ¿Qué modelo, de los descritos, se emplea para estudiar la adopción tecnológica?

En segundo lugar, las tres preguntas de revisión sistemática planteadas en la revisión han sido las siguientes:

- R1) ¿Cuáles son las tendencias más representativas en el diseño de modelos de aceptación de la evaluación basada en tecnologías móviles?
- R2) ¿Qué porcentaje de la varianza se consigue explicar en cada investigación?
- R3) ¿Qué limitaciones presentan las investigaciones?

3.3.1.1. Proceso de selección de publicaciones

Para la selección de publicaciones se establecieron cinco criterios de inclusión/exclusión, que fueron:

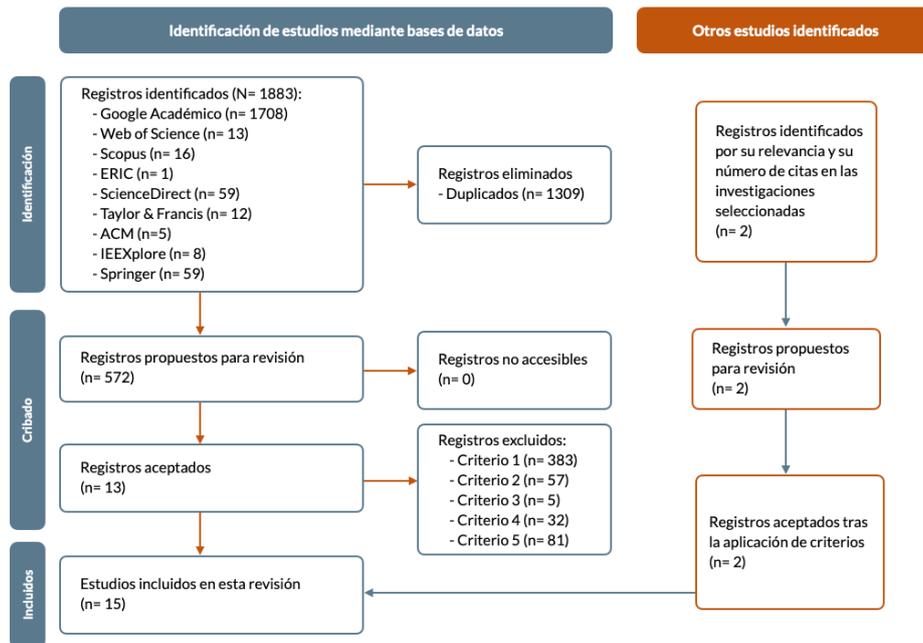
- Se rechazó toda investigación no relacionada con la aceptación tecnológica de las evaluaciones basadas en tecnologías móviles.
- Se rechazaron trabajos de investigación que no aplican TAM, UTAUT, TFT, IDT, TRA, TPB, SDT, MM, DTPB, C-TAM-TPB o MBAAM (o versiones derivadas de los mismos).
- Se excluyeron investigaciones que presentan modelos no aplicables en contextos educativos formales.
- Se rechazaron investigaciones en lengua distinta al inglés.
- Se rechazaron investigaciones no sometidas a un proceso de revisión por pares.

El proceso de búsqueda se llevó a cabo en nueve bases de datos (Google Académico, Springer, ScienceDirect, Scopus, Web of Science, Taylor y Francis, IEEEXplore, ACM y ERIC) mediante once cadenas de búsqueda diferentes orientadas a recuperar trabajos relacionados con los modelos de aceptación tecnológica y las teorías más relevantes.

Tras la selección de publicaciones, se procedió a la aplicación de diez criterios de calidad (Seo y Kim, 2012) para asegurar que las publicaciones seleccionadas cumplieran con una serie de factores e índices que aseguran su rigor científico para formar parte de la revisión.

Una vez concluido el proceso de selección, la aplicación de criterios y la inclusión de dos investigaciones relevantes no incluidas en la extracción (Abdirahman et al., 2018, 2022), se obtuvieron 15 publicaciones comprendidas entre 2014 y 2022. El proceso completo de investigación se encuentra disponible en: <http://bit.ly/3ntr92B>. El diagrama PRISMA de la revisión sistemática puede consultarse en la figura 13.

FIGURA 13. Diagrama PRISMA de la revisión sistemática de la literatura



3.3.2. Resultados de la revisión

3.3.2.1. Resultados de las preguntas de mapeo

Entre los estudios incluidos se identificaron ocho publicaciones en revistas científicas y siete publicaciones en actas de congresos internacionales. Respecto a las revistas más destacadas se encuentran *Computers and Education* y *Computers in Human Behavior* como aquellas con los índices SJR más altos. Por otro lado, se ha extraído el número total de citas en Google Académico a fecha de búsqueda para cada publicación, siendo nuevamente las más relevantes las publicaciones de Nikou y Economides (2017a, 2017b), con un total de 479 citas entre ambas.

La segunda pregunta reflejó una evolución anual limitada en el número de publicaciones debido al estado inicial del campo. En relación con el país de desarrollo, Grecia lidera la lista de países con 3 estudios publicados. La mayoría de las investigaciones se centraron en educación superior (10 estudios) destacando principalmente investigaciones en estudiantes (13) y, en menor medida, en profesores (4).

Respecto a los modelos de adopción detectados, el modelo TAM ha sido descrito en ocho ocasiones, siendo el modelo predominante. Seguidamente, cuatro investigaciones plantean modelos basados en la combinación del Modelo de Aceptación Tecnológica (TAM) y la Teoría de la Autodeterminación (SDT). Por último, un estudio plantea la Teoría de la Autodeterminación como modelo (Nikou y Economides, 2014), otro plantea el Modelo de Aceptación de Evaluaciones mediadas por Dispositivos Móviles (MBAAM) (Nikou y Economides, 2017b) y otro la combinación del Modelo de Aceptación Tecnológica y la Teoría Unificada de la Aceptación y Uso de Tecnología (UTAUT) (Kumar et al., 2020).

3.3.2.2. Resultados de las preguntas de revisión sistemática

Las preguntas propuestas en la revisión sistemática exploraron las tendencias en los estudios de aceptación tecnológica. En total, se identificaron 22 constructos planteados en unos modelos mayoritariamente basados en TAM. Por otro lado, se han extraído el Modelo de Aceptación de la Evaluación Móvil en Profesores (TAMBA) y el Modelo de Aceptación de Evaluaciones mediadas por Dispositivos Móviles (MBAAM) ya descritos como parte de la revisión (Nikou y Economides, 2017b, 2017a).

La segunda pregunta de revisión abordó el porcentaje de varianza explicada por cada modelo. Entre los resultados, son los modelos que combinan TAM con SDT los que explican un mayor porcentaje de varianza, y es el modelo propuesto por Nikou & Economides (Nikou y Economides, 2014) el más explicativo (81,7%) frente al de menor explicación, también de los mismos autores (Nikou y Economides, 2015) que explica un 45,0%.

Por último, la tercera pregunta de revisión identificó las limitaciones en los modelos y estudios analizados. Estas limitaciones se encontraron relacionadas con el alcance de la población y el tamaño muestral, la determinación de las dimensiones de los modelos y la necesidad de reflexión teórica sobre la aceptación de los dispositivos móviles en procesos evaluativos.

3.3.3. Limitaciones de la revisión sistemática

Por un lado, esta revisión ha contado con limitaciones derivadas de sesgos metodológicos, destacando el número de bases de datos seleccionadas. Por otro lado, la revisión se ha centrado en publicaciones únicamente en inglés. La revisión ha podido contar también sesgos de publicación múltiple y de confirmación en los estudios alcanzados (Beltrán, 2005), así como sesgos derivados de la elaboración y aplicación de los criterios de inclusión.

3.3.4. Implicaciones de la revisión sistemática en el desarrollo de la investigación

Una vez analizados los resultados principales de la revisión, se resumen como líneas de investigación futuras de la tesis:

- **El desarrollo de estudios de aceptación tecnológica del uso de dispositivos móviles en evaluación en docentes.** La investigación actual ha centrado su desarrollo en estudios de aceptación en estudiantes, por lo que se abre una nueva vía en el colectivo docente.
- **La búsqueda de nuevas dimensiones y el establecimiento de un modelo consensuado y validado.** Los modelos analizados en la revisión difieren significativamente entre sí, alcanzándose un total de 29 constructos diferentes. Por tanto, es también una vía futura de estudio la búsqueda del consenso interdimensional para el asentamiento de un modelo más amplio, válido y explicativo.

En conclusión, se propone como línea futura, la búsqueda, validación y consolidación de un modelo que combine las distintas teorías y modelos descritos y que trate de explicar un alto porcentaje de varianza.

Conocer de forma específica los condicionantes de la aceptación tecnológica de dispositivos móviles en procesos de evaluación permitirá también el desarrollo de políticas formativas destinadas al profesorado en formación o a la formación del profesorado en activo (formación permanente), lo que mejorará la aceptación tecnológica. Estas investigaciones pueden tomar como referencia estudios de aceptación en campos más avanzados como la evaluación en e-learning (Cruz-Benito et al., 2019; Sánchez-Prieto et al., 2019) o estudios de adopción sobre m-Learning no centrados en evaluación (Alowayr y Al-Azawei, 2021; Gómez-Ramirez et al., 2019; Pratama, 2021).

En base a la revisión sistemática realizada, el próximo estudio presenta la propuesta y validación de un nuevo modelo de aceptación tecnológica de dispositivos móviles en procesos de evaluación para docentes.

Estudio 4:

Diseño y validación de un nuevo modelo de adopción tecnológica

Anexo 4

Ortiz-López, A.; Olmos-Migueláñez, S.; Sánchez-Prieto, J.C. (En prensa). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En E. Chiner (Ed.), *Nuevas tendencias interdisciplinarias en Educación y Conocimiento*. Tirant Lo Blanch

3.4. Un nuevo modelo de adopción tecnológica de dispositivos móviles en e-evaluación

El cuarto estudio presenta la propuesta y validación de un nuevo modelo de aceptación tecnológica de dispositivos móviles en procesos de evaluación para docentes (anexo 4).

3.4.1. Propuesta inicial de modelo: Una reflexión sobre aceptación tecnológica

El nuevo modelo propuesto parte de las revisiones previas ya descritas y de los modelos existentes de adopción tecnológica, considerando las barreras que influyen en la adopción de la tecnología en docentes para su desarrollo. Estas barreras han sido clasificadas en la literatura en barreras externas (o de primer orden) e internas (o de segundo orden) por Ertmer (1999), quien las define como:

- 1) Barreras externas (primer orden): Estas barreras refieren a los obstáculos extrínsecos que puede encontrarse un docente en la implementación tecnológica. Bajo estas barreras se encuentran los recursos (equipos, formación, tiempo) o aquellas vinculadas con el desarrollo de la tecnología. Estas barreras son más fáciles de medir y de superar que las internas (Fisher et al., 1996).

- 2) Barreras internas (segundo orden): Son las barreras que interfieren o impiden el cambio fundamental (Brickner, 1995). Este tipo de barreras guarda relación con las creencias de los docentes sobre el proceso de enseñanza-aprendizaje. Además, son menos observables pese a ser las más comunes (Hannafin y Savenye, 1993) y son más limitantes dado que son más personales y se encuentran más arraigadas a cada profesional (Ertmer, 1999).

Tradicionalmente, la formación en tecnología, tanto para maestros en formación como para maestros en servicio, se ha centrado en ayudar a los educadores a superar obstáculos de primer orden. Más recientemente, los programas de formación han integrado modelos pedagógicos de uso de la tecnología como un medio para abordar obstáculos de segundo orden (Ertmer, 1999).

Superar estas barreras se considera crucial para que los docentes adopten la tecnología en sus procesos de evaluación y fomenten la innovación y la motivación entre los estudiantes, por lo que es necesario un nuevo modelo para la detección de los factores limitantes y para la superación de las barreras en un sistema educativo inmerso en un proceso de transformación digital (Uriarte y Acevedo, 2018).

Con respecto a los modelos de adopción tecnológica, son numerosas las investigaciones que describen el modelo TAM como el más empleado para los estudios desarrollados en materia de educación (García-Peñalvo et al., 2015; Nikou y Economides, 2016). El nuevo modelo presentado ha pretendido dar respuesta a dos preguntas principales:

1. ¿Qué variables pueden ayudar a predecir la intención de uso de tecnologías móviles en evaluación por parte de los docentes?
2. ¿Cuál es la importancia relativa de cada variable en la explicación general de la intención de uso de tecnologías móviles en evaluación?

La propuesta realizada parte del modelo TAM y mantiene sus constructos de intención conductual de uso, utilidad percibida y facilidad de uso percibido. En él, se elimina la actitud hacia el uso por su bajo efecto mediador (Hu et al., 2003) con el objetivo también de reducir los ítems del modelo, una idea sostenida en las evoluciones del modelo TAM (Venkatesh y Bala, 2008; Venkatesh y Davis, 2000).

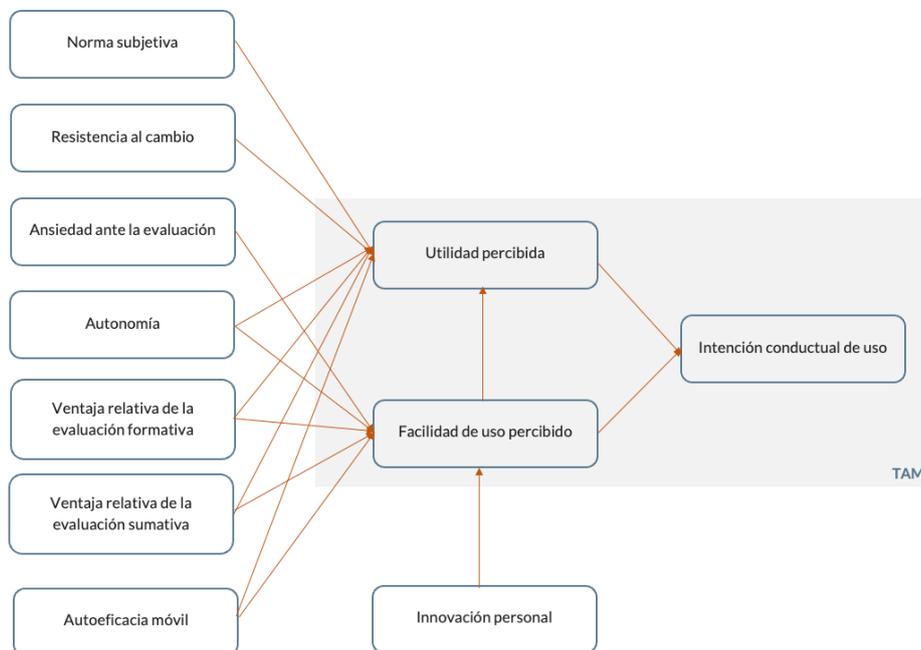
La nueva propuesta introduce constructos empleados en otros modelos de adopción tecnológica pero no en el campo de la evaluación, como es la **norma subjetiva** (derivada de TAM3) (Venkatesh & Bala, 2008) y la **resistencia al cambio**, dado su buen funcionamiento en modelos similares (Schepers y Wetzels, 2007; Tsai et al., 2020). Además, adapta del modelo MBAAM (Nikou y Economides, 2017b) los constructos **autoeficacia en el uso de dispositivos móviles e innovación personal**.

Por último, con el objetivo de profundizar en las modalidades de la evaluación y después de no encontrar modelos hasta la fecha que lo planteen, se incorporan los constructos de **ansiedad ante la evaluación**, la **ventaja relativa de la evaluación formativa** y la **ventaja relativa de la evaluación sumativa**. La totalidad de constructos se definen en el anexo 4.

3.4.2. Proceso de validación de contenido del modelo propuesto

Para la validación del modelo inicial (Figura 14) se diseñó un instrumento de recogida de información con dos secciones: una para la identificación y otra para la evaluación del modelo en una escala Likert de 1 a 7. El instrumento inicial contenía 64 ítems y fue sometido a un proceso de validación por jueces expertos que evaluaron la coherencia, relevancia y claridad de los ítems, así como la adecuación y validez global del modelo.

FIGURA 14. Propuesta inicial del modelo



Esta validación ha sido llevada a cabo por seis evaluadores externos seleccionados por su relación con el ámbito educativo y tecnológico en distintas universidades nacionales e internacionales. El análisis de estas validaciones fue llevado a cabo con el software estadístico SPSS v.26.

Los datos completos del proceso de validación pueden consultarse en la publicación complementaria (Anexo 4) y los indicadores completos se encuentran accesibles en: <http://bit.ly/3YhEbwX>.

3.4.3. Descripción del nuevo modelo validado

Concluida la validación, el modelo cuenta con un total de 50 ítems que evalúan la aceptación mediante los once constructos propuestos. Para la comprensión y justificación de sus relaciones se plantean un total de 17 hipótesis relacionales.

Los constructos propuestos en el modelo responden todos ellos al estudio de barreras internas. En la siguiente tabla (Tabla 4) se presentan las hipótesis planteadas.

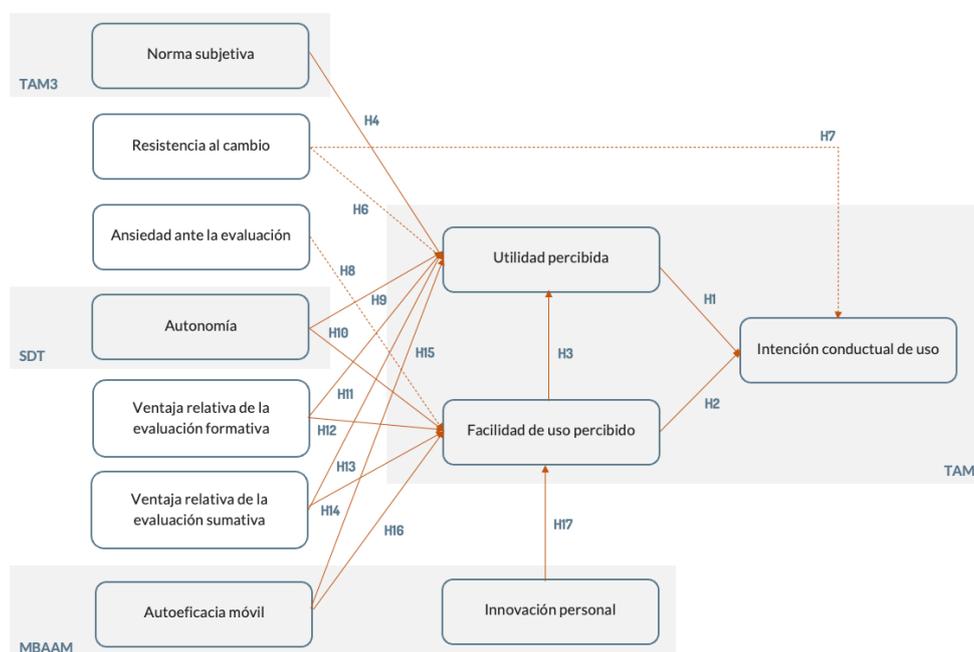
TABLA 4. Hipótesis del modelo validado

Constructo	Hipótesis
Intención de uso, utilidad percibida y facilidad de uso percibido	
H1	La utilidad percibida predice positivamente la intención de uso docente de dispositivos móviles en procesos de evaluación.
H2	La facilidad de uso percibido predecirá positivamente la intención de usar dispositivos móviles en procesos de evaluación.
H3	La facilidad de uso percibido predecirá positivamente la utilidad percibida por los docentes sobre el uso de dispositivos móviles en procesos de evaluación.
Norma subjetiva	
H4	La norma subjetiva predice negativamente la utilidad percibida por docentes en el uso de dispositivos móviles para evaluar.
H5	La norma subjetiva predice negativamente la intención docente de emplear dispositivos móviles en evaluación.
Resistencia al cambio	
H6	La resistencia al cambio predice negativamente la utilidad percibida por los docentes para evaluar a través de dispositivos móviles.
H7	La resistencia al cambio predice negativamente la intención de usar dispositivos móviles en procesos de evaluación.
Ansiedad ante la evaluación	
H8	La ansiedad ante la evaluación tendrá un efecto negativo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
Autonomía	
H9	La autonomía tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
H10	La autonomía tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
Ventaja relativa de la evaluación formativa	
H11	La ventaja relativa de la evaluación formativa tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
H12	La ventaja relativa de la evaluación formativa tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
Ventaja relativa de la evaluación sumativa	
H13	La ventaja relativa de la evaluación sumativa tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
H14	La ventaja relativa de la evaluación sumativa tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
Autoeficacia en el uso de dispositivos móviles	
H15	La autoeficacia móvil tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

Constructo	Hipótesis
H16	La autoeficacia móvil tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
Innovación personal	
H17	La innovación personal tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

La figura 15 presenta el modelo completo validado con las hipótesis que lo justifican y el modelo raíz de aquellos constructos extraídos de otras teorías. Los ítems del modelo completo pueden consultarse en: <http://bit.ly/3MBRwwc>

FIGURA 15. Diseño del modelo validado



3.4.4. Implicaciones del modelo en el desarrollo de la investigación

El nuevo modelo propuesto, en su versión inicial validada, surge de un proceso de reflexión que toma como base otros modelos de adopción y teorías relacionadas con la aceptación probadas en diferentes campos del conocimiento. Además, su construcción ha implicado una profundización en el concepto de la evaluación y sus principales modalidades de desarrollo.

La amplitud de constructos en el modelo es síntoma de la alta variabilidad de un área incipiente en su desarrollo, dado el interés en la búsqueda de la mayor cantidad de factores condicionantes de la aceptación tecnológica posibles. Por tanto, el modelo que se propone es un modelo novedoso que combina constructos ya aplicados con nuevas propuestas que tratan de profundizar en las modalidades formativa y sumativa y determinar la aceptación tecnológica de los docentes en su relación con los dispositivos móviles en procesos de evaluación.

Este modelo será el punto de partida de la fase experimental de la tesis. El próximo estudio describe su aplicación y validación empírica. Esta fase orientará la versión final del modelo y reflejará el estado actual de la aceptación de futuros docente del empleo de dispositivos móviles en procesos de evaluación.

Estudio 5:

Estudio intermedio para la validación del modelo

Anexo 5

Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En prensa). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. *Journal of New Approaches in Educational Research*.

Anexo 7

Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (En revisión). Mobile technology in assessment processes: Comparing its usefulness and antecedents in CB-SEM

3.5. Estudio intermedio: Dos análisis derivados en CB-SEM y PLS-SEM

Una vez realizada la propuesta inicial del modelo para el estudio de la adopción tecnológica de dispositivos móviles en procesos de evaluación en docentes, esta sección presenta el estudio intermedio desarrollado para comprobar el comportamiento de los constructos de manera empírica.

Esta sección presenta dos análisis diferenciados derivados de una misma aplicación. La justificación de la realización de este análisis reside en la experimentación con el modelo diseñado mediante la propuesta de diferentes modelos e hipótesis derivadas.

3.5.1. Decisión previa a la aplicación (I): Constructo Ventaja Relativa de la evaluación

Previo a la realización del estudio, y tras el contraste con expertos en modelos de adopción tecnológica durante el diseño del estudio intermedio, se determinó que los constructos “Ventaja relativa de la evaluación sumativa” y “Ventaja relativa de la evaluación formativa”, tal y como se

encontraban formulados partiendo de Moore y Benbasat (1991), referenciaban en realidad a utilidades percibidas de la evaluación y no a ventajas relativas.

Este hecho condujo a reformular los constructos como “utilidad percibida de la evaluación sumativa” y “utilidad percibida de la evaluación formativa” y a redirigir el modelo hacia un nuevo enfoque específico de la utilidad de la evaluación profundizando en la intención inicial de explorar las implicaciones que tiene la modalidad de evaluación en la adopción tecnológica.

Esta constatación fue reafirmada una vez realizado el estudio intermedio. En él, los datos confirmaron que, en el caso de haber actuado como antecedentes de la utilidad percibida de TAM (planteamiento original como ventaja relativa), se hubiese alcanzado un grave problema de validez discriminante entre estas ventajas y la utilidad, indicando la alta similitud entre los constructos.

De esta forma, los nuevos constructos hacen referencia a la utilidad percibida de los dispositivos móviles para evaluación sumativa y a la utilidad percibida de los dispositivos móviles para evaluaciones formativas.

3.5.2. Decisión previa a la aplicación (II): Constructo Autonomía

El constructo autonomía y su diseño se basan, en el modelo inicial, en las investigaciones de Nikou y Economides (2017b), McAuley et al. (1989) y Van den Broeck (2010). Durante la validación de contenido, un evaluador centró sus recomendaciones en la posible reformulación del constructo, referenciando que determinados ítems parecían no evaluarla desde la teoría de la autodeterminación. Por ello, y en base a sus recomendaciones, se generó para el estudio intermedio un constructo paralelo de autonomía formulada partiendo de los trabajos de Reeve (Reeve y Cheon, 2021; Su y Reeve, 2011) para determinar cuál de las dos mantenía un mejor funcionamiento.

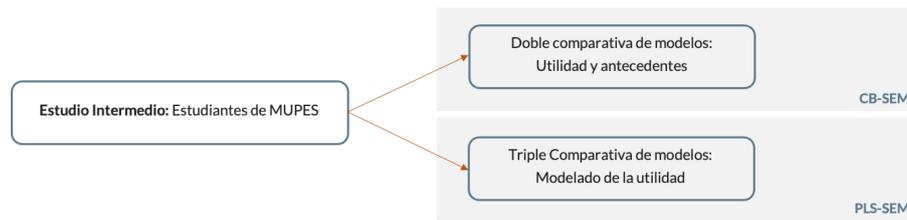
Tras constatar en la validación del modelo que los ítems de la nueva dimensión no superaban las cargas mínimas establecidas, se decidió mantener el constructo inicial de autonomía, ya validado en investigaciones previas y de mejor funcionamiento en el análisis que se describe a continuación.

3.5.3. Análisis del estudio intermedio

Este estudio intermedio comprende dos análisis (Figura 16) de una misma aplicación, consistentes en:

- **Análisis en CB-SEM:** Doble comparativa para validar el funcionamiento de una nueva utilidad percibida de los dispositivos móviles en evaluación con la introducción de los constructos autonomía y autoeficacia.
- **Análisis en PLS-SEM:** Comparativa de tres modelos de adopción tecnológica con diferentes formulaciones de la utilidad percibida en cada uno de ellos (adaptadas a las modalidades evaluativas), con la autoeficacia en el uso de dispositivos móviles como antecedente.

FIGURA 16. Esquema del estudio intermedio



3.5.3. Población y muestra

La población se compone de los estudiantes que cursan el Máster en Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas en la Universidad de Salamanca. El número de cuestionarios recogidos tras la finalización del estudio ha sido de 262 (90,34% de la población, N= 292).

Por otro lado, y pese a que lo ideal para el modelo sería su validación en docentes en ejercicio, la imposibilidad de acceso a una amplia muestra y las limitaciones derivadas de este hecho ha conducido a la toma de docentes en formación como muestra para el desarrollo del estudio intermedio dada su similitud ya contrastada en investigaciones previas (Teo, 2015).

De los participantes, el 41,2% fueron hombres (n=108), el 58,4% mujeres (n=153), y un participante (0,4%) no reflejó género en sus respuestas. En cuanto al grado cursado previamente, se han contabilizado un total de 54 grados universitarios diferenciados, siendo el grado en Filología Hispánica y el grado en Estudios Ingleses el más representado (n=21).

La edad media de los participantes se sitúa en 25 años (DT=4,79). Finalmente, en lo referido a su experiencia docente, un 74,4% (n=195) indicaron no contar con experiencia hasta el momento, un 11% (n=29) contaban con experiencia en un periodo de prácticas, nueve participantes con experiencia inferior a un año, un participante (0,38%) con experiencia entre tres y cinco años de docencia y cuatro de ellos (1,53%) con experiencia superior a cinco años.

3.5.5. Instrumento de medida empleado en el estudio intermedio

Para llevar a cabo el estudio se ha diseñado un instrumento conformado por seis variables de identificación y los ítems propuestos. Las variables de identificación empleadas han sido la edad, el género, el grado académico de acceso al máster, la experiencia docente, el número medio de horas de uso de ordenador y el número medio de horas de uso de dispositivos móviles al día.

La recogida de datos, su tratamiento y el proceso de investigación han sido valorados favorablemente y aprobados por el Comité de Ética de la Investigación de la Universidad de Salamanca con número de registro 903 (anexo A). El proceso de recogida de datos fue llevado a cabo

en noviembre de 2022 mediante un cuestionario difundido a través de la aplicación LimeSurvey (Engard, 2009), accesible en: <http://bit.ly/46b75mQ>

3.5.6. Análisis CB-SEM

3.5.6.1. Objetivo

En vista de los recientes cambios en el modelo, que han reemplazado la **ventaja relativa de la evaluación por utilidades específicas de la misma**, este primer análisis plantea como objetivo la búsqueda de obtención de conocimiento sobre el modelado de la utilidad percibida. Para ello se presentan dos modelos: un modelo basado en TAM y un nuevo modelo con la utilidad percibida de la evaluación.

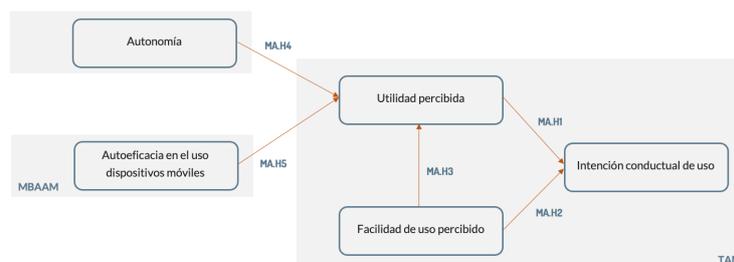
3.5.6.2. Modelos de investigación

a) **Modelo A:** Evaluación sin diferenciación en sus modalidades

El primer modelo propuesto se basa en la perspectiva clásica de la evaluación y toma una utilidad percibida genérica (modelo TAM) para su estudio. Este modelo asume que, si los futuros docentes conciben útiles los dispositivos móviles para la evaluación, tendrán una intención de uso mayor (Davis, 1989).

Esta propuesta inicial no distingue entre utilidad para la evaluación formativa o sumativa y plantea cinco constructos y cuatro hipótesis relacionales (Figura 17).

FIGURA 17. Modelo A del análisis en CB-SEM

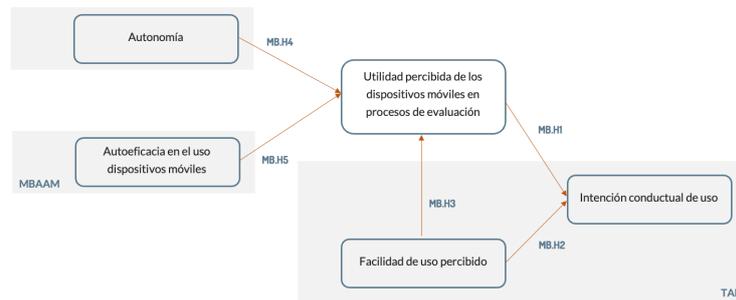


b) **Modelo B:** Evaluación como un conjunto de procesos sumativos y formativos

El segundo modelo de investigación (Figura 18) propone un constructo que comprende ítems relativos a la utilidad percibida de los dispositivos móviles haciendo referencia específica a las dos modalidades de evaluación (dos ítems de utilidad sumativa y dos ítems de utilidad formativa).

Este modelo propone cinco constructos y cinco hipótesis relacionales para la evaluación de la intención de uso.

FIGURA 18. Modelo B del análisis en CB-SEM



3.5.6.3. Metodología de análisis

En esta investigación se han aplicado modelos de ecuaciones estructurales basados en la covarianza (CB-SEM). El análisis se ha llevado a cabo a través del software estadístico JASP (v. 0.17.1).

3.5.6.4. Resumen de los resultados más destacados

En primer lugar, el análisis de la fiabilidad compuesta (CR) y de la varianza media extraída (AVE) ha superado el punto de corte para los ítems y los constructos propuestos (Fornell y Larcker, 1981). Seguidamente, se ha confirmado la validez discriminante de ambos modelos (a través del ratio de correlaciones Heterotrait-Monotrait y la prueba de Fornell-Larcker), posibilitando la continuidad del estudio.

La validez estructural de los modelos ha sido evaluada a través del índice Chi-Cuadrado (χ^2), el índice de ajuste comparativo (CFI), el índice de Tucker-Lewis (TLI), la raíz cuadrada media estandarizada (SRMR) (Hair et al., 2012), el índice de ajuste normalizado de Bentler-Bonett (NFI) (Awang, 2012) y el índice de bondad de ajuste GFI (Hair et al., 2010); encontrándose todos dentro de los valores establecidos.

Los resultados de los coeficientes *path* (Hair et al., 2021) reflejan la influencia que tienen la autonomía y la autoeficacia en el uso de dispositivos móviles para evaluaciones sobre el nuevo constructo relativo a la evaluación, mientras que solamente es la autonomía la que tiene influencia en la utilidad percibida en el modelo A (MAH4).

Por otro lado, la utilidad percibida de la evaluación es un buen antecedente de la intención conductual de uso de dispositivos móviles en ambos modelos (teniendo mejor funcionamiento en la nueva propuesta). Por ello, es importante atender a este nuevo constructo en investigación y contemplar de forma específica las características de la evaluación en estudios de adopción tecnológica.

En el estudio, la facilidad de uso percibida no ha sido un antecedente determinante de la intención de uso ni de la utilidad percibida (se rechazan MAH1, MAH3, MBH1, MBH3). Este hecho

guarda consonancia con investigaciones previas que afirman que la competencia digital nativa condiciona hoy que la facilidad de uso no sea una variable determinante en su aceptación tecnológica (Evans y Robertson, 2020; Recio-Muñoz et al., 2020).

Por último, se ha evaluado la parsimonia de los modelos (Tabla 5) siguiendo el criterio de información de Akaike (AIC) y el criterio de información bayesiano (BIC). En lo referido al criterio de Akaike (AIC), el modelo inicial es ligeramente más parsimonioso (Bozdogan, 1987), mismo caso que para el criterio de información bayesiano (BIC).

TABLA 5. Índices de ajuste para los modelos en CB-SEM

Dimensión	Modelo A	Modelo B
R ²	0.535	0.629
Criterio de información de Akaike (AIC)	13130.453	13348.069
Criterio de información bayesiano (BIC)	13330.280	13547.897

El análisis de la varianza explicada refleja que el nuevo modelo explica un 63% de la intención conductual de uso, mientras que el primero explica un 9% menos (54%). Por tanto, la nueva propuesta se consolida como mejor modelo explicativo.

3.5.6.5. Implicaciones del análisis CB-SEM para el modelo final

El primer análisis realizado ha confirmado que una utilidad formulada con referencias específicas al tipo de evaluación tiene un mayor poder explicativo frente a la formulación de una utilidad genérica basada en el modelo TAM, pese a ser menos parsimonioso. En segundo lugar, el estudio confirma el funcionamiento de la autonomía y la autoeficacia en el uso de dispositivos móviles para evaluar como factores condicionantes de la aceptación. En la propuesta inicial (modelo A), solamente es la autonomía el constructo determinante.

En tercer lugar, la facilidad de uso percibida no ha sido un factor significativo en los modelos propuestos en esta primera aplicación, teniendo solamente en el primer modelo un factor mediador y un efecto indirecto. Esta afirmación puede encontrarse condicionada por las características de la muestra, ya que los participantes mantienen un uso medio de dispositivos tecnológicos de 5,82 horas al día. Este uso elevado guarda relación con la facilidad de uso y disminuye su efecto (Siew et al., 2017).

Por tanto, esta investigación mantiene el constructo para la segunda versión del modelo, pero se recomienda la búsqueda de un nuevo constructo que considere su reemplazo por otros antecedentes vinculados menos explorados hasta el momento.

3.5.7. Análisis PLS-SEM

Los resultados de este segundo análisis se describen de forma sintética. El total de tablas, indicadores y resultados puede ser consultado en la publicación derivada (anexo 5).

3.5.7.1. Objetivo

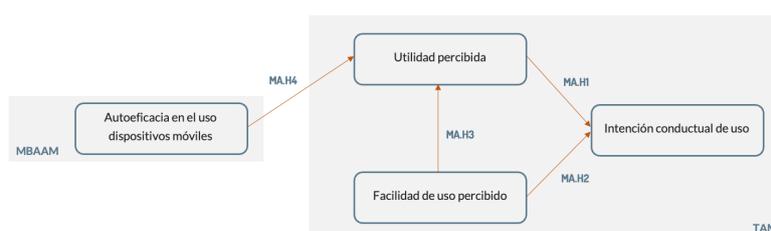
En segundo lugar, se plantea una comparativa de tres diseños, basados nuevamente en TAM, que modelan la utilidad percibida desde las tres perspectivas de evaluación asumibles por los docentes. Estas perspectivas ya descritas se basan en la evaluación formativa y sumativa (Scriven, 1967) y se resumen en entender la evaluación como un proceso único sin diferenciar entre ambas modalidades (Shepard, 2006; Tyler, 1949), en entenderlas como procesos diferenciados tratándolas de forma separada en la práctica (Black y Wiliam, 1998; Harlen y James, 1997; Patton, 1996) o en concebir la total complementariedad de ambas modalidades (Ahmad y Bhat, 2019; Buchholtz et al., 2018).

3.5.7.2. Modelos de investigación

a) Modelo A: Evaluación sin diferenciación en sus modalidades

El primer modelo es el modelo TAM ampliado con el constructo autoeficacia (Figura 19). Este modelo propone cuatro constructos y cuatro hipótesis relacionales para la evaluación de la intención de uso. Se asume el constructo de utilidad percibida de Davis diseñado en el modelo inicial (1989) como corriente genérica dado que no establece diferencias, en sus ítems, entre evaluación formativa y sumativa.

FIGURA 19. Modelo A del análisis en PLS-SEM

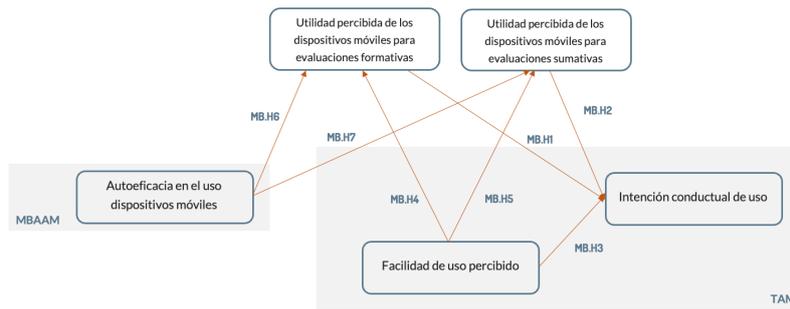


b) Modelo B: Evaluación formativa y sumativa como procesos separados

El segundo modelo asume la evaluación sumativa y la evaluación formativa como dos dimensiones diferenciadas (MacLellan, 2001), y convierte la utilidad percibida original en dos utilidades percibidas específicas.

La construcción del modelo ha sido realizada en base al modelo TAM y al análisis de las características de la evaluación formativa y sumativa (Scriven, 1967). Este modelo plantea cinco constructos y siete hipótesis relacionales (Figura 20).

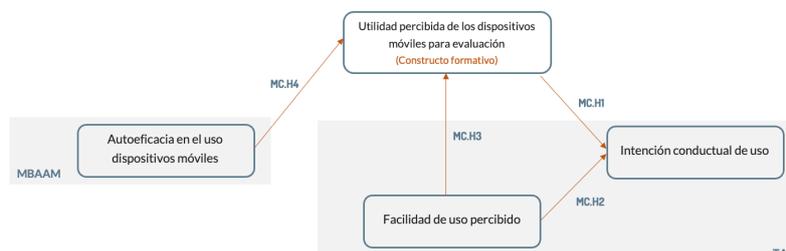
FIGURA 20. Modelo B del análisis en PLS-SEM



c) **Modelo C:** Evaluación sumativa y formativa como un proceso único y combinado

Por último, el tercer modelo concibe la complementariedad de la evaluación formativa y sumativa, asumiendo que el proceso de evaluación es la suma de ellas y de sus componentes (Dolin et al., 2018). Este tercer modelo (Figura 21) presenta una dimensión única de utilidad percibida de los dispositivos móviles en procesos de evaluación que aúna los indicadores propuestos para la evaluación formativa y sumativa, asumiéndose que la suma de sus indicadores refleja la medida general de la dimensión (Simonetto, 2012).

FIGURA 21. Modelo C del análisis en PLS-SEM



3.5.7.3. Metodología de análisis

El segundo análisis basa su desarrollo en los modelos de ecuaciones estructurales basados en mínimos cuadrados parciales (PLS-SEM). El empleo de esta técnica se justifica en el constructo formativo del modelo C, un análisis que solo es posible a través de PLS-SEM (Ajms, 2015; Bollen, 2011).

3.5.7.4. Resultados más destacados

Una vez concluido el análisis de resultados se ha constatado que los tres modelos cumplen con los valores estipulados de validez convergente, fiabilidad, y validez discriminante (Fornell y Larcker, 1981) tras la eliminación de tres ítems que no superan el valor mínimo de carga (Hair et al., 2010). Posteriormente, se realizó el análisis de los modelos estructurales que derivó en el rechazo de una hipótesis en el modelo A (MAH3), dos en el modelo B (MBH4 y MBH5) y otras dos en el modelo C (MCH2 y MCH3).

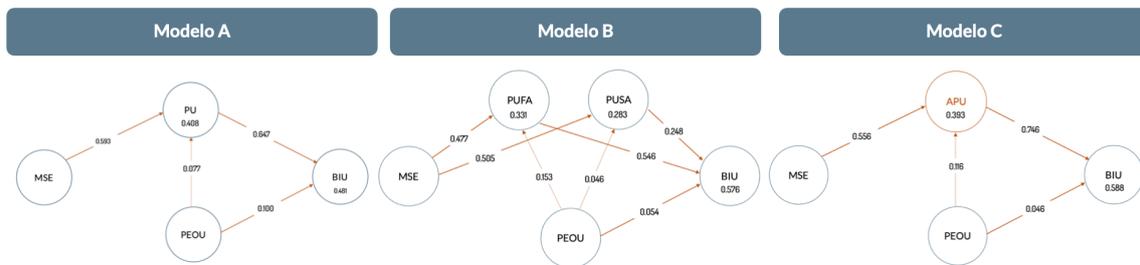
Por tanto, se puede afirmar que los modelos propuestos predicen la intención conductual de uso a través de las distintas utilidades percibidas planteadas. De los resultados se extrae también la capacidad predictiva de las tres propuestas (Q^2) y, bajo este criterio, se determinan la segunda y tercera propuesta como los modelo más parsimoniosos y generalizables (Sharma y Kim, 2012). Teniendo en cuenta el Criterio de Información de Akaike (AIC) (Bozdogan, 1987), el modelo C es significativamente mejor que el modelo B, al igual que los modelos B y C con respecto al modelo A. Por último, el criterio de Información Bayesiano (BIC), que tiene un alto poder predictivo y es estable para la comparación de modelos (Sharma et al., 2022), también confirma la tercera propuesta como el modelo más funcional (Tabla 6).

TABLA 6. Índices de ajuste para los modelos en PLS-SEM

Índice de ajuste	MA	MB	MC
Q^2	0,410	0,492	0,498
Criterio de información de Akaike (AIC)	92,276	46,723	37,065
Criterio de Información Bayesiano (BIC)	-156,174	-203,238	-216,386

Finalmente, con respecto a la explicación de la varianza relativa a la intención conductual de uso en los tres modelos (Figura 22), se puede afirmar que son el segundo y el tercero los que predice un mayor porcentaje de varianza (57% y 58%, respectivamente). Además, se confirma la autoeficacia como un buen antecedente en los tres modelos propuestos.

FIGURA 22. Resultados del análisis de los modelos estructurales de los modelos en PLS-SEM



3.5.7.5. Implicaciones del análisis PLS-SEM para el modelo final

El análisis de resultados pone de manifiesto que los dos nuevos modelos propuestos con la reformulación de la utilidad percibida explican un mayor porcentaje de varianza sobre la intención de uso. Además, las dos propuestas son modelos más parsimoniosos y generalizables que el modelo TAM original, en este caso.

En segundo lugar, el modelo B arroja diferentes valores para las utilidades y sugiere que los futuros docentes conceden una mayor importancia al uso de la tecnología móvil en las evaluaciones formativas (dado el tamaño del efecto), actuando también como un mejor predictor de la intención

de uso. Esta idea está en consonancia con la literatura, ya que las evaluaciones formativas se perciben como más flexibles y abiertas a las innovaciones (Nikou, 2021), mientras que las evaluaciones sumativas siguen siendo una limitación para los docentes en la inclusión de nuevas tecnologías y técnicas en sus procesos (Shepard, 2006).

En tercer lugar, el modelo C obtiene pesos bajos para los indicadores del constructo formativo tras la realización de técnicas de *bootstrapping* (Hair et al., 2021). Pese a que la recomendación de la literatura para estos casos es mantener los ítems (Hair et al., 2021; Ramayah et al., 2017), éstos no contribuyen a la explicación de la dimensión. Bajo este paradigma, la dimensión utilidad percibida de la evaluación se mantiene, pero sus ítems no alcanzan a definir su totalidad.

3.5.8. Orientaciones y consideraciones finales

Por tanto, tras la realización de dos análisis diferenciados donde las propuestas de la **utilidad percibida formativa y sumativa** por separado y el constructo reflectivo de la **utilidad percibida de la evaluación** han tenido mayor poder explicativo (modelo CB-SEM) y predictivo (modelo PLS-SEM), se toma la decisión de continuar el diseño del modelo final hacia una propuesta de dos utilidades específicas y diferenciadas de la evaluación.

En segundo lugar, los estudios han confirmado el funcionamiento de la autonomía como factor condicionante de la aceptación. Con respecto a la autoeficacia en el uso de dispositivos móviles para evaluar, su funcionamiento no se ha confirmado en uno de los cinco modelos propuestos. Por este motivo, y dado su buen funcionamiento en investigaciones previas como la de Castiblanco-Jiménez et al. (2020), se recomienda su utilización en estudios futuros. Los siguientes estudios deberán describir nuevos factores, ya validados en el modelo inicial, que sean condicionantes de la aceptación tecnológica.

En la próxima sección se presentan las conclusiones, la propuesta final de modelo para el estudio de la aceptación tecnológica de dispositivos móviles en procesos de evaluación, las limitaciones y las vías de investigación futuras.

Conclusiones, limitaciones y líneas futuras de investigación

Conclusiones:

Conclusiones, limitaciones y líneas futuras de investigación

La última sección presenta las conclusiones más relevantes de la investigación, incluyendo una versión final del modelo inicialmente diseñado a partir de los resultados obtenidos en el estudio intermedio. Del mismo modo, se muestran las limitaciones y las líneas de investigación futuras.

4.1. Conclusiones

Se describen, a continuación, las principales conclusiones estructuradas de acuerdo con los objetivos inicialmente planteados. Dado el formato de Tesis Doctoral por compendio de publicaciones, estas conclusiones se encuentran desarrolladas en mayor profundidad en los estudios derivados.

OE1: Analizar la relación entre evaluación y calidad, y su relevancia en el proceso de enseñanza-aprendizaje mediado por la tecnología.

Para dar respuesta al primer objetivo específico, y dada la relación bidireccional entre la evaluación y la calidad, se ha llevado a cabo una revisión sistemática sobre los procesos de evaluación de la calidad en la actualidad que ha sintetizado los principales modelos disponibles.

A este respecto, es importante destacar la conexión intrínseca entre calidad y evaluación en el contexto educativo (Koslowski, 2006). Las evidencias alcanzadas en la investigación permiten poner de relieve la importancia crítica de determinar la calidad de las prácticas evaluativas, no solo para garantizar la confianza en el proceso, sino también para respaldar la implementación exitosa de los dispositivos tecnológicos en el nuevo paradigma educativo (Connolly et al., 2005).

Este enfoque sienta las bases para comprender con mayor profundidad los hallazgos de la revisión realizada y del modelo propuesto, justificando su relevancia en el impulso de prácticas evaluativas de calidad en un entorno tecnológico.

En este sentido, la revisión sistemática ha permitido identificar doce dimensiones que afectan a la calidad de las enseñanzas mediadas por la tecnología propuestas en las investigaciones analizadas. Las dimensiones más representadas se centran en la calidad del sistema, de la función docente, del acceso para el usuario y de los contenidos.

En una comparativa con una revisión similar previa (Marciniak y Gairín-Sallán, 2018), se constata la existencia de nueve dimensiones en común. Por el contrario, cinco dimensiones no concuerdan, lo que refleja la amplitud de los componentes de la calidad y la dificultad en su delimitación (Fernández, 2005). Otra revisión relevante centrada en la calidad de los cursos en línea masivos y abiertos (MOOC) (Duart et al., 2017) estableció como dimensiones evaluables la planificación, la metodología, los contenidos, la comunicación, los recursos y la motivación. Por consiguiente, se observa que solamente dos de ellas guardan relación con las identificadas en la revisión sistemática y se reafirma la variabilidad de los modelos de calidad.

Derivado del volumen de los modelos coexistentes (Mejía y López, 2016), autores como Park y Lim (2015) establecen como línea de investigación prioritaria el diseño de un modelo común y consensuado que tome los constructos de las principales propuestas y establezca un nuevo marco para la evaluación de la calidad en enseñanzas virtuales. Esta nueva propuesta deberá abordar un mayor número de factores y tratar de evaluar todas las dimensiones vinculadas con la calidad.

Por este motivo, y tomando esta perspectiva como punto de partida, se ha desarrollado la propuesta de un modelo para la evaluación de la calidad en e-Learning en Educación Superior. Este modelo tomará como base los modelos extraídos en la revisión sistemática de la literatura.

Estudios previos han profundizado en una única dimensión de la calidad (Kurilovas y Vinogradova, 2016) o planteado el estudio de dimensiones específicas (La Rotta et al., 2020; Serrano et al., 2018). En contraposición, el instrumento diseñado en esta investigación presenta una más amplia delimitación de la calidad.

El nuevo modelo propuesto permite la evaluación de todas las fases del proceso de enseñanza-aprendizaje en entornos virtuales, evaluando desde la calidad de la fase de preparación inicial o los recursos empleados hasta la calidad de la evaluación final del proceso. El modelo ha sido validado por un grupo de expertos de evaluación y calidad. Investigaciones futuras pueden llevar a cabo su validación empírica mediante la evaluación de la calidad de una formación en línea.

Este modelo debe garantizar estándares altos de calidad en los procesos de evaluación adaptados a las tecnologías actuales en educación, haciendo de la evaluación un proceso válido y fiable. Por tanto, los dos estudios realizados han mostrado la total vinculación entre calidad y su

evaluación en la necesidad de comprender y mejorar la efectividad de las prácticas evaluativas en el ámbito de la educación mediada por la tecnología.

En base a esto, la evaluación de la calidad se posiciona como un componente crucial para evaluar el impacto de los cambios educativos y las innovaciones tecnológicas en el proceso de enseñanza.

En consecuencia, esta evaluación es esencial para asegurar no solo la fiabilidad de los resultados, sino también para respaldar una implementación efectiva de la tecnología en el proceso de evaluación (García-Soto et al., 2020).

OE2: Describir el estado actual de las investigaciones en materia de e-evaluación en los últimos años, sus tendencias y las líneas futuras de investigación.

Para dar respuesta al segundo objetivo específico se ha llevado a cabo un mapeo de la literatura sobre la evaluación mediada por la tecnología que comprende las publicaciones relativas a la temática realizadas entre 2017 y 2021.

Los resultados permiten concluir la existencia de un elevado interés en la comunidad científica e investigadora sobre la evaluación en entornos digitales, una conclusión alcanzada tras constatar la tendencia creciente de investigaciones. Esta tendencia creciente puede tener su explicación en los resultados obtenidos en estudios previos que confirman los beneficios de la tecnología en procesos de evaluación (Huda et al., 2020).

En segundo lugar, la educación superior presenta un mayor número de estudios que el resto de los niveles educativos abordados y son los estudiantes el colectivo objeto de un mayor número de investigaciones. Esta afirmación guarda relación con las conclusiones alcanzadas por Tonbulglu (2023) en un análisis bibliométrico reciente sobre la materia, quien determinó que las principales investigaciones en evaluación se encuentran centradas, en dicho colectivo, en estudios que analizan su rendimiento, sus experiencias y su satisfacción con el proceso evaluativo.

Entre las investigaciones que abordaron la e-evaluación en educación superior desde la perspectiva de los estudiantes, 36 analizaron los resultados de los procesos de e-evaluación, 23 estudios la actitud y la percepción y 12 la satisfacción. Por el contrario, sólo 11 estudios se centraron en la e-evaluación con profesores, y las tendencias más repetidas fueron la actitud del docente (5), la satisfacción (4) y la aceptación (3). Éstas pueden ser también áreas potenciales de desarrollo en investigaciones próximas.

Por otro lado, los sistemas LMS se manifiestan como las herramientas más descritas en estudios de evaluación en entornos digitales (60,8%) y Moodle como la plataforma más empleada. Este hecho refleja la relevancia y el uso generalizado de los sistemas LMS en educación, lo que también es coherente con su papel clave durante la pandemia como principal método de continuidad del

proceso de enseñanza-aprendizaje (Dascalu et al., 2021; Grigoryeva et al., 2021; Raza et al., 2021). Esta conclusión se encuentra en consonancia con la revisión sistemática desarrollada por Gamage et al. (2022), quienes constataron que Moodle es el principal soporte tecnológico de las enseñanzas mediadas por la tecnología.

Por tanto, la investigación tecnológica y la investigación evaluativa comparten un desarrollo paralelo y complementario en el sistema educativo actual (Yangari e Inga, 2021). En este escenario, el desarrollo de la tecnología se posiciona como un factor impulsor esencial, no solo promoviendo la adopción de herramientas innovadoras, sino también generando cambios fundamentales en las prácticas educativas. Por otro lado, los avances en evaluación suponen un papel crucial en este proceso, desarrollando nuevos marcos que posibilitan su inclusión en el entorno digital actual, y favorecen tanto las innovaciones docentes como la mejora del proceso de enseñanza-aprendizaje.

OE3: Identificar los principales modelos de adopción tecnológica de dispositivos móviles en procesos de evaluación.

En relación con el tercer objetivo específico, se han identificado los principales modelos de adopción tecnológica relativos a la evaluación mediada por dispositivos móviles, un ámbito en desarrollo dado el potencial de esta tecnología en educación (Atiaja, 2016), mediante una revisión sistemática de la literatura.

En este sentido, la revisión llevada a cabo contrasta con la única realizada hasta el momento por Alrofouh et al. (2019). Esta revisión mencionada seleccionó ocho modelos basados exclusivamente en TAM (Davis, 1989) centrados en la evaluación mediada por dispositivos móviles, todos ellos incluidos en este estudio. Por tanto, la presente investigación ha tratado de profundizar en un número más significativo de modelos de adopción y de repositorios para reflejar la casi totalidad de las publicaciones sobre el tema. Una totalidad de publicaciones que, dado el volumen de artículos extraídos (15), demuestra el estado inicial en el que se encuentran los estudios de adopción de esta modalidad evaluativa. Pese a este bajo número de resultados, la revisión también ha confirmado la relevancia del campo, reflejada en el volumen de citas alcanzadas por las investigaciones (631 desde 2015).

En contraposición a la revisión de Alrofouh et al. (2019), esta investigación ha encontrado modelos basados en teorías distintas (como la SDT, la UTAUT o el modelo MBAAM). Por tanto, se muestra la tendencia creciente a ampliar los modelos de adopción para el uso de dispositivos móviles en el campo de la evaluación.

Las dos revisiones han constatado que los estudios de adopción tecnológica en el campo han sido desarrollados mayoritariamente con estudiantes universitarios. En consecuencia, se reafirma la necesidad de un modelo de adopción tecnológica para docentes, ya que son ellos los que tienen que

diseñar, implementar y desarrollar evaluaciones mediadas por dispositivos móviles en sus programaciones (Sáez-López, 2015). Esta revisión ha puesto también de manifiesto la disparidad y el bajo tamaño de las muestras utilizadas en los procesos de validación de los modelos, por lo que se recomienda que futuras investigaciones intenten contar con muestras más amplias para otorgar una mayor validez a las propuestas.

Entre los modelos encontrados en las dos revisiones, destacan los desarrollados por Nikou y Economides (Nikou y Economides, 2014, 2017a, 2017b, 2019). En el modelo más relevante hasta el momento (MBAAM), los autores plantean un total de siete constructos para explorar los factores que condicionan la adopción tecnológica. Este modelo se basa en el modelo TAM, el más desarrollado en estudios de aceptación (Zaineldeen et al., 2020), y se complementa con dimensiones procedentes de diferentes modelos y teorías (Agarwal y Prasad, 1998; Ong et al., 2004; Venkatesh et al., 2003; entre otros).

En la revisión, se identificaron un total de 29 constructos diferentes, siendo los constructos de la teoría TAM y SDT los más frecuentemente abordados: utilidad percibida (14), facilidad de uso percibida (13), actitud hacia el uso (8), autonomía (5), competencia percibida (5) y relación percibida (5). La varianza media explicada entre el total de publicaciones es del 59,18%, lo que sugiere que la mayoría de los modelos actuales explican una varianza media-alta de la intención de uso de dispositivos móviles en procesos de evaluación. Esta varianza media es relevante para la comparación con el nuevo modelo diseñado en la investigación, ya que posibilitará un estudio comparativo en la determinación del modelo más explicativo y predictivo.

Finalmente, los resultados de la revisión muestran el estado inicial y creciente de los estudios de adopción tecnológica de dispositivos móviles en procesos de evaluación, un campo que mantiene abiertas líneas de investigación de amplio recorrido futuro, como la búsqueda de un nuevo modelo de adopción que aúne los constructos descritos o el desarrollo de estudios en el colectivo docente. Además, las investigaciones consideran relevante la introducción de distintas teorías a TAM para los estudios de este campo, la introducción de factores como el entretenimiento y el disfrute percibido o constructos como la retroalimentación cognitiva y la ubicuidad percibida en modelos futuros.

OE4: Validar una propuesta inicial de modelo de adopción tecnológica de dispositivos móviles en evaluación.

El cuarto objetivo específico ha implicado el diseño y validación de una primera propuesta de modelo de adopción tecnológica de dispositivos móviles en procesos de evaluación, unos dispositivos de amplio recorrido en su desarrollo metodológico en el aula (Fombona-Cadavieco y Rodil-Pérez, 2018; Villavicencio y Paredes, 2020; Williams-Muller et al., 2020), pero escaso en procesos evaluativos.

En relación con este objetivo, se concluye que la evaluación a través de dispositivos móviles afronta un nuevo reto en relación con las metodologías que hoy se implementan en las aulas, situando al alumno en el centro del proceso y al docente como responsable de las innovaciones, las transiciones y los cambios que acontecen (Coulby et al., 2011).

En estos procesos de transformación, la aceptación docente tiene que actuar como elemento mediador entre las prácticas de evaluación y la tecnología. Por este motivo, es fundamental establecer políticas de formación, capacitación y preparación docente que conduzcan a una integración efectiva de estas nuevas formas de evaluación en la práctica (Basurto y Navareño, 2020; Cabero-Almenara et al., 2018; Jiménez-Martínez et al., 2022; Oke y Fernandes, 2020; Zabala-Guirado et al., 2020).

La propuesta inicial, que parte de los modelos más relevantes obtenidos en la revisión sistemática y del modelo TAM, es un modelo teórico de doce dimensiones que se relacionan mediante 17 hipótesis y 50 ítems. Este modelo ha sido sometido a un proceso de validación de contenido por seis expertos nacionales e internacionales.

En una comparativa con el modelo más relevante (MBAAM; Nikou y Economides, 2017b), se detectan cinco constructos comunes y doce constructos diferenciados entre las propuestas (Tabla 7). Esta variabilidad deriva de la amplitud de factores relacionados con la integración de estos dispositivos en el proceso de evaluación.

TABLA 7. Comparativa de dimensiones entre el modelo MBAAM y el modelo propuesto

Dimensiones comunes	Dimensiones comunes adaptadas	Dimensiones no comunes (Nikou y Economides, 2017b)	Dimensiones no comunes (Propuestas en Tesis Doctoral)
Facilidad de uso percibido	Innovación personal	Condiciones facilitantes	Autonomía
Intención conductual de uso	Autoeficacia	Confianza percibida	Norma subjetiva
Utilidad percibida		Contenidos	Resistencia al cambio
		Feedback Cognitivo	Ventaja relativa de la evaluación formativa
		Influencia social	Ventaja relativa de la evaluación sumativa
		Interfaz	
		Ubicuidad percibida	

En base a esta comparación, solamente se identifican los tres constructos propios del modelo TAM en común y dos dimensiones similares adaptadas a cada contexto (innovación personal y autoeficacia). Esta investigación ha llevado a cabo la propuesta de un nuevo modelo y no ha aplicado el modelo MBAAM con el objetivo de ampliar la búsqueda de antecedentes de la intención de uso no abordados previamente en el campo de la evaluación mediada por dispositivos móviles. Además, una segunda aportación del nuevo modelo es la inclusión de elementos referidos a la evaluación formativa y sumativa en estudios de adopción tecnológica.

Por otro lado, el modelo MBAAM combina dimensiones vinculadas con factores internos (feedback, confianza percibida, influencia social, ubicuidad percibida) y externos (condiciones facilitantes, contenidos o interfaz) como condicionantes de la adopción. En contraposición, la propuesta que aquí se realiza presenta constructos únicamente vinculados con factores condicionantes internos.

Esta elección se justifica en virtud de considerar que los aspectos psicológicos y motivacionales intrínsecos pueden desempeñar un papel crítico en la adopción de la tecnología en el ámbito educativo (Ertmer, 1999). Al limitar el enfoque a variables internas, se busca una comprensión más profunda de cómo estas afectan a la adopción de los docentes. Además, estas barreras son más difíciles de medir, pero son también las más comunes entre docentes, por lo que se ha optado por su estudio en profundidad. Adicionalmente, esta simplificación se fundamenta en el objetivo de proporcionar un modelo más sencillo y práctico.

Por tanto, la aplicación del nuevo modelo tras su validación permitirá realizar la validación estructural de la propuesta y contrastar las dimensiones con el resto de los modelos actuales, ampliando el estudio de los factores vinculados con la adopción tecnológica en este ámbito y posibilitará también la comparativa en su ajuste y funcionamiento con el modelo MBAAM.

Una vez diseñado el modelo se comprobaron empíricamente los factores incluidos dando como resultado una segunda versión que corrige algunas carencias y limitaciones del modelo aquí propuesto.

OE5: Determinar, de manera empírica, el funcionamiento de los factores incluidos en el modelo inicial.

Para dar respuesta al objetivo, se ha llevado a cabo la validación empírica del modelo descrito previamente a través de dos análisis diferenciados en una muestra de futuros docentes de la Universidad de Salamanca.

Estos análisis han propuesto diferentes versiones del modelo diseñado que reformulan la utilidad percibida en diferentes utilidades específicas de los dispositivos móviles en evaluación formativa y sumativa. Los resultados alcanzados en estudios previos han mostrado la relevancia de la utilidad percibida de TAM en los modelos de adopción tecnológica en el ámbito del m-Learning (Harchay et al., 2019), pero hasta la fecha no se ha publicado literatura sobre la utilidad percibida que haga alusión a las modalidades de la evaluación (formativa/sumativa).

En primer término, se confirma que la propuesta de un modelo con dos utilidades percibidas específicas explica un mayor porcentaje de la varianza de la intención de uso (49,2%) que un constructo de utilidad genérico basado en TAM (41%) en el análisis realizado en PLS-SEM. Además, estas propuestas se confirman también como más parsimoniosas y generalizables (Bozdogan, 1987) al alcanzar valores más bajos en el Criterio De Información de Akaike (AIC) (valor de 46,72 en el

modelo específico frente a un 92,27 en el modelo general) y en el Criterio de Información Bayesiano (BIC) (-203,24 frente a -156,17).

Por otro lado, en el análisis llevado a cabo en CB-SEM (incluyendo la autoeficacia como antecedente), el modelo con un constructo específico de utilidad percibida de la evaluación ha explicado un 9,4% más de varianza de la intención conductual de uso que el modelo genérico basado en TAM (62,9% frente a 53,5%). En lo referido al AIC y BIC de las dos propuestas, ambas alcanzan valores muy similares, siendo ligeramente más parsimonioso el modelo basado en TAM (13130,45 frente a 13348,06 y 13330,28 frente a 13747,90; respectivamente).

En relación con el resto de los antecedentes propuestos, la autonomía ha sido un factor determinante en los dos nuevos modelos de covarianza propuestos. Este hallazgo confirma investigaciones previas que han incorporado la autonomía al Modelo de Aceptación Tecnológica (TAM) y han determinado que es un constructo relevante en modelos de adopción (Roca y Gagné, 2008).

En segundo lugar, la autoeficacia en el uso de dispositivos móviles también se ha posicionado como un constructo vinculante en los modelos específicos generados, pero su funcionamiento no se ha confirmado en uno de los cinco modelos propuestos. El funcionamiento de la autoeficacia en los modelos que sí lo han demostrado guarda relación con investigaciones previas (Chau, 2001; Hasan, 2007) que confirman la influencia de la autoeficacia en el uso tecnológico sobre la utilidad percibida en estudios de adopción de ordenadores.

A continuación, se presenta una comparativa de las dos propuestas más explicativas de la investigación (una extraída del análisis CB-SEM con un constructo único de utilidad y una derivada del análisis PLS-SEM con dos constructos específicos) con el modelo más relevante del campo (MBAAM), anteriormente descrito.

Esta comparación es similar en términos muestrales, ya que el estudio desarrollado con Nikou y Economides (2017b) para la validación de su modelo contó con 145 estudiantes universitarios frente a los 262 empleados en esta investigación. El análisis del modelo MBAAM fue realizado mediante PLS-SEM.

La principal diferencia se encuentra vinculada con las características de la muestra, ya que los estudiantes aquí abordados son futuros docentes, lo que derivará en un mayor nivel de procesamiento de la respuesta dada su futura vinculación con el proceso.

En primer lugar, los dos modelos explorados en esta investigación explican un mayor porcentaje de varianza que el modelo propuesto por Nikou y Economides (2017b) (Tabla 8). Esta comparativa no describe ni el Criterio de Información de Akaike ni el Criterio de Información Bayesiano dado que los autores no introducen estos criterios en el estudio analizado.

TABLA 8. Constructos y varianza explicada del modelo MBAAM y los modelos de la investigación

Modelo	Constructos	Varianza (R ²)
Modelo MBAAM	Intención de uso, Facilidad de uso percibida, Utilidad percibida, Contenido, Feedback cognitivo, Condiciones facilitantes, Ansiedad ante la evaluación con dispositivo móvil, Autoeficacia, Innovación personal, Confianza percibida, Ubicuidad percibida, Influencia social, Interfaz de usuario.	46,8%
Modelo propio (CB-SEM): Modelo B, estudio 1	Intención de uso, Utilidad percibida de los dispositivos móviles en procesos de evaluación, Facilidad de uso percibida, Autonomía, Autoeficacia en el uso de dispositivos móviles.	62,9%
Modelo propio (PLS-SEM): Modelo B, estudio 2	Intención de uso, Utilidad percibida de los dispositivos móviles en evaluación formativa, Utilidad percibida de los dispositivos móviles en evaluación sumativa, Facilidad de uso percibida, Autoeficacia en el uso de dispositivos móviles.	49,2%

Del análisis se concluye que los dos modelos derivados de los estudios en CB-SEM y PLS-SEM explican un mayor porcentaje de varianza que el modelo MBAAM, siendo además modelos más sencillos al contar con un menor número de constructos (13 en MBAAM frente a 6 constructos en el modelo evaluado en CB-SEM y 5 en el modelo evaluado en PLS-SEM), un menor número de hipótesis (12 frente a 5 el modelo CB-SEM y 7 en el modelo PLS-SEM), y un menor número de ítems (47 indicadores frente a 27 y 21, respectivamente).

Por tanto, la comparativa con el modelo más relevante del campo valida las propuestas realizadas al generarse dos versiones de un modelo más sintético, predictivo (modelo PLS-SEM) y explicativo (modelo CB-SEM).

En la literatura, la comparativa del modelo TAM con otros modelos y teorías es una práctica extendida entre investigadores. Entre ellas, destaca su comparación con la Teoría del Comportamiento Planeado (Mathieson, 1991) o con su versión Descompuesta (Chau y Hu, 2001). Otros estudios han confirmado los beneficios de adaptar el modelo de Davis (1989) a un ámbito específico del conocimiento, explicando un mayor porcentaje de varianza y generando modelos más parsimoniosos y generalizables (Dadayan y Ferro, 2005).

Esta conclusión se alcanza también en el estudio realizado en esta investigación, donde un modelo específico para el campo de la e-evaluación tiene un mejor ajuste que un modelo genérico basado en TAM.

Por tanto, esta investigación ha comprobado el funcionamiento de algunos de los factores incluidos en el modelo inicial y ha validado dos propuestas que reformulan la utilidad percibida para modelos de adopción: una propuesta que establece dos constructos de utilidad percibida de los dispositivos móviles en evaluación (uno formativo y otro sumativo), y una segunda propuesta que presenta un constructo único referido a la utilidad percibida de los dispositivos móviles en procesos de evaluación.

Se describen, a continuación, las conclusiones relativas al objetivo general de la investigación, las cuales abordan la propuesta de un modelo final para el estudio de la adopción tecnológica de dispositivos móviles en evaluaciones formativas y sumativas en docentes.

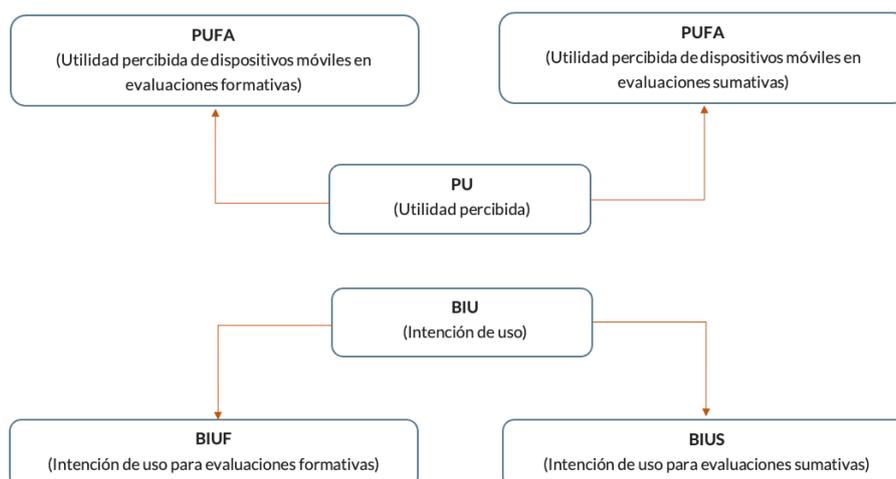
OBJETIVO GENERAL: Diseñar un modelo de adopción tecnológica para el estudio de los factores que condicionan el uso de dispositivos móviles en evaluación por parte del profesorado.

En primer lugar, los estudios realizados en la investigación han confirmado que la evaluación es un proceso heterogéneo en el que coexisten una elevada cantidad de factores que la condicionan, sujetos a la modalidad de evaluación (formativa o sumativa). Por tanto, es importante atender a estas características evaluativas en los modelos de adopción tecnológica.

Dicha cuestión se ha evidenciado en el estudio intermedio realizado, donde se observa una menor capacidad predictiva en los modelos genéricos, como el modelo TAM, en comparación con modelos específicos adaptados a la evaluación (formativa/sumativa). Por otro lado, este estudio intermedio ha confirmado también que el diseño de dos utilidades diferenciadas deriva, a su vez, en un modelo más específico y parsimonioso que un constructo único de utilidad de la evaluación.

Estos resultados indican, por tanto, que el camino a seguir es la propuesta de dos variantes diferenciadas de un mismo modelo adaptadas a las particularidades de cada modalidad evaluativa. Por este motivo, se propone la reformulación de la intención de uso percibida en las versiones finales del modelo, siguiendo el mismo proceso realizado con la utilidad en los estudios previos de la investigación (Figura 23).

FIGURA 23. Reformulación de la utilidad percibida y la intención conductual de uso



La facilidad de uso percibida no ha sido un constructo antecedente de la intención de uso ni de la utilidad percibida en ningún modelo propuesto durante la fase experimental. Entre las posibles

causas se encuentra la alta competencia digital de los participantes, la cual ha podido condicionar su funcionamiento. Investigaciones previas afirman que la facilidad de uso no es una variable relevante para sujetos con un nivel competencial tecnológico elevado (Evans y Robertson, 2020; Recio-Muñoz et al., 2020).

Pese a esto, la determinación final es mantener el constructo en las propuestas que se presentan a continuación. Esta decisión se toma en base a su adecuado funcionamiento en modelos similares con muestras distintas (Nikou y Economides, 2017a, 2017b) y en estudios que afirman que este constructo es parte fundamental de los modelos de adopción (Venkatesh y Davis, 1996). Por tanto, se mantiene la dimensión y se recomienda la búsqueda de nuevos antecedentes de la facilidad de uso que contribuyan a un mejor funcionamiento en modelos posteriores.

De esta forma, se describe a continuación un modelo para el estudio de la adopción de dispositivos móviles en evaluación con una versión para evaluación formativa y una segunda propuesta para evaluaciones sumativas.

Modelo F/S-MAAM: Adopción de dispositivos móviles en evaluaciones formativas y sumativas

El modelo propuesto (F/S-MAAM, Formative/Summative – Mobile-based Assessment Acceptance Model) aborda la adopción tecnológica del uso de dispositivos móviles en evaluaciones de tipo formativo o sumativo. El modelo establece, como ya se ha indicado, dos versiones que difieren en los constructos de utilidad percibida de los dispositivos para evaluación (formativa o sumativa) y en la intención de uso para evaluar (formativa o sumativamente) mediante dispositivos móviles.

El nuevo modelo mantiene los constructos ya definidos en su diseño inicial, reformula la utilidad en utilidades percibidas de la evaluación sumativa/formativa, y sustituye las ventajas relativas de la evaluación por dos nuevos antecedentes, el rechazo ante la incertidumbre y la compatibilidad con el estilo de trabajo.

Respecto al rechazo ante la incertidumbre como antecedente de la intención de uso, este constructo evalúa la forma en la que un individuo, en una sociedad específica, tolera la falta de certezas en sus tareas. Investigaciones previas han demostrado que culturas con niveles bajos de tolerancia ante situaciones inciertas tienden a depender de normas específicas para su correcto funcionamiento (Sadeghi et al., 2014), actuando como una variable moderadora con la norma subjetiva. Las hipótesis diseñadas en el modelo se basan en las propuestas de Hofstede (2001) y Tarhini et al. (2017).

El segundo constructo que se incluye en el modelo final refiere a la compatibilidad con el estilo de trabajo. Este constructo, derivado de la Teoría de Difusión de Innovaciones, se define como “the degree to which an innovation is perceived as being consistent with the existing values, past

experiences, and needs of potential adopters” [grado en que una innovación se percibe como consistente con los valores existentes, experiencias pasadas y necesidades de posibles adoptantes] (Rogers, 1962, p. 15).

En el modelo diseñado, este constructo multidimensional hace referencia al grado en que los cambios que conlleva la incorporación tecnológica a procesos evaluativos son compatibles con la forma de trabajar del docente. Así, se propone que este constructo tiene influencia sobre la utilidad percibida y sobre la intención de uso (Karahanna et al., 2006).

Por tanto, el modelo final se sustenta en 15 hipótesis relacionales, que se presentan a continuación y que definen el modelo F/S-MAAM y sus dos versiones (Tabla 9).

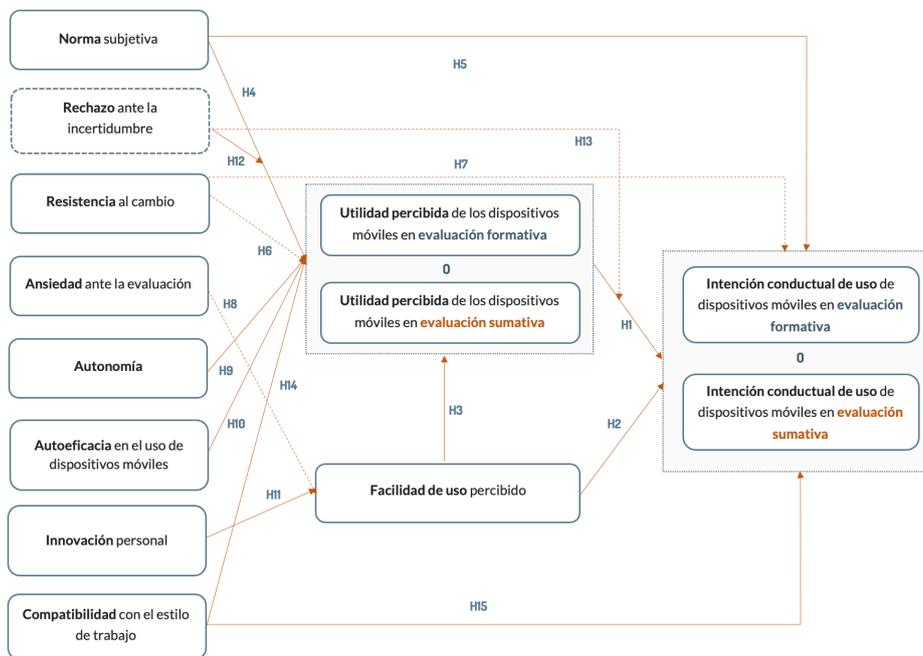
TABLA 9. Dimensiones del modelo F/S-MAAM

Constructo	Hipótesis
Intención de uso, utilidad percibida y facilidad de uso percibido	
H1	La utilidad percibida de los dispositivos móviles para evaluación (formativa o sumativa) predice positivamente la intención docente de usar dispositivos móviles en evaluaciones de modalidad formativa o sumativa en la práctica docente.
H2	La facilidad de uso percibida de los dispositivos móviles para evaluación predice positivamente la intención docente de usarlos en evaluaciones (formativas o sumativas) en la práctica docente.
H3	La facilidad de uso percibida predice positivamente la utilidad percibida de los dispositivos móviles para evaluaciones formativas o sumativas en la práctica docente.
Norma subjetiva	
H4	La norma subjetiva tiene una relación positiva con la utilidad percibida por los docentes de los dispositivos móviles en evaluación formativa o en evaluación sumativa.
H5	La norma subjetiva tiene una relación positiva con la intención de utilizar tecnologías móviles en procesos de evaluación por docentes.
Resistencia al cambio	
H6	La resistencia al cambio predice negativamente la utilidad percibida (formativa o sumativa) de los dispositivos móviles para evaluar por parte de los docentes.
H7	La resistencia al cambio predice negativamente la intención de uso docente (formativa o sumativa) de dispositivos móviles en procesos de evaluación.
Ansiedad ante la evaluación	
H8	La ansiedad ante la evaluación tendrá un efecto negativo en la facilidad de uso percibida por los docentes de los dispositivos móviles en evaluación.
Autonomía	
H9	La autonomía tendrá un efecto positivo en la utilidad percibida (formativa o sumativa) por los docentes en el uso de dispositivos móviles para evaluar.
Autoeficacia en el uso de dispositivos móviles	
H10	La autoeficacia en el uso de dispositivos móviles tendrá un efecto positivo en la utilidad percibida (formativa o sumativa) por los docentes para evaluar.
Innovación personal	
H11	La innovación personal tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para evaluar.
Rechazo ante la incertidumbre	
H12	El rechazo ante la incertidumbre modera positivamente la relación predictiva entre la norma subjetiva y la utilidad percibida (formativa o sumativa) de los dispositivos móviles para evaluación.

Constructo	Hipótesis
H13	El rechazo ante la incertidumbre modera negativamente la relación entre la utilidad percibida (formativa o sumativa) de los dispositivos móviles para evaluación y su intención conductual de uso (formativa o sumativa).
Compatibilidad con el estilo de trabajo	
H14	La compatibilidad percibida influye positivamente en la utilidad percibida de dispositivos móviles en evaluaciones formativas o sumativas.
H15	La compatibilidad percibida se encuentra positivamente relacionada con la intención de uso de dispositivos móviles en evaluaciones formativas o sumativas.

Los ítems relativos al modelo F/S-MAAM (Figura 24) y sus versiones pueden ser consultados en el anexo B de la Tesis Doctoral.

FIGURA 24. Modelo F/S-MAAM



Finalizadas las conclusiones y propuesto el modelo F/S-MAAM en sus dos versiones, se presenta a continuación la implicación que estos hallazgos tendrán en la práctica educativa y en la relación de los docentes con los dispositivos móviles. Posteriormente, se presentan las limitaciones de la Tesis Doctoral y las líneas futuras de investigación.

4.2. Implicaciones educativas

Las contribuciones de la Tesis Doctoral no solo trascienden la propuesta de un modelo que distingue la intención de uso de dispositivos móviles en evaluaciones formativas y sumativas, y amplía el campo de la investigación sobre adopción tecnológica, sino que poseen también una relevancia en el ámbito formativo de los docentes en ejercicio y los futuros profesionales de la educación.

Comprender los factores que condicionan la aceptación final de la tecnología en los procesos de evaluación educativa es esencial para la actualización y modernización de la formación de los futuros docentes. En este sentido, los resultados de la investigación facilitan las dimensiones específicas condicionantes de la adopción tecnológica, un componente esencial para la formación básica de los profesionales educativos, proporcionando una base sólida para la integración efectiva de dispositivos móviles en procesos de evaluación. Entender cómo los docentes perciben y adoptan estas tecnologías, qué barreras encuentran y cómo superarlas, facilitará también a las instituciones educativas las líneas de acción necesarias para la adaptación y mejora de sus programas formativos, posibilitando el abordaje de las necesidades cambiantes en la educación mediada por la tecnología.

La utilidad del modelo F/S-MAAM se extiende más allá de la formación inicial ya descrita. La formación continua de los docentes en activo es parte esencial en el diseño de programas y actuaciones específicas orientadas a la introducción de las innovaciones tecnológicas y el diseño de nuevas prácticas pedagógicas. Al considerar los factores que afectan la adopción de dispositivos móviles en evaluación, se proporcionará un marco útil para el desarrollo profesional continuo de los educadores que facilitará la integración efectiva de la tecnología en sus prácticas diarias.

Por último, las implicaciones de la Tesis Doctoral presentada se extienden también a nivel institucional, ya que conocer los factores de adopción podrá influir, de forma directa o indirecta, en las políticas formativas en materia de inclusión tecnológica y evaluación adoptadas por las instituciones educativas. Los resultados derivados de las aplicaciones futuras del modelo podrán suponer una guía útil para la formulación de políticas que promuevan la integración efectiva de dispositivos móviles en los procesos de evaluación. Este cambio debe dejar atrás la formación en el uso de la tecnología y su facilidad, un componente ya desarrollado en el futuro profesorado (Moyano y Spinelli, 2020) y apostar por la formación en recursos, herramientas e innovaciones que diferencien entre procesos de evaluación formativa y sumativa, facilitando al docente un amplio marco de actuación.

4.3. Limitaciones

En términos generales, las limitaciones de las fases de revisión se centran en la dificultad para alcanzar un número de bases de datos superior, así como la limitación de abordar publicaciones en lenguas distintas al inglés o publicadas en un mayor marco temporal. Estas revisiones son susceptibles también de contar con sesgos de publicación múltiple y de confirmación en los estudios seleccionados (Beltrán, 2005) así como sesgos derivados del proceso de selección de publicaciones y aplicación de criterios (García-Peñalvo, 2022).

Por otro lado, las limitaciones derivadas del proceso de validación y aplicación del modelo son principalmente de corte metodológico. En relación con su validación, futuras propuestas pueden contar con un número mayor de expertos para tal proceso, obteniendo una mayor validez de los

resultados. Pese a ello, las dos validaciones desarrolladas han superado el número mínimo de jueces establecidos en la literatura (Hernández-Nieto, 2002; Pedrosa et al., 2013).

En relación con la aplicación y el análisis, los resultados obtenidos en la investigación podrían encontrarse comprometidos por las características de la muestra y su acceso. Pese al intento por garantizar una alta heterogeneidad de grupos y la mayor muestra posible, los participantes compartían una formación bajo el mismo programa académico en la Universidad de Salamanca. Futuros estudios pueden ampliar el acceso a participantes de otras instituciones y ámbitos geográficos para obtener unos resultados de mayor validez.

4.4. Líneas de investigación futuras

Las líneas futuras de investigación relacionadas con las conclusiones y avances obtenidos de la Tesis Doctoral se presentan en dos categorías: aquellas líneas de avance ya comenzadas y las líneas futuras de exploración.

4.4.1. Líneas de investigación en exploración actual

En el momento de presentación de la Tesis Doctoral, son dos las líneas de investigación que se encuentran abiertas en relación con el modelo último propuesto. Estas líneas tratan validar las nuevas versiones y determinar sus principales características.

1) Línea de investigación activa 1: Aplicación del modelo en titulaciones de grados en maestro

Como primera línea de investigación ya comenzada, se ha aplicado el modelo F/S-MAAM en los Grados en Maestro en Educación Infantil, Maestro en Educación Primaria y el Doble Grado en Maestro en Educación Infantil y Maestro en Educación Primaria de la Universidad de Salamanca en sus tres campus universitarios (Ávila, Salamanca y Zamora).

Esta investigación se encuentra, en este momento, en fase de análisis de datos tras un proceso de recogida de información realizado durante el mes de mayo de 2023 con un periodo de encuestación de 23 días tras contar con la aprobación del Comité de Ética de la Investigación de la Universidad de Salamanca. La muestra del estudio se encuentra conformada por 760 estudiantes (184 estudiantes de Ávila, 366 de Salamanca y 210 de Zamora).

2) Línea de investigación activa 2: Aplicación del modelo F/S-MAAM ampliado en Portugal

En esta segunda línea de investigación ya iniciada, se está llevando a cabo la experimentación con el modelo F-MAAM, complementados con nuevos antecedentes para tratar de mejorar su funcionamiento.

La recogida de información ha dado comienzo durante la estancia realizada en Leiria (Portugal) en junio de 2023, en un proceso que sigue abierto a fecha de depósito de la Tesis Doctoral y ha sido

distribuido a rectores, decanos y coordinadores de diferentes universidades y politécnicos portugueses a través de la aplicación web *LimeSurvey*. El cuestionario se encuentra completo y accesible en: <https://bit.ly/3PKZXIw>

4.4.2. Líneas de investigación de exploración futura

Para finalizar, se proponen siete líneas de investigación no exploradas hasta el momento. El total de líneas prospectivamente abordables se encuentra desarrollado en las conclusiones de los estudios derivados. Estas líneas se sintetizan en:

1. **Aplicación del modelo de evaluación de la calidad propuesto:** Investigaciones futuras pueden aplicar el modelo propuesto de calidad en un contexto específico para su validación empírica. Los resultados de esta aplicación serán relevantes para la evolución del modelo y su consolidación en el estudio de la calidad de las enseñanzas mediadas por la tecnología.
2. **Modelado de la utilidad percibida de la evaluación en estudios de adopción:** Investigaciones futuras pueden profundizar en la exploración del modelado de la utilidad percibida en modelos de adopción próximos. En el estado actual del modelo F/S-MAAM, la utilidad percibida ha sido modelada a partir de la propuesta de Moore y Benbasat (1991) para el constructo de las ventajas relativas.

Investigaciones futuras pueden reformular la utilidad original propuesta por Davis (1989) y retomar el modelo TAM como base completa de dos modelos diferenciados, permitiendo la comparación entre ambas utilidades para la determinación de la más efectiva en los modelos. Esta línea de investigación busca proporcionar una comprensión más precisa de los factores que influyen en la adopción de evaluaciones tecnológicas, fortaleciendo así la base teórica de los modelos existentes.

3. **Modelado de la facilidad de uso percibida en estudios de adopción:** Los resultados del estudio intermedio han reflejado una discrepancia con las hipótesis relativas a la facilidad de uso percibida en comparación con el modelo TAM original. En este caso, se recomienda la apertura de nuevas vías de investigación vinculadas con la reformulación de los ítems, del constructo de la facilidad de uso, o su sustitución por algún otro constructo análogo de cara a la mejora de los modelos.
4. **Aplicación del modelo F/S-MAAM en docentes en ejercicio:** La investigación propone extender la aplicación del modelo a docentes en ejercicio. La participación de estos en la validación del modelo proporcionará una perspectiva fundamental sobre el estado actual de adopción, sobre los condicionantes de la aceptación y sobre la diferenciación entre evaluación formativa y sumativa; obteniéndose unos resultados que podrán derivar en un marco de recomendaciones específicas sobre los factores condicionantes orientado a las

instituciones educativas para la inclusión efectiva de los dispositivos móviles en la evaluación educativa.

5. **Diseño de dos modelos con constructos diferenciados adaptados a las particularidades de cada modalidad de evaluación (formativa y sumativa):** El modelo F/S-MAAM propuesto establece un paralelismo de constructos en las dos propuestas realizadas que solamente difieren en la utilidad percibida y la intención de uso.

La aplicación de este modelo en docentes y su análisis posibilitará la clasificación de los constructos en aquellos más relacionados con la evaluación formativa y aquellos más vinculados con la evaluación sumativa. Este estudio permitirá el desarrollo de dos propuestas diferenciadas en los antecedentes de la intención conductual de uso: el modelo **F-MAAM** para evaluaciones formativas y el modelo **S-MAAM** para evaluaciones sumativas.

6. **Ampliación de los constructos de los modelos F-MAAM y S-MAAM:** Los resultados de un metaanálisis previo sobre dispositivos móviles y el modelo TAM (Abdullah y Ward, 2016) han mostrado que la autoeficacia, la norma subjetiva, la ansiedad ante la tecnología, el disfrute percibido y la experiencia son los factores externos al modelo TAM más utilizados. Los modelos propuestos en esta investigación contemplan los tres primeros constructos en sus delimitaciones. Dado que el disfrute percibido y la experiencia han sido también factores relevantes en el campo de la investigación, se recomienda su inclusión y validación en estudios futuros.
7. **Apertura de una línea de trabajo conjunta e internacional:** Tras los resultados alcanzados, la presente Tesis Doctoral pretende integrarse es un campo de investigación que cuenta, como autores más relevantes, con Economides y Nikou. Por tanto, se buscará la colaboración para la ampliación de los estudios de adopción y el avance de la e-evaluación mediada por dispositivos móviles.

Referencias Bibliográficas

- Abdirahman, A. A., Jabar, M. A. y Hashi, A. O. (2018). Attitude towards intention to use Mobile-Based Teaching Assessment in relation to Technology Acceptance Model. *International Journal of Advances in Electronics and Computer Science*, 5(2), 35-40.
- Abdirahman, A. A., Jabar, M. A., Hashi, A. O., Elmi, M. A. y Rodriguez, O. E. (2022). Attitude Towards Intention to Use Mobile-Based Teaching Assessment Based on TAM and a Novel Extended TAM Model. En F. Saeed, F. Mohammed, y F. Ghaleb (Eds.), *Advances on Intelligent Informatics and Computing* (pp. 583-593). Springer International Publishing. https://doi.org/10.1007/978-3-030-98741-1_48
- Abdou, R. A. (2020). Effects of e-Assessment Tools in academic achievement and motivation towards learning among pre-service kindergarten teachers in Turaif. *Amazonia Investiga*, 9(28), Artículo 28. <https://doi.org/10.34069/AI/2020.28.04.24>
- Abdullah, F. y Ward, R. (2016). Developing a General Extended Technology Acceptance Model for E-Learning (GETAMEL) by analysing commonly used external factors. *Computers in Human Behavior*, 56, 238-256. <https://doi.org/10.1016/j.chb.2015.11.036>
- Abma, T. y Stake, R. (2001). Stake's responsive evaluation: Core ideas and evolution. *New Directions for Evaluation*, 2001, 7-22. <https://doi.org/10.1002/ev.31>
- Acosta-Gonzaga, E. y Walet, N. R. (2018). The role of attitudinal factors in mathematical on-line assessments: A study of undergraduate STEM students. *Assessment & Evaluation in Higher Education*, 43(5), 710-726. <https://doi.org/10.1080/02602938.2017.1401976>
- Adams, C. y King, K. (1995). Towards a Framework for Student Self-Assessment. *Innovations in Education and Training International*, 32(4), 336-343. <https://doi.org/10.1080/1355800950320405>
- Adejo, O., Ewuzie, I., Usoro, A. y Connolly, T. (2018). E-Learning to m-Learning: Framework for Data Protection and Security in Cloud Infrastructure. *International Journal of Information Technology and Computer Science*, 10(4), 1-9. <https://doi.org/10.5815/ijitcs.2018.04.01>
- Agarwal, R. y Prasad, J. (1998). A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research*, 9(2), 204-215. <https://doi.org/10.1287/isre.9.2.204>
- Aguiar, B., Velazquez, R. y Aguiar, J. (2019). Innovación docente y empleo de las TIC en la Educación Superior. *Revista Espacios*, 40(2), 8-20.
- Ahmad, B. y Bhat, G. J. (2019). Formative and Summative Evaluation Techniques for Improvement of Learning Process. *European Journal of Business & Social Sciences*, 7(5), 776-785.
- Ajms, E. (2015). Structure Equation Modeling Basic Assumptions and Concepts: A Novices Guide. *Asian Journal of Management Sciences*, 3(1). <https://www.ajmsjournal.com/index.php/ajms/article/view/70>
- Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. En J. Kuhl y J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11-39). Springer. https://doi.org/10.1007/978-3-642-69746-3_2
- Aldunate, R. y Nussbaum, M. (2013). Teacher adoption of technology. *Computers in Human Behavior*, 29(3), 519-524. <https://doi.org/10.1016/j.chb.2012.10.017>
- Alowayr, A. y Al-Azawei, A. (2021). Predicting mobile learning acceptance: An integrated model and empirical study based on higher education students' perceptions. *Australasian Journal of Educational Technology*, 37(3), 38-55. <https://doi.org/10.14742/ajet.6154>
- Alqahtani, A. Y. y Rajkhan, A. A. (2020). E-Learning Critical Success Factors during the COVID-19 Pandemic: A Comprehensive Analysis of E-Learning Managerial Perspectives. *Education Sciences*, 10(9), Artículo 9. <https://doi.org/10.3390/educsci10090216>

- Alrofouh, A., Modi-Lakulu, M. y Almaiah, M. (2019). A Systematic Review of Mobile-Based Assessment Acceptance Studies From 2009 to 2019. *Journal of Theoretical and Applied Information Technology*, 97(20), 1-25.
- Alsharida, R., Hammood, M. y Al-Emran, M. (2021). Mobile Learning Adoption: A Systematic Review of the Technology Acceptance Model from 2017 to 2020. *International Journal of Emerging Technologies in Learning (iJET)*, 16(5), 147-162.
- Álvarez-Herrero, J. F. (2018). Aprendizaje móvil fuera del aula en Educación Secundaria. En E. López-Meneses, D. Cobos-Sanchiz, A.H. Martín-Padilla, L. Molina-García y A. Jaén-Martínez (Eds.), *Experiencias pedagógicas e innovación educativa. Aportaciones desde la praxis docente e investigadora* (pp. 937-943). Ediciones Octaedro.
- Alvira, F. (1991). *Metodología de la evaluación de programas*. Centro de Investigaciones Sociológicas (CIS).
- Anderson, L. W. (2005). Objectives, evaluation, and the improvement of education. *Studies in Educational Evaluation*, 31(2), 102-113. <https://doi.org/10.1016/j.stueduc.2005.05.004>
- Anderson, L. W. (2003). *Classroom assessment. Enhancing the Quality of Teacher Decision Making*. Erlbaum Associates.
- Area-Moreira, M. (2000). Bajo el efecto 2000. Líneas de investigación sobre Tecnología Educativa en España. *Revista Interuniversitaria de Tecnología Educativa*, 1, 98-113.
- Area-Moreira, M., Alonso, C., Gorospe, J., Pérez, M., Pons, J., Paredes, J., Chacón, J., Sanabria Mesa, A., Alonso, Á. y Valverde-Berrocoso, J. (2014). Las políticas educativas TIC en España después del Programa Escuela 2.0: Las tendencias que emergen / ICT education policies in Spain after School Program 2.0: Emerging Trends. *Revista Latinoamericana de Tecnología Educativa*, 13, 11-33.
- Atiája, L. N. (2016). Los dispositivos móviles, un potencial para el aprendizaje móvil en el contexto educativo. *Contenido. Cultura y Ciencias Sociales*, 7, 115-125.
- Awang, Z. (2012). *Structural Equation Modeling Using Amos Graphic*. UiTM Press.
- Bacca-Acosta, J. y Ávila-Garzón, C. (2020). Student engagement with mobile-based assessment systems: A survival analysis. *Journal of Computer Assisted Learning*, 37(1), 158-171. <https://doi.org/10.1111/jcal.12475>
- Barrientos-Hernán, E. J., López-Pastor, V. M. y Pérez-Brunicardi, D. (2020). Evaluación Auténtica y Evaluación Orientada al Aprendizaje en Educación Superior. Una Revisión en Bases de Datos Internacionales. *Revista Iberoamericana de Evaluación Educativa*, 13(2), 67-83.
- Bashkireva, T., Bashkireva, A., Morozov, A., Tsvetkov, S. y Popov, A. (2020). Problems of the formation of digital competence in the modern educational space. *Journal of Physics*, 1. <https://doi.org/10.1088/1742-6596/1691/1/012130>
- Bejar, I. I. (1984). Educational Diagnostic Assessment. *Journal of Educational Measurement*, 21(2), 175-189. <https://doi.org/10.1111/j.1745-3984.1984.tb00228.x>
- Bell, B. y Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85(5), 536-553. <https://doi.org/10.1002/sce.1022.abs>
- Beltrán, Ó. A. (2005). Revisiones sistemáticas de la literatura. *Revista Colombiana de Gastroenterología*, 20(1), 60-69.
- Biggs, J. (1998). Assessment and Classroom Learning: A role for summative assessment?, *Assessment in Education: Principles, Policy & Practice*, 103-110. <https://doi.org/10.1080/0969595980050106>
- Black, H. D. (1983). Introducing Diagnostic Assessment. *Programmed Learning and Educational Technology*,

- 20(1), 58-63. <https://doi.org/10.1080/0033039830200109>
- Black, P., Harrison, C., Lee, C., Marshall, B. y Wiliam, D. (2003). *Formative and summative assessment: Can they serve learning together?* Annual meeting of the American Educational Research Association, Chicago.
- Black, P. y Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Bogdanovic, Z., Barac, D., Jovanic, B., Popovic, S. y Radenkovic, B. (2014). Evaluation of Mobile Assessment in a Learning Management System. *British Journal of Educational Technology*, 45(2), 231-244. <https://doi.org/10.1111/bjet.12015>
- Bollen, K. A. (2011). Evaluating Effect, Composite, and Causal Indicators in Structural Equation Models. *MIS Quarterly*, 35(2), 359-372. <https://doi.org/10.2307/23044047>
- Boote, S. K., Boote, D. N. y Williamson, S. (2021). Assessing Graph Comprehension on Paper and Computer with MBA Students: A Crossover Experimental Study. *Cogent Education*, 8(1), 1960247. <https://doi.org/10.1080/2331186X.2021.1960247>
- Boud, D. (Ed.). (1995). *Enhancing learning through self assessment*. Routledge Falmer.
- Boud, D. y Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400-413. <https://doi.org/10.1080/02602938.2015.1018133>
- Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3), 345-370. <https://doi.org/10.1007/BF02294361>
- Brazuelo-Grund, F., Gallego-Gil, D. J. y Cacheiro-González, M. L. (2017). Los docentes ante la integración educativa del teléfono móvil en el aula. *Revista de Educación a Distancia (RED)*, 52.
- Brickner, D. L. (1995). *The effects of first and second-order barriers to change on the degree and nature of computer usage of mathematics teachers: A case study*. West Lafayette, IN: Doctoral Dissertation, Purdue University.
- Brown, V. S., Toussaint, M. y Lewis, D. (2018). Students' Perceptions of Quality across Four Course Development Models. *Online Learning*, 22(2), 173-195.
- Buchholtz, N. F., Krosanke, N., Orschulik, A. B. y Vorhölter, K. (2018). Combining and integrating formative and summative assessment in mathematics teacher education. *ZDM*, 50(4), 715-728. <https://doi.org/10.1007/s11858-018-0948-y>
- Cabero-Almenara, J. (1999). Tecnología Educativa: Diversas formas de definirla. En J. Cabero-Almenara (Ed.), *Tecnología Educativa* (pp. 17-34). Síntesis Educación.
- Cabero-Almenara, J., Gallego, O., Puentes, Á. y Jiménez, T. (2018). La "Aceptación de la Tecnología de la Formación Virtual" y su relación con la capacitación docente en formación virtual. *EDMETIC*, 7(1), 225-241.
- Caldeiro-Pedreira, M. C., Castro-Zubizarreta, A. y Havránková, T. (2021). Móviles y pantallas en edades tempranas: convivencia digital, derechos de la infancia y responsabilidad adulta. *Research in Education and Learning Innovation Archives*, 26, 1-17. <https://doi.org/10.7203/realia.26.15936>
- Casanova, M. A. (1995). *Manual de evaluación educativa*. La Muralla.
- Casanova, M. A. (1998). *La evaluación educativa*. SEP.
- Castiblanco-Jimenez, I. A., Cepeda-García, L. C., Grazia-Violante, M., Marcolin, F. y Vezzetti, E. (2020). Commonly Used External TAM Variables in e-Learning, Agriculture and Virtual Reality Applications. *Future Internet*, 13(1), 1-21.
- Castillo-Arredondo, S. y Cabrerizo-Diago, J. (2010). *Evaluación educativa de aprendizaje y competencias*. Pearson

Prentice Hall.

- Chandra, P., Tomitsch, M. y Large, M. (2020). Innovation education programs: A review of definitions, pedagogy, frameworks and evaluation measures. *European Journal of Innovation Management*, 24(4), 1268-1291. <https://doi.org/10.1108/EJIM-02-2020-0043>
- Chau, P. Y. (2001). Influence of Computer Attitude and Self-Efficacy on IT Usage Behavior. *Journal of Organizational and End User Computing (JOEUC)*, 13(1), 26-33. <https://doi.org/10.4018/joeuc.2001010103>
- Chau, P. Y. K. y Hu, P. J.-H. (2001). Information Technology Acceptance by Individual Professionals: A Model Comparison Approach. *Decision Sciences*, 32(4), 699-719. <https://doi.org/10.1111/j.1540-5915.2001.tb00978.x>
- Cheng, X., Chan, L. K., Cai, H., Zhou, D. y Yang, X. (2021). Adaptions and perceptions on histology and embryology teaching practice in China during the Covid-19 pandemic. *Translational Research in Anatomy*, 24. <https://doi.org/10.1016/j.tria.2021.100115>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S. y Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4. <https://doi.org/10.1016/j.caeai.2022.100118>
- Colás-Bravo, M. P. C., de Pablos, J. y Ballesta, J. (2018). Incidencia de las TIC en la enseñanza en el sistema educativo español: Una revisión de la investigación. *Revista de Educación a Distancia (RED)*, 56.
- Connolly, M., Jones, N. y O'Shea, J. (2005). Quality assurance and e-learning: Reflections from the front line. *Quality in Higher Education*, 11(1), 59-67. <https://doi.org/10.1080/13538320500077660>
- Cooper, I. D. (2016). What is a "mapping study?", *Journal of the Medical Library Association: JMLA*, 104(1), 76-78. <https://doi.org/10.3163/1536-5050.104.1.013>
- Córica, J. L. y García-Aretio, L. G. (2018). Teachers' reluctance to ICT in Argentina. A qualitative approach. *Educación Superior*, 25.
- Coulby, C., Hennessey, S., Davies, N. y Fuller, R. (2011). The use of mobile technology for work-based assessment: The student experience. *British Journal of Educational Technology*, 42(2), 251-265. <https://doi.org/10.1111/j.1467-8535.2009.01022.x>
- Cronbach, L. J. (1963). Course Improvement through Evaluation. *Teachers College Record*, 64(8), 1-13. <https://doi.org/10.1177/016146816306400802>
- Cruz-Benito, J., Sánchez-Prieto, J. C., Therón, R. y García-Peñalvo, F. J. (2019). Measuring Students' Acceptance to AI-Driven Assessment in eLearning: Proposing a First TAM-Based Research Model. En P. Zaphiris y A. Ioannou (Eds.), *Learning and Collaboration Technologies. Designing Learning Experiences* (pp. 15-25). Springer International Publishing. https://doi.org/10.1007/978-3-030-21814-0_2
- Cubero-Ibáñez, J., Ibarra-Sáiz, M. S. y Rodríguez-Gómez, G. (2018). Propuesta metodológica de evaluación para evaluar competencias a través de tareas complejas en entornos virtuales de aprendizaje. *Revista de Investigación Educativa*, 36(1). <https://doi.org/10.6018/rie.36.1.278301>
- Cutumisu, M. (2018). The informational value of feedback choices for performance and revision in a digital assessment game. *Interactive Technology and Smart Education*, 15(4), 363-380. <https://doi.org/10.1108/ITSE-03-2018-0017>
- Dadayan, L. y Ferro, E. (2005). When Technology Meets the Mind: A Comparative Study of the Technology Acceptance Model. *Electronic Government*, 137-144. https://doi.org/10.1007/11545156_13
- Dascalu, M.-D., Ruseti, S., Dascalu, M., McNamara, D. S., Carabas, M., Rebedea, T. y Trausan-Matu, S. (2021).

- Before and during COVID-19: A Cohesion Network Analysis of students' online participation in Moodle courses. *Computers in Human Behavior*, 121. <https://doi.org/10.1016/j.chb.2021.106780>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Davis, F. D. y Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: Three experiments. *International Journal of Human-Computer Studies*, 45(1), 19-45. <https://doi.org/10.1006/ijhc.1996.0040>
- De la Orden, A. (1982). *La evaluación educativa*. Docencia.
- DeMara, R., Chen, B., Hartshorne, R. y Thripp, R. (2017). Elevating participation and outcomes with computer-based assessments: An immersive development workshop for engineering faculty. *Computers in Education*, 8(3), 1-12.
- Dixon, D. D. y Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory Into Practice*, 55(2), 153-159. <https://doi.org/10.1080/00405841.2016.1148989>
- Dochy, F., Segers, M. y Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3), 331-350. <https://doi.org/10.1080/03075079912331379935>
- Dolin, J., Black, P., Harlen, W. y Tiberghien, A. (2018). Exploring Relations Between Formative and Summative Assessment. En J. Dolin and R. Evans (Eds.), *Transforming Assessment: Through an Interplay Between Practice, Research and Policy* (pp. 53-80). Springer International Publishing.
- Domingo-Coscollola, M., Bosco-Paniagua, A., Carrasco-Segovia, S. y Sánchez-Valero, J.-A. (2020). Fomentando la competencia digital docente en la universidad: Percepción de estudiantes y docentes. *Revista de Investigación Educativa*, 38(1). <https://doi.org/10.6018/rie.340551>
- Duart, J. M., Roig-Vila, S., Mengual-Andrés, S. y Maseda-Durán, M. Á. (2017). La calidad pedagógica de los MOOC a partir de la revisión sistemática de las publicaciones JCR y Scopus (2013-2015). *Revista Española de Pedagogía*. <https://doi.org/10.22550/REP75-1-2017-02>
- Eisner, E. W. (1981). On the Differences Between Scientific and Artistic Approaches to Qualitative Research. *Review of Research in Visual Arts Education*, 7(1), 1-8.
- Elkheir, Z. y Abdul-Mutalib, A. (2015). Mobile Learning Applications Designing Concepts and Challenges: Survey. *Research Journal of Applied Sciences, Engineering and Technology*, 10, 438-442. <https://doi.org/10.19026/rjaset.10.2509>
- Engard, N. C. (2009). LimeSurvey. *Public Services Quarterly*, 5(4), 272-273. <https://doi.org/10.1080/15228950903288728>
- Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: Strategies for technology integration. *Educational Technology Research and Development*, 47(4), 47-61. <https://doi.org/10.1007/BF02299597>
- Escobar-Pérez, J. y Cuervo, A. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 6, 27-36.
- Escudero, T. (2003). Desde los test hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 9(1). <https://doi.org/10.7203/relieve.9.1.4348>
- Eurboonyanun, C., Wittayapairoch, J., Aphinives, P., Petrusa, E., Gee, D. W. y Phitayakorn, R. (2021). Adaptation to Open-Book Online Examination During the COVID-19 Pandemic. *Journal of Surgical Education*, 78(3), 737-739. <https://doi.org/10.1016/j.jsurg.2020.08.046>

- Evans, C. y Robertson, W. (2020). The four phases of the digital natives debate. *Human Behavior and Emerging Technologies*, 2(3), 269-277. <https://doi.org/10.1002/hbe2.196>
- Fernández, V. (2005). Calidad y e-Learning. En F. Martínez (Coord.), *E-aprendizaje en bibliotecología. Perspectivas globales*. Universidad Nacional Autónoma de México.
- Fernández-Ramírez, B., Alonso-Morillejo, E., Reboloso-Pacheco, E. y Pozo-Muñoz, C. (1999). Evaluación de la calidad en la educación superior. *Papeles del psicólogo*, 74(2).
- Fetterman, D. M., Kaftarian, S. J. y Wandersman, A. (2015). *Empowerment Evaluation: Knowledge and Tools for Self-Assessment, Evaluation Capacity Building, and Accountability*. SAGE Publications.
- Fetterman, D. M. y Wandersman, A. (2004). *Empowerment Evaluation Principles in Practice*. Guilford Press.
- Fink, A. (1998). *Conducting literature research reviews: From paper to the Internet*. SAGE Publications.
- Fishbein, M. y Ajzen, I. (1975). *Belief, Attitude, Intention And Behavior: An introduction to theory and research*. Addison-Wesley.
- Fisher, C., Dwyer, D. C. y Yocam, K. (1996). *Education & Technology: Reflections on Computing in Classrooms*. Jossey-Bass Publishers.
- Fombona-Cadavieco, J. y Rodil-Pérez, F. J. (2018). Niveles de uso y aceptación de los dispositivos móviles en el aula. *Pixel-Bit. Revista de Medios y Educación*, 52. <https://doi.org/10.12795/pixelbit.2018.i52.02>
- Fonseca, J. G. (2007). Modelos cualitativos de evaluación. *Educere*, 11(38), 427-432.
- Fornell, C. y Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- Gabriel-Pedro, L. F. M., de Oliveira-Barbosa, C. M. M. y Neves-Santos, C. M. das. (2018). A critical review of mobile learning integration in formal educational contexts. *International Journal of Educational Technology in Higher Education*, 15(1). <https://doi.org/10.1186/s41239-018-0091-4>
- Galicia, L. A., Balderrama, J. A. y Edel, R. (2017). Validez de contenido por juicio de expertos: Propuesta de una herramienta virtual. *Apertura: Revista de Innovación Educativa*, 9(2), 42-53.
- Gamage, S. H. P. W., Ayres, J. R. y Behrend, M. B. (2022). A systematic review on trends in using Moodle for teaching and learning. *International Journal of STEM Education*, 9(1). <https://doi.org/10.1186/s40594-021-00323-x>
- García-Aretio, L. (2021). ¿Podemos fiarnos de la evaluación en los sistemas de educación a distancia y digitales? *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2). <https://doi.org/10.5944/ried.24.2.30223>
- García-Holgado, A. y García-Peñalvo, F. J. (2017). A Metamodel Proposal for Developing Learning Ecosystems. En P. Zaphiris y A. Ioannou (Eds.), *Learning and Collaboration Technologies. Novel Learning Ecosystems* (pp. 100-109). Springer International Publishing.
- García-Holgado, A., Marcos-Pablos, S., Therón-Sánchez, R. y García-Peñalvo, F. J. (2019). Technological Ecosystems in the Health Sector: A Mapping Study of European Research Projects. *Journal of Medical Systems*, 43(4). <https://doi.org/10.1007/S10916-019-1241-5>
- García-Llorente, H. J., Martínez-Abad, F. y Rodríguez-Conde, M. J. (2019). Validación de un instrumento de evaluación de competencias informacionales autopercibidas en educación secundaria obligatoria. *Anales de Documentación*, 22(1). <https://doi.org/10.6018/analesdoc.22.1.305641>
- García-Peñalvo, F. J. (2014). Formación en la sociedad del conocimiento, un programa de doctorado con una perspectiva interdisciplinar. *Education in the Knowledge Society (EKS)*, 15(1). <https://doi.org/10.14201/eks.11641>

- García-Peñalvo, F. J. (2017). *Revisión sistemática de literatura en los Trabajos de Final de Máster y en las Tesis Doctorales*. <https://knowledgesociety.usal.es/sites/default/files/20170316%20-%20Seminario%20SLR.pdf>
- García-Peñalvo, F. J. (2020). Modelo de referencia para la enseñanza no presencial en universidades presenciales. *Campus Virtuales*, 9(1), 41-56.
- García-Peñalvo, F. J. (2022). Desarrollo de estados de la cuestión robustos: Revisiones Sistemáticas de Literatura. *Education in the Knowledge Society (EKS)*, 23. <https://doi.org/10.14201/eks.28600>
- García-Peñalvo, F. J., Hernández-García, Á., Conde, M. Á., Fidalgo-Blanco, Á., Sein-Echaluce, M. L., Llorens Largo, F. e Iglesias-Pradas, S. (2015). *Mirando hacia el futuro: Ecosistemas tecnológicos de aprendizaje basados en servicios*. III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad (CINAIC 2015), Madrid. <http://rua.ua.es/dspace/handle/10045/51427>
- García-Peñalvo, F. J., Olmos-Migueláñez, S. y Sánchez-Prieto, J. C. (2015). Las tecnologías móviles en educación primaria: Estudio sobre la actitud de los futuros docentes. *Série-Estudos - Periódico do Programa de Pós-Graduação em Educação da UCDB*, 40, 35-55.
- García-Peñalvo, F. J., Rivero-Ortega, R., Rodríguez-Conde, M. J. y Rodríguez-García, N. (2020). The institutional decisions to support remote learning and teaching during the COVID-19 pandemic. *2020 X International Conference on Virtual Campus (JICV)*, 1-5. <https://doi.org/10.1109/JICV51605.2020.9375683>
- García-Peñalvo, F. J., Rodríguez-Conde, M. J., Therón, R., García-Holgado, A., Martínez-Abad, F. y Benito, A. (2019). Grupo GRIAL. *IE Comunicaciones: Revista Iberoamericana de Informática Educativa*, 30, 33-48.
- García-Peñalvo, F. J. y Seoane-Pardo, A. M. (2010). *Introducción al e-Learning*. Universidad de Salamanca. <https://repositorio.grial.eu/bitstream/123456789/27/6/introelearning.pdf>
- García-Peñalvo, F. J. y Seoane-Pardo, A. M. (2015). Una revisión actualizada del concepto de eLearning. Décimo Aniversario. *Education in the Knowledge Society (EKS)*, 16(1). <https://doi.org/10.14201/eks2015161119144>
- García-Soto, G. Y., García-López, R. I. y Lozano-Rodríguez, A. (2020). Calidad en la educación superior en línea: Un análisis teórico. *Revista de Educación*, 44(2).
- Gisbert, J. P. y Bonfill, X. (2004). ¿Cómo realizar, evaluar y utilizar revisiones sistemáticas y metaanálisis? *Gastroenterología y Hepatología*, 27(3), 129-149. [https://doi.org/10.1016/S0210-5705\(03\)79110-9](https://doi.org/10.1016/S0210-5705(03)79110-9)
- Gómez-Ramírez, I., Valencia-Arias, A. y Duque, L. (2019). Approach to M-learning Acceptance Among University Students: An Integrated Model of TPB and TAM. *International Review of Research in Open and Distributed Learning*, 20(3). <https://doi.org/10.19173/irrodl.v20i4.4061>
- González-Monteagudo, J. (2001). John Dewey y la pedagogía progresista. En J. Trilla (Coord.), *El legado pedagógico del siglo XX para la escuela del siglo XXI*. Grao.
- Goodhue, D. y Thompson, R. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, 19(2), 213-236. <https://doi.org/10.2307/249689>
- Grande-de-Prado, M., García-Peñalvo, F. J., Almuzara, A. C. y Abella-García, V. (2021). Evaluación en Educación Superior durante la pandemia de la COVID-19. *Campus Virtuales*, 10(1).
- Green, R. H. (2004). *Evaluación formativa: Algunas ideas prácticas*. I Jornadas de Innovación Universitaria, Villaviciosa de Odón, Madrid. <https://abacus.universidadeuropea.com/handle/11268/1485>
- Grigoryeva, N. V., Melikov, I. M., Palanchuk, N. V., Kokhanovskaya, I. I. y Aralova, E. (2021). Opportunities for Organizing Distance Learning Presented by the Moodle Platform: Experience in the Conditions of the

- COVID-19 Pandemic. *Propósitos y Representaciones*, 9. <https://doi.org/10.20511/pyr2021.v9nSPE3.1259>
- Gros-Salvat, B. (2018). La evolución del e-learning: Del aula virtual a la red. *Revista Iberoamericana de Educación a Distancia (RIED)*, 21(2), 69-82. <https://doi.org/10.5944/ried.21.2.20577>
- Guba, E. G. y Lincoln, Y. S. (1989). *Fourth generation evaluation*. SAGE Publications.
- Guerrero-Roldán, A.-E., Rodríguez, M., Karadeniz, A., Kocdar, S., Aleksieva, L. y Peytcheva-Forsyth, R. (2020). Students' Experiences on Using an Authentication and Authorship Checking System in E-Assessment. *Eğitim Fakültesi Dergisi*, 35, 6-24. <https://doi.org/10.16986/HUJE.2020063670>
- Guskey, T. R. (2009). Using assessments to improve teaching and learning. En B. Douglas (Ed.), *Ahead of the Curve: The Power of Assessment to Transform Teaching and Learning*. Solution Tree Press.
- Hair, J., Black, W., Babin, B. y Anderson, R. (2010). *Multivariate Data Analysis: A Global Perspective*. Prentice-Hall.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P. y Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. Springer International Publishing.
- Hair, J. F., Matthews, L. M., Matthews, R. L. y Sarstedt, M. (2017). PLS-SEM or CB-SEM: Updated guidelines on which method to use. *International Journal of Multivariate Data Analysis*, 1(2), 107-123. <https://doi.org/10.1504/IJMDA.2017.087624>
- Hair, J. F., Risher, J. J., Sarstedt, M. y Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2-24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hair, J. F., Sarstedt, M., Ringle, C. M. y Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, 40(3), 414-433. <https://doi.org/10.1007/s11747-011-0261-6>
- Hair, J., Sarstedt, M., Hopkins, L. y Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*, 26(2), 106-121. <https://doi.org/10.1108/EBR-10-2013-0128>
- Hannafin, R. D. y Savenye, W. C. (1993). Technology in the Classroom: The Teacher's New Role and Resistance to It. *Educational Technology*, 33(6), 26-31.
- Harchay, A., Berguiga, A., Cheniti-Belcadhi, L. y Braham, R. (2019). Student Perception of Mobile Self-assessment: An Evaluation of the Technology Acceptance Model. *Interaction Design and Architecture(s)*, 109-124. <https://doi.org/10.55612/s-5002-041-008>
- Harlen, W. y James, M. (1997). Assessment and Learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365-379. <https://doi.org/10.1080/0969594970040304>
- Harper, B. (2018). Technology and Teacher-Student Interactions: A Review of Empirical Research. *Journal of Research on Technology in Education*, 50(3), 214-225. <https://doi.org/10.1080/15391523.2018.1450690>
- Hasan, B. (2007). Examining the Effects of Computer Self-Efficacy and System Complexity on Technology Acceptance. *Information Resources Management Journal (IRMJ)*, 20(3), 76-88. <https://doi.org/10.4018/irmj.2007070106>
- Hattie, J. y Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Havnes, A., Smith, K., Dysthe, O. y Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*, 38(1), 21-27. <https://doi.org/10.1016/j.stueduc.2012.04.001>

- Heil, J. y Ifenthaler, D. (2023). Online Assessment in Higher Education: A Systematic Review. *Online Learning*, 27(1), Article 1. <https://doi.org/10.24059/olj.v27i1.3398>
- Hernández, P. (1998). *Diseñar y Enseñar*. Narcea.
- Hernández-García, Á. (2012). *Desarrollo de un modelo unificado de adopción del comercio electrónico entre empresas y consumidores finales. Aplicación al mercado español* [Tesis Doctoral, E.T.S.I. Telecomunicación (UPM)].
- Hernández-Nieto, R. A. (2002). *Contributions to Statistical Analysis*. Universidad de Los Andes.
- Hernández-Sellés, N., Muñoz-Carril, P. C. y Sanmamed, M. G. (2018). La e-evaluación en el trabajo colaborativo en entornos virtuales: Análisis de la percepción de los estudiantes. *EduTec. Revista Electrónica de Tecnología Educativa*, 65. <https://doi.org/10.21556/edutec.2018.65.997>
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. SAGE Publications.
- Hu, P. J.-H., Clark, T. H. K. y Ma, W. W. (2003). Examining technology acceptance by school teachers: A longitudinal study. *Information & Management*, 41(2), 227-241. [https://doi.org/10.1016/S0378-7206\(03\)00050-8](https://doi.org/10.1016/S0378-7206(03)00050-8)
- Huang, J., Lin, Y. y Chuang, S. (2007). Elucidating user behavior of mobile learning: A perspective of the extended technology acceptance model. *The Electronic Library*, 25(5), 585-598. <https://doi.org/10.1108/02640470710829569>
- Huda, S. S. M., Kabir, M. y Siddiq, T. (2020). E-Assessment in Higher Education: Students' Perspective. *International Journal of Education and Development Using Information and Communication Technology*, 16(2), 250-258.
- Ibarra-Sáiz, M. S. y Rodríguez-Gómez, G. (2019a). Una evaluación como aprendizaje. En J. Paricio, A. Fernández y I. Fernández (Eds.), *Cartografía de la buena docencia universitaria. Un marco para el desarrollo del profesorado basado en la investigación* (pp. 175-196). Narcea.
- Ibarra-Sáiz, M. S. y Rodríguez-Gómez, G. (2019b). *FLOASS - Resultados y analíticas de aprendizaje en la educación superior: Un marco de acción desde la evaluación sostenible*. XIX Congreso Internacional de Investigación Educativa. AIDIPE, Madrid. <https://doi.org/10.13140/RG.2.2.23303.65446>
- Ibarra-Sáiz, M. S. y Rodríguez-Gómez, G. (2020). Aprendiendo a Evaluar para Aprender en la Educación Superior. *Revista Iberoamericana de Evaluación Educativa*, 13(1).
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., Boud, D., Rotsaert, T., Brown, S., Salinas-Salazar, M. L. y Rodríguez-Gómez, H. M. (2020). El futuro de la evaluación en la Educación Superior. *Revista electrónica de investigación y evaluación educativa*. <https://doi.org/10.7203/relieve.26.1.17323>
- Ifanov, Jessica, P., Salim, S., Syahputra, M. E. y Suri, P. A. (2023). A Systematic literature review on implementation of virtual reality for learning. *Procedia Computer Science*, 216, 260-265. <https://doi.org/10.1016/j.procs.2022.12.135>
- James, M. (2006). Assessment, Teaching and Theories of Learning. En J. Gardner (Ed.). *En Assessment and Learning* (pp. 47-60). Sage Publications. <https://doi.org/10.13140/2.1.5090.8960>
- Jia, K., Wang, P., Li, Y., Chen, Z., Jiang, X., Lin, C.-L. y Chin, T. (2022). Research Landscape of Artificial Intelligence and e-Learning: A Bibliometric Research. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.795039>
- Jiménez-Martínez, K. A., Zamudio-Rodríguez, B. R. y Martínez-Moreno, M. (2022). Evaluación de la aceptación de las herramientas digitales para la enseñanza a través del modelo de aceptación tecnológica. *Revista*

Internacional de Desarrollo Regional Sustentable, 6(1-2).

- Kaisara, G. y Bwalya, K. J. (2023). Strategies for Enhancing Assessment Information Integrity in Mobile Learning. *Informatics*, 10(1). <https://doi.org/10.3390/informatics10010029>
- Karahanna, E., Agarwal, R. y Angst, C. M. (2006). Reconceptualizing compatability beliefs in technology acceptance research. *MIS Quarterly*, 30(4), 781-804.
- Kellaghan, T. y Greaney, V. (2001). *Using Assessment To Improve the Quality of Education. Fundamentals of Educational Planning*. United Nations Educational, Scientific, and Cultural Organization, International Institute for Educational Planning.
- Kemmis, S. (1982). Seven Principles for Programme Evaluation in Curriculum Development and Innovation. *Journal of Curriculum Studies*, 14(3), 221-240. <https://doi.org/10.1080/0022027820140302>
- Kibble, J. D. (2017). Best practices in summative assessment. *Advances in Physiology Education*, 41(1), 110-119. <https://doi.org/10.1152/advan.00116.2016>
- King, W. R. y He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, 43(6), 740-755. <https://doi.org/10.1016/j.im.2006.05.003>
- Kitchenham, B. y Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering* (EBSE Technical Report Version 2.3; pp. 1-57). Keele University.
- Knight, P. T. (2002). Summative Assessment in Higher Education: Practices in disarray. *Studies in Higher Education*, 27(3), 275-286. <https://doi.org/10.1080/03075070220000662>
- Koslowski, F. A. (2006). Quality and assessment in context: A brief review. *Quality Assurance in Education*, 14(3), 277-288. <https://doi.org/10.1108/09684880610678586>
- Kühbeck, F., Berberat, P. O., Engelhardt, S. y Sarikas, A. (2019). Correlation of online assessment parameters with summative exam performance in undergraduate medical education of pharmacology: A prospective cohort study. *BMC Medical Education*, 19(1). <https://doi.org/10.1186/s12909-019-1814-5>
- Kuklinski, H. P. y Cobo, C. (2021). Expandir la universidad más allá de la enseñanza remota de emergencia. Ideas hacia un modelo híbrido post-pandemia. *Outliers School*, 30(3), 246-248.
- Kumar, K., Raghuwaiya, K., Sharma, B. y Dakuidreketi, M. (2020). *Factors that influence academics' intention to use mobile-based assessment in Higher Education in South Pacific*. Frontiers in Education (FIE) Conference, 1-8. <https://doi.org/10.1109/FIE44824.2020.9274091>
- Kurilovas, E. y Vinogradova, I. (2016). Improved fuzzy AHP methodology for evaluating quality of distance learning courses. *The International Journal of Engineering Education*, 32(4), 1618-1624.
- La Rotta, D., Usuga, O. C. y Clavijo, V. (2020). Perceived service quality factors in online higher education. *Learning Environments Research*, 23(2), 251-267. <https://doi.org/10.1007/s10984-019-09299-6>
- Lai, Y. (2020). The social network analysis on the behavioral intention to use cloud sphygmomanometer. *Health and Technology*, 10, 787-794. <https://doi.org/10.1007/s12553-019-00339-y>
- Lau, A. M. S. (2016). 'Formative good, summative bad?' – A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*, 40(4), 509-525. <https://doi.org/10.1080/0309877X.2014.984600>
- Lee, K. y Fanguy, M. (2022). Online exam proctoring technologies: Educational innovation or deterioration? *British Journal of Educational Technology*, 53(3). <https://doi.org/10.1111/bjet.13182>
- Legris, P., Ingham, J. y Collette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management*, 40(3), 191-204.

[https://doi.org/10.1016/S0378-7206\(01\)00143-4](https://doi.org/10.1016/S0378-7206(01)00143-4)

- Li, G., Sun, Z. y Jee, Y. (2019). The more technology the better? A comparison of teacher-student interaction in high and low technology use elementary EFL classrooms in China. *System*, 84, 24-40. <https://doi.org/10.1016/j.system.2019.05.003>
- Liu, Y., Li, H. y Carlsson, C. (2010). Factors driving the adoption of m-learning: An empirical study. *Computers & Education*, 55(3), 1211-1219. <https://doi.org/10.1016/j.compedu.2010.05.018>
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D. y Wagenmakers, E.J. (2019). JASP: Graphical Statistical Software for Common Statistical Designs. *Journal of Statistical Software*, 88, 1-17. <https://doi.org/10.18637/jss.v088.i02>
- MacLellan, E. (2001). Assessment for Learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education*, 26(4), 307-318. <https://doi.org/10.1080/02602930120063466>
- Mailizar, M., Almanthari, A. y Maulina, S. (2021). Examining Teachers' Behavioral Intention to Use E-learning in Teaching of Mathematics: An Extended TAM Model. *Contemporary Educational Technology*, 13(2), ep298. <https://doi.org/10.30935/cedtech/9709>
- Manterola, C., Astudillo, P., Arias, E. y Claros, N. (2013). Revisión sistemática de la literatura. Qué se debe saber acerca de ellas. *Cirugía Española*, 91(3), 149-155. <https://doi.org/10.1016/j.ciresp.2011.07.009>
- Máñez-Sáez, I., Vidal-Abarca, E. y Martínez-Giménez, T. (2019). Does Computer-based Elaborated Feedback Influence the Students' Question-Answering Process? *Electronic Journal of Research in Education Psychology*, 17(47). <https://doi.org/10.25115/ejrep.v17i47.2156>
- Marchesi, A. y Martín, E. (1998). *Calidad de la enseñanza en tiempos de cambio*. Alianza.
- Marciniak, R. y Gairín-Sallán, J. (2018). Dimensiones de evaluación de calidad de educación virtual: Revisión de modelos referentes. *RIED. Revista Iberoamericana de Educación a Distancia*, 21(1). <https://doi.org/10.5944/ried.21.1.16182>
- Mathieson, K. (1991). Predicting User Intentions: Comparing the Technology Acceptance Model with the Theory of Planned Behavior. *Information Systems Research*, 2(3), 173-191. <https://doi.org/10.1287/isre.2.3.173>
- McAuley, E., Duncan, T. y Tammen, V. V. (1989). Psychometric Properties of the Intrinsic Motivation Inventory in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport*, 60(1), 48-58. <https://doi.org/10.1080/02701367.1989.10607413>
- McDonald, B. y Boud, D. (2003). The Impact of Self-assessment on Achievement: The effects of self-assessment training on performance in external examinations. *Assessment in Education: Principles, Policy & Practice*, 10(2), 209-220. <https://doi.org/10.1080/0969594032000121289>
- Mejía, J. y López, D. (2016). Modelo de Calidad de E-learning para Instituciones de Educación Superior en Colombia. *Formación Universitaria*, 9(2), 59-72. <https://doi.org/10.4067/S0718-50062016000200007>
- Mejía-Pérez, O. (2012). De la evaluación tradicional a una nueva evaluación basada en competencias. *Educare*, 16(1). <https://doi.org/10.15359/ree.16-1.3>
- Miras, M. y Solé, I. (1990). La evaluación del aprendizaje y la evaluación del proceso de enseñanza -aprendizaje. En C. Coll, J. Palacios, y A. Marchesi (Eds.), *Desarrollo psicológico y educación, II Psicología de la Educación* (pp. 419-434). Alianza.
- Moore, G. C. y Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, 2(3), 192-222.

- Morris, R., Perry, T. y Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3). <https://doi.org/10.1002/rev3.3292>
- Moyano, G. C. M. y Spinelli, M. del R. (2020). El uso de dispositivos móviles como estrategia educativa en la universidad. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(1), 201-222.
- Municio, P. (2004). Evaluación de la calidad. En R. Pérez-Juste, F. López, M.D. Peralta y P. Municio (Eds.), *Hacia una educación de calidad. Gestión, instrumentos y evaluación* (pp. 101-156). Narcea.
- Murray, H. G. (1984). The Impact of Formative and Summative Evaluation of Teaching in North American Universities. *Assessment & Evaluation in Higher Education*, 9(2), 117-132. <https://doi.org/10.1080/0260293840090204>
- Nikou, S. (2021). Web-based videoconferencing for teaching online: Continuance intention to use in the post-COVID-19 period. *Interaction Design and Architecture(s)*, 47, 123-143. <https://doi.org/10.55612/s-5002-047-006>
- Nikou, S. y Economides, A. (2015). The impact of paper-based, computer-based and mobile-based self-assessment on students' science motivation and achievement. *Computers in Human Behavior*, 55(Part B). <https://doi.org/10.1016/j.chb.2015.09.025>
- Nikou, S. y Economides, A. (2016). An Outdoor Mobile-Based Assessment Activity: Measuring Students' Motivation and Acceptance. *International Journal of Interactive Mobile Technologies (IJIM)*, 10, 11-17. <https://doi.org/10.3991/ijim.v10i4.5541>
- Nikou, S. y Economides, A. (2017a). Mobile-Based Assessment: Integrating acceptance and motivational factors into a combined model of Self-Determination Theory and Technology Acceptance. *Computers in Human Behavior*, 68, 83-95. <https://doi.org/10.1016/j.chb.2016.11.020>
- Nikou, S. y Economides, A. (2017b). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 109, 56-73. <https://doi.org/10.1016/j.compedu.2017.02.005>
- Nikou, S. y Economides, A. (2018). Mobile-Based micro-Learning and Assessment: Impact on learning performance and motivation of high school students. *Journal of Computer Assisted Learning*, 34(3), 269-278. <https://doi.org/10.1111/jcal.12240>
- Nikou, S. y Economides, A. (2019). Factors that influence behavioral intention to use mobile-based assessment: A STEM teachers' perspective. *British Journal of Educational Technology*, 50(2). <https://doi.org/10.1111/bjet.12609>
- Nikou, S. y Economides, A. (2014). A model for Mobile-based Assessment adoption based on Self-Determination Theory of Motivation. *Proceedings of 2014 International Conference on Interactive Mobile Communication Technologies and Learning, IMCL 2014*. International Conference on Interactive Mobile Communication, Thessaloniki. <https://doi.org/10.1109/IMCTL.2014.7011111>
- Noah, N. y Das, S. (2021). Exploring evolution of augmented and virtual reality education space in 2020 through systematic literature review. *Computer Animation and Virtual Worlds*, 32(3-4). <https://doi.org/10.1002/cav.2020>
- Olmos-Migueláñez, S. (2008). *Evaluación formativa y sumativa de estudiantes universitarios: Aplicación de las tecnologías a la evaluación educativa* [Tesis Doctoral, Universidad de Salamanca].
- Ong, C.-S., Lai, J.-Y. y Wang, Y.-S. (2004). Factors affecting engineers' acceptance of asynchronous e-learning systems in high-tech companies. *Information & Management*, 41(6), 795-804. <https://doi.org/10.1016/j.im.2003.08.012>
- Ortiz-López, A., Olmos-Migueláñez, S. y Sánchez-Prieto, J. C. (2021). Calidad en e-Learning: Identificación de

- sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2). <https://doi.org/10.5944/ried.24.2.29073>
- Ortiz-López, A., Olmos-Migueláñez, S. y Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-Learning en Educación Superior: Una revisión sistemática de la literatura. *Education in the Knowledge Society (EKS)*, 23. <https://doi.org/10.14201/eks.26986>
- Ortiz-López, A., Olmos-Migueláñez, S. y Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12117-5>
- Ortiz-López, A.; Olmos-Migueláñez, S. y Sánchez-Prieto, J.C. (En prensa). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En E. Chiner (Ed.), *Nuevas tendencias interdisciplinarias en Educación y Conocimiento*. Tirant Lo Blanch
- Ortiz-López, A.; Sánchez-Prieto, J.C. y Olmos-Migueláñez, S. (En prensa). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. *Journal of New Approaches in Educational Research*.
- Ozturk, E. (2019). Students' expectations from mobile devices for mobile learning. *International Journal of Mobile Communications*, 17(4), 409-421. <https://doi.org/10.1504/IJMC.2019.100496>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(71). <https://doi.org/10.1136/bmj.n71>
- Park, S. Y. (2009). An Analysis of the Technology Acceptance Model in Understanding University Students' Behavioral Intention to Use e-Learning. *Educational Technology & Society*, 12(3), 150-162.
- Park, Y. y Lim, K. (2015). Effects of Environmental and Human Constructs on e-learning Effectiveness in Online University Settings. *Indian Journal of Science and Technology*, 8(1), 1-9. <https://doi.org/10.17485/ijst/2015/v8iS1/57729>
- Parlakiliç, A. (2020). Transition From E-Learning to U-Learning: Basic Characteristics, Media, and Researches. En D. Gürhan Durak and C. Aerkan (Eds.), *Managing and Designing Online Courses in Ubiquitous Learning Environments* (pp. 296-310). IGI Global. <https://doi.org/10.4018/978-1-5225-9779-7.ch016>
- Parlett, M. y Hamilton, D. (1972). «*Evaluation as Illumination: A New Approach to the Study of Innovative Programs*». Occasional Paper. <https://eric.ed.gov/?id=ED167634>
- Patton, M. Q. (1996). A World Larger than Formative and Summative. *Evaluation Practice*, 17(2), 131-144. <https://doi.org/10.1177/109821409601700205>
- Paul, J. y Barari, M. (2022). Meta-analysis and traditional systematic literature reviews—What, why, when, where, and how? *Psychology & Marketing*, 39(6), 1099-1115. <https://doi.org/10.1002/mar.21657>
- Pearse-Romera, C. R. y Ruiz-Cecilia, R. (2019). A Motivation Case Study of Students Learning English at a Secondary School in Granada, Spain. *The International Journal of Interdisciplinary Educational Studies*, 14(1), 31-45. <https://doi.org/10.18848/2327-011X/CGP/v14i01/31-45>
- Pedrosa, I., Suárez-Álvarez, J. y García-Cueto, E. (2013). Evidencias sobre la validez de contenido: Avances teóricos y métodos para su estimación. *Acción Psicológica*, 10(2), 3-18. <https://doi.org/10.5944/ap.10.2.11820>
- Pérez-Gómez, Á. I. (1989). Modelos contemporáneos de evaluación. En J. Gimeno-Sacristán (Ed.), *La enseñanza: Su teoría y su práctica* (pp. 426-449). Akal.

- Pérez-Juste, R. (2000). La evaluación de programas educativos: Conceptos básicos, planteamientos generales y problemática. *Revista de Investigación Educativa*, 18(2).
- Pérez-Juste, R. y García-Ramos, J. M. (1989). *Diagnóstico, evaluación y toma de decisiones*. Rialp.
- Pérez-Sanz, A. (2011). Escuela 2.0: Educación para el mundo digital. *Revista de Estudios de Juventud*, 92, 63-86.
- Perrenoud, P. (1991). *La evaluación de los alumnos: De la producción de la excelencia a la regulación de los aprendizajes*. Alternativa Pedagógica.
- Posner, G. (1998). *Análisis del curriculum*. McGraw Hill.
- Pratama, A. R. (2021). Fun first, useful later: Mobile learning acceptance among secondary school students in Indonesia. *Education and Information Technologies*, 26(2), 1737-1753. <https://doi.org/10.1007/s10639-020-10334-w>
- Pu, S. y Xu, H. (2021). Examining Changing Assessment Practices in Online Teaching: A Multiple-Case Study of EFL School Teachers in China. *The Asia-Pacific Education Researcher*, 30(6), 553-561. <https://doi.org/10.1007/s40299-021-00605-6>
- Ramayah, T., Hwa, C., Chuah, F., Ting, H. y Memon, M. (2017). PLS-SEM using SmartPLS 3.0: Chapter 8: Assessment of Formative Measurement Models. En *Partial least squares structural equation modeling (PLS-SEM) using smartPLS 3.0: An Updated and Practical Guide to Statistical Analysis*. Pearson.
- Rangel-Baca, A. (2018). Tecnologías de la Información y la Comunicación en la Educación Superior: Una revisión documental. *Pixel-Bit: Revista de Medios y Educación*, 52, 125-137. <https://doi.org/10.12795/pixelbit.2018.i52.09>
- Raza, S., Khan, K. y Salam, J. (2021). Impact of environmental triggers on students' behavior to use ride-sharing services: The moderating role of perceived risk. *Current Psychology*, 42, 11329-11343. <https://doi.org/10.1007/s12144-021-02405-z>
- Recio-Muñoz, F., Silva-Quiroz, J. E. y Abricot-Marchant, N. (2020). Análisis de la competencia digital en la formación inicial de estudiantes universitarios: Un estudio de meta-análisis en la Web of Science. *Pixel-Bit*, 59. <https://doi.org/10.12795/pixelbit.77759>
- Reeve, J. y Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist*, 56(1), 54-77. <https://doi.org/10.1080/00461520.2020.1862657>
- Rivlin, A. M. (1971). *Systematic Thinking for Social Action*. Brooking Institution.
- Roberts, T. S. (2006). The use of multiple choice tests for formative and summative assessment. *Proceedings of the 8th Australasian Conference on Computing Education - Volume 52*, 175-180.
- Roca, J. C. y Gagné, M. (2008). Understanding e-learning continuance intention in the workplace: A self-determination theory perspective. *Computers in Human Behavior*, 24(4), 1585-1604. <https://doi.org/10.1016/j.chb.2007.06.001>
- Rodríguez-Conde, M. J. (2005). Aplicación de las TIC a la evaluación de alumnos universitarios. *Education in the Knowledge Society (EKS)*, 6(2). <https://doi.org/10.14201/eks.18185>
- Rodríguez-Espinar, S. (2001). La calidad en la Enseñanza Universitaria. *Ágora Digital*, 2.
- Rodríguez-Gómez, G., Ibarra-Saiz, M. S. y García-Jiménez, E. (2013). Autoevaluación, evaluación entre iguales y coevaluación: Conceptualización y práctica en las universidades españolas. *Revista de Investigación en Educación*, 2(11), 198-210.
- Rodríguez-Neira, T., Álvarez-Pérez, L., Cadrecha y Caparrós, M. Á., Hernández-García, J., Luengo-García, M. A.,

- Muñiz-Fernández, J., Ordóñez-Alvarez, J. J. y Soler-Vázquez, E. (2001). *La evaluación en el aula*. Ediciones Nobel.
- Rogers, E. M. (1962). *Diffusion of innovations*. Free Press of Glencoe.
- Rossi, P. H., Freeman, H. E. y Wright, S. R. (1979). *Evaluation. A systematic approach*. SAGE Publications.
- Ruíz-Morales, Y. A., García-García, M. y Biencinto-López, C. (2018). Evaluación de competencias genéricas en la universidad. Estudio comparativo en entorno b-learning y presencial. *Acción Pedagógica*, 27(1), 6-21.
- Sadeghi, K., Saribagloo, J. A., Aghdam, S. H. y Mahmoudi, H. (2014). The Impact of Iranian Teachers Cultural Values on Computer Technology Acceptance. *Turkish Online Journal of Educational Technology*, 13(4), 124-136.
- Sáenz, A. (2001). Leer e interpretar una revisión sistemática. *Boletín de la Sociedad de Pediatría de Asturias, Cantabria, Castilla y León*, 41(177), 215-221.
- Sáez-López, J. M. (2015). Actitudes de los docentes respecto a las tic, a partir del desarrollo de una práctica reflexiva. *Escuela Abierta*, 13, 37-54.
- Samala, A. D., Usmeldi, Taali, Ambiyar, Bojic, L., Indarta, Y., Tsoy, D., Denden, M., Tas, N. y Dewi, I. P. (2023). Metaverse Technologies in Education: A Systematic Literature Review Using PRISMA. *International Journal of Emerging Technologies in Learning (IJET)*, 18(5). <https://doi.org/10.3991/ijet.v18i05.35501>
- Sánchez-Prieto, J. C., Hernández-García, Á., García-Peñalvo, F. J., Chaparro-Peláez, J. y Olmos-Migueláñez, S. (2019). Break the Walls! Second-Order Barriers and the Acceptance of mLearning by First-Year Pre-Service Teachers. *Computers in Human Behavior*, 95. <https://doi.org/10.1016/j.chb.2019.01.019>
- Sánchez-Prieto, J. C., Olmos-Migueláñez, S. y García-Peñalvo, F. J. (2016). Do Mobile Technologies Have a Place in Universities?: The TAM Model in Higher Education. En *En L. Briz-Ponce, J.A. Juanes y F.J. García-Peñalvo (Eds.), Handbook of Research on Mobile Devices and Applications in Higher Education Settings* (pp. 25-52). IGI Global.
- Sánchez-Prieto, J. C., Olmos-Migueláñez, S. y García-Peñalvo, F. J. (2017a). MLearning and pre-service teachers: An assessment of the behavioral intention using an expanded TAM model. *Computers in Human Behavior*, 72, 644-654. <https://doi.org/10.1016/j.chb.2016.09.061>
- Sánchez-Prieto, J. C., Olmos-Migueláñez, S. y García-Peñalvo, F. J. (2017b). ¿Utilizarán los futuros docentes las tecnologías móviles? Validación de una propuesta de modelo TAM extendido. *Revista de Educación a Distancia (RED)*, 17(52).
- Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O. y Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, 69(10), 3998-4010. <https://doi.org/10.1016/j.jbusres.2016.06.007>
- Schepers, J. y Wetzels, M. (2007). A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. *Information & Management*, 44(1), 90-103. <https://doi.org/10.1016/j.im.2006.10.007>
- Scherer, R., Siddiq, F. y Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Scriven, M. (1967). The Methodology of Evaluation. En R. W. Tyler, R. M. Gagne, y M. Scriven (Eds.), *Perspectives of Curriculum Evaluation* (pp. 39-83). Rand McNally.
- Scriven, M. (1991a). Beyond Formative and Summative Evaluation. *Teachers College Record*, 92(6), 19-64. <https://doi.org/10.1177/016146819109200603>

- Scriven, M. (1991b). Prose and cons about goal-free evaluation. *Evaluation Practice*, 12(1), 55-63. [https://doi.org/10.1016/0886-1633\(91\)90024-R](https://doi.org/10.1016/0886-1633(91)90024-R)
- Seo, H.-J. y Kim, K. U. (2012). Quality assessment of systematic reviews or meta-analyses of nursing interventions conducted by Korean reviewers. *BMC Medical Research Methodology*, 12(1), 129. <https://doi.org/10.1186/1471-2288-12-129>
- Serrano, E. L., Ceballos, S. P., Arroyo, G. C. y Cisneros-Cohernour, E. (2018). Marco para evaluar las condiciones institucionales de la enseñanza en línea. *Revista Electrónica de Investigación Educativa*, 20(2). <https://doi.org/10.24320/redie.2018.20.2.2072>
- Serrano, M. S. (1990). *El Proceso de Enseñanza Aprendizaje*. Universidad de Los Andes Consejo de Estudios de Postgrado. Consejo Editorial.
- Sharma, P. y Kim, K. (2012). *Model Selection in Information Systems Research Using Partial Least Squares Based Structural Equation Modeling*. International Conference on Interaction Sciences.
- Sharma, P., Liengard, B., Hair, J., Sarstedt, M. y Ringle, C. (2022). Predictive model assessment and selection in composite-based modeling using PLS-SEM: Extensions and guidelines for using CVPAT. *European Journal of Marketing*. <https://doi.org/10.1108/EJM-08-2020-0636>
- Shepard, L. (2006). La evaluación en el aula. En R. Brennan (Ed.), *Educational Measurement* (pp. 623-646). Praeger Westport.
- Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing*, 18(4), 373-391. <https://doi.org/10.1177/026553220101800404>
- Simonetto, A. (2012). Formative and reflective models: State of the art. *Electronic Journal of Applied Statistical Analysis*, 5(3). <https://doi.org/10.1285/i20705948v5n3p452>
- Sirianni, J. M., Ng, Y. J. y Vishwanath, A. (2017). Adopting Computer-Based Assessments: The Role of Perceived Value in Classroom Technology Acceptance. *Online Journal of Communication and Media Technologies*, 7(4), 1-23. <https://doi.org/10.29333/ojcm/2607>
- Slade, C., Lawrie, G., Taptamat, N., Browne, E., Sheppard, K. y Matthews, K. E. (2022). Insights into how academics reframed their assessment during a pandemic: Disciplinary variation and assessment as afterthought. *Assessment & Evaluation in Higher Education*, 47(4), 588-605. <https://doi.org/10.1080/02602938.2021.1933379>
- Souabi, S., Retbi, A., Idrissi, M. K. y Bennani, S. (2021). Towards an Evolution of E-Learning Recommendation Systems: From 2000 to Nowadays. *International Journal of Emerging Technologies in Learning (IJET)*, 16(6). <https://doi.org/10.3991/ijet.v16i06.18159>
- Stake, R. E. (1967). The Countenance of Educational Evaluation. *Teachers College Record*, 68(7), 1-15. <https://doi.org/10.1177/016146816706800707>
- Stake, R. E. (1982). La evaluación de programas; en especial la evaluación de réplica. En W.B. Dockrell y D. Hamilton (Eds.), *Nuevas reflexiones sobre la investigación educativa* (pp. 91-108). Narcea.
- Stenhouse, L. (1975). *An introduction to curriculum research an development*. Heinemann.
- Stufflebeam, D. L. (1987). Professional standars for assuring the quality of educational program and personal evaluations. *International Journal of Educational Research*, 11(1).
- Stufflebeam, D. L. (2000). The CIPP Model for Evaluation. En D. L. Stufflebeam, G. F. Madaus, y T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (pp. 279-317). Springer Netherlands.

- Stufflebeam, D. L. y Coryn, C. L. S. (2014). *Evaluation Theory, Models, and Applications*. John Wiley & Sons.
- Stufflebeam, D. L. y Shinkfield, A. J. (1987). *Evaluación sistemática: Guía teórica y práctica*. Paidós Ibérica.
- Su, Y.-L. y Reeve, J. (2011). A Meta-analysis of the Effectiveness of Intervention Programs Designed to Support Autonomy. *Educational Psychology Review*, 23(1), 159-188. <https://doi.org/10.1007/s10648-010-9142-7>
- Sukendro, S., Habibi, A., Khaeruddin, K., Indrayana, B., Syahrudin, S., Makadada, F. A. y Hakim, H. (2020). Using an extended Technology Acceptance Model to understand students' use of e-learning during Covid-19: Indonesian sport science education context. *Heliyon*, 6(11). <https://doi.org/10.1016/j.heliyon.2020.e05410>
- Sullivan, K. y Hall, C. (1997). Introducing Students to Self-assessment. *Assessment & Evaluation in Higher Education*, 22(3), 289-305. <https://doi.org/10.1080/0260293970220303>
- Taherdoost, H. (2019). Importance of Technology Acceptance Assessment for Successful Implementation and Development of New Technologies. *Global Journal of Engineering Sciences*, 1. <https://doi.org/10.33552/GJES.2019.01.000511>
- Taras, M. (2010). Student self-assessment: Processes and consequences. *Teaching in Higher Education*, 15(2), 199-209. <https://doi.org/10.1080/13562511003620027>
- Tarhini, A., Hone, K., Liu, X. y Tarhini, T. (2017). Examining the moderating effect of individual-level cultural values on users' acceptance of E-learning in developing countries: A structural equation modeling of an extended technology acceptance model. *Interactive Learning Environments*, 25(3), 306-328. <https://doi.org/10.1080/10494820.2015.1122635>
- Taylor, S. y Todd, P. (1995). Understanding information technology usage: A test of competing models. *Information Systems Research*, 6(2), 144-176.
- Tejada, J. (1997). La evaluación. En J. Gairin y A. Ferrández (Eds.), *Planificación y gestión de instituciones de formación* (pp. 243-268). Praxis.
- Teo, T. (2015). Comparing pre-service and in-service teachers' acceptance of technology: Assessment of measurement invariance and latent mean differences. *Computers & Education*, 83. <https://doi.org/10.1016/j.compedu.2014.11.015>
- Tomas, C., Borg, M. y McNeil, J. (2015). E-assessment: Institutional development strategies and the assessment life cycle. *British Journal of Educational Technology*, 46(3), 588-596. <https://doi.org/10.1111/bjet.12153>
- Tonbuloglu, B. (2023). State of research on e-assessment in education: A bibliometric analysis. *Turkish Online Journal of Distance Education*, 24(4). <https://doi.org/10.17718/tojde.1184218>
- Topuz, A. C., Saka, E., Fatsa, Ö. F. y Kurşun, E. (2022). Emerging trends of online assessment systems in the emergency remote teaching period. *Smart Learning Environments*, 9(1). <https://doi.org/10.1186/s40561-022-00199-6>
- Tsai, T.-H., Lin, W.-Y., Chang, Y.-S., Chang, P.-C. y Lee, M.-Y. (2020). Technology anxiety and resistance to change behavioral study of a wearable cardiac warming system using an extended TAM for older adults. *PLOS ONE*, 15(1). <https://doi.org/10.1371/journal.pone.0227270>
- Tyler, R. (1949). *Basic principle of curriculum and instruction*. Chicago University.
- Van den Broeck, A., Vansteenkiste, M., De Witte, H., Soenens, B. y Lens, W. (2010). Capturing autonomy, competence, and relatedness at work: Construction and initial validation of the Work-related Basic Need Satisfaction scale. *Journal of Occupational and Organizational Psychology*, 83(4), 981-1002. <https://doi.org/10.1348/096317909X481382>
- Venkatesh, V. (2000). Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and

- Emotion into the Technology Acceptance Model. *Information Systems Research*, 11(4), 342-365. <https://doi.org/10.1287/isre.11.4.342.11872>
- Venkatesh, V. y Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39(2), 273-315. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>
- Venkatesh, V. y Davis, F. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46(2), 186-204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, V. y Davis, F. D. (1996). A Model of the Antecedents of Perceived Ease of Use: Development and Test. *Decision Sciences*, 27(3), 451-481. <https://doi.org/10.1111/j.1540-5915.1996.tb00860.x>
- Venkatesh, V., Morris, M. G., Davis, G. B. y Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Venkatesh, V., Thong, J. y Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, 36, 157-178. <https://doi.org/10.2307/41410412>
- Vergés, A. (2018). From piloting e-submission to electronic management of assessment (EMA): Mapping grading journeys. *British Journal of Educational Technology*, 49(3), 463-478. <https://doi.org/10.1111/bjet.12547>
- Vidal, M. del P. (2006). Investigación de las TIC en la educación. *Revista Latinoamericana de Tecnología Educativa*, 5(2).
- Villamar-Coronel, J. del R. (2022). Heteroevaluación vs coevaluación y su incidencia en el plan de mejora para alcanzar la calidad educativa. *Ibero-American Journal of Education & Society Research*, 2(2). <https://doi.org/10.56183/iberoeds.v2i2.581>
- Villavicencio, A. M. V. y Paredes, E. V. P. (2020). Relación del uso del teléfono celular y los niveles de atención en el proceso de enseñanza – aprendizaje. *Revista Encuentros*, 18(1), 11-22.
- Warf, B. (2019). Teaching Digital Divides. *Journal of Geography*, 118(2), 77-87. <https://doi.org/10.1080/00221341.2018.1518990>
- Wentworth, D. K. y Middleton, J. H. (2014). Technology use and academic performance. *Computers & Education*, 78, 306-311. <https://doi.org/10.1016/j.compedu.2014.06.012>
- Williams, M. D., Rana, N. P. y Dwivedi, Y. K. (2015). The unified theory of acceptance and use of technology (UTAUT): A literature review. *Journal of Enterprise Information Management*, 28(3), 443-488. <https://doi.org/10.1108/JEIM-09-2014-0088>
- Williams-Muller, E., Castro-Galviz, M. Y., Cruz-López, M. D. la, Pinillas, L. N., García, L. R. y Carvalho, J. L. (2020). Uso de los teléfonos móviles en el aula de educación primaria. *Revista Electrónica de Conocimientos, Saberes y Prácticas*, 3(2). <https://doi.org/10.5377/recsp.v3i2.10688>
- Williamson, B., Eynon, R. y Potter, J. (2020). Pandemic politics, pedagogies and practices: Digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, 45(2), 107-114. <https://doi.org/10.1080/17439884.2020.1761641>
- Wu, P. y Wang, Y. (2021). Investigating Business English Teachers' Belief About Online Assessment: Q Methodology Conducted During COVID-19 Period. *The Asia-Pacific Education Researcher*, 30(6), 621-630. <https://doi.org/10.1007/s40299-021-00604-7>
- Yangari, M. e Inga, E. (2021). Educational Innovation in the Evaluation Processes within the Flipped and Blended Learning Models. *Education Sciences*, 11(9). <https://doi.org/10.3390/educsci11090487>

- Zaineldeen, S., Li, H., Koffi, A. y Mohammed, B. (2020). Technology Acceptance Model' Concepts, Contribution, Limitation, and Adoption in Education. *Universal Journal of Educational Research*, 8(11), 5061-5071. <https://doi.org/10.13189/ujer.2020.081106>
- Zamora-Delgado, R. (2019). El M-Learning, las ventajas de la utilización de dispositivos móviles en el proceso autónomo de aprendizaje. *ReHuSo: Revista de Ciencias Humanísticas y Sociales*, 4(3), 29-38.
- Zhang, C., Yan, X. y Wang, J. (2021). EFL Teachers' Online Assessment Practices During the COVID-19 Pandemic: Changes and Mediating Factors. *The Asia-Pacific Education Researcher*, 30(6), 499-507. <https://doi.org/10.1007/s40299-021-00589-3>

Anexos de la Tesis Doctoral

ÍNDICE DE ANEXOS:

1. **ANEXOS 1-7: Publicaciones científicas derivadas de la Tesis Doctoral**
 - **Anexo 1:** Artículo publicado en Revista Iberoamericana de Educación a Distancia (RIED)
 - **Anexo 2:** Artículo publicado en Education in the Knowledge Society (EKS)
 - **Anexo 3:** Artículo aceptado en Education and Information Technologies (EAIT)
 - **Anexo 4:** Artículo en Education and Information Technology (RIED)
 - **Anexo 5:** Capítulo de libro aceptado en editorial Tirant Lo Blanch
 - **Anexo 6:** Artículo en revisión: Revisión sistemática de la literatura sobre e-evaluación
 - **Anexo 7:** Artículo en revisión: Análisis en CB-SEM del estudio intermedio

2. **ANEXOS 8A-8E: Comunicaciones a Congreso derivadas de la Tesis Doctoral**
 - **Anexo 8A:** XI Conference on Virtual Campus (2021)
 - **Anexo 8B:** Technological Ecosystems for Enhancing Multiculturality (2022)
 - **Anexo 8C:** Congreso Internacional: Educación y Conocimiento (2023)
 - **Anexo 8D:** Congreso Internacional Aprendizaje, Innovación y Cooperación (2023)
 - **Anexo 8E:** Technological Ecosystems for Enhancing Multiculturality (2023)

3. **ANEXOS A-C: Anexos de la Tesis Doctoral (no vinculados con publicación)**
 - **Anexo A:** Informe del Comité de Ética en la Investigación de la Universidad de Salamanca
 - **Anexo B:** Ítems del modelo F/S-MAAM
 - **Anexo C:** Extended abstract of the Ph.D. Dissertation

Anexos 1-7
Publicaciones derivadas de la Tesis Doctoral

Anexo 1

Publicación en revista científica: Revista Iberoamericana de Educación a Distancia

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2021). Calidad en e-Learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2). <https://doi.org/10.5944/ried.24.2.29073>

INDICADORES:

Revista Q2 – JIF en año de publicación (2021): Factor de impacto 2,689. Pos. 111/270. Percentil 59,07
Revista Q1 – JCI en año de publicación (2021): Factor de impacto 2,57. Pos. 93/743. Percentil 87,55

Calidad en e-Learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior

(e-Learning Quality: Identification of its Dimensions, Proposal and Validation of a Model for its Assessment in Higher Education)

Alberto Ortiz-López
Susana Olmos-Migueláñez
José Carlos Sánchez-Prieto
Universidad de Salamanca, USAL (España)

DOI: <https://doi.org/10.5944/ried.24.2.29073>

Cómo referenciar este artículo:

Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2021). Calidad en e-Learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2), pp. 225-244. <https://doi.org/10.5944/ried.24.2.29073>

Resumen

La evaluación de la calidad en instituciones de Educación Superior es un proceso no estandarizado y dependiente de cada institución universitaria. La digitalización de la docencia y las situaciones acontecidas en los últimos tiempos han provocado un aumento notable en el interés y la relevancia del e-Learning en los procesos de enseñanza/aprendizaje; entendiéndose también la calidad como un elemento fundamental en su evaluación y mejora.

Esta investigación propone la identificación de las dimensiones relativas a la calidad del e-Learning, denotando la necesidad de un nuevo marco evaluativo tras el análisis de la literatura y generando un nuevo modelo. Tras su construcción, el estudio se centró en la validación de su contenido por cuatro jueces expertos en materia de evaluación de la calidad y de e-Learning. Este proceso de validación incluyó la elaboración de una herramienta para la evaluación de dimensiones e ítems y una presentación del modelo. Por otro lado, para la evaluación de resultados se realizó una evaluación cualitativa (análisis de contenidos y sugerencias) y una evaluación cuantitativa (analizada mediante estadísticos descriptivos y el cálculo del CVC). Esta validación mostró un alto nivel de aceptación y conformidad de los jueces

con el modelo, las dimensiones y los ítems propuestos. Partiendo de esta validación se elaboró el modelo definitivo que consta de cuatro dimensiones principales, 14 subdimensiones y un total de 80 indicadores que tratan de evaluar cada factor de calidad relevante en e-Learning. Las limitaciones y futuras vías de investigación derivadas de la propuesta se presentan en las conclusiones finales.

Palabras clave: evaluación; calidad de la enseñanza; tecnología de la educación; enseñanza superior; método de enseñanza; construcción de modelos.

Abstract

Quality assessment in higher education institutions is a non-standardised process and it depends on each university. The digitalisation of teaching and the situations, which have occurred in recent times, have led to a notable increase in the interest and relevance of e-Learning in both the teaching as well as the learning processes, with quality also being understood as a fundamental element in its evaluation and improvement.

This research proposes the identification of the dimensions related to the e-Learning quality, noting the need of a new evaluative framework according to the analysis of the literature and, also, generating a new model. After its construction, this study focused on the validation of its content by four expert judges in the quality evaluation and e-Learning fields. This validation process included the development of a tool for the evaluation of its dimensions and items, as well as a presentation of the model. On the other hand, both a qualitative evaluation (content analysis and suggestions) as well as a quantitative evaluation (analysed through descriptive statistics and the CVC calculation) were carried out for evaluating the outcomes. This validation showed a high level of acceptance and conformity of the judges with the model, the dimensions and the proposed items. Based on this validation, the final model was developed consisting of four main dimensions, 14 sub-dimensions and a total of 80 indicators that attempt to assess each relevant quality factor in e-Learning. Limitations and future avenues of research derived from the proposal are presented in the final conclusions.

Keywords: assessment; quality of teaching; education technology; higher education; teaching method; model building.

La Educación Superior se encuentra hoy inmersa en una amplia transformación hacia la digitalización de sus enseñanzas, servicios y procesos; y hacia la mejora de la calidad tanto de su docencia como de todos aquellos factores externos que la condicionan y que dependen de la propia institución. Esta constante reflexión y mejora de los procesos hacia la consecución de altos niveles de calidad es hoy una de las principales hojas de ruta para cualquier entidad y profesional que la conforma, y es aquí donde la intervención de los nuevos escenarios y marcos de aprendizaje condiciona y cuestiona todo el proceso de consecución de la calidad, de su relevancia y sus diferentes interpretaciones (Fernández et al., 1999; Córdova et al., 2017).

En la transformación hacia las enseñanzas en red son dos las vertientes bibliográficas que se posicionan en función de la concepción de la calidad en e-Learning. Así, mientras que parte de la comunidad científica considera fundamental partir de una reestructuración y replanteamiento de las dimensiones clásicas de evaluación de la calidad en el ámbito presencial, valorando su adecuación y su nueva dimensionalización de cara a una correcta evaluación en este ámbito; otros consideran que la evaluación ha de ser exactamente la misma (sin cambiar la dimensionalización de la calidad), solamente difiriendo en los pesos de sus dimensiones (Marciniak y Gairín-Sallán, 2017).

Estos cambios que hoy se producen en la mayoría de enseñanzas tienen como principal sustento las Tecnologías de la Información y la Comunicación (TIC), las cuales permiten la configuración de procesos de enseñanza a las características del estudiante, la adaptación de estilos de aprendizaje, la posibilidad de una interacción mayor, más rápida y más sincrónica; y la implementación de modelos de calidad, no solo centrados en la variable tecnológica sino también en las variables didácticas, organizativas y pedagógicas de la enseñanza en su término más amplio (Almenara, 2006).

Los avances que la Educación Superior dirigía en los últimos años hacia esta modalidad se impulsan hoy a causa de las repercusiones educativas de la crisis sanitaria del COVID-19, la cual ha obligado a realizar una fugaz reestructuración de los marcos de enseñanza (derivado del cese forzado de la presencialidad) y una rápida adaptación a los entornos virtuales de aprendizaje, en la que el e-Learning se presenta como la única alternativa viable a la continuación del proceso de enseñanza-aprendizaje (Aguas-Díaz et al., 2020). El problema que esta adaptación ha tenido deriva de la falta de consenso y la heterogeneidad de las características encontradas en función de la institución, la plataforma o el perfil de cada docente; entre otros. Además, la rapidez de cambio e implementación de nuevos modelos requiere también de la evaluación de la calidad de dichos cambios, la cual es esencial para la justificación y el correcto desarrollo del e-Learning. Por otro lado, es importante que el docente y la institución cuente con ayudas en el proceso de migración a la educación virtual (Moreno-Correa, 2020).

El e-Learning y las grandes readaptaciones que su implementación supone requiere también una gran flexibilidad académica e institucional, ya que todavía son escasos los protocolos concretos para la planificación y ejecución de esta nueva modalidad de enseñanza en muchas instituciones de Educación Superior (Cerdas-Montano et al., 2020), por lo que es fundamental evaluar la calidad de cada paso y cada avance. Esta calidad e-Learning, según Seoane y García-Peñalvo (2010), es definida como la “adecuación y relación directa de recursos técnicos, humanos, tecnológicos, métodos y herramientas utilizados en el proceso de enseñanza-aprendizaje virtual, acorde a las necesidades y expectativas de los estudiantes con el fin de garantizar la optimización del proceso de aprendizaje” (p. 156).

En lo que a las diferentes perspectivas respecta, Colás et al. (2005) presentan que e-Learning ha de considerar, para una evaluación completa; la perspectiva tecnológica (valoración de la calidad de las plataformas, orientadas a la adecuación de los cursos y enseñanzas), la perspectiva pedagógica (necesaria evaluación y readaptación dada la implementación de modelos provenientes de la formación presencial, con el replanteamiento de la “e-Pedagogía” y la revisión de modelos evaluativos pedagógicos acordes a los nuevos pensamientos actuales y a las posibilidades que el e-Learning ofrece) y, la evaluación desde una perspectiva psicológica (repensando las teorías del aprendizaje, las teorías cognitivas y el constructivismo).

García-Peñalvo (2005) presenta las dimensiones que constituyen el e-Learning y justifica que en función de los pesos de una u otra dimensión se obtienen diferentes modelos formativos y evaluativos. Estas dimensiones son los contenidos (calidad y estructuración), la tecnológica (plataformas y campus virtuales de formación), la tutoría (nueva concepción), la evaluación/acreditación (cambios en su planteamiento y en su función) y los servicios (más amplio y menos delimitado, abarca acción docente, gestión, elementos de evaluación, etc.).

Las diferentes perspectivas y dimensiones tratan de abarcar el e-Learning y su calidad desde un ámbito muy amplio y con distintas bases (tecnológicas, pedagógicas, económicas y de gestión; entre otras), sin lograr cubrir aún una metodología tan compleja y cambiante. Por ello, es importante apostar por la creación de estándares de calidad y por la evaluación de la calidad en todos sus contextos, fases y niveles; explorando nuevos métodos de evaluación con el objetivo de presentar niveles de calidad acordes a las nuevas realidades (Choi y Jeong, 2019).

Esta situación provoca que hoy se cuente con diversas perspectivas para la evaluación del e-Learning y su calidad. Por un lado, se encuentran modelos centrados y especializados en elementos parciales de la actividad formativa (materiales, plataformas, costos-beneficios, eficacia...), y por otro lado modelos con perspectivas globales de los que destacan dos tendencias: los modelos centrados en normas y estándares de calidad, y aquellos centrados en la práctica del benchmarking (Martínez, 2015).

Sin embargo, pese a los avances en la literatura y entrando ya en los modelos específicos de la evaluación de la calidad en instituciones de educación superior, éstas siguen presentando planes de calidad sin contar con los avances en la investigación, tomando aspectos de infinidad de fuentes y sin alcanzar criterios comunes para la evaluación. Además, es necesario abogar también por la construcción de estándares e indicadores para la evaluación del e-Learning que vayan en consonancia con las características propias y los diferentes modelos existentes; y que no solo tenga en consideración aspectos formativos y tecnológicos, sino que también ahonde en las variables personales y organizacionales, poco presentes en algunos modelos y parte esencial también de una completa evaluación de la calidad (Mejía y López, 2016).

En un metaanálisis llevado a cabo por Marciniak y Gairín-Sallán (2017) sobre diferentes modelos propuestos para la evaluación de la calidad en e-Learning, fueron

catorce las dimensiones resultantes como relevantes para la calidad y su evaluación: Justificación del programa virtual, objetivos formativos del programa virtual, perfil de ingreso y egreso, contenido temático del programa virtual, actividades de aprendizaje, perfil del docente en línea, materiales y recursos didácticos (guía, unidad y otros materiales), estrategias didácticas, tutoría, evaluación del aprendizaje, plataforma virtual, evaluación de la fase inicial del programa, evaluación del desarrollo del programa y evaluación final del programa.

La calidad en Educación Superior toma diferentes perspectivas y propuestas para su evaluación, siendo muy diversas las herramientas y los marcos disponibles para ello, y encontrando infinidad de modelos en función de los autores, las dimensiones y el objetivo global de la evaluación. En un breve recorrido por estos modelos y propuestas, destaca el modelo ELQ, implementado por la Agencia Nacional Sueca de Educación Superior con el objetivo de determinar los componentes de la calidad en entornos virtuales. Este modelo fue resultado de prácticas educativas en materia digital en Europa, incluyendo aspectos y criterios para la evaluación de la calidad en e-Learning en Educación Superior (Mixnahuatl et al., 2013). Este modelo avanza y se perfecciona con la práctica, y tras él se plantean los parámetros de calidad para la evaluación de las distintas fases propuestas, facilitando en gran medida su evaluación (Hansson et al., 2009).

En la práctica evaluativa en las Instituciones de Educación Superior, la calidad no sigue un estándar en su evaluación, tomando cada entidad distintas perspectivas en función de los intereses que sus responsables dictaminen, priorizando la evaluación de unas dimensiones frente a otras e incluso no contando con todas las dimensiones que componen la calidad en este proceso; y denotando la falta clara de un marco evaluativo común y ampliamente aceptado.

Este artículo se compone de tres secciones que se exponen seguidamente y que incluyen la presentación de los modelos analizados, la construcción del nuevo modelo y el proceso de validación de contenido (metodología); el análisis de la validación realizada por los jueces expertos y la adaptación y presentación final del modelo resultante (resultados) y el análisis de las aportaciones que este nuevo modelo supone en el campo del e-Learning y en la evaluación de la calidad, así como las vías futuras de investigación que este nuevo marco evaluativo supone (discusión).

METODOLOGÍA: DISEÑO Y VALIDACIÓN DEL MODELO

Se presenta a continuación el proceso seguido para la construcción del nuevo modelo y el diseño del instrumento para la evaluación de la calidad. Para ello, se partirá del análisis de varios modelos empleados en la actualidad y la categorización de sus dimensiones, continuando con el proceso de validación de contenido al que este nuevo modelo ha sido sometido para alcanzar el modelo final con las nuevas dimensiones.

Análisis de modelos de evaluación de calidad en e-Learning

Esta investigación toma como punto de partida las aportaciones realizadas durante los últimos cinco años en materia de evaluación de la calidad en e-Learning y toma la referencia de los modelos propuestos por los autores más relevantes en este ámbito y que hoy en día son herramientas en diferentes instituciones de Educación Superior para la evaluación de la calidad en sus modalidades e-Learning, seleccionados tras la realización de una revisión sistemática de la literatura (Ortiz-López et al., 2020).

El nuevo modelo se propone con el objetivo de aunar los criterios y dimensiones planteadas por distintos investigadores en este campo, presentando un instrumento más exhaustivo y profundo al contemplar distintas perspectivas e indicadores sobre la calidad, permitiendo así una evaluación más completa tras la integración de las dimensiones más comunes y el estudio de su relevancia, y tratando de minimizar las debilidades que los modelos actuales presentan al no contemplar la totalidad de dimensiones que la calidad tiene y que influyen sobre el resultado de cada evaluación (Rubio, 2003).

Para el análisis de las propuestas actuales se ha realizado una categorización de las dimensiones genéricas que los estudios proponen, tratando de delimitarlas para cada uno de ellos bajo una denominación común y corroborando cuáles son objeto de estudio y cuáles no en cada instrumento de evaluación, y presentando una muestra de todas las dimensiones que afectan a la evaluación de la calidad (tabla 1).

Tabla 1

Análisis de las dimensiones relevantes en la evaluación de la calidad en e-Learning según cada autor

	Dimensiones ¹											
	SQ	CX	CQ	AT	SP	DP	EV	AS	EN	IP	ES	MR
Choi y Jeong (2019)	▪	▪		▪								
García y González (2015)		▪				▪	▪					
Kurilovas y Vinogradova (2016)	▪		▪		▪	▪						
La Rotta et al. (2019)	▪				▪	▪	▪	▪	▪			
Mahdiuon et al. (2017)		▪						▪				▪
Marciniak (2018)	▪	▪	▪	▪	▪	▪	▪	▪	▪		▪	▪
Marciniak (2015)	▪		▪		▪	▪				▪	▪	
Margalina et al. (2017)	▪				▪	▪		▪				

	Dimensiones ¹											
	SQ	CX	CQ	AT	SP	DP	EV	AS	EN	IP	ES	MR
Martínez-Argüelles y Batalla-Busquets (2016)	▪		▪			▪		▪				
Martínez-Caro et al. (2015)	▪				▪	▪					▪	
Park y Lim (2015)	▪		▪		▪	▪	▪		▪			
Pecori et al. (2019)					▪				▪			
Pham et al. (2019)	▪		▪	▪	▪	▪		▪				▪
Prieto-Rodríguez et al. (2016)	▪				▪	▪		▪				▪
Raspopovic y Jankulovic (2017)					▪		▪					▪
Serrano et al. (2018)	▪					▪				▪	▪	
Torres-Barbazal et al. (2018)		▪	▪		▪	▪	▪				▪	
Waheed et al. (2016)	▪	▪	▪	▪	▪					▪		
Waheed y Kaur (2017)	▪	▪	▪	▪			▪	▪				

Esta variedad en cuanto a las dimensiones de los distintos modelos, así como la variabilidad presente en sus ítems, hacen que la evaluación de la calidad carezca todavía de una herramienta fiable y aceptada de forma global, por lo que es importante focalizar la investigación en la detección de las actuales dimensiones, subdimensiones e ítems y tratar de aunar todas ellas en un marco común y una herramienta que contemple las distintas perspectivas. (Choi y Jeong, 2019). Las dimensiones y sus múltiples denominaciones se aúnan de forma posterior en cuatro grandes subdimensiones (facilitando la comprensión del modelo), disgregadas en subdimensiones posteriormente. Estas cuatro dimensiones son la calidad de la institución, la calidad docente, la calidad del sistema o plataforma y la evaluación del programa. En la búsqueda de un estándar a la hora de categorizar las dimensiones de la calidad y atendiendo también a los pesos relativos que los estudios analizados dan a cada dimensión, se alcanza la conclusión unánime de los autores de la identificación de cuatro dimensiones principales. Para contar con un mayor nivel de conformidad, el análisis de la propuesta validada de Marciniak y Gairín-Sallán (2017) para la autoevaluación de la calidad confirmó la presencia de cuatro dimensiones principales: organización del programa, diseño del programa (metodológico y docente), desarrollo del programa (cronológico y tecnológico) y evaluación del programa. Por ello, se mantiene la decisión de asumir cuatro dimensiones principales que conformarán el primer modelo diseñado de cara al proceso de validación.

Una vez analizadas estas propuestas y detectadas las dimensiones que afectan a la calidad, se diseña y formula la hipótesis relacional que es sobre la que trabaja este modelo propuesto. Tras ello, se diseña el nuevo modelo atendiendo a las dimensiones más relevantes y a los cuatro grupos dimensionales propuestos (mencionados anteriormente) para poder realizar la evaluación de la calidad en e-Learning. El modelo resultante consta de dos secciones diferenciadas, presentando un carácter mixto (recopilación cuantitativa y cualitativa de la información). La primera parte del modelo se destina a la evaluación de las cuatro dimensiones (indicando previamente el objeto de evaluación) y de sus dimensiones mediante la categorización de los indicadores. La segunda parte del mismo responde a una evaluación cualitativa del evaluador, completando y enriqueciendo los resultados del proceso.

La escala que acompaña a cada ítem en esta evaluación de la calidad es una escala nominal dicotómica y excluyente, en la que indicar el cumplimiento o no cumplimiento del indicador y en la que poder incluir una pequeña justificación. En este planteamiento se ha optado por el empleo de la dicotomía para dictaminar si el criterio se cumple o no se cumple, y no entrar en interpretaciones del grado de cumplimiento que complicarían en gran medida la evaluación final de los resultados. Se incorpora además una evaluación cualitativa de cada ítem, de cada dimensión y del constructo global; en la que el evaluador puede incluir un comentario escrito para poder realizar cualquier indicación que considere oportuna y de cara también a contar con una evaluación más completa y orientada a la mejora.

El instrumento elaborado cuenta con un total de 74 ítems agrupados en cuatro dimensiones (tabla 2). Estos ítems proceden de la reflexión y adaptación de las distintas propuestas y también de la elaboración propia de cara a una correcta comprensión total del constructo.

Tabla 2
Estructura del modelo diseñado

Dimensión	Número de subdimensiones	Número de ítems
Calidad de la institución	3	14
Calidad docente	4	30
Calidad del sistema de aprendizaje	2	15
Calidad de la evaluación del programa	3	15

Los constructos planteados, con el objetivo de facilitar y profundizar en la evaluación, se dividen a su vez en un total de 12 subdimensiones que profundizan en un criterio de evaluación determinado para cada uno de ellos (tabla 3).

Tabla 3
Composición inicial del modelo

	Número de ítems
Dimensión: Calidad de la institución	
Política institucional	5
Contexto institucional	4
Respuesta y ayuda institucional	5
Dimensión: Calidad Docente	
Función docente	11
Estrategias docentes	7
Materiales y recursos docentes	7
Tutoría	5
Dimensión: Calidad del sistema de aprendizaje o plataforma	
Funcionamiento	9
Accesibilidad	6
Dimensión: Calidad de la evaluación del programa	
Evaluación de la preparación previa del programa	3
Evaluación procesual	4
Evaluación final	8

Proceso de validación de contenido

Para el proceso de validación de contenido del modelo diseñado, se planteó un instrumento de validación adaptado de las propuestas de Chiang et al. (2013), Galicia et al. (2017) y García-Llorente et al. (2019). La herramienta propuesta para la validación contiene una descripción del modelo e instrumento y su finalidad, una descripción de las dimensiones y subdimensiones descritas y un planteamiento de la evaluación a realizar dividida en tres grandes módulos. En las instrucciones se indica además qué variables valorar de cada subdimensión/dimensión (suficiencia de ítems y relevancia para la dimensión/constructo) y de cada ítem (coherencia, relevancia y claridad). La escala para la validación de contenido es una escala tipo Likert de 1 a 4, en la que 1 es la falta total de coherencia / relevancia / claridad / insuficiencia de ítems / irrelevancia de la dimensión, y 4 la completa relación / relevancia / relación / excesividad de ítems.

El primer módulo corresponde, para cada dimensión, a la evaluación de la coherencia, relevancia y claridad de cada ítem de forma totalmente independiente al resto y a las dimensiones, analizando exclusivamente la redacción y el sentido del ítem. En segundo lugar, los evaluadores tuvieron que analizar la suficiencia de los ítems y la representatividad para el conjunto (ambas también en escala de 1 a 4) para

las subdimensiones planteadas para cada constructo. Este proceso fue repetido de la misma forma para las cuatro dimensiones, incluyendo espacios para comentarios adicionales del evaluador al final de cada dimensión.

Por último, se incluye también una evaluación para las subdimensiones presentes en el cuestionario en relación con su suficiencia y representatividad para el constructo global (también evaluable en una escala Likert 1-4). Esta escala se complementa con una valoración cualitativa de los evaluadores, con espacios al final de cada dimensión y un espacio final para que los evaluadores puedan emitir sus propios juicios cualitativos sobre el instrumento y la validación de contenido realizada.

Tras el diseño del instrumento para la validación de contenido, esta investigación seleccionó a un total de cuatro evaluadores expertos nacionales que destacan por ser profesionales del campo educativo y tecnológico e investigadores tanto en el ámbito de la evaluación de la calidad como en el ámbito del e-Learning y contar con publicaciones relevantes de carácter internacional sobre el tema de estudio. Una vez establecido el contacto vía mail con los mismos, se les facilitaron todas las instrucciones, un resumen del instrumento original y las dimensiones planteadas y, el documento para la validación. Ya recibidas las evaluaciones, éstas fueron digitalizadas para su análisis e interpretación de resultados. El proceso de validación de contenido del modelo comenzó el día 11 de noviembre (con el envío de la solicitud, instrucciones y modelo a los autores) y la última validación fue recibida por los autores el día 28 del mismo mes; con una duración total de 17 días.

RESULTADOS: UN MODELO UNIFICADO PARA LA EVALUACIÓN DE LA CALIDAD EN E-LEARNING EN EDUCACIÓN SUPERIOR

Como criterio empleado en la validación y en la selección de dimensiones, subdimensiones e ítems; se eliminan aquellas que presentan valores inferiores a 2, se realizan modificaciones profundas en los ítems que comprenden valores entre 2 y 3, y se modifica la formulación de los ítems con valores inferiores a 3,5. Complementariamente también se calculó el coeficiente de validez de contenido (CVC) (Hernández-Nieto, 2002) para la suma de las puntuaciones obtenidas por los ítems en los tres criterios. De esta manera, se admitieron aquellos ítems y dimensiones con un CVC superior a 0,7 (Pedrosa et al., 2014) y se modificaron aquellos que a pesar de tener puntuaciones medias superiores a 3,5 en las tres dimensiones no contasen con un CVC superior al punto de corte.

Finalmente, para la reformulación y la readaptación, se han tomado en cuenta todas aquellas consideraciones realizadas de forma cualitativa por los revisores expertos.

En el análisis de las dimensiones planteadas, los resultados presentan una media de suficiencia de 3,84 (D.T.= 0,289) y una representatividad para el constructo de 3,92 (D.T.= 0,144) sobre 4, por lo que se deduce un buen planteamiento de

dimensiones. En cuanto a las subdimensiones, los resultados también son buenos, con una suficiencia de medida media de 3,63 (D.T.= 0,602) y una representatividad de 3,74 (D.T.= 3,74). Por lo que a la evaluación individual de los ítems respecta, los resultados generales fueron muy positivos. El valor medio de coherencia para los ítems fue de 3,90 (D.T.= 0,172) sobre 4, de relevancia de 3,85 (D.T.= 0,308) y su claridad de 3,68 (D.T.= 0,537). Los resultados estadísticos obtenidos para el conjunto de ítems que componen cada constructo siguiendo los criterios de pertinencia, relevancia y claridad se encuentran disponibles en el enlace: <https://docs.google.com/document/d/1ThDL20-2oVj5SL9V8gVPJ1O3umGjjEXoH8-UJMiXUkk/edit?usp=sharing>

Tras el análisis de resultados y con base en los criterios establecidos para esta validación se determinó, en primer lugar, que las dimensiones presentan una adecuada estructura que no requiere modificación alguna. En cuanto a las dimensiones, se estudió el planteamiento de la suficiencia de ítems en la subdimensión 2 y 3. Respecto a los ítems, se llevó a cabo la realización de modificaciones en cuanto a la coherencia de ellos (63, 64, 65, 66), en su relevancia e interpretación (ítems 63, 64, 66) y en la redacción y claridad que presentan (ítems 9, 30, 40, 45, 64, 72, 74). Además, se determinó la reestructuración y modificación profunda por problemas de claridad de los ítems: 14, 29, 64, 65, 66 y 71. Por último, el ítem 35 fue modificado profundamente por no alcanzar la puntuación umbral del CVC (0,7).

En una perspectiva cualitativa, los cuatro jueces incluyeron comentarios tanto para las dimensiones como para el constructo general. Las principales evaluaciones refieren a la sustitución terminológica en alguno de los ítems de distintos sustantivos, atributos o tiempos verbales; y también al planteamiento de incorporación de alguna subdimensión e ítem para mejorar la medida del constructo, así como la eliminación o movimiento de alguno de ellos de cara al modelo final.

Los puntos fuertes que señala esta evaluación de jueces expertos se centran en la profundidad del modelo para medir los distintos componentes de la calidad, así como su potencial en el momento actual en materia de virtualización de la docencia universitaria. Además, se indica también lo positivo de plantear un nuevo modelo en materia de evaluación de la calidad en un campo aún por explorar, y la buena profundidad que el nuevo marco generado tiene en cuanto al número de ítems que comprenden la calidad.

En cuanto a las limitaciones del modelo que los jueces señalan, se encuentra la complejidad para la evaluación (dada la profundidad de la misma), y la necesaria experiencia y conocimientos del evaluador que la implemente. Además es una limitación también el acceso a las políticas institucionales en materia de calidad y a su entramado administrativo y de soporte, lo cual requiere de una buena predisposición del docente, de los evaluadores y de los agentes implicados en la evaluación para poder constatar el cumplimiento de los ítems.

Teniendo en cuenta las aportaciones realizadas por los expertos, se concluyó el diseño del modelo final manteniendo las dimensiones planteadas (en consonancia con lo planteado en el diseño del modelo y con el número de dimensiones planteado

por Marciniak y Gairín-Sallán, 2017), incluyendo dos nuevas subdimensiones (una relativa a la coordinación pedagógica e institucional y otra a la evaluación de la calidad del programa) y realizando las sustituciones indicadas en los ítems, añadiendo 8 ítems nuevos propuestos por los jueces expertos tras la eliminación de dos previos e introduciendo también las modificaciones en aquellos que así ya se han indicado, siendo un modelo resultante de 4 dimensiones, 14 subdimensiones y 80 ítems para la evaluación de la calidad en e-Learning en Educación Superior (tablas 4 y 5).

Tabla 4
 Estructura del modelo resultante tras la validación

	Número de subdimensiones	Número de ítems
Calidad de la institución	4	19
Calidad docente	4	26
Calidad del sistema de aprendizaje	2	15
Calidad de la evaluación del programa	4	20

Tabla 5
 Ítems validados para el modelo *final*

	Número de ítems
Dimensión: Calidad de la institución	
Política institucional	6
Contexto institucional	4
Respuesta y ayuda institucional	5
Coordinación pedagógica e institucional	4
Dimensión: Calidad Docente	
Función docente	8
Estrategias docentes	5
Materiales y recursos docentes	8
Tutoría	5
Dimensión: Calidad del sistema de aprendizaje o plataforma	
Funcionamiento	10
Accesibilidad	5
Dimensión: Calidad de la evaluación del Programa	
Evaluación de la preparación previa del programa	3
Evaluación procesual	4
Evaluación final	9
Evaluación de la calidad del programa	4

El modelo final se puede consultar en el siguiente enlace: https://docs.google.com/document/d/1BQxjD3xo2eTiQBbWOHP_1yfC61DpsowpSelmpDzFG2M/edit#heading=h.31is5tajphdr

El marco propuesto para la evaluación de la calidad en e-Learning en Educación Superior consta de cuatro dimensiones que evalúan las diferentes perspectivas y componentes de la calidad, tratando de aportar un estudio completo y un indicador fiable sobre la misma. Las cuatro dimensiones:

Calidad de la Institución

La calidad de la institución es entendida como un factor relevante y esencial para la calidad de un programa (Marciniak, 2015). La institución es la base del programa, el sustento del mismo y la entidad que lo ampara, lo desarrolla y lo promueve. Por ello, la institución ha de ser el primer factor en el estudio de la calidad de un programa e-Learning. La institución centra sus bases de la calidad en las políticas institucionales que implementa y desarrolla (SD1), en su organización y el contexto que la rodea y que comprenden sus estructuras (SD2), en la respuesta y en la ayuda que la institución ofrece a sus usuarios en materia de e-Learning y en el acompañamiento de sus enseñanzas (SD3); y en las políticas que la institución tiene y articula en materia de coordinación pedagógica y su consonancia con el e-Learning (SD4).

Calidad Docente

Por otro lado, la figura y la calidad del docente aportan también una representación importante en el valor de la calidad de la institución en el ámbito digital (Santelices-Etchegaray y Valenzuela-Rojas, 2015). Esta calidad está constituida por la función docente, su formación y su preparación en este campo (SD5), las estrategias docentes y las metodologías que éstos implementan en las enseñanzas en red de la institución (SD7), la adaptación de los materiales y de los recursos a las plataformas de enseñanza (SD7) y la acción tutorial, el acompañamiento que los docentes realizan a su alumnado a través de la red (SD8).

Calidad del Sistema de Aprendizaje o Plataforma

La tercera dimensión y componente de la calidad de un programa de educación en línea en Educación Superior es la calidad del sistema o de la plataforma mediante la que se lleva a cabo el aprendizaje. Hoy prácticamente la totalidad de instituciones poseen plataformas de esta índole, contando con suficientes herramientas para dotar también de calidad a los sistemas formativos (Velasteguí, 2017). Esta plataforma es fundamental y mediadora en el aprendizaje, por lo que su estudio es relevante

para la calidad global de la enseñanza. Por un lado, esta dimensión comprende el funcionamiento del sistema institucional –analizando su funcionamiento, respuesta, sincronía, consistencia, personal supervisor; entre otras– (SD9), y la accesibilidad de la misma para la consulta y la interacción con los materiales, las funciones, las guías docentes (SD10).

Calidad de la evaluación del Programa

Finalmente, la calidad ha de comprender también la propia evaluación de ella que se realiza en el Programa de enseñanza y de los resultados y avances obtenidos. En los programas actuales y en la formación en línea es hoy fundamental la implementación y la evaluación continuada de la calidad y de los avances, contando con órganos específicos y planes de seguimiento para ello (Tari et al., 2007). Esta evaluación se estudia desde la perspectiva de la calidad desde una visión completa y profunda: desde la evaluación de la preparación previa del programa (preparación de recursos y del programa, del entorno virtual) (SD11), la evaluación procesual del mismo (seguimiento del programa, retroalimentación docente e institucional, estudio continuado de la calidad y los resultados) (SD12); la evaluación final (objetivos, evaluaciones aplicadas, niveles de satisfacción...) (SD13) y finalmente también se comprende la propia evaluación de la calidad que el programa implementa, los encargados de dicha evaluación y los ítems de calidad con los que cuenta (SD14).

DISCUSIÓN

La calidad en e-Learning es un aspecto fundamental en el desarrollo, la continuidad y la garantía de las instituciones, enseñanzas y programas educativos que hoy se implementan en red en el ámbito de la Educación Superior. Esta transformación de la enseñanza y los cambios que afrontan las universidades, docentes, alumnado y todo el personal implicado suponen grandes avances y adaptaciones a un nuevo contexto más dinámico, adaptado y cambiante que la presencialidad. Por ello, y ya que todavía no se dispone de un modelo profundo que evalúe todas las fases de un proceso completo de enseñanza en red, es importante plantear un nuevo marco evaluativo que contemple todo el proceso y todos los agentes implicados, pudiendo realizar una amplia y profunda evaluación para entender la calidad como un factor relevante y esencial en la enseñanza digital (Verdezoto-Rodríguez y Chávez-Vaca, 2018).

Con el objetivo de asentar las bases de un nuevo modelo validado, profundo y longitudinal para la evaluación de un curso en modalidad e-Learning, se propuso una validación de contenido por jueces expertos sobre el modelo generado. El proceso de validación de contenido fue llevado a cabo por cuatro jueces expertos en e-Learning y en evaluación de la calidad y el resultado de dicho proceso arrojó valores muy elevados de suficiencia y representatividad de las dimensiones y subdimensiones

planteadas y también altos valores de coherencia, relevancia y representatividad para los ítems propuestos. Además, la evaluación cualitativa de los jueces expertos contribuyó a la puntualización y mejora final del modelo con la adicción de ocho nuevos ítems para una más completa evaluación de la calidad.

Mientras que los estudios actuales profundizan en una única dimensión de la calidad, o plantean el conjunto de una forma limitada; el instrumento diseñado en esta investigación para la evaluación de la calidad plantea una nueva y más amplia perspectiva en el campo, presentando una nueva visión de la calidad y tratando además de profundizar más en las dimensiones que ya han sido estudiadas por otros autores. Por ello, este nuevo modelo plantea la unión de diferentes perspectivas y dimensiones, la continuidad en la búsqueda de indicadores de calidad en ellas y también una nueva visión de la calidad como un continuo de acciones compartidas y complementarias que comprenden desde la acción institucional y las políticas de coordinación, hasta la acción docente, la preparación previa del programa, la respuesta de la plataforma o la evaluación realizada.

Este modelo no solo pretende ser de utilidad en las continuas evaluaciones que las instituciones de Educación Superior realizan sobre sus enseñanzas en la modalidad en red, sino que también es un modelo que tiene también cabida en la autoevaluación docente sobre sus enseñanzas y cursos de cara a la mejora y a la detección de necesidades. Además, el modelo diseñado es un modelo versátil que admite una amplia multidisciplinariedad de evaluadores (dada la fácil comprensión y medición de los indicadores) y la orientación a la mejora al detectar en qué dimensión y subdimensión se encuentran aquellos ítems que presenten déficits en la evaluación. Además, la evaluación propuesta en este nuevo modelo permite la evaluación de todas las fases del proceso (no solo la centrada en el proceso de enseñanza-aprendizaje en sí), evaluando desde la preparación inicial de la plataforma o los recursos hasta la evaluación final del modelo; ofreciendo así datos mucho más amplios y específicos de cada fase y del proceso completo.

La principal limitación de esta investigación es la fase en la que el estudio se encuentra aún (derivando de la misma su falta de validez de constructo); lo cual abre ya una nueva vía en la investigación para continuar el estudio con una validación empírica del modelo y su aplicación en un programa e-Learning de un campus en línea de educación superior. Este estudio piloto, el cual se encuentra ya planteado de cara a la evaluación de la docencia en un curso blended-Learning en el año académico 2021/2022, permitirá comprobar la fiabilidad y la validez del modelo, ofreciendo profundos resultados sobre los indicadores de calidad de la institución, de la evaluación y del programa en sí, y haciendo de la calidad uno de los factores más relevantes en las enseñanzas e-Learning en Educación Superior.

Otra limitación adicional ha sido la dificultad para realizar el proceso de validación de contenido teniendo en cuenta las circunstancias provocadas por la pandemia durante las fechas en las que se llevó a cabo y que dificultaron la obtención de jueces dispuestos a revisar la propuesta debido a la saturación de trabajo y las

dificultades de índole personal. Esto ha provocado que el número de jueces fuera más reducido del deseable, aunque se supera el número mínimo establecido en la literatura (Hernández-Nieto, 2002; Pedrosa et al., 2014) y el grado de satisfacción de los jueces con el nuevo modelo es alta, pudiendo continuar con su estudio y aplicación para una validación en la práctica educativa.

NOTAS

1. SQ: System Quality, CX: Context / External Variables, CQ: Content Quality, At: Attractiveness, SP: Student Profile, TP: Teacher Profile, EV: Evaluation, AS: Administrative Support, EN: Enrollment, IP: Institutional Policies, ES: Educative Strategies, MR: Material Resources.

REFERENCIAS

- Aguas-Díaz, C. J., Flores-Flores, J. L., Sarmiento-Sarmiento, I. K., y Aguirre-Robalino, D. F. (2020). Aprendizaje móvil (m-learning) como método educativo en Educación Superior. *Revista Arbitrada Interdisciplinaria Koinonía*, 5(Extra 1), 867-879. <https://doi.org/10.35381/r.k.v5i1.887>
- Almenara, J. (2006). La calidad educativa en el e-Learning: Sus bases pedagógicas. *Educación Médica*, 9, 9-14. <https://doi.org/10.4321/S1575-18132006000700003>
- Cerdas-Montano, V., Mora-Espinoza, Á., y Salas-Soto, S. E. (2020). Educación remota en el contexto universitario: Necesidad del trabajo colaborativo para la mediación pedagógica docente en tiempos de COVID. *Revista Electrónica Educare*, 24(Extra 0), 3. <https://doi.org/10.15359/ree.24-S.9>
- Chiang, M. T., Díaz, C., Rivas, A., y Martínez, P. (2013). Validación del cuestionario estilos de enseñanza (CEE): Un instrumento para el docente de Educación Superior. *Revista de estilos de aprendizaje*, 6(12), 30-45.
- Choi, C.-R., y Jeong, H.-Y. (2019). Quality evaluation for multimedia contents of e-learning systems using the ANP approach on high speed network. *Multimedia Tools and Applications*, 78(4), 28853–28875. <https://doi.org/10.1007/s11042-019-7351-8>
- Colás, M. P., Jiménez, R., y Rodríguez, M. (2005). Evaluación de e-learning. Indicadores de calidad desde el enfoque sociocultural. *Education in the knowledge society (EKS)*, 6(2), 3.
- Córdova, L. C., Solis, M. E., y Solis, L. A. (2017). *Aplicación de los sistemas e-Learning en las Universidades* [Tesis Doctoral, Universidad Estatal de Milagro]. <http://repositorio.unemi.edu.ec/handle/123456789/3605>
- Fernández, B., Alonso, E., Rebolloso, E., y Pozo, C. (1999). Evaluación de la calidad en la educación superior. *Papeles del psicólogo*, 74, 1886-1415.
- Galicia, L. A., Balderrama, J. A., y Edel, R. (2017). Validez de contenido por juicio de expertos: Propuesta de una herramienta virtual. *Apertura: Revista de Innovación Educativa*, 9(2), 42-53. <https://doi.org/10.32870/Ap.v9n2.993>
- García, P. A., y González, M. S. (2015). Validación de escala para evaluación de la calidad docente en entornos virtuales. *Opción: Revista de Ciencias Humanas y Sociales*, 31(5), 394-406.
- García-Llorente, H. J., Martínez-Abad, F., y Rodríguez-Conde, M. J. (2019). Validación

- de un instrumento de evaluación de competencias informacionales autopercebidas en educación secundaria obligatoria. *Anales de Documentación*, 22(1), Article 1. <https://doi.org/10.6018/analesdoc.22.1.305641>
- García-Peñalvo, F. J. (2005). Estado actual de los sistemas e-learning. *Education in the knowledge society (EKS)*, 6(2).
- Hansson, H., Westman, P., Åström, E., y Johansson, M. (2009). Un modelo para la evaluación de la calidad e-learning. *Revista de Investigaciones UNAD*, 8(1), 11-29. <https://doi.org/10.22490/25391887.619>
- Hernández-Nieto, R. A. (2002). *Contributions to Statistical Analysis*. Mérida, Venezuela: Universidad de Los Andes.
- Kurilovas, E., y Vinogradova, I. (2016). Improved fuzzy AHP methodology for evaluating quality of distance learning courses. *The International journal of engineering education*, 32(4), 1618-1624.
- La Rotta, D., Usuga, O. C., y Clavijo, V. (2019). Perceived service quality factors in online higher education. *Learning Environments Research*, 23, 251-267. <https://doi.org/10.1007/s10984-019-09299-6>
- Mahdiun, R., Masoumi, D., y Farasatkah, M. (2017). Quality improvement in virtual higher education: A grounded theory approach. *Turkish Online Journal of Distance Education*, 18(1), 111-131. <https://doi.org/10.17718/tojde.285720>
- Marciniak, R. (2015). Methodological proposal for the application of international benchmarking in order to assess the quality of virtual higher education. *RUSC. Universities and Knowledge Society Journal*, 12(3), 46-60. <https://doi.org/10.7238/rusc.v12i3.2163>
- Marciniak, R. (2018). Quality Assurance for Online Higher Education Programmes: Design and Validation of an Integrative Assessment Model Applicable to Spanish Universities. *The International Review of Research in Open and Distributed Learning*, 19(2), 126-154. <https://doi.org/10.19173/irrodl.v19i2.3443>
- Marciniak, R., y Gairín-Sallán, J. (2017). Un modelo para la autoevaluación de la calidad de programas de educación universitaria virtual. *Revista de Educación a Distancia (RED)*, 54. <https://doi.org/10.6018/red/54/2>
- Margalina, V., De Pablos-Heredero, C., y Montes, J. L. (2017). Achieving quality in e-Learning through relational coordination. *Studies in Higher Education*, 42(9), 1655-1670. <https://doi.org/10.1080/03075079.2015.1113953>
- Martínez, N. (2015). Aprendizaje y evaluación con TIC: Un estado del arte. *Revista Científica Universidad Don Bosco*, 12.
- Martínez-Argüelles, M., y Batalla-Busquets, J.-M. (2016). Perceived Service Quality and Student Loyalty in an Online University. *The International Review of Research in Open and Distributed Learning*, 17(4), 264-276. <https://doi.org/10.19173/irrodl.v17i4.2518>
- Martínez-Caro, E., Cegarra-Navarro, J. G., y Cepeda-Carrión, G. (2015). An application of the performance-evaluation model for e-learning quality in higher education. *Total Quality Management & Business Excellence*, 26(5-6), 632-647. <https://doi.org/10.1080/14783363.2013.867607>
- Mejía, J. F., y López, D. (2016). Modelo de Calidad de E-learning para Instituciones de Educación Superior en Colombia. *Formación Universitaria*, 9(2), 59-72. <https://doi.org/10.4067/S0718-50062016000200007>
- Mixnahuatl, A., Santamaría, J., Urbina, A., Calleja, J. de la, y Auxilio, M. (2013). Diseño del curso análisis de sistemas aplicando el modelo de calidad para e-learning. *Educación Handbook T-I: Congreso Interdisciplinario de Cuerpos Académicos*, 143-152.
- Moreno-Correa, S. M. (2020). La innovación educativa en los tiempos del Coronavirus. *Salutem Scientia Spiritus*, 6(1), 14-26.

- Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2020). e-Learning quality assessment in higher education: A mapping study. *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality*, 833-838. <https://doi.org/10.1145/3434780.3436602>
- Park, Y., y Lim, K. (2015). Effects of Environmental and Human Constructs on E-learning Effectiveness in Online University Settings. *Indian Journal of Science and Technology*, 8(1), 103-109. <https://doi.org/10.17485/ijst/2015/v8iS1/57729>
- Pecori, R., Suraci, V., y Ducange, P. (2019). Efficient computation of key performance indicators in a distance learning university. *Information Discovery and Delivery*, 47(2), 96-105. <https://doi.org/10.1108/IDD-09-2018-0050>
- Pedrosa, I., Suárez-Álvarez y García-Cueto, E. (2014). Evidencias sobre la Validez de Contenido: Avances Teóricos y Métodos para su Estimación. *Acción Psicológica*, 10(2), 3-20. <https://doi.org/10.5944/ap.10.2.11820>
- Pham, L., Limbu, Y. B., Bui, T. K., Nguyen, H. T., y Pham, H. T. (2019). Does e-learning service quality influence e-learning student satisfaction and loyalty? Evidence from Vietnam. *International Journal of Educational Technology in Higher Education*, 16(7), 1-26. <https://doi.org/10.1186/s41239-019-0136-3>
- Prieto-Rodríguez, E., Gore, J., & Holmes, K. (2016). Exploring Quality Teaching in the Online Environment Using an Evidence-Based Approach. *Australian Journal of Teacher Education*, 41(8), 22-39. <https://10.14221/ajte.2016v41n8.2>
- Raspopovic, M., y Jankulovic, A. (2017). Performance measurement of e-learning using student satisfaction analysis. *Information Systems Frontiers*, 19(4), 869-880. <https://doi.org/10.1007/s10796-016-9636-z>
- Rubio, M. J. (2003). Enfoques y modelos de evaluación del e-learning. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 9(2), Article 2. <https://doi.org/10.7203/relieve.9.2.4332>
- Santelices-Etchegaray, M. V., y Valenzuela-Rojas, F. (2015). Importancia de las características del profesor y de la escuela en la calidad docente: Una aproximación desde la Teoría de Respuesta del Ítem. *Estudios pedagógicos (Valdivia)*, 41(2), 233-254. <https://doi.org/10.4067/S0718-07052015000200014>
- Seoane Pardo, A. M., y García-Peñalvo, F. (2010). *Introducción al eLearning*. Universidad de Salamanca.
- Serrano, E. L., Ceballos, S. P., Cordero, G., y Cisneros-Cohernour, E. (2018). Marco para evaluar las condiciones institucionales de la enseñanza en línea. *Revista Electrónica de Investigación Educativa*, 20(2), 1-14. <https://doi.org/10.24320/redie.2018.20.2.2072>
- Tarí, J. J., de Juana, S. A., y Mora, J. (2007). Evaluación de la calidad: Importancia de la evaluación externa y del seguimiento. *Avances en Supervisión Educativa*, 5, Article 5.
- Torres-Barzabal, L., Ortiz, P., y Barcia-Tirado, D. (2018). Quality Indicators for Auditing on-Line Teaching in European Universities. *TechTrends*, 63(3), 330-340. <https://doi.org/10.1007/s11528-018-0365-7>
- Velasteguí, P. (2017). Plataformas virtuales y su impacto en la Educación Superior. *Explorador Digital*, 1(2), 5-21 <https://doi.org/10.33262/exploradordigital.v1i2.318>
- Verdezoto-Rodríguez, R. H., y Chávez-Vaca, V. A. (2018). Importancia de las herramientas y entornos de aprendizaje dentro de la plataforma e-learning en las universidades del Ecuador. *EduTec. Revista Electrónica de Tecnología Educativa*, 65, 68-92. <https://doi.org/10.21556/edutec.2018.65.1067>

- Waheed, M., y Kaur, K. (2017). Students' perceptual quality standards for judging knowledge quality: Development and validation of a knowledge quality scale. *Information Development*, 35(2), 319-332. <https://doi.org/10.1177/0266666917744370>
- Waheed, M., Kaur, K., y Qazi, A. (2016). Students' Perspective on Knowledge Quality in eLearning Context: A Qualitative Assessment. *Internet Research*, 26(1), 120-145. <https://doi.org/10.1108/IntR-08-2014-0199>

PERFIL ACADÉMICO Y PROFESIONAL DE LOS AUTORES

Alberto Ortiz-López. Graduado en Pedagogía y estudiante del máster universitario en “Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas” en la modalidad de orientación educativa en la Universidad de Salamanca. Ha recibido una beca de colaboración en el Instituto Universitario de Ciencias de la Educación en Salamanca, donde se encuentra desarrollando una línea de investigación centrada en la evaluación de la calidad y su implementación en sistemas e-Learning en educación superior. <https://orcid.org/0000-0003-3164-5680>
E-mail: aortiz@usal.es

Susana Olmos-Migueláñez. Titular de Universidad del Área de Métodos de Investigación y Diagnóstico en Educación, de la Universidad de Salamanca. Directora del Instituto Universitario de Ciencias de la Educación. Miembro del Grupo de Investigación en Evaluación Educativa y Orientación, dentro del Grupo GRIAL, InterAcción y eLearning. Docente universitaria de Grado y Máster en Metodología de Evaluación de Programas, Técnicas e instrumentos de evaluación. Su línea de investigación se centra en evaluación educativa, metodología de investigación y de evaluación, y en evaluación en contexto de formación virtual. <https://orcid.org/0000-0002-0816-4179>
E-mail: solmos@usal.es

José Carlos Sánchez-Prieto. Licenciado en Pedagogía y máster en TIC aplicadas a la educación por la Universidad de Salamanca (España), donde también presentó su tesis doctoral sobre la adopción de las tecnologías móviles entre el profesorado dentro del Programa de Formación en la Sociedad del Conocimiento. Actualmente es profesor asociado en la Facultad de Educación de dicha universidad. Su área de investigación es la evaluación de las actitudes de los profesores y estudiantes en activo y en formación. Es autor de varios artículos sobre el tema publicados en prestigiosas revistas internacionales. <https://orcid.org/0000-0002-8917-9814>
E-mail: josecarlos.sp@usal.es

Anexo 2

Publicación en revista científica: Education in the Knowledge Society

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-Learning en Educación Superior: Una revisión sistemática de la literatura. *Education in the Knowledge Society (EKS)*, 23. <https://doi.org/10.14201/eks.26986>

INDICADORES:

Revista Q1 – JCI en año de publicación (2022): Factor de impacto 1,86. Posición 68/742. Percentil 90,90.

Revista Q2 – SJR último año de índice publicado a fecha de depósito artículo (2022): Factor de impacto 0,664

Incluido en índice JIF en 2023, factor de impacto 3,9. Se asignará cuartil en 2024.



e-Learning quality assessment in Higher Education: A systematic literature review

Evaluación de la calidad en e-Learning en Educación Superior: Una revisión sistemática de la literatura

Alberto Ortiz-López^{a*}, Susana Olmos-Migueláñez^b, José Carlos Sánchez-Prieto^c

^a Instituto Universitario de Ciencias de la Educación (IUCE), Universidad de Salamanca, España.

<https://orcid.org/0000-0003-3164-5680> aortiz@usal.es

^b Instituto Universitario de Ciencias de la Educación (IUCE), Universidad de Salamanca, España.

<http://orcid.org/0000-0002-0816-4179> solmos@usal.es

^c Instituto Universitario de Ciencias de la Educación (IUCE), Universidad de Salamanca, España.

<https://orcid.org/0000-0002-8917-9814> josecarlos.sp@usal.es

(*) Autor de Correspondencia / Corresponding author

ARTICLE INFO

Keywords

quality assurance, quality of education, e-Learning, higher education, systematic literature review.

Palabras clave:

evaluación de la calidad, calidad de la educación, educación en línea, enseñanza superior, revisión sistemática de la literatura.

ABSTRACT

Quality is nowadays a critical factor in Higher Education institutions, where evaluation processes and quality standards are reaffirmed as a guarantee of success and effectiveness of their online teaching. The proposed systematic review attempts to answer four questions that address the studied dimensions of quality, the processes followed for its evaluation, the contributions made, and the future lines of research. To this end, 715 publications were initially analyzed, which, after applying six inclusion criteria, resulted in a total of 25 contributions made in the last seven years to offer an in-depth view of quality assessment in e-Learning. The results reflect the breadth of the field in terms of its dimensions and the continuous generation of evaluative models. This field has not yet managed to unify the criteria and dimensions to generate a standard model to establish quality as a stable and validated factor in Higher Education institutions.

RESUMEN

La calidad es hoy un factor clave en las instituciones de Educación Superior, dónde los procesos de evaluación y los estándares de calidad se reafirman como garantía de éxito y eficacia de sus enseñanzas en línea. La revisión sistemática que se propone trata de dar respuesta a cuatro preguntas que abordan las dimensiones estudiadas de la calidad, los procesos seguidos para su evaluación, las aportaciones realizadas y las líneas futuras de investigación. Para ello, se analizaron inicialmente 715 publicaciones que, tras la aplicación de seis criterios de inclusión, derivaron en un total de 25 aportaciones realizadas en los últimos siete años con el objetivo de ofrecer una visión profunda de la evaluación de la calidad en e-Learning. Los resultados reflejan la amplitud del campo en cuanto a sus dimensiones y la continua generación de modelos evaluativos, un campo que no consigue unificar aún criterios y dimensiones de cara a la generación de un modelo común para asentar la calidad como un factor estable, continuo y validado en las instituciones de Educación Superior.

1. Introducción

La Educación Superior es, en el momento actual, un ámbito en constante transformación inmerso en un proceso continuo hacia la digitalización, la implantación en red de la enseñanza tradicional y la implementación de las Tecnologías de la Información y la Comunicación (TICs) en todos sus procesos, tanto de gestión como de enseñanza (Guerrero, 2018). Este proceso implica un cambio metodológico, conceptual y procedimental en las metodologías y en los procesos de enseñanza-aprendizaje e implica también situar el e-Learning como foco central en las instituciones de Educación Superior (Parlakkilic, 2017).

En este nuevo contexto universitario, apostar por generar marcos de calidad ha de ser un objetivo claro en toda institución educativa. El aumento en el interés y el estudio de las variables relacionadas con el e-Learning no sólo se refleja en el creciente número de agencias de evaluación de la calidad (Olmos-Migueláñez, 2008), sino también en los modelos para su evaluación y evolución, o en el ámbito más normativo en nuevas políticas como la Resolución de 6 de abril de 2021, de la Secretaría General de Universidades, por la que se aprueban recomendaciones en relación con los criterios y estándares de evaluación para la verificación, modificación, seguimiento y renovación de la acreditación de títulos universitarios oficiales de Grado y de Máster ofertados en modalidades de enseñanzas virtuales e híbridas; una resolución que establece los criterios y estándares de calidad para la continuidad y renovación de la docencia en la modalidad *e-Learning* e híbrida, y que sienta las bases de la nueva dirección que están tomando las instituciones de Enseñanza Superior que apuestan por el *e-Learning* en nuestro país (BOE, 2021).

Realizando una breve aproximación a los mencionados modelos, son algunos como el *Quality Matters* o el *e-Learning Quality* (ELQ), los que centran hoy este campo de la evaluación. El primero de ellos, el *Quality Matters*, es un modelo centrado en la evaluación de la calidad en Educación Superior que cuenta con ocho estándares generales que evalúan la descripción del curso (1), los objetivos (2), la evaluación (3), los materiales (4), la interacción (5), la tecnología empleada (6), el apoyo al estudiante (7) y la accesibilidad (8); y es hoy uno de los modelos más conocidos y empleados en las enseñanzas en línea (Brown et al., 2018).

Un segundo modelo representativo y asentado en el ámbito de la evaluación de la calidad en *e-Learning* es el modelo ELQ, desarrollado por la Agencia Nacional Sueca que comprende la evaluación de los materiales (1), la estructura (2), la comunicación y cooperación (3), la evaluación (4), la flexibilidad y adaptabilidad (5), el apoyo (6), la cualificación y experiencia docente (7), la visión y liderazgo (8), los recursos (9) y los procesos integrales (10) (Hansson et al., 2009; Mixnahuatl et al., 2013).

En los últimos años, los modelos privados de acreditación y certificación de calidad en *e-Learning* se han especializado y diseminado en modelos como el *ODLQC Standards* (evaluación de la calidad en aprendizajes abiertos o a distancia, incluyendo el *e-Learning*) o el *BAOL Quality Mark* (con el objetivo de promover la calidad en prácticas de aprendizaje flexible a distancia) (Fernández, 2005).

Por otro lado, la evaluación de la calidad en *e-Learning* toma también recientemente como referencia modelos provenientes del ámbito empresarial y de la evaluación de la calidad de los servicios ofrecidos, adaptándose a la calidad de los entornos virtuales y entendiendo a los usuarios de estos como clientes y destinatarios de un servicio que ha de ser evaluado y mejorado en la práctica, como los modelos *Servqual* y *Servperf* (Jain y Gupta, 2004).

En la actualidad, la coexistencia multitud de modelos para evaluar la calidad del *e-Learning* y aspectos concretos de sus procesos es una realidad, lo que pone de manifiesto la importancia y relevancia que tienen los avances en el campo de la educación virtual en los últimos años (Marciniak y Gairín-Sallán, 2017). Pese a esta multitud, la calidad no alcanza aún un modelo que logre cubrir las necesidades de un ámbito tan complejo, y refleja la necesidad de la creación de estándares y modelos de evaluación de calidad en la formación *on-line*.

Por ello, el elevado número de agencias de evaluación de la calidad y el creciente número de estudios que también relacionan la evaluación de la calidad y la mejora de la enseñanza no hacen más que subrayar la importancia de esta investigación en su afán por conocer el estado del arte sobre este tema. Además, la situación epidemiológica generada por la COVID-19 y la transición a la no presencialidad que ha supuesto en la totalidad de las enseñanzas y sistemas educativos (García-Peñalvo et al., 2020; Gómez-Gómez et al., 2022) ha posicionado al *e-Learning* como epicentro de los cambios y de la continuidad de estos procesos educativos (Gazca-Herrera, 2020). Dentro de esta adaptación, la calidad se ha posicionado como un factor relevante y una garantía de la nueva modalidad de enseñanza adoptada por la mayoría de las instituciones educativas, entre ellas las de Educación Superior.

Dada la importancia que el e-Learning y su calidad toman en el momento actual, en este estudio se presenta una revisión sistemática de la literatura (SLR) de las aportaciones sobre evaluación de la calidad en *e-Learning* en Educación Superior publicadas en los últimos siete años (2015-2021), pretendiendo llenar un vacío en la

investigación proporcionando una perspectiva global de este campo. En la presente investigación se profundizará en las dimensiones de la calidad, los procedimientos de evaluación y las aportaciones y nuevas vías de investigación que se sugieren para el campo.

A continuación, se presenta la metodología seguida en la revisión realizada, incluyendo las preguntas de investigación, los criterios de inclusión y exclusión, y la descripción completa del proceso de búsqueda y selección. Posteriormente, se presentan los resultados que responden a cada una de las preguntas de investigación y la discusión de los resultados a partir del proceso completo de revisión y las aportaciones de los autores. Esta revisión es ampliación y continuidad de un estudio de *mapping* realizado sobre el estado de la cuestión (Ortiz-López et al., 2021), del que solo se ha tomado como referencia el proceso de búsqueda documental.

2. Método

La revisión literaria debe resumir el estado de la cuestión de las investigaciones pertinentes que contextualizan el trabajo en el panorama internacional, y explicar qué conclusiones de otros autores, si los hubiere, están siendo cuestionadas o extendidas. Debe incluir la explicación general del estudio, su objetivo central y el diseño metodológico seguido. En la presente investigación se ha seguido la declaración PRISMA para revisiones sistemáticas de literatura (Liberati et al., 2009), la cual plantea los ítems y las directrices necesarias sobre aspectos clave a seguir en este tipo de investigaciones (desde la construcción del título, resumen, justificación y objetivos; pasando por el detalle del protocolo metodológico seguido o la selección de estudios para los resultados y la elaboración de discusiones) (González et al., 2011).

2.1. Preguntas de Investigación

Las preguntas planteadas para la revisión sistemática de la evaluación de la calidad en el e-Learning en la Educación Superior son las siguientes:

1. ¿En qué dimensión/dimensiones del *e-Learning* se centra la evaluación de la calidad en Educación Superior?
2. ¿Qué proceso se propone para evaluar la calidad en *e-Learning* en Educación Superior?
3. ¿Cuáles son las aportaciones principales que cada autor considera que plantea su investigación en materia de evaluación de la calidad en *e-Learning*?
4. ¿Qué nuevas vías se abren en la evaluación de la calidad en *e-Learning* para los autores?

2.2. Criterios de inclusión / exclusión

Se plantean a continuación los criterios de inclusión/exclusión que determinarán las publicaciones que pondrán el análisis y la respuesta a las preguntas de investigación planteadas:

1. No se incluye toda publicación que no sea artículo científico, y que no esté publicado en *Web of Science* (WoS) y Scopus.
2. Los artículos que no traten sobre evaluación de la calidad del eLearning en Educación Superior (entendiendo como tal las definiciones que se presentan al finalizar estos criterios) han sido excluidos de la revisión.
3. Los artículos publicados con anterioridad a los últimos siete años han sido excluidos (2015-2021).
4. Los artículos que se encuentran en lenguas distintas al español o el inglés han sido excluidos.
5. Los artículos que no sean estudios empíricos ni incluyan un proceso de validación de contenido por parte de jueces expertos
6. Los artículos que no traten del *e-Learning* en su totalidad (*b-Learning*, *m-Learning*, *u-Learning*, etc.) o que refieran a los cursos abiertos.

Para la puntualización de los criterios, se aceptará todo aquel artículo que se refiera al *e-Learning* como “una capacitación no presencial que, a través de plataformas tecnológicas, posibilita y flexibiliza el acceso y el tiempo en el proceso de enseñanza-aprendizaje” (García-Peñalvo, 2005, p. 2). Por tanto, será válido todo aquel artículo que refiera al *e-Learning* como enseñanza no presencial, flexible y que emplee herramientas de comunicación

tanto síncrona como asíncrona. Los MOOC y la educación abierta no están incluidos, ya que no se puede garantizar que tengan como destinatarios exclusivamente población de Educación Superior.

Asimismo, se considerarán válidos solamente aquellos artículos que conciban la calidad como el grado en que un curso, programa o enseñanza cumple unos requisitos establecidos y que implique una correcta y profunda planificación, evaluación y mejora en los procesos de docencia e investigación y en las áreas de gestión y administración (Fernández et al., 1999). Adicionalmente, la calidad deberá garantizar también la optimización y maximización del proceso de enseñanza-aprendizaje en cuanto a los conocimientos y habilidades adquiridas por los usuarios/destinatarios del *e-Learning* y la forma en la que estos se transmiten.

2.3. Cadena de búsqueda y base de datos

Inicialmente, la búsqueda se llevó a cabo el 26/11/2019 en las bases de datos *Web of Science* y *Scopus*, dos de las que presentan mayor relevancia en el ámbito de la investigación científica (Pérez-Escoda, 2017), limitándose los resultados a los últimos cinco años, de manera que el estudio ofreciera información centrada en la situación actual del campo. La cadena de búsqueda empleada fue: ("Evaluación Calidad" OR "Quality Assessment" OR "Quality Evaluation") AND ("eLearning" OR "e-Learning") AND ("Educación Superior" OR "Universidad" OR "Higher Education" OR "University").

Posteriormente, para mantener el estudio actualizado se repitió la búsqueda utilizando la misma cadena en los mismos repositorios el 28/05/2021, empleando la herramienta de filtro de los repositorios para limitar los resultados al periodo de tiempo transcurrido desde la búsqueda original. Adicionalmente también se utilizaron las herramientas de filtro para limitar los resultados en función del tipo de publicación de manera que solamente se incluyeran los artículos publicados en revistas científicas. La base de datos con los artículos, el proceso completo y las revisiones realizadas se encuentra disponible para su acceso libre en el enlace: <https://bit.ly/3kFz7lM>.

2.4. Proceso de selección de publicaciones

La búsqueda inicial completa arrojó un total de 753 artículos que conformaron la primera fase de la Investigación. Tras la eliminación de duplicados, la cifra descendió a 715. En la segunda fase de la investigación se aplicaron los criterios de exclusión ya referidos en este trabajo. El resultado tras la primera aplicación de criterios fue de 52 artículos aceptados y 27 propuestos para revisión posterior (79 artículos). A continuación, los artículos marcados para revisión y los ya aceptados fueron también solicitados para ser revisados y discutidos con un profesor e investigador experto en la materia y en el desarrollo de estudios de tipo *mapping*/SLR.

Además, con el claro objetivo de garantizar la validez, la objetividad y mejorar la fiabilidad de la investigación, los autores de la investigación solicitaron a dos jueces expertos en evaluación y en calidad en Educación Superior la revisión de cincuenta artículos aleatorios para comprobar la validez de la aplicación de los criterios hasta la fecha. Los jueces determinaron que el 100% de los artículos excluidos/incluidos revisados estaban correctamente dictaminados, lo que permite que esta investigación continúe con un mayor grado de validez.

En relación con el número de publicaciones que promocionaron en cada fase, el número final de aceptadas fue de 25 (con una tasa de aceptación del 3,49%), y un total de 690 artículos fueron rechazados. Las cifras restantes se presentan a continuación (Figura 1).

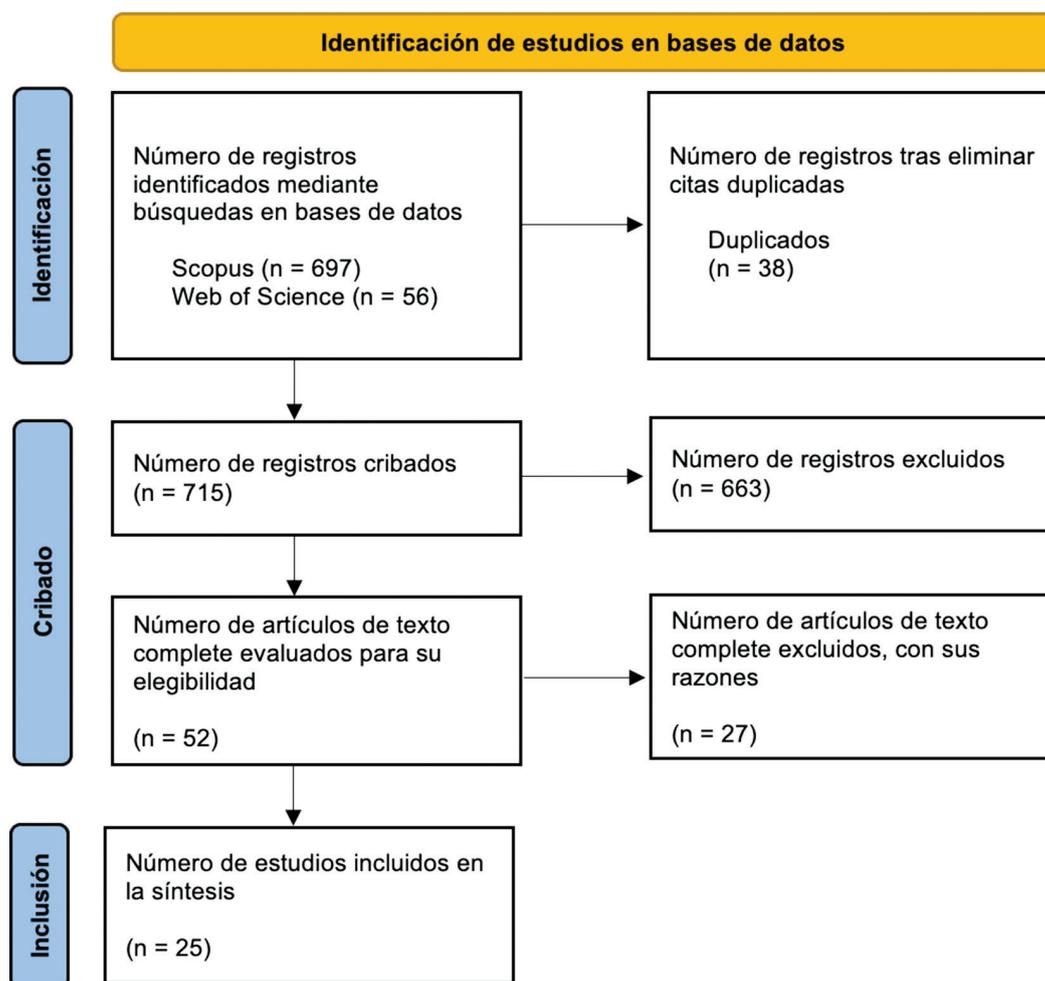
3. Resultados

3.1. ¿En qué dimensión/dimensiones del *e-Learning* se centra la evaluación de la calidad en Educación Superior?

La primera pregunta hace referencia directa a conocer en qué dimensión o dimensiones del *e-Learning* se centra la evaluación de la calidad para los investigadores. A la hora de abordar esta temática, existe una gran heterogeneidad de dimensiones, entre las que se incluyen (Marciniak y Gairín-Sallán, 2018):

- Calidad del contexto.
- Calidad de la formación.

Figura 1. Diagrama de flujo PRISMA de la revisión.



- Características del estudiante.
- Características de la función docente.
- Calidad los apoyos y las entradas.
- Calidad de la infraestructura tecnológica.
- Calidad del diseño y desarrollo del curso.
- Accesibilidad y gestión del entorno.
- Organización y ética.
- *Feedback*.
- Calidad de la información del programa.
- Pedagogía y estrategias de enseñanza.
- Calidad de los objetivos, los materiales, recursos y las actividades.
- Calidad de la evaluación.

Dada esta situación, y de cara a realizar la presente revisión, se optó por no establecer un sistema de dimensiones previo, sino que estas fueron extraídas a partir de los artículos incluidos, de manera que el proceso permitiera conocer qué dimensiones presentan una mayor relevancia y cuales son aquellas que están aún por evaluar en función de los criterios de los autores de las publicaciones. En primer lugar, los resultados observados demuestran que son muy relativas las dimensiones que cada investigación plantea (Tabla 1). La tabla extendida con las dimensiones planteadas por cada autor se encuentra disponible en el archivo complementario a esta revisión.

Tabla 1. Dimensiones de la evaluación de la calidad en e-Learning.

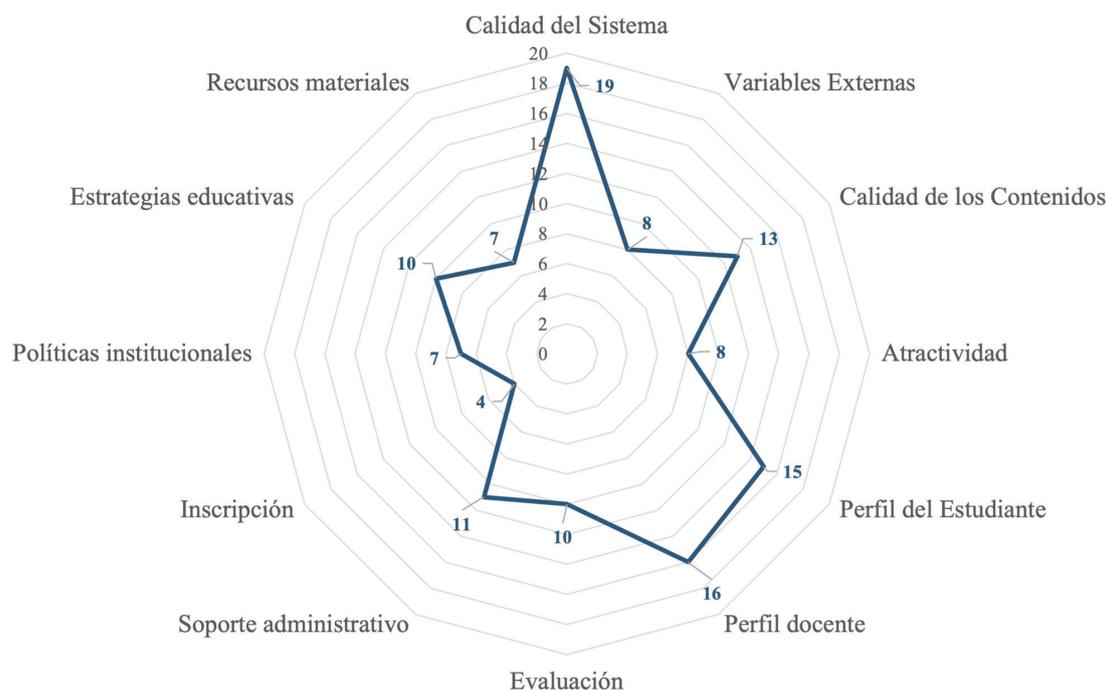
Dimensión	Descripción	Autores
Calidad del sistema	Relativo al sistema o plataforma de aprendizaje (LMS) sobre el que se realiza la evaluación de la calidad.	Alam et al. (2021), Choi y Jeong (2019), Kurilovas y Vinogradova (2016), La Rotta et al. (2019), Ling y Pei (2020), Marciniak (2015, 2018), Margalina et al. (2015), Martínez-Arguelles y Batalla-Busquets (2016), Martínez-Caro et al. (2015), Ortiz-López et al. (2021), Park y Lim (2015), Pham et al. (2019), Pontoriero (2021), Prieto-Rodríguez et al. (2016), Radu et al. (2020), Serrano et al. (2018), Waheed & Kaur (2017), Waheed et al. (2016).
Perfil docente	Relativo a las características del docente y al desarrollo de sus procesos evaluativos.	Alam et al. (2021), García y González (2015), Kurilovas y Vinogradova (2016), La Rotta et al. (2019), Marciniak (2015, 2018), Margalina et al. (2015), Martínez-Arguelles y Batalla-Busquets (2016), Martínez-Caro et al. (2015), Ortiz-López et al. (2021), Park y Lim (2015), Pham et al. (2019), Pontoriero (2021), Prieto-Rodríguez et al. (2016), Serrano et al. (2018), Torres-Barbazal et al. (2018).
Perfil del estudiante	Relativo a las características y personalización que se alcanza en el estudiante.	Kurilovas y Vinogradova (2016), La Rotta et al. (2019), Ling y Pei (2020), Marciniak (2015, 2018), Margalina et al. (2015), Martínez-Caro et al. (2015), Park y Lim (2015), Pecori et al. (2019), Pham et al. (2019), Pontoriero (2021), Prieto-Rodríguez et al. (2016), Raspopovic y Jankulovic (2017), Torres-Barbazal et al. (2018), Waheed et al. (2016).
Calidad de los contenidos	Relativo a las características de los contenidos, su acceso, su presentación, formato...	Alam et al. (2021), Kurilovas y Vinogradova (2016), Ling y Pei (2020), Marciniak (2015, 2018), Martínez-Argüelles y Batalla-Busquets (2016), Park y Lim (2015), Pham et al. (2019), Pontoriero (2021), Radu et al. (2020), Torres-Barbazal et al. (2018), Waheed et al. (2016), Waheed y Kaur (2017).
Soporte administrativo	Relativo a la atención que desde la Institución se facilita a dar soporte administrativo a la enseñanza.	La Rotta et al. (2019), Ling y Pei (2020), Mahdiuon et al. (2017), Marciniak (2018), Margalina et al. (2015), Martínez-Argüelles y Batalla-Busquets (2016), Ortiz-López et al. (2021), Pham et al. (2019), Prieto-Rodríguez et al. (2016), Radu et al. (2020), Waheed & Kaur (2017).
Estrategias educativas	Relativo a la evaluación de la calidad de las metodologías y procesos educativos seguidos.	Alam et al. (2021), Ling y Pei (2020), Mahdiuon et al. (2017), Marciniak (2018), Marciniak (2015), Martínez-Caro et al. (2015), Pontoriero (2021), Radu et al. (2020), Serrano et al. (2018), Torres-Barbazal et al. (2018).
Evaluación	Relativo a la evaluación de la calidad de los procesos evaluativos seguidos	García y González (2015), La Rotta et al. (2019), Marciniak (2018), Ortiz-López et al. (2021), Park y Lim (2015), Pontoriero (2021), Radu et al. (2020), Raspopovic y Jankulovic (2017), Torres-Barbazal et al. (2018), Waheed y Kaur (2017).
VARIABLES EXTERNAS	Relativo a la atención, evaluación y control de las variables externas del proceso.	Choi y Jeong (2019), García y González (2015), Mahdiuon et al. (2017), Marciniak (2018), Radu et al. (2020), Torres-Barbazal et al. (2018), Waheed et al. (2016), Waheed y Kaur (2017).
Atractividad y diseño	Relativo a la evaluación del diseño del curso, plataforma, etc.	Choi y Jeong (2019), Marciniak (2018), Ortiz-López et al. (2021), Pham et al. (2019), Pontoriero (2021), Radu et al. (2020), Waheed et al. (2016), Waheed y Kaur (2017).
Políticas institucionales	Relativo a la evaluación de las políticas institucionales en materia de calidad en e-Learning.	Alam et al. (2021), Ling y Pei (2020), Marciniak (2015), Ortiz-López et al. (2021), Pontoriero (2021), Serrano et al. (2018), Waheed et al. (2016).
Recursos materiales	Relativo a la evaluación del número, calidad y contenidos de los recursos materiales facilitados.	Ling y Pei (2020), Marciniak (2018), Ortiz-López et al. (2021), Pham et al. (2019), Pontoriero (2021), Prieto-Rodríguez et al. (2016), Raspopovic y Jankulovic (2017).
Inscripción y acceso	Relativo a la calidad de los procesos de inscripción y acceso.	La Rotta et al. (2019), Marciniak (2018), Park y Lim (2015), Pecori et al. (2019).

Fuente: elaboración propia.

La mayoría de las investigaciones han tratado de conocer las dimensiones relevantes para la evaluación. La profundización y categorización de estas es muy variada, ya que estudios como el de Marciniak (2018) plantea 14 dimensiones, otros como el de Ling y Pei (2020) plantean seis y algunos como el de Prieto-Rodríguez et al. (2016); solamente tres. Por tanto, dicha variabilidad en el número y en su delimitación refleja estar ante un campo muy difícilmente abarcable y heterogéneo, compuesto por un gran número de dimensiones e indicadores que influyen y condicionan la evaluación de la calidad en este ámbito.

Pese a la heterogeneidad encontrada, sí se detecta la presencia de un número de dimensiones que se repite con mayor frecuencia en los estudios, mientras que otras se encuentran aún menos abordadas (Figura 2).

Figura 2. Agrupación de dimensiones en función del número de autores que las abordan.



Con relación a las dimensiones más estudiadas, teniendo en cuenta las planteadas por Marciniak y Gairín-Sallán (2018) en su estudio, sí se observa una cierta correlación entre las dos revisiones, alcanzando hasta nueve dimensiones presentes en ambos. Pese a ello, se detecta también una alta variabilidad en cuanto a la delimitación de las dimensiones en cada estudio y la presencia también de diferentes subdimensiones; por lo que se vuelve a poner de manifiesto la amplitud de este campo evaluativo.

3.2. ¿Qué proceso se propone para evaluar la calidad en e-Learning en Educación Superior?

El proceso mediante el que se evalúa la calidad difiere en función del estudio, de las dimensiones contempladas y de la metodología propuesta. Conocer cómo los autores llevan a cabo su evaluación es fundamental para conocer los procedimientos que están siendo seguidos en este campo. Por ello, con el fin de categorizar el proceso realizado en cada estudio, se han establecido tres criterios para determinar si las publicaciones analizadas generan un modelo propio, explícito y complejo para la evaluación de la calidad (en consecuencia, amplían el campo); si aplican un modelo ya existente y evalúan la calidad en un contexto determinado, o si por el contrario realizan una evaluación de la calidad sin la formulación explícita y compleja de un modelo. Dichos criterios son:

1. Si la publicación hace mención explícita a un modelo. Es decir, lo define modelo y define también sus dimensiones y:
 - a. Si la publicación refiere la aplicación o validación de un modelo propio, generado por los autores y tiene dos o más dimensiones (complejo): Generación de un nuevo modelo.

- b. Si la publicación refiere a un modelo ya existente y a su aplicación en un contexto determinado: Aplicación de un modelo existente.
- 2. Si los autores plantean una definición y unas características para el desarrollo de un marco evaluativo, sin presentar en ningún momento dimensiones ni indicadores para la categorización como modelo: Marco evaluativo.
- 3. Si los autores evalúan la calidad sin la formulación explícita (definido y dimensionalizado) y compleja de un modelo propio (dos o más dimensiones): Evaluación de una dimensión de la calidad sin la formulación explícita y compleja de un modelo.

Por tanto, siguiendo la clasificación propuesta para la categorización de los procesos evaluativos, se presenta la siguiente tipificación (Tabla 2).

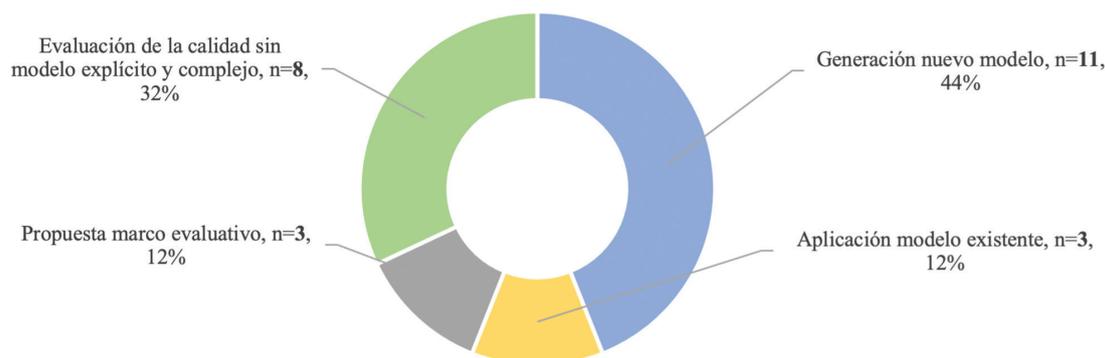
Tabla 2. Proceso seguido para la evaluación de la calidad en e-Learning.

Proceso seguido	Autores
Generación de un modelo.	Alam et al. (2021), Choi y Jeong (2019), Hou (2021), Kurilovas y Vinogradova (2016), Ling y Pei (2020), Marciniak (2018), Margalina et al. (2015), Ortiz-López et al. (2021), Pontoriero (2021), Waheed et al. (2016), Waheed y Kaur (2017).
Aplicación de un modelo ya existente.	Martínez-Argüelles y Batalla-Busquets (2016), Martínez-Caro et al. (2015), Prieto-Rodríguez et al. (2016).
Propuesta de un marco evaluativo.	Marciniak (2015), Pecori et al. (2019), Serrano et al. (2018).
Evaluación de la calidad sin la formulación de un modelo explícito y complejo.	García y González (2015), La Rotta et al. (2019), Mahdiun et al. (2017), Park y Lim (2015), Pham et al. (2019), Radu et al. (2020), Raspopovic & Jankulovic (2017), Torres-Barbazal et al. (2018).

Fuente: elaboración propia.

Del análisis se extrae que son varias corrientes las que los investigadores siguen en sus estudios. Por un lado, destacan aquellos autores que plantean la generación de un modelo propio para evaluar la calidad (11 de los 25). En segundo lugar, un total de ocho autores evalúan la calidad sin un modelo único y complejo (es decir, o evalúan sin modelo definido o evalúan una única dimensión). Y, en tercer lugar, son tres los autores que aplican un modelo ya existente y tres autores también los que proponen un marco de evaluación para la calidad en e-Learning en Educación Superior (proponiendo condiciones, no dimensiones ni referencias explícitas a un modelo único y complejo) (Figura 3).

Figura 3. Distribución de los procesos seguidos por los autores con relación al total.



Entre aquellas publicaciones que proponen su modelo, se comprueba la variabilidad y las diferencias entre unos modelos y otros (no abordan las mismas dimensiones y no lo hacen de la misma forma), y se detecta también que el proceso seguido para cada propuesta difiere entre las once publicaciones (unos parten del análisis teórico, otros de reformulaciones de modelos previos y otros de estudios empíricos sobre la calidad), por lo que no se detecta una única dirección en cuanto a los nuevos modelos evaluativos.

Por otro lado, las ocho publicaciones que evalúan una única dimensión de la calidad, o que lo hacen sin un modelo, plantean estudios centrados en la satisfacción del usuario (García y González, 2015; La Rotta et al., 2019; Pham et al., 2019; Raspopovic y Jankulovic, 2017; Torres-Barbazal et al., 2018), en la calidad docente (García y González, 2015), en la percepción del proceso de enseñanza e-Learning (Mahdiuon et al., 2017; Park y Lim, 2015) o en el impacto de la COVID en la calidad de la enseñanza (Radu et al., 2020).

Entre aquellos estudios que aplican un modelo ya existente, lo hacen mediante la validación del cuestionario *e-Servqual* para la calidad (Martínez-Argüelles y Batalla-Busquets, 2016), mediante la aplicación de *Servqual*, *Serperf* e *IPA* (Martínez-Caro et al., 2015) o empleando el *Quality Teaching Model* (Prieto-Rodríguez et al., 2016).

Finalmente, son tres autores también quienes proponen un marco evaluativo para la evaluación de la calidad en *e-Learning*, estableciendo una serie de condiciones y características de la evaluación a realizar. De estos, unos se centran más en un marco comparativo de protocolos en *e-Learning* (Marciniak, 2015), otros en marcos para el análisis de datos mediante *Big Data* (Pecori et al., 2019) y otros centrados en las condiciones institucionales (Serrano et al., 2018).

3.3. ¿Cuáles son las aportaciones principales que cada autor considera que plantea su investigación en materia de evaluación de la calidad en e-Learning?

Se abordan las aportaciones que los autores consideran que su investigación alcanza. Estas aportaciones centran su desarrollo, principalmente, en las aportaciones del modelo, marco o proceso evaluativo diseñado. Por ello, se presentan las principales categorías sobre lo que consideran que aporta su investigación para la evaluación de la calidad (Tabla 3). Estas categorías han sido extraídas de cada publicación, y se encuentran ampliadas en el archivo complementario a la investigación (en la hoja 3.3).

Tabla 3. Aportaciones de las investigaciones analizadas.

Conclusión	Nº	Autores
El proceso seguido o planteado profundiza y amplía dimensiones específicas del <i>e-Learning</i> .	12	Alam et al. (2021), La Rotta et al. (2019), Marciniak (2018), Margalina et al. (2015), Martínez-Argüelles y Batalla-Busquets (2016), Martínez-Caro et al. (2015), Park y Lim (2015), Pham et al. (2019), Pontoriero (2021), Serrano et al. (2018), Waheed et al. (2016), Waheed y Kaur (2017).
Se plantea un modelo o marco fiable y eficaz para la evaluación.	6	Alam et al. (2021), Kurilovas y Vinogradova (2016), La Rotta et al. (2019), Mahdiuon et al. (2017), Ortiz-López et al. (2021).
El modelo que se propone es único hasta el momento dadas sus características.	5	García y González (2015), Hou (2021), Mahdiuon et al. (2017), Pecori et al. (2019), Prieto-Rodríguez et al. (2016), Torres-Barbazal et al. (2018).
Se presenta un mecanismo nuevo que permite comparar sistemas <i>e-Learning</i> .	3	Choi y Jeong (2019), Marciniak (2015), Raspopovic y Jankulovic (2017).
El modelo propuesto unifica dimensiones de otros modelos.	1	Ortiz-López et al. (2021).

Fuente: elaboración propia.

Las publicaciones realizadas hasta la fecha se centran, en su mayoría, en continuar el estudio de las dimensiones, generando nuevos modelos evaluativos y ampliando el campo a estudio. Esta tendencia refleja que aún se está lejos de alcanzar un único modelo en la evaluación de la calidad, ya que la exploración en los mismos varía en función del autor, del proceso y de las dimensiones; teniendo 12 propuestas que destacan por profundizar en dimensiones particulares. Además, seis autores indican que su herramienta mejora la fiabilidad y eficacia de su evaluación frente a las evaluaciones ya existentes, e incluso cinco de ellos la plantean como una herramienta única hasta la fecha dadas sus características.

Por otro lado, tres autores destacan la posibilidad de sus modelos para comparar sistemas de enseñanza, y solamente uno de ellos plantea la necesidad de unificar dimensiones de modelos vigentes con el objetivo de tratar de establecer un marco común, consensuado y validado por la comunidad científica, dada la elevada diseminación actual de los modelos.

3.4. ¿Qué nuevas vías se abren en la evaluación de la calidad en línea para los autores?

Por último, se abordan las nuevas vías de investigación que los autores plantean tras su estudio. Estas nuevas vías han de centrar el devenir futuro de las investigaciones, y son los campos hacia los que dicha evaluación debe avanzar para consolidarse como una garantía y un campo relevante en el ámbito del *e-Learning* (Tabla 4). En el archivo auxiliar se encuentran las líneas de investigación completas planteadas en cada estudio (3.4).

Tabla 4. Líneas de investigación propuestas en los estudios analizados.

Línea Propuesta	Nº	Autores
Identificar dimensiones estables para otros contextos y Universidades.	7	Marciniak (2018), Martínez-Argüelles y Batalla-Busquets (2016), Martínez-Caro et al. (2015), Park y Lim (2015), Pham et al. (2019), Waheed y Kaur (2017), Ling y Pei (2020).
Incluir más usuarios en la evaluación.	6	Alam et al. (2021), Choi y Jeong (2019), García y González (2015), Radu et al. (2020), Waheed y Kaur (2017), Ling y Pei (2020).
Examinar más variables y factores externos.	6	Alam et al. (2021), Park y Lim (2015), Pham et al. (2019), Radu et al. (2020), Raspopovic y Jankulovic (2017), Ling y Pei (2020).
Explicar más porcentaje de la calidad.	4	Alam et al. (2021), Choi y Jeong (2019), Park y Lim (2015), Waheed et al. (2016).
Validar el instrumento propuesto.	4	La Rotta et al. (2019), Ortiz-López et al. (2021), Radu et al. (2020), Waheed et al. (2016).
Ampliar contexto geográfico.	2	Alam et al. (2021), García y González (2015).
Ampliar variables relacionadas estudiante.	2	Alam et al. (2021), Mahdiun et al. (2017).
Extender el instrumento propuesto / diseñado.	2	Marciniak (2015), Serrano et al. (2018).
Avanzar en la investigación de las características del sistema educativo.	1	Mahdiun et al. (2017).
Analizar impacto organizativo.	1	Margalina et al. (2015).
Aplicar a otras modalidades de enseñanza (<i>b-Learning</i> , <i>u-Learning</i> , <i>m-Learning</i>).	1	Martínez-Caro et al. (2015).

Fuente: elaboración propia.

Parece claro que las principales corrientes abogan por identificar y aunar dimensiones, tratando así de extrapolar y validar el modelo que cada autor genera en otros marcos universitarios, estableciendo un marco amplio y consensado. Además, se apuesta por la necesidad de incluir un mayor número de usuarios en los procesos evaluativos y por examinar los factores externos que influyen y condicionan la calidad del *e-Learning*, pudiendo establecer así un listado completo de factores condicionantes que forman parte también de esta evaluación.

Destacan también corrientes que defienden la necesidad de analizar no solo la calidad, si no la vinculación de esta con el impacto que una mejora en ella supone para la institución de Educación Superior y; por último, corrientes que plantean la idea de aplicar los modelos de evaluación de la calidad de *e-Learning* a otras modalidades, como pueden ser el *b-Learning*, el *u-Learning* o el *m-Learning*, atendiendo a cada una de sus peculiaridades.

4. Conclusiones

La revisión realizada muestra el interés existente en la comunidad investigadora por la evaluación de la calidad en *e-Learning* en Educación Superior, la cual centra sus esfuerzos y avances en la profundización sobre las dimensiones que componen y determinan la calidad, sin denotarse aún la generación de un modelo único para dicho proceso, lo que condiciona la evaluación para las instituciones de Educación Superior.

4.1. Conclusiones sobre las dimensiones de la calidad

La revisión ha identificado un total de doce dimensiones diferenciadas que afectan a la calidad y que se proponen en los estudios analizados. De ellas, destaca la variabilidad en cuanto a la presencia en unos u otros estudios, dónde se priorizan, alternan y eliminan en función de los intereses y hallazgos particulares. De ellas, predominan en la revisión la calidad del sistema, la función docente, la calidad de acceso para el usuario y la calidad de los contenidos (Figura 2). Aún así, no se detecta una relación directa que justifique si una mayor presencia conlleva una mayor explicación de la calidad; y más dada la variabilidad en cuanto a sus múltiples delimitaciones, pesos e indicadores en función de cada estudio.

En una comparativa con la revisión de Marciniak y Gairín-Sallán (2018) de modelos evaluativos en educación virtual, se constata la relación directa de nueve dimensiones en ambas revisiones, mientras que cinco de ellas no se correlacionan; lo que refleja nuevamente la amplitud y dificultad en la definición del campo (Fernández, 2005). Esta amplitud se observa también en la comparación con la revisión de Duarte et al. (2017) sobre las dimensiones de la calidad en MOOCs, quien plantea que para esta modalidad son la planificación, los contenidos, la metodología, los recursos, la motivación, la comunicación y la evaluación las dimensiones más representativas; sin alcanzar un alto grado de similitud a las dimensiones aquí alcanzadas para la calidad en *e-Learning*.

Por tanto, la dimensionalización de la calidad en este campo ha de ser un objetivo principal para la investigación, la cual se encuentra lejos aún de la estabilidad necesaria para un modelo común.

4.2. Conclusiones sobre los procesos seguidos en los estudios

Esta revisión confirma el elevado número de propuestas particulares y diferenciadas en la apuesta por la evaluación de la calidad, con un total de once nuevos modelos entre los veinticinco estudios analizados. Además, otros tres estudios proponen un nuevo marco evaluativo, y ocho más evalúan la calidad sin generar un modelo mediante un proceso propio. Es importante resaltar que solamente un estudio trata de aglutinar las ya existentes en la generación de un modelo común y consensuado (partiendo de las dimensiones más relevantes y de aquellas más influyentes en la calidad para sus autores).

Dado el elevado volumen de modelos coexistentes (Mejía y López, 2016), consideramos necesario que investigaciones futuras aglutinen las propuestas y se centren en la generación de un modelo validado, aceptado y consensuado (Park y Lim, 2015); permitiendo así la evaluación de calidad en e-Learning en Educación Superior a gran escala y posibilitando la comparación interuniversitaria y la generación de nuevos espacios de colaboración conjunta.

Por otro lado, resaltar también que aquellos modelos de evaluación de la calidad citados en el comienzo de la publicación (ELQ, ODLQC, BAOL; entre otros), y relevantes en el desarrollo de la evaluación de la calidad, no han sido detectados en las investigaciones analizadas, por lo que se abre una nueva y posible discusión sobre la practicidad de estos modelos frente a modelos individuales y propuestas de autores; así como su aplicación en ámbitos de calidad en e-Learning en Educación Superior.

4.3. Conclusiones sobre las aportaciones de los estudios

En cuanto a las aportaciones analizadas, prácticamente la totalidad de las investigaciones valoran su estudio por profundizar en dimensiones específicas que afectan a la calidad del *e-Learning*, y por presentar también un modelo novedoso, fiable y eficaz para dicha evaluación. De los estudios, solamente una aportación trata la necesidad de unificar dimensiones, mientras que el resto presentan nuevos modelos y marcos, o aplicaciones de modelos ya existentes para avanzar en el campo de la evaluación.

Se concluye, por tanto, que la investigación avanza en la exploración de dimensiones como fase previa al establecimiento de un acuerdo entre las mismas para la generación de modelos más relacionados. En el análisis de estos se ha de tener en cuenta también el posible sesgo de análisis de datos (destacando el sesgo de autoconfirmación y el sesgo de publicación) los cuales pueden manifestarse en las conclusiones alcanzadas por los autores sobre sus propuestas (Palma & Delgado, 2006), dada la elevada eficacia y validez en todos los modelos analizados.

4.4. Conclusiones sobre las líneas de investigación futuras

La revisión ha tratado de plasmar la situación actual del campo, teniendo en cuenta también las líneas futuras propuestas por los estudios. De ellas, se extrae como principal objetivo la necesaria continuidad en la identificación de dimensiones estables para todos los contextos y la necesidad de ampliar la evaluación a un mayor número de usuarios y entidades.

Tras el estudio, consideramos que futuras investigaciones deben explorar aquellas dimensiones menos estudiadas, como los procedimientos de inscripción, el acceso a plataformas, las políticas institucionales y la calidad de los recursos (Figura 2). El establecimiento de un modelo evaluativo para la calidad del *e-Learning* en Educación Superior debe pasar por la identificación de dimensiones estables para todos los contextos (mediante validaciones interuniversitarias), analizando el máximo número de variables posibles que en dichos contextos influyan y haciendo partícipes a más usuarios de estos procesos evaluativos. Además, consideramos necesario también el trabajo coordinado interuniversitario para la validación de dichos modelos, y el metaanálisis de los modelos ya existentes para la generación de un modelo común partiendo de los mismos.

4.5. Consideraciones finales

Esta revisión ha tratado de aunar diferentes propuestas para centrar los avances que se han de seguir en materia de calidad en *e-Learning*, y ha denotado la falta de un modelo común en la práctica educativa de la Educación Superior, por lo que se abre una fuerte vía en la investigación para alcanzar la validación futura de un modelo único y el establecimiento de dimensiones ampliamente consensuadas para evaluar la calidad de forma válida y fiable, haciendo así que el *e-Learning* sea una realidad educativa con un alto nivel de calidad en la nueva realidad educativa que se abre en nuestros días.

Entre las limitaciones de esta investigación, destacan la imposibilidad de abarcar un mayor número de bases de datos dados los recursos de la investigación (García-Peñalvo, 2022) y el marco temporal comprendido, sobre el que estudios futuros pueden plantear su ampliación para analizar las tendencias y la evolución en el estudio de la calidad en *e-Learning*; el cual no era el objetivo de esta revisión, que ha tratado de reflejar la situación actual del campo en la Educación Superior para posicionar la calidad como un factor fundamental en las enseñanzas en red.

Referencias

- Alam, M. M., Ahmad, N., Naveed, Q. N., Patel, A., Abohashrh, M. y Khaleel, M. A. (2021). E-Learning Services to Achieve Sustainable Learning and Academic Performance: An Empirical Study. *Sustainability*, 13(5), 2653. <https://doi.org/10.3390/su13052653>
- BOE. (2021). Resolución de 6 de abril de 2021, de la Secretaría General de Universidades, por la que se aprueban recomendaciones en relación con los criterios y estándares de evaluación para la verificación, modificación, seguimiento y renovación de la acreditación de títulos universitarios oficiales de Grado y de Máster ofertados en modalidades de enseñanzas virtuales e híbridas. *Boletín Oficial del Estado*, 90, sec. III, de 15 de abril de 2021, 43455-43467.
- Brown, V. S., Toussaint, M. y Lewis, D. (2018). Students' Perceptions of Quality across Four Course Development Models. *Online Learning*, 22(2), 173-195. <https://doi.org/10.24059/olj.v22i2.1213>
- Choi, C.-R. y Jeong, H.-Y. (2019). Quality evaluation for multimedia contents of e-learning systems using the ANP approach on high-speed network. *Multimedia Tools and Applications*, 78(4), 28853-28875. <https://doi.org/10.1007/s11042-019-7351-8>
- Duart, J. M., Roig Vila, S., Mengual Andrés, S. y Maseda Durán, M. Á. (2017). La calidad pedagógica de los MOOC a partir de la revisión sistemática de las publicaciones JCR y Scopus (2013-2015). *Revista española de pedagogía*, 22. <https://doi.org/10.22550/REP75-1-2017-02>
- Fernández, B., Alonso, E., Reboloso, E. y Pozo, C. (1999). Evaluación de la calidad en la educación superior. *Papeles del psicólogo*, 74, 1886-1415.
- Fernández, V. (2005). Calidad y e-learning. En F. F. Martínez (Coord). En *E-aprendizaje en bibliotecología: Perspectivas globales* (pp. 74-91). Universidad Nacional Autónoma de México.
- García, P. A. y González, M. S. (2015). Validación de escala para evaluación de la calidad docente en entornos virtuales. *Opción: Revista de Ciencias Humanas y Sociales*, 31(5), 394-406.

- García-Peñalvo, F. J. (2005). Estado actual de los sistemas e-learning. *Education in the knowledge society*, 6(2), 1. <https://doi.org/10.14201/eks.18184>
- García-Peñalvo, F. J. (2022). Developing robust state-of-the-art reports: Systematic Literature Reviews. *Education in the Knowledge Society*, 23, Article e28600. <https://doi.org/10.14201/eks.28600>
- García-Peñalvo, F. J., Corell, A., Abella-García, V. y Grande-de-Prado, M. (2020). Online Assessment in Higher Education in the Time of COVID-19. *Education in the Knowledge Society*, 21, Article 12. <https://doi.org/10.14201/eks.23013>
- Gazca-Herrera, L. A. (2020). Implicaciones del coronavirus covid-19 en los procesos de enseñanza en la educación superior. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 11(21). <https://doi.org/10.23913/ride.v11i21.753>
- Gómez-Gómez, M., Hijón-Neira, R., Santacruz-Valencia, L. y Pérez-Marín, D. (2022). Impact of the Emergency Remote Teaching and Learning Process on Digital Competence and Mood in Teacher Training. *Education in the Knowledge Society*, 23, Article e27037. <https://doi.org/10.14201/eks.27037>
- González, J., Buñuel, J. C. y Aparicio, M. (2011). Listas guía de comprobación de revisiones sistemáticas y metaanálisis: Declaración PRISMA. *Evidencias en Pediatría*, 7(4), 97-102.
- Guerrero, C. (2018). Innovación y Transformación Digital en la Educación Superior. En C. Cáceres, N. Esteban, M. d. C. Gálvez y B. Rivas (Eds.). En *Innovación y transformación digital: Estrategias y metodologías docentes en Educación Superior* (pp. 19-26). Dykinson.
- Hansson, H., Westman, P., Åström, E. y Johansson, M. (2009). Un modelo para la evaluación de la calidad e-Learning. *Revista de Investigaciones UNAD*, 8(1), 11-29. <https://doi.org/10.22490/25391887.619>
- Hou, J. (2021). Online teaching quality evaluation model based on support vector machine and decision tree. *Journal of Intelligent & Fuzzy Systems*, 40, 2193-2203. <https://doi.org/10.3233/JIFS-189218>
- Jain, S. K. y Gupta, G. (2004). Measuring Service Quality: Servqual vs. Servperf Scales. *Vikalpa*, 29(2). <https://doi.org/10.1177/0256090920040203>
- Kurilovas, E. y Vinogradova, I. (2016). Improved fuzzy AHP methodology for evaluating quality of distance learning courses. *The International journal of engineering education*, 32(4), 1618-1624.
- La Rotta, D., Usuga, O. C. y Clavijo, V. (2019). Perceived service quality factors in online higher education. *Learning Environments Research*, 23, 251-267. <https://doi.org/10.1007/s10984-019-09299-6>
- Liberati, A., Altman, D., Tetzlaff, J., Mulrow, C., Gotzsche, P., Ioannidis, J. P. y Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *British Medical Journal*, 339. <https://doi.org/10.1136/bmj.b2700>
- Ling, J. H. y Pei, R. Y. (2020). Perceptions of learning management system quality, satisfaction, and usage: Differences among students of the arts. *Australasian Journal of Educational Technology*, 36(3), 26-40. <https://doi.org/10.14742/ajet.5187>
- Mahdiuon, R., Masoumi, D. y Farasatkah, M. (2017). Quality improvement in virtual higher education: A grounded theory approach. *Turkish Online Journal of Distance Education*, 18(1), 111-131. <https://doi.org/10.17718/tojde.285720>
- Marciniak, R. (2015). Methodological proposal for the application of international benchmarking in order to assess the quality of virtual higher education. *RUSC. Universities and Knowledge Society Journal*, 12(3), 46-60. <https://doi.org/10.7238/rusc.v12i3.2163>
- Marciniak, R. (2018). Quality Assurance for Online Higher Education Programmes: Design and Validation of an Integrative Assessment Model Applicable to Spanish Universities. *The International Review of Research in Open and Distributed Learning*, 19(2), 126-154. <https://doi.org/10.19173/irrodl.v19i2.3443>
- Marciniak, R. y Gairín-Sallán, J. (2017). Un modelo para la autoevaluación de la calidad de programas de educación universitaria virtual. *Revista de Educación a Distancia (RED)*, 54, Article 54. <https://doi.org/10.5944/ried.21.1.16182>
- Marciniak, R. y Gairín-Sallán, J. (2018). Dimensiones de evaluación de calidad de educación virtual: Revisión de modelos referentes. *RIED. Revista Iberoamericana de Educación a Distancia*, 21(1), 217-238. <https://doi.org/10.5944/ried.21.1.16182>
- Margalina, V., De Pablos-Heredero, C. y Montes, J. L. (2015). Achieving quality in e-Learning through relational coordination. *Studies in Higher Education*, 42(9), 1655-1670. <https://doi.org/10.1080/03075079.2015.1113953>
- Martínez-Argüelles, M. y Batalla-Busquets, J.-M. (2016). Perceived Service Quality and Student Loyalty in an Online University. *The International Review of Research in Open and Distributed Learning*, 17(4), 264-276. <https://doi.org/10.19173/irrodl.v17i4.2518>

- Martínez-Caro, E., Cegarra-Navarro, J. G. y Cepeda-Carrión, G. (2015). An application of the performance-evaluation model for e-learning quality in higher education. *Total Quality Management & Business Excellence*, 26(5-6), 632-647. <https://doi.org/10.1080/14783363.2013.867607>
- Mejía, J. F. y López, D. (2016). Modelo de Calidad de E-learning para Instituciones de Educación Superior en Colombia. *Formación Universitaria*, 9(2), 59-72. <https://doi.org/10.4067/S0718-50062016000200007>
- Mixnahuatl, A., Santamaría, J., Urbina, A., Calleja, J. de la y Auxilio, M. (2013). Diseño del curso análisis de sistemas aplicando el modelo de calidad para e-learning. *Educación Handbook T-I: Congreso Interdisciplinario de Cuerpos Académicos* (pp. 143-152).
- Olmos-Migueláñez, S. (2008). *Evaluación Formativa y Sumativa de estudiantes universitarios: Aplicación de las Tecnologías a la Evaluación Educativa* [Tesis Doctoral]. Universidad de Salamanca.
- Ortiz-López, A., Olmos-Migueláñez, S. y Sánchez-Prieto, J. C. (2021). Calidad en e-Learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2), 225-244. <https://doi.org/10.5944/ried.24.2.29073>
- Ortiz-López, A., Olmos-Migueláñez, S. y Sánchez-Prieto, J. C. (2021). *eLearning, assessment and Higher Education: A mapping of the literature* [Comunicación escrita]. 2021 XI International Conference on Virtual Campus. Salamanca, Spain.
- Palma, S. y Delgado, M. (2006). Consideraciones prácticas acerca de la detección del sesgo de publicación. *Gaceta Sanitaria*, 20(3), 10-16. <https://doi.org/10.1157/13101085>
- Park, Y. y Lim, K. (2015). Effects of Environmental and Human Constructs on E-learning Effectiveness in Online University Settings. *Indian Journal of Science and Technology*, 8(1), 103-109. <https://doi.org/10.17485/ijst/2015/v8iS1/57729>
- Parlakkiliç, A. (2017). Change Management in Transition To E-learning System. *Qualitative and Quantitative Methods in Libraries*, 3(3), 637-651.
- Pecori, R., Suraci, V. y Ducange, P. (2019). Efficient computation of key performance indicators in a distance learning university. *Information Discovery and Delivery*, 47(2), 96-105. <https://doi.org/10.1108/IDD-09-2018-0050>
- Pérez-Escoda, A. (2017). WoS y Scopus: Los grandes aliados de todo investigador. *Comunicar*, 5. <https://doi.org/10.3916/escuela-de-autores-031>
- Pham, L., Limbu, Y. B., Bui, T. K., Nguyen, H. T. y Pham, H. T. (2019). Does e-learning service quality influence e-learning student satisfaction and loyalty? Evidence from Vietnam. *International Journal of Educational Technology in Higher Education*, 16(7), 1-26. <https://doi.org/10.1186/s41239-019-0136-3>
- Pontoriero, F. A. (2021). E-learning en la educación superior argentina—Modelo de evaluación de calidad a partir del aporte de referentes clave. *Virtualidad, Educación y Ciencia*, 12(22), 22-45.
- Prieto-Rodríguez, E., Gore, J. y Holmes, K. (2016). Exploring Quality Teaching in the Online Environment Using an Evidence-Based Approach. *Australian Journal of Teacher Education*, 41(8), 22-39. <https://doi.org/10.14221/ajte.2016v41n8.2>
- Radu, M.-C., Schnakovszky, C., Herghelegiu, E., Ciubotariu, V.-A. y Cristea, I. (2020). The Impact of the COVID-19 Pandemic on the Quality of Educational Process: A Student Survey. *International Journal of Environmental Research and Public Health*, 17(21), E7770. <https://doi.org/10.3390/ijerph17217770>
- Raspopovic, M. y Jankulovic, A. (2017). Performance measurement of e-learning using student satisfaction analysis. *Information Systems Frontiers*, 19(4), 869-880. <https://doi.org/10.1007/s10796-016-9636-z>
- Serrano, E. L., Ceballos, S. P., Cordero, G. y Cisneros-Cohernour, E. (2018). Marco para evaluar las condiciones institucionales de la enseñanza en línea. *Revista Electrónica de Investigación Educativa*, 20(2), 1-14. <https://doi.org/10.24320/redie.2018.20.2.2072>
- Torres Barzabal, L., Ortiz, P. y Barcia-Tirado, D. (2018). Quality Indicators for Auditing On-Line Teaching in European Universities. *TechTrends*, 63(3), 330-340. <https://doi.org/10.1007/s11528-018-0365-7>
- Waheed, M. y Kaur, K. (2017). Students' perceptual quality standards for judging knowledge quality: Development and validation of a knowledge quality scale. *Information Development*, 35(2), 319-332. <https://doi.org/10.1177/0266666917744370>
- Waheed, M., Kaur, K. y Qazi, A. (2016). Students' Perspective on Knowledge Quality in eLearning Context: A Qualitative Assessment. *Internet Research*, 26(1), 120-145. <https://doi.org/10.1108/IntR-08-2014-0199>

Anexo 3

Publicación en revista científica: Education and Information Technologies

VERSIÓN POST-PRINT

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12117-5>

INDICADORES:

Revista Q1 – JIF en último índice publicado (2022): Factor de impacto 5,5. Pos. 19/269. Percentil 93,1

Revista Q1 – JCI en último índice publicado (2022): Factor de impacto 2,56. Pos. 21/742. Percentil 97,24

Revista Q1 – SJR en último índice publicado (2022): Factor de impacto 1,249

Enlace a la publicación completa:

<https://link.springer.com/article/10.1007/s10639-023-12117-5>

Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context

Alberto Ortiz-López¹  · Susana Olmos-Migueláñez¹  · José Carlos Sánchez-Prieto¹ 

Abstract

Today, education is facing a new reality in which technology and new teaching methods are being quickly introduced into educational systems and institutions. Educational institutions are now dealing with the challenge of providing continuity to e-learning, turning it into a more flexible and up-to-date field, and considering assessment as a quality element in this transition. Therefore, with the aim of determining the current state of the research focused on assessment in digital environments (e-assessment), a mapping of the literature has been carried out. After examining 1,771 results extracted from Web of Science and Scopus and after the application of seven inclusion criteria, a total of 159 publications from the period of the past five years were read. The answer four research questions on the evolution of publications, the authors, the tools used, the contexts, the objects of study, and the future avenues of research, among others. The results show the increasing importance of e-assessment in this new context, moving toward a new reality in which technology plays a decisive and fundamental role in the teaching and learning processes. Thus, educational systems are heading towards a new context in which both teachers and students should rethink their roles and functions leading education to a more flexible, current, and digitally mediated context.

Keywords E-assessment · Mapping · Review · Education · Trends

✉ Alberto Ortiz-López
aortiz@usal.es

Susana Olmos-Migueláñez
solmos@usal.es

José Carlos Sánchez-Prieto
josecarlos.sp@usal.es

¹ Instituto Universitario de Ciencias de la Educación (IUCE), GRIAL Group, Universidad de Salamanca, Salamanca, Spain

1 Introduction

The educational transition toward the online environment is an area of particular interest in the field of education, with virtual environments and information and communication technologies already being positioned as the main lines of educational change. This change in the educational perspective leads education to a new scenario in which there are only references of a few specific experiences; therefore, a new stage for education arises, one that has been brought about by the coronavirus crisis and digitalization of teaching throughout the world (Schwartzman et al., 2021; St-Onge et al., 2022).

In this broad field, it is essential to pay attention to one of the most complex activities (Cabero-Almenara & Palacios-Rodríguez, 2021; Chacín, 2021)—the assessment—to better understand the state of its transformation and the guarantees it offers in the new teaching contexts. An e-assessment has been defined by Ruiz-Morales (2013) as the planning, design, and development of an assessment in the digital environment to guarantee the same training competencies as in face-to-face scenarios and promote student participation in their learning.

In this new context, the transition of assessment toward digital environments makes it possible to study how the processes related to teaching are developed and implemented in this new setting, paying special attention to the assessment processes (Alruwais, 2018; Doğan et al., 2020). Morales et al. (2018) found that different competencies were prioritized in the two learning environments. Tomas et al. (2015) defended the idea of the necessary reconstruction in e-assessment, here as conceived from the prioritization of rapid adoption and not from a deep understanding of it; therefore, its transformation is still a major challenge.

Sellés et al. (2018) suggested that this new conception and reconstruction of e-assessment must be based on a holistic character, enhancing meaningful learning, promoting collaboration between students, and promoting self-reflection of learning. In addition to being supported by information and communication technologies (ICT), it enables more flexible scenarios that allow for a greater volume and management of data, less effort for the student, and faster and more objective correction (Ghouali et al., 2020). This concept contrasts with the rigidity of traditional assessment methods, which are excessively focused on marking, lack creativity, and are not very motivating for students (Pearse-Romera & Ruiz-Cecilia, 2019). This perspective on an assessment today represents a complete challenge for educational institutions because it must overcome the inflexibility maintained by specific teaching contexts, improve its articulation with training, and achieve an effective change (Jiménez et al., 2021), hence understanding assessment as an orientation toward improvement and guarantee of the quality of teaching.

The new digital contexts in education have presented limitations and doubts on the part of society in their development regarding the accessibility and training of both teachers and students in digital matters (Domingo-Coscollola et al., 2020), along with the terms of security and quality assurance (García-Aretio, 2021; Lee & Fanguy, 2022). Field research, such as that of Dermo (2009), with 130 university students, or that of Sellés et al. (2018), with 106 students at the same level,

and other research (Burgos-Videla, 2021; Geomar, 2020; Humanante-Ramos et al., 2019; Palacios et al., 2021), have shown that the advantages outweigh their limitations, being an opportunity for the future of teaching organizations.

Working in this context, the present study aims to focus on the literature that has been researched on e-assessments over the past five years, covering all those investigations published in Web of Science and Scopus in English. The current research is presented under a mapping study, given that it will summarize and focus the current state of research in the e-assessment field. The importance of mapping research lies in its ability to structure and classify an area of research: e-evaluation (Ramírez, 2020). Among its advantages, this methodology is a more open-ended systematic review that can facilitate a global view of the field by evaluating existing quantitative evidence (Marshall & Brereton, 2013).

The remainder of the present article is divided as follows: Section 2 depicts the theoretical framework followed by the related works (Section 3); Section 4 describes the methodology of the review; the Section 5 presents the research results; and Sections 6 and 7 include the discussion and conclusion derived from the results, respectively.

2 Theoretical framework

E-assessment is an increasingly important field in both online and face-to-face teaching–learning processes, and, as happens with most technology-mediated activities, it is subject to a process of continuous change (Ghouali et al., 2020).

Currently, most online assessment systems are based on learning management systems (LMS) for their development, here contemplating the use of multiple-choice questions (MCQs), essays, forums, virtual portfolios, or wikis, among others (Gupta et al., 2019). Technological development in assessment is also focused on countless applications or resources, depending on the purpose (Blackboard, Kahoot, Posterlet, Google Classroom, among others) (Jaiswal, 2020) that are adapted according to the assessment modality and based on the teacher's preferences for formative assessment or summative assessment (Black, 1993; Dixson & Worrell, 2016; Nasab, 2015).

The first one, formative assessment, focuses on information and feedback, the understanding of the content, and continued performance (Daly et al., 2010; Stanković et al., 2017). Technology-based formative assessments focus on different applications and software that allow for the continuous monitoring of the student and implementation of feedback activities (portfolios, dossiers, practical activities, etc.), which have been claimed to be the most developed modality in its relationship with technology in practice and in the literature (McCallum & Milner, 2021; Mimirinis, 2019). Summative assessment, on the other hand, focuses on the certification and measurement of the knowledge acquired once the process has been completed, after the evaluation of products, and on the performance of final tests (Knight, 2002). In this modality, technology is used in the form of the implementation of tools that make it possible to perform objective tests in a safe and reliable manner (Atoum et al., 2017) and on the systematization of these for the systematic

certification of performance (Riera et al., 2018), a lesser route given the quality, control, and safety standards that these processes require (Timmis et al., 2015).

The development of e-assessment has been accelerated in the last few years by two main factors: the technological advances experienced in recent years (González-González & Silveira-Bonilla, 2022; Muñoz-Guevara et al., 2021) and the impact of the COVID-19 pandemic (Ozdamli & Karagozlu, 2022).

The technological changes experienced by the inclusion of new technologies in teaching offer new opportunities for the assessment in new educational scenarios (use of mobile devices), the automation of results and feedback (development of artificial intelligence in education), the identification of patterns in assessment and the detection of learning difficulties (implementation of learning analysis strategies), and the assessment in new immersive and technology-enriched environments (use of extended reality in the classroom) (Sembey et al., 2023).

Moreover, e-assessment has been completely disrupted by COVID-19, which forced all educational systems to adapt their teaching to nonvirtual learning environments, taking on the challenge of assessment in alternative spaces and leading to new strategies, policies, and rapid training aimed at enabling the educational community to carry out assessments (formative and summative) on institutional platforms (Zharova et al., 2020). As a result, the volume of resources, new tools, and derivative publications on e-assessment have drastically increased (Almossa, 2021; Kundu & Bej, 2021; Montenegro-Rueda et al., 2021; Seraj et al., 2022), suggesting the focus on technology and flexibilization of educational institutions regarding what is termed the new educational normality (García-Peñalvo, 2021; Grande-de-Prado et al., 2021).

The combination of these two elements has created a new context presenting a wide spectrum of possibilities for developing e-assessments (Kotiash et al., 2022) and a significant increase in its popularity (Szymkowiak et al., 2021).

3 Related works

Systematic literature reviews and their variant of systematic mapping studies have been increasingly used in investigations focused on analyzing the state of the art in a particular area or context, being one of the most relevant techniques in research (Kitchenham et al., 2011).

There are examples of previous reviews that have addressed technological implementation in educational processes (Lai & Bower, 2019; Martínez-Soto & Prendes-Espinosa, 2023), in education during the pandemic in higher education (Muhaimin et al., 2023), in lower levels (Crompton et al., 2021), or in the inclusion of mobile devices for the development of new teaching methodologies (Dorris et al., 2021).

The current review has combined two aspects: the pandemic and technological development. Regarding assessment during the pandemic, the review conducted by Montenegro-Rueda et al. (2021), which was an analysis of 13 studies, highlighted the challenges faced by teachers and students, the lack of training, the honesty of assessment processes, and the differentiation between formative and summative assessment during the pandemic. Second, on technological development in

e-assessment, reviews such as the one by Heil and Ifenthaler (2023) focused on the characteristics of these environments and the evaluative modalities in higher education, here determining the potential of these processes and the need for training and capacity building; the review by Sembey et al. (2023) on assessment, emerging technologies and feedback practices through an analysis of 38 accepted articles highlighted the advantages of the inclusion of technologies in assessment and provided a list of technological tools.

Alrofou et al. (2019) conducted a review on the acceptance and inclusion of mobile devices in assessment processes; the study analyzed 53 publications over a 10-year period and concluded the need to overcome the limitations of the educational community for the effective incorporation of mobiles in assessment processes. Other reviews looked at the incorporation of artificial intelligence (Chiu et al., 2023; Samala et al., 2023) and virtual reality in educational and assessment processes (Ifanov et al., 2023).

The mapping described in the present study has aimed to achieve a broader vision than the reviews above. For this reason, after not detecting any updated study for the field and given the volume of systematic reviews and related publications, a mapping is justified as a broad synthesis, allowing us to map the current publications, detect new avenues of research and identify, in a generic way, the current state of knowledge after the latest changes and advances in research (Cooper, 2016). In addition, this review approaches the field of e-assessment without prior restrictions, not limiting itself to a specific educational level, a particular tool, or an assessment modality so that the results achieved summarize all the publications on the field in the past five years, enabling the mapping analysis in a synthetic and accurate way.

4 Methods

The method followed in the present research was a mapping of the literature, which summarizes the state of the art of relevant research and contextualizes the work in the international landscape while explaining which findings if any, are being challenged or extended. In this mapping research, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement for systematic reviews was followed (Page et al., 2021).

5 Research questions

To provide an overview of the complete state of the field, research questions have been designed to describe the main data of the publications (country, annual evolution, authors, journals, etc.) and to describe the general lines of each study (objective, educational level, sample, methodology followed). The research questions posed in the mapping study are as follows:

MQ1) What is the relevance of the included studies?

MQ2) What has been the current state of e-assessment research in the past five years?

MQ3) What is the studies' approach to e-assessment?

MQ4) What are the future trends in the e-assessment field?

5.1 Exclusion criteria

The inclusion/exclusion criteria that will determine the publications making up the analysis and answers to the research questions posed are as follows:

- 1 Any publication that is not published in a scientific journal indexed in Web of Science or Scopus and that has not undergone a peer review process is not accepted.
- 2 Articles that did not focus on the educational field were excluded.
- 3 Articles that did not have e-assessment as their object of study were excluded.
- 4 Articles that did not present empirical methodologies (theoretical studies) were excluded.
- 5 Articles written in languages other than English were excluded.
- 6 Articles that did not directly refer to the e-assessment tool under study were excluded.
- 7 Articles that did not describe the population and sample participating in their research were excluded.

5.2 Search string and database

The search was conducted on December 10, 2021, in the Scopus and Web of Science databases. The results have been limited in time to offer a broad perspective of the past five years. A search string was used (“E-ASSESSMENT” OR “EASSESSMENT” OR “COMPUTER BASED ASSESSMENT” OR “MOBILE BASED ASSESSMENT”), obtaining all those articles that addressed assessment in virtual environments or those that addressed it through technological devices. The datasets generated during the current study are available in the Google Sheets repository at the following: bit.ly/3N8xBXf.

5.3 Publication selection process

The initial number of publications was 1,771 (907 from Scopus and 891 from Web of Science). After removing duplicates, the figure dropped to 1,136. After applying the exclusion criteria, 159 publications were accepted as part of the present research (acceptance rate of 13.99%) (Fig. 1). The complete list of accepted references is available at the following: bit.ly/3Noh55Q.

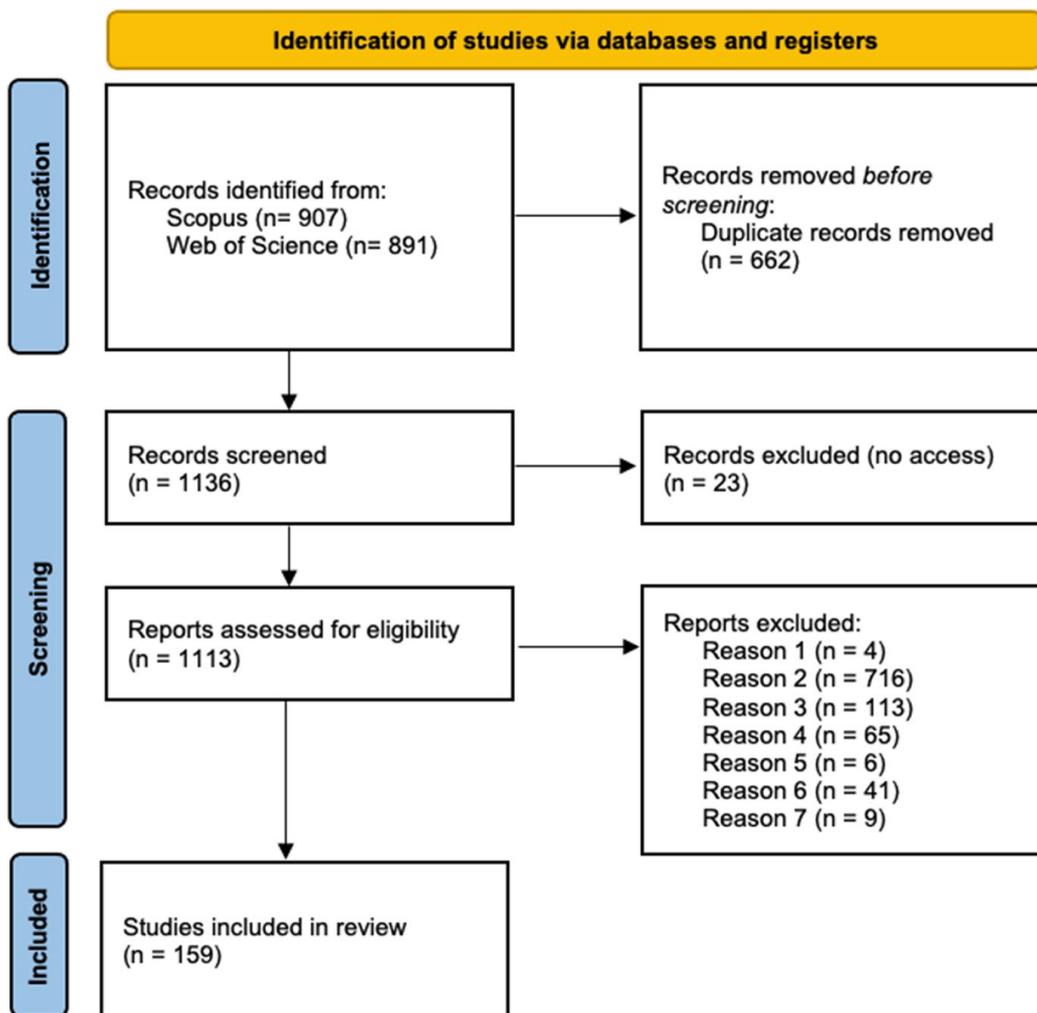


Fig. 1 PRISMA flow chart of the review

6 Results

6.1 What is the relevance of the included studies?

To study relevance (MQ1), the number of citations of each article and the bibliometric indicators of the journals were considered. For citation extraction, Scopus was used as the search database on January 10, 2023. The results indicate an average citation level of 14.21 (145 valid articles, 15 not included in Scopus). The two most relevant articles focused on the technological acceptance of e-assessment and were developed by the same authors (Nikou & Economides, 2017a, b), accumulating a total of 354 citations between them since 2017, followed by references such as Nguyen et al. (2017) and Faber et al. (2017), at 79 and 74, respectively.

Table 1 shows the results; the item identification number is available in the supplementary database under the tabs “Final Phase” and “Metrics.” The

Table 1 Number of citations of the selected publications

Number of citations	Identification number
181	42
173	40
80	43
79	37
74	117
73	79
70	96
48	31
43	19
42	107
38	28
37	17
36	148
34	103
30	48
29	36, 100
26	89
24	102
23	29, 126
22	3, 18, 49
21	30, 55
19	25, 65
18	47, 121
17	86
15	10, 159
14	22, 90, 101, 104, 113, 153
13	58, 60, 64, 105, 129
12	39, 114, 139, 146, 155
11	16, 137
10	4, 20, 27, 75, 116, 138, 152, 154, 156
9	24, 54, 106, 112
8	6, 35, 53, 78, 97
7	38, 84, 115, 132, 136
6	14, 76, 88, 125
5	15, 41, 56, 70, 72, 80, 98, 99, 158
4	8, 26, 34, 45, 46, 50, 51, 59, 74, 87, 91, 94, 95, 111, 118, 128, 134, 135
3	13, 21, 23, 32, 61, 69, 77, 131, 141
2	2, 7, 12, 52, 81, 85, 93, 108, 109, 120, 124, 144, 150
1	1, 33, 44, 57, 62, 63, 68, 82, 92, 122, 130, 142, 145, 149
0	11, 73, 119, 123
No data	5, 9, 66, 67, 71, 83, 110, 127, 133, 140, 143, 147, 151, 157

Fig. 2 Publications metrics

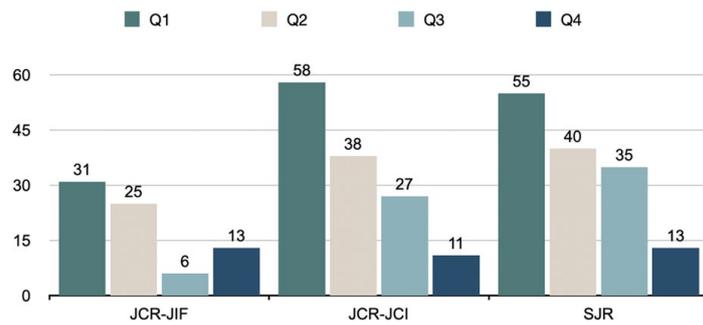
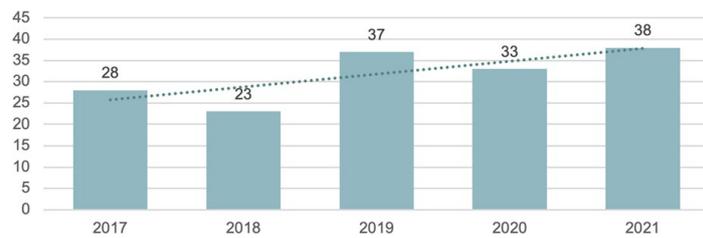


Fig. 3 Annual evolution of publications



extended table with each publication reference and complete list can be found at the following: bit.ly/3oU1Wzy.

For further analysis of the publication metrics, the data relating to SJR, JIF, JCI, and quartiles (Q) corresponding to the year of publication of the article in each journal were extracted. The results indicate that, of the 159 results, 75 were included in JIF (31 in Q1, 25 in Q2, 6 in Q3 and 13 in Q4) and 134 in the JCI index (58 in Q1, 38 in Q2, 27 in Q3 and 11 in Q4). Finally, 143 publications have an SJR index (55 in Q1, 40 in Q2, 35 in Q3 and 13 in Q4) (Fig. 2).

Full information about the publication metrics including impact factors, quartiles, and categories in JIF, JCI, and SJR can be found in the “Mapping Question 1 (MQ1) sheet” of the complementary database.

6.2 What has been the current state of e-assessment research in the past five years?

To position the e-assessment field (MQ2), it is important to first describe the publication trend occurring. For this purpose, the publication date of each study has been taken, as presented in Fig. 3.

The data show a growing trend in the number of publications, from 28 in the first year analyzed (2017) to 38 in 2021. This trend indicates the efforts and relevance of e-assessment in current educational trends, as reflected in the increase of studies linked to the subject. This growing trend also indicates an interest in the field of e-assessment. This must also be supported by theoretical research into the means, methodologies, limitations, and possibilities of this assessment form.

Once the annual trends have been identified, it is important to identify the trends followed according to the publication location and identify research leaders in terms of their geographic nature. The analyzed studies have been carried out in a total of 49 countries, led by the United States (24) and followed by Spain (14), the United

Table 2 Country where the studies were carried out

<i>N</i>	Country
24	United States
14	Spain
13	U.K
12	China
8	Saudi Arabia
6	Greece
5	Australia, Malaysia, Turkey
4	Germany, Indonesia, Portugal
3	Finland, Israel, Morocco, Pakistan, Taiwan, The Netherlands
2	Canada, Hungary, Iran, Ireland, Singapur, South Africa, Switzerland
1	Argelia, Ajman, Austria, Chile, Croatia, Dubai, Egypt, France, Greece, India, Iraq, Jordan, Mexico, Namibia, New Zealand, Norway, Qatar, Ruanda, Serbia, Slovakia, Thailand, Tunisia Yemen

Table 3 Authors according to the number of publications

<i>N</i>	Author
6	Economides, A.A.; Nikou, S.A
4	Kurniawan, D.A
3	Babo R., Suhonen J., Astalini A., DeMara R.F., Guerrero-Roldan, AE, Kocdar, S, Lin, JW, Whitelock, D
2	Chemsi, G., Chen B., Darmaji D., Hartshorne R., Karadeniz, A., Maison, Ninkovic, S., Noguera, I., Okada, A., Omar N.N., Perdana, R., Peytcheva-Forsyth, R., Radid, M., Sa'di, RA., Sancho-Vinuesa, T., Sharadgah, TA., Tukiainen M., Yilmaz, R
1	457 authors (accessible in the attached database)

Kingdom (13), and China (12). The countries mentioned above are the only ones with more than 10 studies carried out in them. The rest of the countries can be seen in the following table (Table 2).

Regarding the most relevant authors in the literature, a total of 486 contributors have been listed. Of these, Anastasios A. Economides and Stavros A. Nikou each stood out with six publications, followed by Dwi A. Kurniawan with four. Eight authors followed with three publications, 18 with two, and 457 with only one (Table 3).

Finally, the field is positioned by describing the publication source in which the research has been published. The analysis extracted a total of 106 journals that published 159 publications (Table 4). Among them, the *International Journal of Emerging Technologies in Learning* stood out as the leading journal in this research field, with a total of seven publications, followed by *Assessment and*

Table 4 Journals by number of publications

<i>N</i>	Journal
7	International Journal Of Emerging Technologies In Learning
6	Assessment & Evaluation In Higher Education
5	British Journal Of Educational Technology, Computers & Education
4	The Asia–Pacific Education Researcher
3	Advances in Physiology Education, BMC Medical Education, Education And Information Technologies
2	Asian EFL Journal, Computer Applications In Engineering Education, Computers In Human Behavior, Eurasia Journal Of Mathematics, Science and Technology Education, Frontiers In Education, Frontiers In Psychology, Hacettepe University Journal Of Education, Interactive Learning Environments, International Journal For Educational Integrity, International Journal Of Stem Education, Irish Educational Studies, Journal of Information Technology Education: Research
1	86 journals (accessible in the supplementary database)

Evaluation in Higher Education (6) and the *British Journal of Educational Technology and Computers and Education* (5).

The publications are next identified in terms of their main characteristics, helping determine the key aspects of current e-assessment research. Concerning the educational level at which e-assessment has been studied, higher education was the main field (133 research studies), followed by secondary education (with only 12) and then the other educational levels with lower figures, as can be seen in Table 5.

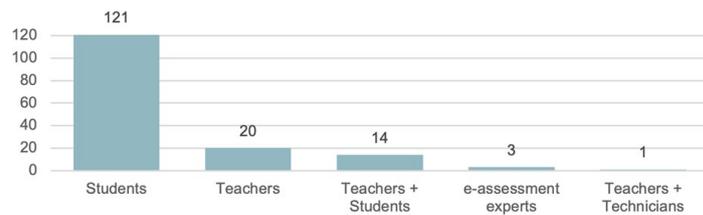
Once the educational levels have been established, it is necessary to identify the target group of the research studies. In this analysis, student samples (120 of 159), followed by teachers (20), and combined studies with teachers and students (13) have stood out by a large margin. The remaining figures are presented below (Fig. 4).

After studying the groups, the present research examined the methodological processes of each one. The results show that mainly quantitative studies were

Table 5 Educational fields of the analyzed studies

<i>N</i>	Field
133	University Degree
12	Secondary Education
5	Primary Education
4	Vocational Education and Training (VET)
2	Certificate of Higher Education (HNC)
1	Preschool, Software Development, mixed (Primary Education + Secondary Education)

Fig. 4 Publications according to each study group



conducted (66%), followed by qualitative studies (18%), and, finally, mixed studies (16%) (Fig. 5).

Looking further into these methodological processes and addressing the techniques used in each study to collect information, 26 studies used two techniques in their studies. In contrast, the remaining 133 applied only one.

Of the 185 total techniques applied, the survey (78 studies) was the most frequently used. This was followed by an analysis of the results of e-assessments carried out through questionnaires and objective tests (68 studies) and then the rest of the techniques (interviews, validation questionnaires, focus groups, and observation processes) (Fig. 6).

Finally, the population, methodology, and instruments used were explored. For this purpose, we have maintained the grouping categories proposed by the authors in their research data collection instruments. Given the samples' dispersion and the nonrepresentativeness of means, the information is presented according to the different groups and sample sizes (Table 6).

Regarding the 18 studies that carried out interviews, their classification according to their sample sizes is presented below (Table 7).

Fig. 5 Methodology of the analyzed publications

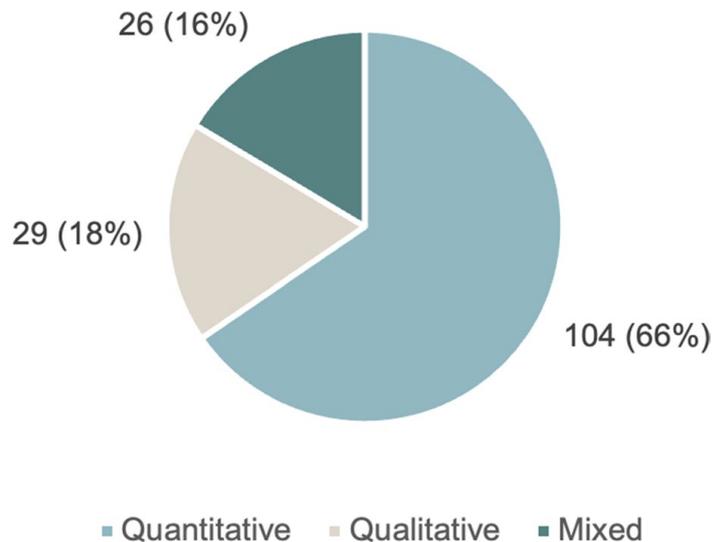


Fig. 6 Techniques used in studies by frequency

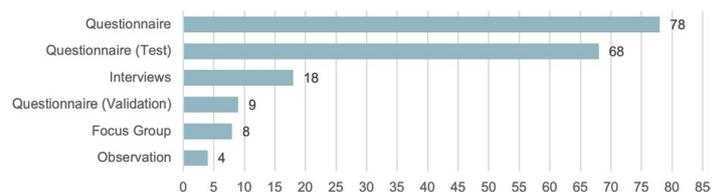


Table 6 Sample sizes for questionnaire studies

	<i>N</i> = 1–100	<i>N</i> = 101–200	<i>N</i> = 201–300	<i>N</i> = 301–400	<i>N</i> = 401–500	<i>N</i> = 501–600	<i>N</i> = 601–700	<i>N</i> = 701–800	<i>N</i> = 801–900	<i>N</i> = 901–1000	<i>N</i> = > 1001
Questionnaire	25	23	8	5	3	1	5	1	2	2	4
Questionnaire (Test)	26	10	6	6	5	1	1	1	0	2	10
Questionnaire (Validation)	3	3	1	-	-	1	-	-	-	-	1

Considering the eight studies that used a focus group, their classification in terms of ranges for their sample sizes is given below (Table 8).

Of the articles using an interview process, the article by Washburn et al. (2017) presented a sample of 138 participants that was carried out in a total of 12 focus groups. Finally, the analysis of sample sizes is presented for those studies using observation processes (Table 9).

6.3 What is the studies' approach to e-assessment?

The third research question (MQ3) explores the way in which researchers have approached e-assessment, where we describe the variable with which it is approached (object of study). These objects were generated following an inductive method after the categorization of each of the investigations (Table 10).

These categories are exhaustive and exclusive and have been validated after ascertaining that each selected publication is part of a single category (Fig. 7).

Regarding the technological tools on which the e-assessment processes have been based, these have been classified into LMS systems (integrated learning management systems), mobile technologies (tools or applications for exclusive use through mobile devices), and other software (i.e., all those accessible through the network). Of the 159 articles, three (Nikou & Economides, 2019; Rolim & Isaias, 2018; Smith, 2021) addressed three different e-assessment tools, three studies looked at two different tools (Egarter et al., 2021; Liu et al., 2021; Ninković et al., 2021), and the remaining 164 focused on only one (total of 170 tools found). These tools, classified according to their typology, are presented in Fig. 8.

The LMS platforms were a main focus of the studies (59%). Of these, the presence of Moodle ($n=42$) and educational institutions' own LMS platforms ($n=35$) has stood out, as well as Blackboard ($n=14$) and Canvas ($n=3$). Moodle and Blackboard can be seen as the critical representations of this type of platform, accounting for 59.6% of the category.

Second, other publications focused on software and tools accessible through the web (they are not only access devices, and their operation does not allow for the integrated management of teaching as an LMS). This category represented 31% of the 159 publications and included tools from different fields, such as security and authentication (Proctorio, Tinfollec); web-based e-assessment tools (Kahoot, Dewis, JCQuest, Posterlet; among others); the Google suite for virtual teaching (Google Forms, Google Classroom); and communication and social networking tools (Microsoft Teams, Zoom, Yammer, WebEx). The rest of the educational software and digital tools can be seen in the complementary database.

Regarding mobile technologies, some research looked at applications aimed at teaching and group management (Snappet, Socrative, Tronclass), applications aimed at communication and team management (Slack, WeChat, Viber), applications focused on assessment (ExamSoft, tExam), and applications focused on the teaching and assessment of specific subjects (GeoGebra, GotItLanguage), among other applications with lower representations. These apps represented only 14% of the total.

Table 7 Sample sizes for interviewing studies

	<i>N</i> =1-10	<i>N</i> =11-20	<i>N</i> =21-30	<i>N</i> =31-40	<i>N</i> =41-50	<i>N</i> =51-60	<i>N</i> =61-70	<i>N</i> =71-80	<i>N</i> =81-90	<i>N</i> =91-100	<i>N</i> > 101
Interviews	8	2	3	2	-	-	1	-	-	-	2

Table 8 Sample sizes for focus group studies

	<i>N</i> =1–5	<i>N</i> =6–10	<i>N</i> =11–15	<i>N</i> =16–20	<i>N</i> =21–25	<i>N</i> =26–30	<i>N</i> > 30
Focus Group	1	4	-	-	1	-	1 (138)

Table 9 Sample sizes for observational studies

	<i>N</i> =1–2	<i>N</i> =3–4	<i>N</i> =5–6	<i>N</i> =7–9	<i>N</i> > 10
Observation	1	-	1	1	1 (120)

Table 10 Object of study for each study

Object	Description
Attitude	Studies focus on attitudes and perceptions towards e-assessment tasks
Acceptance	Studies focus on the technological acceptance of e-assessment
Efficacy	Studies focus on the effectiveness of e-assessment processes
Innovation	Studies focus on describing innovation processes in e-assessment without focusing the study on a specific variable
Monitoring	Studies focus on monitoring and surveillance in e-assessment processes
Performance	Studies focus on outcomes and performance in e-assessment processes
Quality	Studies focus on the quality of e-assessment processes
Satisfaction	Studies focus on studying satisfaction after an e-assessment process

Fig. 7 Publications categorized according to the object of study and their frequency

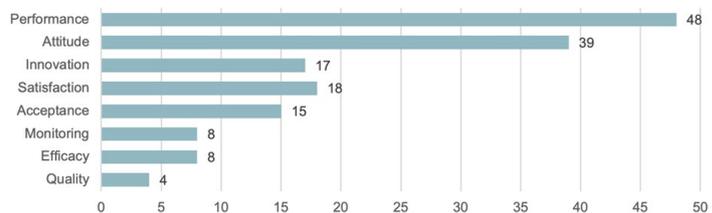
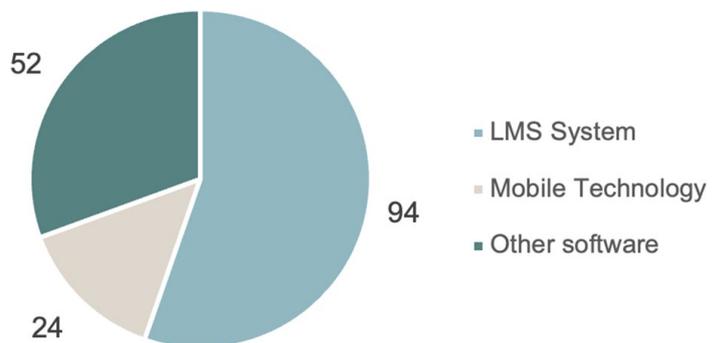
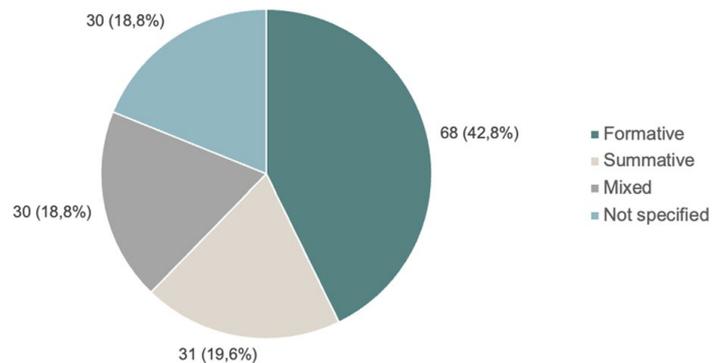


Fig. 8 E-assessment tools grouped according to their typology



Lastly, we examined how the research conceived of e-assessment, that is, distinguishing between formative, summative, and mixed assessments (Fig. 9). Of the 159 studies, 42.8% ($n=68$) focused on the analysis of e-assessment in formative

Fig. 9 Studies according to the assessment's purpose



processes, while only 19.6% ($n=31$) did so for summative assessments. A total of 30 (18.8%) studies addressed e-assessment as a tool for comprehensive assessment (formative and summative), and another 30 (18.8%) studies did not specify the e-assessment method (focusing on other aspects such as plagiarism, supervision, or risks, among others).

6.4 What are the future trends in the e-assessment field?

Finally, the aim of the present review was not only to focus and unify the advances in the field, but also to bring together the main lines of research arising in the field of e-assessment; this was done by considering the lines proposed by the authors as future paths of research (MQ4). Here, 11 of the 159 studies did not propose lines of future research, focusing only on describing the progress and results achieved. Twelve others do not focus on future areas. Regarding this transfer, eight authors suggested the need to promote institutional development and collaboration to promote e-assessment (e.g., Cheng et al., 2021; DeMara et al., 2017; Slade et al., 2022), and four focused on promoting teacher education and training as an element of quality within these assessment processes (Abdou, 2020; Cutumisu, 2018; Wu & Wang, 2021; Zhang et al., 2021).

Therefore, a total of 136 studies suggested into future lines of research in the e-Learning field. After their complete extraction and analysis, these lines were categorized into six categories, as shown in Table 11.

The main future line focused on the development of new methodologies, instruments, and software related to e-assessment processes ($n=51$, 37.5%). This line synthesized the proposals made by the authors in relation, on the one hand, to the need for continuity in developing instruments and software that support e-assessment processes, which was one of the main pillars of this assessment modality (Lin et al., 2020; Liu et al., 2021). On the other hand, another group of authors highlighted not only the need to introduce new software, but also to improve the reliability of technology-mediated assessment processes, hence improving the reliability and stability of instruments (Bickerton & Sangwin, 2021), tools (Nyland et al., 2017), and assessment processes (Khlaisang & Koraneekij, 2019).

On the other hand, another stream of researchers raised the importance of introducing processes and mechanisms for supervision and monitoring in online

Table 11 Proposed lines of research

Line	<i>n</i>	Objective
1	51	To develop new methodologies, tools and software related to e-assessment
2	40	To replicate the study to confirm the findings (populations, samples and platforms)
3	27	To explore further conditioning factors for the e-assessment acceptance
4	11	To extend the study with other research methodologies (qualitative, quantitative, mixed)
5	5	To study the effects of long-term e-assessment on academic performance
6	2	To expand on the rationale and theoretical underpinning of e-assessment

assessments, highlighting the importance of having a supervisor and valid e-authentication practices (e.g., Mellar et al., 2018; Okada et al., 2019a, b), and with systems of safeguards and policies in place against plagiarism and cheating (e.g., Hassan et al., 2020; Snekalatha et al., 2021; Vázquez et al., 2021). Finally, another group of studies advocated for improving the instruments from the feedback given (e.g., Judge, 2021; Misut & Misutova, 2017; Wang et al., 2019).

Another group of studies showed interest in replicating each study in different populations, samples, and platforms to confirm the findings ($n=40$, 29.4%). Of these 40 studies, 38 proposed their methodological and instrumental replication in different populations (different educational levels, different contexts) or different samples. These studies focused on confirming the results by comparing different populations and samples to certify the performance of online assessment processes. An example is the study carried out by AlTameemy et al. (2020), who proposed replicating their investigation on writing assessments through platforms at other educational levels, or that of Witchel et al. (2018), who proposed the replication in a different course with a different population.

Third, some studies explored the conditioning factors of teachers' e-assessment acceptance ($n=27$, 19.9%). These studies sought to continue the search for different factors that may help the inclusion of e-assessments in teaching processes through teachers. These studies proposed searching for new dimensions in established acceptance models (Bacca-Acosta & Avila-Garzon, 2021; Balta et al., 2018; Nikou & Economides, 2018) or the search for other factors, such as social and cultural ones (Pu & Xu, 2021).

Next, 11 publications (8.1%) addressed the importance of extending the research through a new methodology. Most proposed complementing the existing study with qualitative data (Babo et al., 2020; Cakiroglu et al., 2017; Delaney et al., 2019, among others) or conducting multicenter studies by changing the methodology (Dutton et al., 2017).

Finally, from a more theoretical point of view, five (3.7%) publications proposed addressing the effects that this assessment modality has on the student's future academic performance, such as Kühbeck et al. (2019) or Boote et al. (2021), and two (1.4%) proposed expanding the theoretical justification of this type of technology-mediated assessments (Guerrero-Roldán & Noguera, 2018; Sirianni et al., 2017).

The lines of research proposed by the authors can guide future research in this field, so attending to them should be a fundamental task for experts and researchers, institutions, and teaching professionals.

7 Discussion

7.1 Evolution, relevance, and methodology trends of the publications

First, addressing the number of publications over time, the scientific community has been increasingly interested in e-assessment in digital environments, with the number of research works published showing a clear upward trend, even though the average number of citations is low at 14.21 (this may be because most articles are very recent and do not yet have enough time to be cited in other research). This trend may be a result of studies being conducted because of the coronavirus pandemic, which was a significant educational change and methodological breakthrough in education. Moreover, given the temporal proximity of the review to the pandemic (still ongoing) and the publication deadlines of the journals and volume of articles written in all fields about COVID-19 and its impact (Forero-Peña et al., 2020), a substantial increase in the number of future publications is expected.

The research questions provided insights into the research trends over the past five years, anticipating where the research in e-assessment may be heading. First, higher education was the educational level with the highest number of studies (83.7%), and students were the most frequently addressed group (76.1%). Here, 104 out of 159 studies dealt with evaluation in higher education with students (65.4%), while only 15 (9.4%) did so with university teachers and 11 (6.9%) from a joint perspective. Therefore, students in higher education has been one of the most explored areas of research, while there were a small number of investigations focused on the preprimary, primary, and high school levels (Al-Emran et al., 2018).

Of these 104 studies that investigated e-assessment in higher education from the perspective of the students, 36 analyzed the results of e-assessment processes, 23 on the study of attitude and perception, and 12 on satisfaction. In contrast, only 11 studies focused on e-assessment with teachers, and the most repeated trends were the teaching attitude toward perspective (5), satisfaction (4), and acceptance (3). These can be potential areas of research in higher education.

7.2 Target population trends in the research studies achieved

The scarcity of publications has greatly conditioned the analysis of other levels of education. For example, the combination of compulsory schooling (primary and secondary) accounted for barely 10% of the total number of studies ($n = 17$), meaning that this was almost a new field in e-assessment and its development. Therefore, these fields still require more research and, as an example, can model the research carried out in higher education because the few studies have focused on the study

of attitudes toward e-assessment (7 of the 17) and on the results of e-assessment processes (5).

Additionally, this review has tried to synthesize the sample sizes of the current research to reflect the trend in the number of participants for each study and for each of the research areas. These sample sizes presented a large variability for each technique in the different groupings carried out, and factors such as the capacity of each research, the personnel available, and the funding largely conditioned the sample sizes of the research analyzed. On the other hand, polls and tests were the most used instruments used (91.8%) because of their rapid deployment and application among students and the ease of collecting and analyzing the results obtained in the digital environment (Pozzo et al., 2018).

7.3 Technological support: A key factor in the development of e-assessments

The review has shown a high variability in the tools used for e-assessment, which makes it difficult to synthesize and group some of them. The results have also shown the LMS as the predominant tool, approached in a generic way or through built-in tools such as Stack or WIRIS. At the same time, mobile technologies can be a new avenue of research at all educational levels (already reflected in higher education), given their possibilities and scarce research addressing them (Nikou & Economides, 2017a, b 2018, 2019).

As for the technological systems used for e-assessment in higher education, 79 of the 130 studies in this category analyzed the LMS systems of the institutions (60.8%), 37 on available tools and software (28.5%), and only 15 on mobile technologies (11.5%); thus, new research should focus on the study of mobile technologies (Alkiş & Coşkunçay, 2018; Sánchez-Prieto et al., 2017). This indicates the importance and widespread use of LMSs in education, which is also consistent with their key role during the pandemic as the main method of continuity and assurance of education (Dascalu et al., 2021; Grigoryeva et al., 2021; Raza et al., 2021).

Finally, the present review has confirmed the greater development of e-assessment research for formative assessments (McCallum & Milner, 2021; Mimirinis, 2019), as represented by the 42.8% of the studies included, compared with summative (19.6%) and mixed (18.8%). Therefore, the current research has focused on formative e-assessment developments, opening a new avenue for research in overcoming the limitations of these processes for summative assessments and expanding the field in this direction.

7.4 Postpandemic trends, limitations, and future avenues for research

The mapping has shown a clear trend and continuity of prepandemic research with research carried out during the pandemic, maintaining the same trends. It is important that future studies, once the pandemic has ended and full normality has been achieved, provide comparative frameworks for evaluation in virtual environments pre- and postpandemic and that they can analyze changes in publications, systems used, and study populations in the two sets of research.

Looking to the future, three lines of research stand out: innovation in terms of technology and means for online assessment (1), consolidation of the findings achieved so far through the replication of the study in different populations and platforms (2), and studies focused on technology adoption among educational agents (3).

In addition, the fact that only two studies in the past five years have proposed extending the theoretical justification of e-assessment shows that the research on e-assessment is no longer in its initial state but transitioning to a new phase and, hence, is starting to be a more consolidated practice in the educational fields, especially in higher education (Mastan et al., 2022). Also, it has been shown that the pedagogical design, innovation, and development of e-assessments are more important today than the theoretical explanation of it, so these must be the main lines and future challenges that this field must address in the coming years for the transformation of teaching in this new educational reality.

The mapping has also shown the importance of progressing in the research field to support this e-learning assessment modality. A recommendation is developing more specific systematic literature reviews in the future (e.g., a review for each technology detected here), in which a smaller number of publications would be included and a more thorough analysis of them could be carried out, thus obtaining more in-depth methods of advancing.

8 Limitations

It is essential to note the limitations. First, the number of databases selected (two in this case) should be highlighted because, although they are two of the most relevant in the field of research, they do not bring together all the existing publications on the subject in other databases, a limitation derived from the high volume of publications available in the field (García-Peñalvo, 2022).

Second, the time limit of the mapping (five years) does not clearly represent the evolution of the field. However, the high volume of publications indicates this time period was sufficient. Third, to reflect on the field's heterogeneity, the few classifications available on e-assessment to date and volume of publications analyzed may have generated problems in interpreting the data, for which the authors have generated specific analysis categories.

9 Conclusion

The current mapping has underlined the importance of the e-assessment field, trying to focus on the current state of research and synthesize the trends and lines currently addressed in the literature. In addition, this mapping has proposed analyzing the e-assessment transition during the pandemic and before, comprising years of normal education (2017–2019) and technology-mediated education (2020–2021). The present review sets the future direction of the field and advances in the return to educational normality, so it is recommended future

studies verify the progress of the e-assessment. Thus, to conclude, we would like to propose some possible research areas:

- **Higher education teachers:** The present review has shown the high volume of publications focused on higher education, but most focused on students. Therefore, future research can address the assessment from the perspective of university professors because studies in this population were quite scarce.
- **Other educational levels:** Similarly, the lower levels of education (secondary and primary education) or nonformal contexts were the focus of fewer studies, so future investigations may seek to explore these fields and their relationship with e-assessments.
- **Mobile technologies:** Currently, the research has mostly dealt with evaluation studies on Moodle and LMS platforms, as well as different software and educational tools based on the web. In contrast, far fewer studies have focused on assessment through mobile devices and electronic tablets (m-assessment) and through the specific applications for these devices. Therefore, the use of mobile technologies in electronic assessments opens a new avenue of research.
- **Acceptance, efficacy, and quality studies:** Most published studies on e-assessment have focused on the motivation, attitude, and results of students in these processes. The authors of the present research consider relevant to open new lines of research focused on technological acceptance models and the quality of these new assessment designs in this new educational reality.

Data availability The datasets generated during and/or analysed during the current study are available in the Google Sheets repository: bit.ly/3N8xBXf.

Declarations

Competing interests The authors have no relevant financial or nonfinancial interests to disclose.

References

- Abdou, R. A. E. (2020). Effects of e-assessment tools in academic achievement and motivation towards learning among preservice kindergarten teachers in Turaif. *Amazonia Investiga*, 9(28). <https://doi.org/10.34069/AI/2020.28.04.24>
- Al-Emran, M., Mezghuyev, V., & Kamaludin, A. (2018). Technology Acceptance Model in M-learning context: A systematic review. *Computers & Education*, 124, 389–412. <https://doi.org/10.1016/j.compedu.2018.06.008>
- Alkiş, N., & Coşkunçay, D. F. (2018). Mobile learning acceptance: A systematic literature review. *Erzincan University Journal of Education Faculty*, 20(2), 571–589.
- Almossa, S. Y. (2021). University students' perspectives toward learning and assessment during COVID-19. *Education and Information Technologies*, 26(6), 7163–7181. <https://doi.org/10.1007/s10639-021-10554-8>

- Alrofou, A., Lakulu, M. M., & Almaiah, M. (2019). A systematic review of mobile-based assessment acceptance studies from 2009 to 2019. *Journal of Theoretical and Applied Information Technology*, 97, 1–25.
- Alruwais, N. (2018). Advantages and challenges of using e-assessment. *International Journal of Information and Education Technology*, 8(1), 34–37. <https://doi.org/10.18178/ijiet.2018.8.1.1008>
- AlTameemy, F. A. Q., Alrefaee, Y., & Alalwi, F. S. S. (2020). *Using blackboard as a tool of e-assessment in testing writing skill in saudi arabia*. Retrieved January 1, 2023, from <http://repository.psau.edu.sa:80/jspui/handle/123456789/1229784>. Accessed 27 Jan 2023.
- Atoum, Y., Chen, L., Liu, A. X., Hsu, S. D. H., & Liu, X. (2017). Automated online exam proctoring. *IEEE Transactions on Multimedia*, 19(7), 1609–1624. <https://doi.org/10.1109/TMM.2017.2656064>
- Babo, R., Babo, L., Suhonen, J. T., & Tukiainen, M. (2020). E- Assessment with multiple-choice questions: A 5 Year study of students' opinions and experience. *Journal of Information Technology Education: Innovations in Practice*, 19, 001–029.
- Bacca-Acosta, J., & Avila-Garzon, C. (2021). Student engagement with mobile-based assessment systems: A survival analysis. *Journal of Computer Assisted Learning*, 37(1), 158–171. <https://doi.org/10.1111/jcal.12475>
- Balta, N., Perera-Rodríguez, V.-H., & Hervás-Gómez, C. (2018). Using Socrative as an online homework platform to increase students' exam scores. *Education and Information Technologies*, 23(2), 837–850. <https://doi.org/10.1007/s10639-017-9638-6>
- Bickerton, R., & Sangwin, C. (2021). Practical online assessment of mathematical proof. *International Journal of Mathematical Education in Science and Technology*, 0(0), 1–24. <https://doi.org/10.1080/0020739X.2021.1896813>
- Black, P. J. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, 21(1), 49–97. <https://doi.org/10.1080/03057269308560014>
- Boote, S. K., Boote, D. N., & Williamson, S. (2021). Assessing graph comprehension on paper and computer with MBA students: A crossover experimental study. *Cogent Education*, 8(1), 1960247. <https://doi.org/10.1080/2331186X.2021.1960247>
- Burgos-Videla, C. (2021). Innovación y aprendizajes flexibles en entornos formativos universitarios. *Innovación y Aprendizajes Flexibles En Entornos Formativos Universitarios*, 1–152.
- Cabero-Almenara, J., & Palacios-Rodríguez, A. (2021). La evaluación de la educación virtual: Las e-actividades. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2), Art. 2. <https://doi.org/10.5944/ried.24.2.28994>
- Cakiroglu, U., Erdogdu, F., Kokoc, M., & Atabay, M. (2017). Students' preferences in online assessment process: Influences on academic performances. *Turkish Online Journal of Distance Education*, 18(1), 132–142.
- Chacín, R. (2021). Los docentes y la evaluación en la modalidad virtual. *Interacciones*, 1(1), Art. 1. Retrieved March 15, 2023, from <https://p3.usal.edu.ar/index.php/interacciones/article/view/5531>. Accessed 15 Mar 2023.
- Cheng, X., Chan, L. K., Cai, H., Zhou, D., & Yang, X. (2021). Adaptions and perceptions on histology and embryology teaching practice in China during the Covid-19 pandemic. *Translational Research in Anatomy*, 24, 100115. <https://doi.org/10.1016/j.tria.2021.100115>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Cooper, I. D. (2016). What is a “mapping study?” *Journal of the Medical Library Association: JMLA*, 104(1), 76–78. <https://doi.org/10.3163/1536-5050.104.1.013>
- Crompton, H., Burke, D., Jordan, K., & Wilson, S. W. G. (2021). Learning with technology during emergencies: A systematic review of K-12 education. *British Journal of Educational Technology*, 52(4), 1554–1575. <https://doi.org/10.1111/bjet.13114>
- Cutumisu, M. (2018). The informational value of feedback choices for performance and revision in a digital assessment game. *Interactive Technology and Smart Education*, 15(4), 363–380. <https://doi.org/10.1108/ITSE-03-2018-0017>
- Daly, C., Pachler, N., Mor, Y., & Mellar, H. (2010). Exploring formative e-assessment: Using case stories and design patterns. *Assessment & Evaluation in Higher Education*, 35, 619–636. <https://doi.org/10.1080/02602931003650052>
- Dascalu, M.-D., Ruseti, S., Dascalu, M., McNamara, D. S., Carabas, M., Rebedea, T., & Trausan-Matu, S. (2021). Before and during COVID-19: A Cohesion Network Analysis of students' online participation in moodle courses. *Computers in Human Behavior*, 121, 106780. <https://doi.org/10.1016/j.chb.2021.106780>

- Delaney, D., Kummer, T.-F., & Singh, K. (2019). Evaluating the impact of online discussion boards on student engagement with group work. *British Journal of Educational Technology*, *50*(2), 902–920. <https://doi.org/10.1111/bjet.12614>
- DeMara, R., Chen, B., Hartshorne, R., & Thripp, R. (2017). Elevating participation and outcomes with computer-based assessments: An immersive development workshop for engineering faculty. *Computers in Education Journal*, *8*, 1–12.
- Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, *40*(2), 203–214. <https://doi.org/10.1111/j.1467-8535.2008.00915.x>
- Dixson, D. D., & Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory into Practice*, *55*(2), 153–159. <https://doi.org/10.1080/00405841.2016.1148989>
- Doğan, N., Kibrislioglu Uysal, N., Kelecioğlu, H., & Hambleton, R. (2020). An Overview of E-Assessment. *35*, 1–5. <https://doi.org/10.16986/HUJE.2020063669>
- Domingo-Coscollola, M., Bosco, A., Carrasco-Segovia, S., & Sánchez, J. A. (2020). Fomentando la competencia digital docente en la universidad: Percepción de estudiantes y docentes. <https://doi.org/10.6018/rie.340551>
- Dorris, C., Winter, K., O'Hare, L., & Lwoga, E. T. (2021). PROTOCOL: A systematic review of mobile device use in the primary school classroom and its impact on pupil literacy and numeracy attainment. *Campbell Systematic Reviews*, *17*(2). <https://doi.org/10.1002/cl2.1155>
- Dutton, P. J., Bickerstaff, H. E., Rymer, J. M., Webb, M. E., Ballinger-Mills, D., Greenough, A., & Reynolds, P. A. (2017). Investigation of Formative Assessment of Learning (INFORMAL): The Performance Indicator Tool (PIT). *Technology, Knowledge and Learning*, *22*(2), 161–171. <https://doi.org/10.1007/s10758-017-9307-2>
- Egarter, S., Mutschler, A., & Brass, K. (2021). Impact of COVID-19 on digital medical education: Compatibility of digital teaching and examinations with integrity and ethical principles. *International Journal for Educational Integrity*, *17*. <https://doi.org/10.1007/s40979-021-00084-8>
- Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & Education*, *106*, 83–96. <https://doi.org/10.1016/j.compedu.2016.12.001>
- Forero-Peña, D. A., Carrión-Nessi, F. S., Camejo-Ávila, N. A., & Forero-Peña, M. J. (2020). COVID-19 en Latinoamérica: Una revisión sistemática de la literatura y análisis bibliométrico. *Revista de Salud Pública*, *22*(2), 246–252. <https://doi.org/10.15446/rsap.v22n2.86878>
- García-Aretio, L. (2021). ¿Podemos fiarnos de la evaluación en los sistemas de educación a distancia y digitales? *RIED. Revista Iberoamericana de Educación a Distancia*, *24*(2), Art. 2. <https://doi.org/10.5944/ried.24.2.30223>
- García-Peñalvo, F. J. (2021). Transformación digital en las universidades: Implicaciones de la pandemia de la COVID-19. *Education in the Knowledge Society: EKS*, *22*, 1–6. <https://doi.org/10.14201/eks.25465>
- García-Peñalvo, F. J. (2022). Developing robust state-of-the-art reports: Systematic Literature Reviews. *Education in the Knowledge Society*, *23*. <https://doi.org/10.14201/eks.28600>
- Geomar, E. (2020). Percepciones de los estudiantes pre-universitarios sobre el perfil del profesor virtual. *Conference proceedings CIVINEDU 2020: 4th International Virtual Conference on Educational Research and Innovation September 23–24, 2020, 2020, ISBN 978–84–09–22966–6*, 271–272. Retrieved March 15, 2023, from <https://dialnet.unirioja.es/servlet/articulo?codigo=7837919>. Accessed 15 Mar 2023.
- Ghouali, K., Benmoussat, S., & Ruiz-Cecilia, R. (2020). E-assessment on the spotlight: Present and future prospects. *ReiDoCrea: Revista electrónica de Investigación Docencia Creativa*, *9*, 52–62. <https://doi.org/10.30827/Digibug.59151>
- González-González, R. A., & Silveira-Bonilla, M. H. (2022). Educación e inteligencia artificial: Nodos temáticos de inmersión. *EduTec. Revista Electrónica de Tecnología Educativa*, *82*, Article 82.
- Grande-de-Prado, M., García-Peñalvo, F. J., Almuzara, A. C., & Abella-García, V. (2021). Evaluación en Educación Superior durante la pandemia de la COVID-19. *Campus Virtuales*, *10*(1), Article 1.
- Grigoryeva, N. V., Melikov, I. M., Palanchuk, N. V., Kokhanovskaya, I. I., & Aralova, E. (2021). Opportunities for organizing distance learning presented by the moodle platform: experience in the conditions of the COVID-19 pandemic. *Propósitos y Representaciones*, *9*, e1259. <https://doi.org/10.20511/pyr2021.v9nSPE3.1259>
- Guerrero-Roldán, A.-E., & Noguera, I. (2018). A model for aligning assessment with competences and learning activities in online courses. *The Internet and Higher Education*, *38*, 36–46. <https://doi.org/10.1016/j.iheduc.2018.04.005>

- Gupta, A., Gupta, K., Joshi, A., & Sharma, D. (2019). Tools for E-Assessment Techniques in Education: A Review. En *Handbook of Research on E-Assessment in Higher Education* (pp. 28–52). IGI Global. <https://doi.org/10.4018/978-1-5225-5936-8.ch002>
- Hassan, B., Shati, A. A., Alamri, A., Patel, A., Asseri, A. A., Abid, M., Al-Qahatani, S. M., & Satti, I. (2020). Online assessment for the final year medical students during covid-19 pandemics; the exam quality and students' performance. *Onkologia i Radioterapia*, 1–6.
- Heil, J., & Ifenthaler, D. (2023). Online assessment in higher education: A systematic review. *Online Learning*, 27(1). <https://doi.org/10.24059/olj.v27i1.3398>
- Humanante-Ramos, P., Fernández-Acevedo, J., & Jiménez, C. (2019). Aulas virtuales en contextos universitarios: Percepciones de uso por parte de los estudiantes. *Revista ESPACIOS*, 40(2), 3–20
- Ifanov, Jessica, P., Salim, S., Syahputra, M. E., & Suri, P. A. (2023). A Systematic literature review on implementation of virtual reality for learning. *Procedia Computer Science*, 216, 260–265. <https://doi.org/10.1016/j.procs.2022.12.135>
- Jaiswal, P. (2020). Integrating educational technologies to augment learners' academic achievements. *International Journal of Emerging Technologies in Learning (IJET)*, 15(02), Article 02. <https://doi.org/10.3991/ijet.v15i02.11809>
- Jiménez, Y. I., Hernández, J., & Rodríguez, E. (2021). Educación en línea y evaluación del aprendizaje: De lo presencial a lo virtual. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 12(23). <https://doi.org/10.23913/ride.v12i23.1005>
- Judge, M. (2021). Covid 19, school closures and the uptake of a digital assessment for learning pilot project during Ireland's national lockdown. *Irish Educational Studies*, 40(2), 419–429. <https://doi.org/10.1080/03323315.2021.1917443>
- Kitchenham, B., Budgen, D., & Brereton, P. (2011). Using mapping studies as the basis for further research: A participant-observer case study. *Information & Software Technology*, 53, 638–651. <https://doi.org/10.1016/j.infsof.2010.12.011>
- Khlaisang, J., & Koraneekij, P. (2019). Open online assessment management system platform and instrument to enhance the information, media, and ICT literacy skills of 21st century learners. *International Journal of Emerging Technologies in Learning (iJET)*, 14, 111. <https://doi.org/10.3991/ijet.v14i07.9953>
- Knight, P. T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education*, 27(3), 275–286. <https://doi.org/10.1080/03075070220000662>
- Kotiash, I., Shevchuk, I., Borysonok, M., Matviienko, I., Popov, M., Terekhov, V., & Kuchai, O. (2022). *Possibilities of Using Multimedia Technologies in Education*. 22, 727-732. <https://doi.org/10.22937/IJCSNS.2022.22.6.91>
- Kühbeck, F., Berberat, P. O., Engelhardt, S., & Sarikas, A. (2019). Correlation of online assessment parameters with summative exam performance in undergraduate medical education of pharmacology: A prospective cohort study. *BMC Medical Education*, 19(1), 412. <https://doi.org/10.1186/s12909-019-1814-5>
- Kundu, A., & Bej, T. (2021). Experiencing e-assessment during COVID-19: An analysis of Indian students' perception. *Higher Education Evaluation and Development*, 15(2), 114–134. <https://doi.org/10.1108/HEED-03-2021-0032>
- Lai, J., & Bower, M. (2019). How is the use of technology in education evaluated? A systematic review. *Computers & Education*, 133. <https://doi.org/10.1016/j.compedu.2019.01.010>
- Lee, K., & Fanguy, M. (2022). Online exam proctoring technologies: Educational innovation or deterioration? *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13182>
- Lin, J.-W., Tsai, C.-W., & Hsu, C.-C. (2020). A comparison of computer-based and game-based formative assessments: A long-term experiment. *Interactive Learning Environments*, 1–17. <https://doi.org/10.1080/10494820.2020.1815219>
- Liu, Y., Béliveau, A., Besche, H., Wu, A. D., Zhang, X., Stefan, M., Gutlerner, J., & Kim, C. (2021). Bayesian mixed effects model and data visualization for understanding item response time and response order in open online assessment. *Frontiers in Education*, 5. <https://doi.org/10.3389/educ.2020.607260>
- Marshall, C., & Brereton, P. (2013). Tools to Support Systematic Literature Reviews in Software Engineering: A Mapping Study. *Conference: 7th International Symposium on Empirical Software and Measurement (ESEM 2013)*, 296–299. <https://doi.org/10.1109/ESEM.2013.32>
- Mastan, I., Sensuse, D., Suryono, R., & Kautsarina, K. (2022). Evaluation of distance learning system (e-learning): A systematic literature review. *Jurnal Teknoinfo*, 16, 132. <https://doi.org/10.33365/jti.v16i1.1736>
- Martínez-Soto, T., & Prendes-Espinosa, P. (2023). A systematic review on the role of ICT and CLIL in compulsory education. *Education Sciences*, 13(1). <https://doi.org/10.3390/educsci13010073>

- McCallum, S., & Milner, M. M. (2021). The effectiveness of formative assessment: Student views and staff reflections. *Assessment & Evaluation in Higher Education*, 46(1), 1–16. <https://doi.org/10.1080/02602938.2020.1754761>
- Mellar, H., Peytcheva-Forsyth, R., Kocdar, S., Karadeniz, A., & Yovkova, B. (2018). Addressing cheating in e-assessment using student authentication and authorship checking systems: Teachers' perspectives. *International Journal for Educational Integrity*, 14(1), Art. 1. <https://doi.org/10.1007/s40979-018-0025-x>
- Mimirinis, M. (2019). Qualitative differences in academics' conceptions of e-assessment. *Assessment & Evaluation in Higher Education*, 44(2), 233–248. <https://doi.org/10.1080/02602938.2018.1493087>
- Misut, M., & Misutova, M. (2017). Software solution improving productivity and quality for big volume students' group assessment process. *International Journal of Emerging Technologies in Learning (IJET)*, 12(04), Art. 04. <https://doi.org/10.3991/ijet.v12i04.6608>
- Montenegro-Rueda, M., Luque-de la Rosa, A., Sarasola Sánchez-Serrano, J. L., & Fernández-Cerero, J. (2021). Assessment in higher education during the COVID-19 pandemic: A systematic review. *Sustainability*, 13(19), Article 19. <https://doi.org/10.3390/su131910509>
- Morales, Y. A. R., García, M. G., & López, C. B. (2018). Evaluación de competencias genéricas en la universidad. Estudio comparativo en entorno b-learning y presencial. *Estudio comparativo en Acción Pedagógica*, 27(1), 6–21.
- Muñoz-Guevara, E., Velázquez-García, G., & Barragán-López, J. F. (2021). Análisis sobre la evolución tecnológica hacia la Educación 4.0 y la virtualización de la Educación Superior. *Transdigital*, 2(4), Article 4. <https://doi.org/10.56162/transdigital86>
- Muhaimin, M., Habibi, A., Riady, Y., Alqahtani, T. M., Chaerunisaa, A. Y., Wijaya, T. T., Milanda, T., Yusop, F. D., & Albelbisi, N. A. (2023). Covid-19 distance and online learning: A systematic literature review in pharmacy education. *BMC Medical Education*, 23(1), 367. <https://doi.org/10.1186/s12909-023-04346-6>
- Nasab, F. G. (2015). Alternative versus traditional assessment. *Journal of Applied Linguistics and Language Research*, 2(6), Article 6.
- Nguyen, Q., Rienties, B., Toetenel, L., Ferguson, R., & Whitelock, D. (2017). Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates. *Computers in Human Behavior*, 76, 703–714. <https://doi.org/10.1016/j.chb.2017.03.028>
- Nikou, S. A., & Economides, A. A. (2018). Mobile-Based micro-Learning and Assessment: Impact on learning performance and motivation of high school students. *Journal of Computer Assisted Learning*, 34(3), 269–278. <https://doi.org/10.1111/jcal.12240>
- Nikou, S., & Economides, A. (2017a). Mobile-Based Assessment: Integrating acceptance and motivational factors into a combined model of Self-Determination Theory and Technology Acceptance. *Computers in Human Behavior*, 68, 83–95. <https://doi.org/10.1016/j.chb.2016.11.020>
- Nikou, S., & Economides, A. (2017b). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 109, 56–73. <https://doi.org/10.1016/j.compedu.2017.02.005>
- Nikou, S., & Economides, A. (2019). A comparative study between a computer-based and a mobile-based assessment: Usability and user experience. *Interactive Technology and Smart Education*. <https://doi.org/10.1108/ITSE-01-2019-0003>
- Ninković, S., Olić Ninković, S., Lazarević, T., & Adamov, J. (2021). Serbian teachers' perceptions of online assessment during COVID-19 school closure: The role of teachers' self-efficacy. *Educational Studies*, 0(0), 1–13. <https://doi.org/10.1080/03055698.2021.1960151>
- Nyland, R., Davies, R. S., Chapman, J., & Allen, G. (2017). Transaction-level learning analytics in online authentic assessments. *Journal of Computing in Higher Education*, 29(2), 201–217. <https://doi.org/10.1007/s12528-016-9122-0>
- Okada, A., Noguera, I., Alexieva, L., Rozeva, A., Kocdar, S., Brouns, F., Ladonlahti, T., Whitelock, D., & Guerrero-Roldán, A.-E. (2019a). Pedagogical approaches for e-assessment with authentication and authorship verification in Higher Education. *British Journal of Educational Technology*, 50(6), 3264–3282. <https://doi.org/10.1111/bjet.12733>
- Okada, A., Whitelock, D., Holmes, W., & Edwards, C. (2019b). e-Authentication for online assessment: A mixed-method study. *British Journal of Educational Technology*, 50(2), 861–875. <https://doi.org/10.1111/bjet.12608>
- Ozdamli, F., & Karagozlu, D. (2022). Online education during the pandemic: A systematic literature review. *International Journal of Emerging Technologies in Learning (IJET)*, 17(16), Article 16. <https://doi.org/10.3991/ijet.v17i16.32287>

- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ..., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(372). <https://doi.org/10.1136/bmj.n71>
- Palacios, C., Rubio, M., & Arce, M. (2021). Percepciones y experiencias de los estudiantes de Odontología en educación en línea: Una revisión bibliográfica. *Applied Sciences in Dentistry*, 1(0), Art. 0. <https://doi.org/10.22370/asd.2020.1.0.2627>
- Pearse-Romera, C. R., & Ruiz-Cecilia, R. (2019). A motivation case study of students learning english at a secondary school in Granada, Spain. *The International Journal of Interdisciplinary Educational Studies*, 14(1), 31–45. <https://doi.org/10.18848/2327-011X/CGP/v14i01/31-45>
- Pozzo, M. I., Borgobello, A., & Pierella, M. P. (2018). Uso de cuestionarios en investigaciones sobre universidad: Análisis de experiencias desde una perspectiva situada. *Revista Latinoamericana de Metodología de las Ciencias Sociales*, 8(2), e046. Retrieved March 16, 2023, from <https://www.memoria.fahce.unlp.edu.ar/library?a=d&c=arti&d=Jpr9347>. Accessed 16 Mar 2023.
- Pu, S., & Xu, H. (2021). Examining changing assessment practices in online teaching: A multiple-case study of EFL school teachers in China. *The Asia-Pacific Education Researcher*, 30(6), 553–561. <https://doi.org/10.1007/s40299-021-00605-6>
- Ramírez, L. N. R. (2020). Tendencias de la innovación educativa en los contextos sociales. Análisis del mapeo de literatura. *Revista Educación*, 44(1), 381–398. <https://doi.org/10.15517/revedu.v44i1.33222>
- Raza, S. A., Qazi, W., Khan, K. A., & Salam, J. (2021). Social isolation and acceptance of the Learning Management System (LMS) in the time of COVID-19 pandemic: An expansion of the UTAUT model. *Journal of Educational Computing Research*, 59(2), 183–208. <https://doi.org/10.1177/0735633120960421>
- Riera, J., Ardid, M., Gómez, J. A., Vidaurre, A., & Meseguer, J. M. (2018). Students' perception of auto-scored online exams in blended assessment: Feedback for improvement. *Educación XXI: Revista de La Facultad de Educación*, 21(2), 79–103. <https://doi.org/10.5944/educxx1.19559>
- Rolim, C., & Isaias, P. (2018). Examining the use of e-assessment in higher education: Teachers and students' viewpoints. *British Journal of Educational Technology*, 50. <https://doi.org/10.1111/bjet.12669>
- Ruiz-Morales, Y. (2013). Evaluación de competencias genéricas en la Universidad. Estudio comparativo en entorno b-learning y presencial. Tesis [Tesis Doctoral]. Universidad Complutense de Madrid.
- Samala, A. D., Usmeldi, Taali, Ambiyar, Bojic, L., Indarta, Y., Tsoy, D., Denden, M., Tas, N., & Dewi, I. P. (2023). Metaverse technologies in education: A systematic literature review using PRISMA. *International Journal of Emerging Technologies in Learning (IJET)*, 18(5). <https://doi.org/10.3991/ijet.v18i05.35501>
- Sánchez-Prieto, J. C., Olmos-Migueláñez, S., & García-Peñalvo, F. J. (2017). MLearning and pre-service teachers: An assessment of the behavioral intention using an expanded TAM model. *Computers in Human Behavior*, 72, 644–654. <https://doi.org/10.1016/j.chb.2016.09.061>
- Schwartzman, G., Roni, C., Berk, M., Delorenzi, E., Sánchez, M., & Eder, M. L. (2021). Evaluación Remota de Aprendizajes en la Universidad: Decisiones docentes para encarar un nuevo desafío. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2), Art. 2. <https://doi.org/10.5944/ried.24.2.29078>
- Sellés, N. H., Carril, P. C. M., & Sanmamed, M. G. (2018). La e-evaluación en el trabajo colaborativo en entornos virtuales: Análisis de la percepción de los estudiantes. *Edutec. Revista Electrónica de Tecnología Educativa*, 65, Art. 65. <https://doi.org/10.21556/edutec.2018.65.997>
- Sembey, R., Hoda, R., & Grundy, J. (2023). Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *SSRN Electronic Journal*, 1–21. <https://doi.org/10.2139/ssrn.4328075>
- Seraj, P. M., Chakraborty, R., Mehdi, T., & Roshid, M. M. (2022). A systematic review on pedagogical trends and assessment practices during the COVID-19 pandemic: Teachers' and students' perspectives. *Education Research International*, 2022, e1534018. <https://doi.org/10.1155/2022/1534018>
- Sirianni, J. M., Ng, Y. J., & Vishwanath, A. (2017). Adopting computer-based assessments: The role of perceived value in classroom technology acceptance. *Online Journal of Communication and Media Technologies*, 7(4), 1–23. <https://doi.org/10.29333/ojcm/2607>
- Slade, C., Lawrie, G., Taptamat, N., Browne, E., Sheppard, K., & Matthews, K. E. (2022). Insights into how academics reframed their assessment during a pandemic: Disciplinary variation and

- assessment as afterthought. *Assessment & Evaluation in Higher Education*, 47(4), 588–605. <https://doi.org/10.1080/02602938.2021.1933379>
- Smith, C. A. (2021). Development and integration of freely available technology into online stem courses to create a proctored environment during exams. *Journal of Higher Education Theory and Practice*, 21(5). <https://doi.org/10.33423/jhetp.v21i5.4274>
- Snekalatha, S., Marzuk, S. M., Meshram, S. A., Maheswari, K. U., Sugapriya, G., & Sivasharan, K. (2021). Medical students' perception of the reliability, usefulness, and feasibility of unproctored online formative assessment tests. *Advances in Physiology Education*, 45(1), 84–88. <https://doi.org/10.1152/advan.00178.2020>
- St-Onge, C., Ouellet, K., Lakhali, S., Dubé, T., & Marceau, M. (2022). COVID-19 as the tipping point for integrating e-assessment in higher education practices. *British Journal of Educational Technology*, 53(2), 349–366. <https://doi.org/10.1111/bjet.13169>
- Stanković, J., Milovanović, S., & Radović, O. (2017). Applying the moodle platform in online student self-assessment. *Economic Themes*, 55(2), 281–304. <https://doi.org/10.1515/ethemes-2017-0016>
- Szymkowiak, A., Melović, B., Dabić, M., Jeganathan, K., & Kundi, G. S. (2021). Information technology and Gen Z: The role of teachers, the internet, and technology in the education of young people. *Technology in Society*, 65, 101565. <https://doi.org/10.1016/j.techsoc.2021.101565>
- Tomas, C., Borg, M., & McNeil, J. (2015). E-assessment: Institutional development strategies and the assessment life cycle. *British Journal of Educational Technology*, 46(3), 588–596. <https://doi.org/10.1111/bjet.12153>
- Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2015). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal*, 42(3). <https://doi.org/10.1002/berj.3215>
- Vázquez, J. J., Chiang, E. P., & Sarmiento-Barbieri, I. (2021). Can we stay one step ahead of cheaters? A field experiment in proctoring online open book exams. *Journal of Behavioral and Experimental Economics*, 90, 101653. <https://doi.org/10.1016/j.socec.2020.101653>
- Wang, Z., Gong, S.-Y., Xu, S., & Hu, X.-E. (2019). Elaborated feedback and learning: Examining cognitive and motivational influences. *Computers & Education*, 136, 130–140. <https://doi.org/10.1016/j.compedu.2019.04.003>
- Washburn, S., Herman, J., & Stewart, R. (2017). Evaluation of performance and perceptions of electronic vs. Paper multiple-choice exams. *Advances in Physiology Education*, 41, 548–555. <https://doi.org/10.1152/advan.00138.2016>
- Witchel, H. J., Guppy, J. H., & Smith, C. F. (2018). The self-assessment dilemma: An open-source, ethical method using Matlab to formulate multiple-choice quiz questions for online reinforcement. *Advances in Physiology Education*, 42(4), 697–703. <https://doi.org/10.1152/advan.00081.2018>
- Wu, P., & Wang, Y. (2021). Investigating business english teachers' belief about online assessment: Q methodology conducted during COVID-19 period. *The Asia-Pacific Education Researcher*, 30(6), 621–630. <https://doi.org/10.1007/s40299-021-00604-7>
- Zhang, C., Yan, X., & Wang, J. (2021). EFL teachers' online assessment practices during the COVID-19 pandemic: Changes and mediating factors. *The Asia-Pacific Education Researcher*, 30(6), 499–507. <https://doi.org/10.1007/s40299-021-00589-3>
- Zharova, M. V., Trapitsin, S. Yu., Timchenko, V. V., & Skurikhina, A. I. (2020). Problems and Opportunities of Using LMS Moodle before and during COVID-19 Quarantine: Opinion of Teachers and Students. *2020 International Conference Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS)*, 554–557. <https://doi.org/10.1109/ITQMIS51053.2020.9322906>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Anexo 4

Publicación en revista científica: Journal of New Approaches in Educational Research

VERSIÓN PREPRINT

Ortiz-López, A., Sánchez-Prieto, J. C. & Olmos-Migueláñez, S. (En prensa). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM, *Journal of New Approaches in Educational Research*

INDICADORES:

Revista Q1 – SJR en último índice publicado (2022): Factor de impacto 1,034

Revista Q1 – JCI en último índice publicado (2022): Factor de impacto 2,63. Pos. 17/742. Percentil 97,78

Incluido en índice JIF en 2023, factor de impacto 4,5. Se asignará cuartil en 2024

Article Acceptance Certificate

This certificate confirms that the following paper has been accepted for publication in:

**Journal of New Approaches in Educational Research (2023),
13(1).**

Journal e-ISSN: 2254-7339

Title: Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM

Authors: Alberto Ortiz-López, José Carlos Sánchez-Prieto, Susana Olmos-Migueláñez

Received: 2023-05-24

Accepted: 2023-07-19



Rosabel Roig, Editor

Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM

Alberto Ortiz-López^{1*}, José Carlos Sánchez-Prieto², Susana Olmos-Migueláñez³

¹ Instituto Universitario de Ciencias de la Educación (IUCE), Grupo GRIAL, Universidad de Salamanca, Spain {aortiz@usal.es}

² Instituto Universitario de Ciencias de la Educación (IUCE), Grupo GRIAL, Universidad de Salamanca, Spain {josecarlos.sp@usal.es}

³ Instituto Universitario de Ciencias de la Educación (IUCE), Grupo GRIAL, Universidad de Salamanca, Spain {solmos@usal.es}

Instituto Universitario de Ciencias de la Educación, Paseo de Canalejas, 169, 37008, Salamanca, España.

ABSTRACT

The use of digital media in education has already been addressed in numerous technology acceptance models, but there is very little research on establishing a link between acceptance and assessment using mobile devices, a reality in educational institutions. This work aims to extend research by developing the TAM model and studying teachers' perceived usefulness of mobile devices in terms of how they understand assessment: generically, as a summative and a formative assessment, or as the complementarity of these. This study proposes a comparison between three models using the partial least squares structural equation modeling (PLS-SEM) on a sample of 262 master's degree students (pre-service teachers). The results show the validity of the three proposals and confirm the advantages to specifically consider assessment in acceptance models, as well as the importance of addressing its modalities differently after obtaining better results in the two models that do so. The study also confirms the importance of self-efficacy in the use of mobile devices as a predictor of usefulness and intention to use in the three models. The use of a comparative approach and the development of the perceived usefulness construct in assessment represents a new contribution to the field of acceptance studies.

KEYWORDS: EVALUATION METHODS, TECHNOLOGY ASSESSMENT, MATHEMATICAL MODELS, EDUCATIONAL TECHNOLOGY, HIGHER EDUCATION

1 INTRODUCTION

Information and Communication Technologies (ICTs) are nowadays an essential tool in the teaching-learning processes in all educational institutions, from the lowest levels to higher education (Burbules et al., 2020; Valverde-Berrocoso et al., 2021). This digital inclusion has facilitated access to an unprecedented variety of resources in education, leading to the creation of new virtual learning environments (Gómez-Galán, 2020) and the development of new methodologies, such as collaborative work or flipped classrooms (Clark et al., 2022; Sun

et al., 2022; Ciobanu, 2022; Hossain et al., 2021), especially after the pandemic (Sar & Misra, 2020).

Several studies on assessment in digital environments attempt to define assessment frameworks (Gómez-Ruiz et al., 2022; Moccozet et al., 2019) and present a large number of technological tools for this process (Cosi et al., 2020), with mobile devices as one of the main research lines in the development of teaching (Guardia et al., 2019; Moreno-Guerrero et al., 2020), given their universality and potential (Abd-Karim et al., 2018; Criollo et al., 2021). The advantages of these devices in the educational context have been summarised as portability, interactivity, context sensitivity, connectivity, and individuality (Ally & Prieto-Blázquez, 2014). For Al-Emran et al. (2020), their use is an opportunity for ubiquitous access to media and resources, enriched content, new means of interaction and collaboration, and new methodological and assessment procedures (Aljawarneh, 2020; Baier & Kunter, 2020; Reisoğlu & Çebi, 2020).

Therefore, it is necessary to deepen the concept of assessment in education, since it implies describing two modalities in its process, formative assessment, and summative assessment (Black, 1993; Dixon & Worrell, 2016; Scriven, 1967). Consequently, it is possible that, if a teacher has a different understanding of these modalities, their perspective on the use of mobile devices may also differ depending on the assessment made; a distinction that should be considered in the research.

The described potential of mobile devices contrasts, in many cases, with the reluctance of teachers to introduce innovations in a process as rigorous as assessment (Hébert et al., 2021). These tools have been extensively validated in the introduction of new classroom methodologies (Bernacki et al., 2020; Marín-Díaz et al., 2022; Vieira & Ribeiro, 2018), or as supplementary and substitute teaching materials (Wan-Sulaiman & Mustafa, 2020); but there are still very few initiatives that address the use of mobile devices in assessment, a practice that is currently not widespread (Nikou & Economides, 2016).

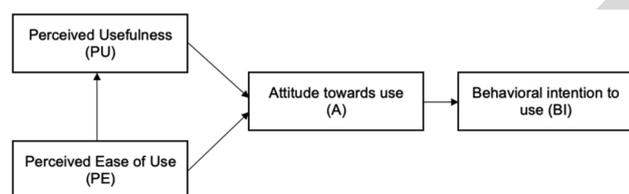
*aortiz@usal.es

* Instituto Universitario de Ciencias de la Educación, Paseo de Canalejas, 169, 37008, Salamanca, España.

It should also be considered that the ultimate decision on the use of technological devices for assessment in their classrooms, whether face-to-face or in technology-mediated instruction, rests with the teachers (Olimov, 2021). The knowledge of the factors influencing the teacher's decision to use technology and to know the current state of its technological acceptance will contribute to the establish appropriate training policies, based on an effective and real diagnosis, which will allow teachers to be trained in those aspects in which they face limitations for the inclusion of technology in the assessment process (García-Aretio, 2019).

The determination of this acceptance and its factors is nowadays addressed through different technology acceptance models described in the literature and validated in practice, with the TAM model (Davis, 1989; Fishbein & Ajzen, 1975). Regarding the variables that the TAM model includes (Figure 1), Davis (1989) described the behavioural intention to use through the attitude toward using and stated that the factors that determine them are usefulness, understood as "the degree to which a person believes that using a particular system would enhance their job performance" (p. 320) and perceived ease of use, defined as "the degree to which a person believes that using a particular system would be free from effort" (p. 320).

Figure 1. Technology Acceptance Model (Davis, 1989)



The TAM model has been widely discussed in studies of technology acceptance of teaching in virtual environments (Bizzo, 2021; Cruz-Benito et al., 2019; Terán-Guerrero, 2019) and in teaching via mobile devices (Hao et al., 2017; Mutambara & Bayaga, 2021; Sánchez-Prieto et al., 2017b), but few studies have addressed the technology acceptance of assessment using mobile devices. Among these models, Nikou y Economides (2017b, 2017a) are the most relevant authors in the study of the field with their proposal of the Mobile-Based Assessment Acceptance Model (MBAAM). This model, which is based on TAM, suggests the influence of ten factors that try to predict the perceived ease of use and usefulness, thus increasing Davis' (1989) explanation of the intention of use in this field.

Research on the acceptance of technology has mainly been carried out on active teachers (Sánchez-Prieto et al., 2019), hence this study proposes to explore the acceptance of technology in the group of future teachers, as they will be the main actors in these new e-learning or m-learning frameworks (Anisimova et al., 2020; Castañeda-Vázquez et al., 2019; Sánchez-Prieto et al., 2019).

Therefore, this research aims to study perceived usefulness and its current definition in this field. Given that the TAM model proposes a generic perceived usefulness, this proposal is to reformulate the construct through two new models that specifically include formative and summative assessments, since, as indicated above, the use of mobile devices may vary depending on the type of assessment to be carried out. The research questions to be addressed are as follows:

- Which of the proposed perceived usefulness explains a higher intention to use mobile devices in e-assessment?
- What are the implications of formative and summative assessments for the construct of perceived usefulness?

A review of the literature and the proposed models are provided in section 2. The methodology is presented in section 3, and the results are in section 4. The discussion of the results and the conclusions are included in section 5 and 6.

1.1 Literature review and models development

In educational practice, assessment is conducted from a dual perspective depending on its purpose, distinguishing between summative and formative assessments (Scriven, 1967). In general, the summative assessment focuses on the final product and accountability, being a form of assessment aimed at obtaining a final grade (Knight, 2002). In contrast, the formative assessment is focused on monitoring the process and all the factors that influence it, being more oriented toward feedback and continuous improvement (Morris et al., 2021).

The debate in research on these assessment modalities has generated three different currents, one assuming the non-differentiation between formative and summative assessments, the second one assuming their existence but independence, and the third one assuming their necessary complementarity for effective assessment (Lau, 2016). The three conceptions and their implications for the assessment process are presented below.

On the one hand, the first current describes assessment in general terms as a single whole, with no distinction between process and product, and understanding its importance in terms of objectives and effectiveness. (Mejía-Pérez, 2012). This current corresponds to the classical conception of assessment, without distinguishing between formative and summative modalities or exploring their possibilities (Shepard, 2006; Tyler, 1950).

On the other hand, the second current assumes their independence and is in favour of differentiating between the effects of the two on learning (Harlen & James, 1997; Patton, 1996), treating both modalities independently and separately (P. Black & Wiliam, 1998). This is consistent with an established idea in assessment, which distinguishes between the effects of the two and considers that a new combined model may be a problem for teachers in the search for the combination of different techniques, objectives, and goals (MacLellan, 2001).

Finally, the third approach considers these assessments jointly, implying that the whole assessment process is considered as the sum of formative and summative assessments by a teacher (Ahmad & Bhat, 2019; Buchholtz et al., 2018), as their purposes are fully related and complementary for them in a complete process (P. Black et al., 2003) mediated by feedback and guidance aimed at the consecution of achievements and their final certification (Dixson & Worrell, 2016; Dolin et al., 2018).

1.2 Perceived Usefulness Modelling: A Triple Comparison

This study reports three TAM-based models with perceived usefulness defined from the three assessment perspectives to find out which one explains, to a greater extent, the perceived usefulness of mobile devices in e-assessment for teachers (Horvat et al., 2012). These models will reveal how most teachers understand assessment, which will provide insight into the actual usefulness of assessment and how best to identify it in acceptance models and may lead to the analysis of the implications for teacher education and training.

1.2.1. Model A: TAM: Perceived Usefulness as a generic construct

The first model understands assessment from a classical perspective (Tyler, 1950), assuming the generic perceived usefulness of the TAM model for its study. This usefulness means that if teachers perceive mobile devices as useful for assessment, they will be more likely to use them in their classrooms (Davis, 1989). According to this assumption, it is proposed that:

- Perceived usefulness (PU) positively predicts teachers' intention to use (BIU) mobile devices for e-assessment in future teaching practice (MAH1).

A second predictor of the intention to use mobile devices is, according to TAM, the perceived ease of use, which assumes that all professionals who innovate in their teaching must make an additional effort in terms of technical knowledge and training, taking on a greater workload for effective inclusion (Thorsteinsson & Niculescu, 2013). This research focuses on teachers in training, a group that is assumed to have greater ease of use of mobile devices given their native digital competence (Evans & Robertson, 2020), but, since this is not true in all scenarios, the assessment of this construct in the new models is recommended (Kimmons et al., 2017). Therefore, it is proposed that:

- Perceived usefulness (PU) positively predicts teachers' intention to use (BIU) mobile devices for e-assessment in future teaching practice (MAH2).

Furthermore, in the early stages of adopting technological tools, perceived ease of use can become an internal barrier to conditions both perceived usefulness and intention to use them (Sánchez-Prieto et al., 2019; Venkatesh & Bala, 2008). That is if teachers find it difficult to use a mobile phone for assessment, not only will it condition the final intention of use, but they will not find it useful in their teaching practice (Al-Gasawneh et al., 2022). The following is therefore proposed:

- Perceived ease of use (PEOU) positively predicts teachers' perceived usefulness (PU) of mobile devices for e-assessment in their future teaching practice (MAH3).

The constructs of perceived usefulness, perceived ease of use, and behavioural intention from the TAM model are preserved in this study, but the construct of attitude toward using has been removed due to its limited moderating effect (Hu et al., 2003), thus simplifying the instrument by reducing the number of items. This elimination is justified by the evolution of the TAM model (Venkatesh et al., 2003) and by acceptance models applied in the educational field (Sánchez-Prieto et al., 2017a).

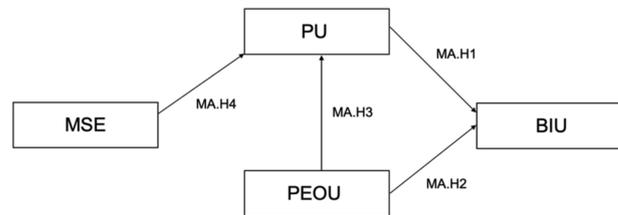
Having defined the three constructs proposed in TAM, the models presented in this paper are extended with the construct of self-efficacy to explore an antecedent of perceived usefulness according to the literature (Nikou & Economides, 2017b). Previous studies have shown that experience is a relevant factor in the use of mobile devices, based on the fact that the higher the level of skill and dexterity in their use, the less effort will be required to use them (Anderson, 1996); which is related to this perceived usefulness.

Self-efficacy is defined in research as an individual's perception of their ability to use mobile devices to perform certain tasks (Nikou & Economides, 2017a). In the educational context, it can be inferred that the higher the level of skill with the use of devices (self-efficacy), the greater the ease of use and perceived usefulness in assessment processes (Pikkarainen et al., 2004; Wang et al., 2020). Based on these premises, it is proposed that:

- Mobile self-efficacy (MSE) will have a positive effect on teachers' perceived usefulness (PU) in using mobile devices for e-assessment (MAH4).

Accordingly, the first model proposed is the TAM model, which presents mobile self-efficacy as an antecedent of perceived usefulness based on Nikou and Economides' (2017b) construction of the construct (Figure 2).

Figure 2. First proposed model (MA)



1.2.2. Model B (MB): Formative perceived usefulness and summative perceived usefulness

A second model is based on the understanding of summative and formative assessments as two distinct dimensions (MacLellan, 2001), and converting the original utility of the TAM model into a specific utility for each of the assessments.

The construction of the model has been based on the perceived usefulness construct of Davis (1989), the theoretical analysis of formative and summative assessments presented here, and following Moore & Benbasat's proposal for modelling (1991). These new utilities are also derived from Scriven's distinction (1991) and the formulation of two parallel and distinct constructs. Therefore, the hypotheses regarding perceived usefulness are as follows:

- Perceived usefulness of formative assessment (PUFA) positively predicts teachers' intention to use (BIU) mobile devices for e-assessment in future teaching practice (MBH1).
- Perceived usefulness of summative assessment (PUSA) positively predicts teachers' intention to use (BIU) mobile devices for e-assessment in future teaching practice (MBH2).

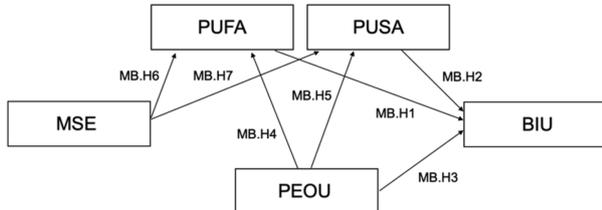
In addition, the constructs of ease of use (PEOU), behavioural intention (BIU), and mobile self-efficacy (MSE) are maintained from the previous model, preserving the hypotheses to test the differences between the models. Having two perceived usefulness, the PEOU and MSE constructs establish a double hypothesis, one for each utility, such that:

- Perceived ease of use (PEOU) positively predicts teachers' intention to use (BIU) mobile devices for e-assessment in their future teaching practice (MBH3).
- Perceived ease of use (PEOU) positively predicts teachers' perceived usefulness of mobile devices in formative assessment (PUFA) for e-assessment in their future teaching practice (MBH4).
- Perceived ease of use (PEOU) positively predicts teachers' perceived usefulness of mobile devices in summative assessment (PUSA) for e-assessment in their future teaching practice (MBH5).
- Mobile self-efficacy (MSE) will have a positive impact on teachers' perceived usefulness of mobile devices in formative assessment (PUFA) for e-assessment (MBH6).

- Mobile self-efficacy (MSE) will have a positive impact on teachers' perceived usefulness of mobile devices in summative assessment (PUSA) for e-assessment (MBH7).

The second model of the study is thus the combination of two constructs of perceived usefulness of the formative and summative assessments, the constructs of the TAM model and mobile self-efficacy (Figure 3).

Figure 3. Second proposed model (MB)



1.2.3. Model C (MC): Perceived usefulness of assessment as a formative construct

A third group of authors considers assessment as a combination of summative and formative assessments, assuming that the assessment process is the sum of its components (Dolin et al., 2018). This implies adopting, in acceptance models, a single construct that brings together the indicators defined in the previous model for the perceived usefulness of formative and summative assessments, establishing a single dimension of assessment, the assessment perceived utility (APU).

This dimension is understood as a formative construct, assuming that the sum of all indicators provides the overall measure of the dimension (Simonetto, 2012), the usefulness of assessment. Consequently, the following hypothesis is formulated for perceived usefulness:

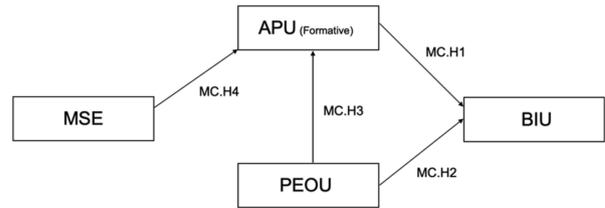
- Perceived usefulness (PU) positively predicts teachers' intention to use (BIU) mobile devices for e-assessment in their future teaching practice (MCH1).

As in the previous model (MB), the constructs of ease of use (PEOU), behavioural intention (BIU), and mobile self-efficacy (MSE) are kept from the original model, and the hypotheses of these constructs are maintained to carry out the triple comparison:

- Perceived ease of use (PEOU) positively predicts teachers' intention to use (BIU) mobile devices for e-assessment in their future teaching practice (MCH2).
- Perceived ease of use (PEOU) positively predicts teachers' perceived usefulness of mobile devices for e-assessment (APU) in their future teaching practice (MCH3).
- Mobile self-efficacy (MSE) will have a positive impact on teachers' perceived usefulness of mobile devices for e-assessment (APU) (MCH4).

The third and final model is the sum of a single construct on assessment (bringing together the items of perceived usefulness of formative and summative assessments), and the constructs of the TAM model with self-efficacy as a predecessor of perceived usefulness (Figure 4)

Figure 4. Third proposed model (MC)



The methodology followed is presented in Section 3. The results are included in section 4, whereas the conclusions, main findings, limitations, and future prospects are in section 5.

2 METHODS

2.1 Measurement tool

To conduct this study, a questionnaire consisting of a set of identity variables and the items that compose the dimensions has been designed. The socio-demographic identity variables included are age, gender, the academic degree of access to the master's degree, teaching experience, and the mean number of hours of daily use of mobile devices. The second section consists of 27 items to assess acceptance through the study constructs using a Likert scale ranging from 1 to 7 (1 = strongly disagree, 7 = strongly agree) (Matas, 2018). The complete list of items can be found in the complementary database to this research, available at: <https://bit.ly/3T3DK8q>.

The measurement tool designed is based on the analysis of current models of technology acceptance assessment (Davis, 1989; Venkatesh & Bala, 2008; Venkatesh & Davis, 2000; Nikou & Economides, 2017b, 2017a) and the evaluation of comparative structural equation models developed in PLS-SEM in the field of m-learning (Alshurideh et al., 2020).

In its construction, the items related to behavioural intention, perceived ease of use, and perceived usefulness have been taken and adapted from the TAM3 model (Venkatesh & Bala, 2008), an evolution of the original proposal by Davis (1989), adapting them for mobile technologies in e-assessment. The items related to the perceived usefulness of formative and summative assessments (Model B) have been developed by the authors based on the considerations of Moore & Benbasat (1991) and the delimitation of the classical concept of formative and summative assessments (Scriven, 1967). Finally, the mobile self-efficacy construct has been adapted from the original self-efficacy proposal (Anderson, 1996) and its adaptation to the field of mobile technologies (Nikou & Economides, 2017a).

The administration of the questionnaires, the entire design process, and results and data processing have been approved by the Research Ethics Committee of the University of Salamanca. Data collection was conducted in November 2022, with a total survey period of 28 days. The questionnaire was distributed through the LimeSurvey application (Engard, 2009).

2.2 Population and sample

The study population is composed of students enrolled in the "Master's Degree in Teaching Secondary and Upper-Secondary Education, Vocational Training and Languages", which is specific training for teachers before access to the teaching profession, responsible for providing pedagogical knowledge to graduates in the areas of knowledge related to teaching in Secondary Education (Vilches & Gil, 2010). This study addresses the group of teachers in training as this is a period which, depending on its development, can be a barrier or a driver in technological changes

and the training of future teachers (Sánchez-Prieto et al., 2019). The questionnaire was administered to the students of the University of Salamanca (N = 292), and 262 responses were collected.

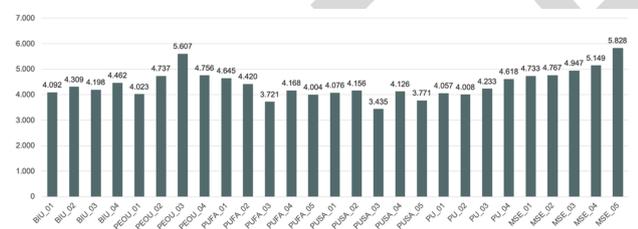
2.3 Data analysis technique

The Partial Least Squares Structural Equation Modeling (PLS-SEM) has been applied in this study to analyse the data collected from the questionnaires. The analysis was carried out using the statistical software *SmartPLS 3.2.9* (Hair et al., 2019; Sarstedt et al., 2016) and consisted of several steps: the evaluation of the measurement model (convergent validity and reliability analysis, collinearity and cross-loading checks), the evaluation of the structural model (study of the effects, effect sizes, and mediations) and the behavioural explanation of the model (predictive power and model fit indicators) (Bayaga & Kyobe, 2021; Hair et al., 2022). This study is based on PLS-SEM given its characteristics, as it presents a triple comparison focused on perceived usefulness in which one of the models considers the usefulness construct as formative, an analysis that is only possible using PLS-SEM (Ajms, 2015; Bollen, 2011).

3 RESULTS

First, a description of the mean scores on the questionnaire items (which will later form the basis for the study of the different models) is presented (Figure 5). The dimensions with the highest mean score were perceived ease of use ($\bar{x} = 4.75$), intention to use ($\bar{x} = 4.27$), and perceived usefulness of mobile devices in e-assessment ($\bar{x} = 4.22$). In contrast, the lowest mean scores were found in the usefulness of mobile devices for summative assessment ($\bar{x} = 3.90$), mobile self-efficacy ($\bar{x} = 4.18$), and perceived usefulness of mobile devices for formative assessments ($\bar{x} = 4.18$).

Figure 5. Descriptive analysis of the items



In an initial approximation to the differences between the perceived usefulness according to the models, it was observed that future teachers rated the generic perceived usefulness higher, as it does not distinguish between summative and formative assessments in its formulation, which may be due to the fact that the formulation of items is different, and the results can hardly be compared in purely descriptive terms. However, the comparison between the perceived usefulness of formative assessment ($\bar{x} = 4.18$) and the perceived usefulness of summative assessment ($\bar{x} = 3.90$) showed higher values for formative assessment. Therefore, it is suggested that future teachers find the use of mobile devices more useful in formative assessments, which will be explored below through the structural equation analysis conducted.

3.1 Evaluation of the measurement model

The evaluation of the measurement model included the study of reliability (composite reliability index), construct validity (convergent validity [average variance extracted, AVE]), and

Cronbach's alpha. Data for the first two models (reflective models) and the reflective part of the third model (MC) are shown in Table 1. Regarding the weights, following the recommendations of Hair et al. (2010), indicators with loadings higher than 0.5 were accepted; and those with a lower value were removed (BIU_03, PEOU_01, and MSE_05 were removed in the three models). After the removal, the reliability of the items was confirmed. The results also confirmed convergent validity through the AVE, the Cronbach's alpha (α), and the composite reliability (CR); with values above 0.5, 0.7, and 0.6; respectively (Fornell & Larcker, 1981)

Table 1. Item reliability and convergent validity analysis (reflective variables)

Latent variable	Indicator	Loading	α	CR	AVE
Model A (MA): PU + PEOU + BIU + MSE					
BIU	BIU_01	0.931	0.924	0.952	0.869
	BIU_02	0.650			
	BIU_04	0.915			
PEOU	PEOU_02	0.829	0.794	0.869	0.690
	PEOU_03	0.739			
	PEOU_04	0.914			
PU	PU_01	0.887	0.909	0.936	0.786
	PU_02	0.908			
	PU_03	0.900			
	PU_04	0.851			
MSE	MSE_01	0.851	0.881	0.916	0.732
	MSE_02	0.903			
	MSE_03	0.863			
	MSE_04	0.803			
Model B (MB): PUFA + PUSA + PEOU + BIU + MSE					
BIU	BIU_01	0.932	0.924	0.952	0.869
	BIU_02	0.948			
	BIU_04	0.916			
PEOU	PEOU_02	0.828	0.794	0.869	0.689
	PEOU_03	0.739			
	PEOU_04	0.915			
PUFA	PUFA_01	0.823	0.885	0.917	0.688
	PUFA_02	0.863			
	PUFA_03	0.868			
	PUFA_04	0.735			
	PUFA_05	0.851			
PUSA	PUSA_01	0.828	0.894	0.922	0.703
	PUSA_02	0.859			
	PUSA_03	0.810			
	PUSA_04	0.816			
	PUSA_05	0.878			
MSE	MSE_01	0.844	0.881	0.917	0.734
	MSE_02	0.903			
	MSE_03	0.865			
	MSE_04	0.811			
Model C (MC): APU (Formative) + PEOU + BIU + MSE					
BIU	BIU_01	0.932	0.950	0.964	0.869
	BIU_02	0.948			
	BIU_04	0.916			
PEOU	PEOU_02	0.829	0.776	0.851	0.689
	PEOU_03	0.735			
	PEOU_04	0.916			
MSE	MSE_01	0.776	0.881	0.917	0.734
	MSE_02	0.864			
	MSE_03	0.892			
	MSE_04	0.863			

For the formative construct of the usefulness of the evaluation of the third model (MC), the values of the variance inflation factor (VIF), which will enable us to dismiss problems of collinearity (Hair et al., 2022), are presented in Table 2, with the acceptance of items with values lower than 5 (Hair et al., 2021), and, as all of

them were below the indicated value, the analysis was therefore continued. Subsequent bootstrapping techniques with 5000 subsamples (Hair et al., 2022) revealed low weights of the indicators. However, the recommendation is to retain items with low weights if the loadings are higher than 0.5 (Hair et al., 2021; Ramayah et al., 2017); hence, all items were retained even if they did not contribute to explaining the dimension. Under this paradigm, the perceived usefulness dimension of the assessment (MC) was maintained, but its items do not fully define it.

Table 2. VIFs, weights, loads and p-values of the formative construct indicators (MC)

	VIF	Weight	Weight CI	Load	p-Value
APU_01 (RAFA1)	2,31	0.366	0.202-0.537	0.810	0.000
APU_02 (RAFA2)	2,64	0.182	0.006-0.377	0.817	0.058
APU_03 (RAFA3)	3,42	-0.119	-0.315-0.081	0.734	0.228
APU_04 (RAFA4)	2,12	0.094	-0.084-0.269	0.673	0.272
APU_05 (RAFA5)	3,14	0.311	0.125-0.503	0.848	0.001
APU_06 (RASA1)	2,75	0.091	-0.092-0.287	0.617	0.323
APU_07 (RASA2)	3,04	0.078	-0.142-0.256	0.761	0.428
APU_08 (RASA3)	2,92	0.162	-0.051-0.315	0.718	0.069
APU_09 (RASA4)	2,65	0.167	-0.028-0.390	0.629	0.101
APU_10 (RASAS)	3,67	-0.030	0.245-0.174	0.737	0.775

Discriminant validity was assessed according to two criteria, the Fornell-Larcker criterion (Fornell & Larcker, 1981) and the heterotrait-monotrait ratio of correlations (HTMT) (Henseler et al., 2015). The results showed optimal discriminant validity for all three models (Table 3).

Table 3. Discriminant validity tests for the three models

Fornell-Larcker					Heterotrait-Monotrait					
Model A (MA): PU + PEOU + BIU										
	BIU	MSE	PEOU	PU	BIU	MSE	PEOU	PU		
BIU	.909									
MSE	.397	.831			0.631					
PEOU	.688	.423	.848		0.368	0.640				
PU	.773	.489	.802	.860	0.744	0.678	0.418			
Model B (MB): PUFA + PUSA + PEOU + BIU										
	BIU	MSE	PEOU	PUFA	PUSA	BIU	MSE	PEOU	PUFA	PUSA
BIU	0.932									
MSE	0.592	0.856				0.631				
PEOU	0.361	0.550	0.830			0.368	0.640			
PUFA	0.733	0.561	0.415	0.829		0.806	0.615	0.440		
PUSA	0.629	0.530	0.324	0.666	0.839	0.687	0.574	0.337	0.741	
Model C (MC): APU (Formative) + PEOU + BIU										
	APU	BIU	MSE	PEOU	BIU	MSE	PEOU			
APU	-									
BIU	0.765	0.932								
MSE	0.620	0.591	0.857		0.631					
PEOU	0.422	0.361	0.5551	0.830	0.368	0.640				

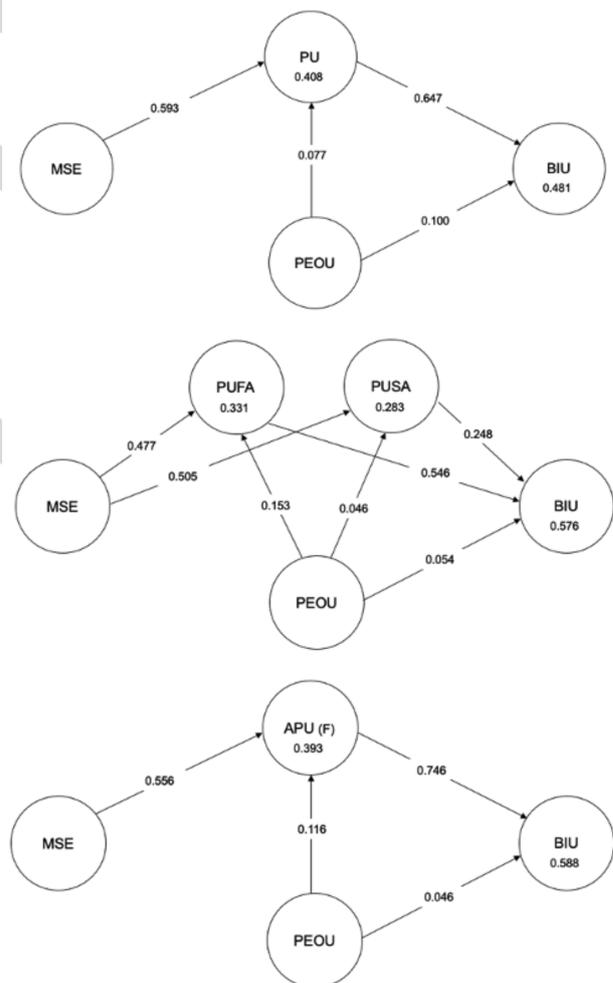
3.2 Structural validity of the model

The analysis of the structural validity of the models included the description of the adjusted R² values, i.e. the explained variance (Cohen, 1988). The first model (MA) explained 0.48 of the behavioural intention to use; while the second model (MB) explained 0.58 of it, and the third one (MC) explained 0.59; being MC the model with the highest percentage of explained variance. Furthermore, the Stone-Geisser test results (Geisser, 1974; Stone, 1974) showed positive values at Q² for each model, confirming their predictive relevance.

The path coefficients (Hair et al., 2022) of the three models are shown in consecutive order in Figure 6. In the first model (MA), usefulness positively predicts intention to use mobile devices (MAH1 is accepted), and ease of use positively predicts intention to use (MAH3 is accepted, but MAH2 is not confirmed). Regarding the second model (MB), the usefulness of formative assessment and usefulness of summative assessment positively predict final intention to use (accepting MBH1 and MBH2), while ease of use is only a predictor of formative usefulness (accepting MBH4 and rejecting MBH5). In the third model (MC), the formative construct “perceived usefulness of assessment” positively predicts behavioural intention to use (MHC1), but in this model ease of use does not predict intention to use nor is it a predictor of the usefulness of mobile devices in assessment (MHC2 and MCH3 are not accepted).

Finally, mobile self-efficacy is confirmed as a predictor of ease of use in all three proposed models, including the perceived usefulness of TAM (MAH4 is accepted), the two differentiated usefulness constructs of model two (MBH5 and MBH6 are accepted), as well as the third model (MCH4 is accepted).

Figure 6. Proposed structural models (from top to bottom: MA, MB and MC)



The bootstrapping results (Davison & Hinkley, 1997) (Table 4) show the significance of the relationships proposed above, as well as the effect sizes of each of these relationships (small: 0.02 ≤ f² < 0.15; median: 0.15 ≤ f² < 0.35; large: ≥ 0.35) (Cohen, 1988).

Table 4. Direct effects

	Path Coeff.	CI	f2	Results
Model A (MA): PU + PEOU + BIU				
PU > BIU	0.647*	0.554-0.731	0.676c	MAH1 Supported
PEOU > PU	0.077ns	-0.044-0.203	0.007	MAH2 not supported
PEOU > BIU	0.100*	0.003-0.204	0.016a	MAH3 supported
MSE > PU	0.593*	0.490-0.697	0.416c	MAH4 supported
Model B (MB): PUFA + PUSA + PEOU + BIU				
PUFA > BIU	0.546*	0.437-0.655	0.359c	MBH1 supported
PUSA > BIU	0.248*	0.122-0.371	0.080a	MBH2 supported
PEOU > BIU	0.054ns	-0.030-0.139	0.006	MBH3 not supported
PEOU > PUFA	0.153*	-0.008-0.289	0.024b	MBH4 Supported
PEOU > PUSA	0.046ns	-0.093-0.181	0.002	MBH5 not supported
MSE > PUFA	0.477*	0.316-0.591	0.238b	MBH6 supported
MSE > PUSA	0.505*	0.335-0.594	0.248b	MBH7 Supported
Model C (MC): APU (Formative) + PEOU + BIU				
APU > BIU	0.746*	0.687-0.820	1.108c	MCH1 supported
PEOU > BIU	0.116ns	-0.025-0.260	0.015	MCH2 not supported
PEOU > APU	0.046ns	-0.042-0.118	0.004	MCH3 not supported
MSE > APU	0.556*	0.448-0.683	0.354c	MCH4 Supported

Path coefficient: * (significant), ns (no significant)
 Size effect (f2): ns: no significant; a: small; b: medium; c: big

Previous research has found that variables such as perceived ease of use can have a direct but also an indirect effect on other study variables (Rothmann, 2015). It is therefore important to evaluate these effects to determine whether perceived ease of use influences behavioural intention to use through other constructs, thus providing a complete understanding of the model. The indirect effects, their size, and significance are shown in Table 5. It follows from this table that ease of use has an indirect effect on behavioural intention to use only in model B, through the perceived usefulness of mobile devices for formative assessment. Moreover, self-efficacy has an indirect effect in all three models reported, acting through the perceived usefulness described in model A, the two perceived usefulness (formative and summative) in model B, and the perceived usefulness of assessment in model C.

Table 5. Indirect effects

	Eff. Coeff.	Conf. Interv.
Model A (MA): PU + PEOU + BIU		
MSE > BIU	0.384*	0.297-0.477
PEOU > BIU	0.050ns	-0.028-0.131
Model B (MB): PUFA + PUSA + PEOU + BIU		
MSE > BIU	0.386*	0.284-0.491
PEOU > BIU	0.095*	-0-001-0-188
Model C (MC): APU (Formative) + PEOU + BIU		
MSE > BIU	0.414*	0.326-0.532
PEOU > BIU	0.087ns	-0-020-0.196

Effect coefficient: * (significant), ns (no significant)

To conclude the analysis of the results, it should be emphasised that this study aimed to carry out a triple comparison between a model based on the generic utility of TAM (MA), a model with two differentiated usefulness in formative and summative assessments (MB) and a model with a single usefulness of formative and summative assessments (MC). After confirming the validity of the models (reliability of the items and validity), we proceed to

describe the differences found in their effect sizes. MA, MB, and MC can predict behavioural intention through perceived usefulness with a large effect size (except MBH2, which predicts it with a small effect size). Ease of use has a medium effect size on the behavioural intention for model A and formative usefulness for model B. Finally, self-efficacy in using mobile devices has a large effect on perceived usefulness (MA) and perceived usefulness of assessment (MC), and a medium effect on perceived usefulness for summative and formative assessments (MB).

The predictive relevance (Q2) of the three models confirms that all three models have predictive power. According to this criterion, the best model is determined to be the one with the highest score in the comparison (Hair et al., 2011, 2013), with MC being the most parsimonious and generalisable model (Sharma & Kim, 2012), practically equal to MB. Considering Akaike Information Criterion (AIC) (Bozdogan, 1987), model C is significantly better than model B, as are models B and C with regard to model A. The criterion, which has high predictive power and is stable for model comparison (Sharma et al., 2022), also confirms MC as the best model, followed by MB (Table 6).

Table 6. Model comparison

	Model A (MA)	Model B (MB)	Model C (MC)
Q2	0.410	0.492	0.498
Akaike Information Criteria (AIC)	97.276	46.723	37.065
Bayesian information criterion (BIC)	-156.174	-203.238	-216.386

After completing the description of the results, the discussion and conclusions drawn from the comparative study, as well as the limitations of the research and future prospects are covered in section 5.

4 DISCUSSION

The study highlights the importance of the usefulness of mobile devices in assessment processes by teachers in training. In response to the research questions posed (R1 and R2), the model that best predicts the intention to use mobile devices is the one that considers perceived usefulness as a formative construct (MC). However, when considering that seven of the ten items of the formative construct of usefulness are not significant for its assessment, it can be stated that the best model in predicting intention to use is model B since it only predicts 1% less than the previous one and maintains all the items in two distinct constructs. Therefore, the usefulness of assessment should be approached as two distinct sources, since both reflective modelling (MB) and formative modelling (MC) that specifically consider the usefulness of assessment perform better than Davis' (1989) generic model. Future acceptance studies should not address usefulness in a non-specific way, which would fail to consider the differences that teachers set for the use of mobile devices between both them and which have been confirmed in this study.

4.1 Implications of Assessment for Technology Adoption Models

Previous studies have shown the relevance of generic perceived usefulness in models of technology adoption in the field of m-learning (Harchay et al., 2019), but to date no literature has been published on perceived usefulness that addresses the particularities of assessment, generating here two new models more specific to the field. This work highlights the relevance of assessment in

teaching processes, interpreted from the perspective of the usefulness of technology for this purpose, making its determination crucial for the study of the intention to use it for teaching.

The results of the second model (MB), which addresses formative and summative assessments in differentiated constructs, suggest that teachers attach greater importance to the use of mobile technology in formative assessments, also acting as a greater predictor of the intention to use mobile technology in assessment. This idea is in line with the literature, as formative assessments are perceived as more flexible and open to innovations (Nikou & Economides, 2021) while traditional assessments are still reluctant to include new techniques in their processes (Shepard, 2006).

The study also supports the association between self-efficacy in the use of mobile devices and perceived usefulness in the three models, which confirms that this is an influential factor related to the perceived usefulness of technology in assessment (Rahmawati, 2019). This statement confirms the potential of ICT in the teaching-learning process, as a teacher who is prepared and effective in the use of technology will perceive the usefulness for assessment processes in a much higher way, and his or her intention to use technology and change methodologies towards virtual environments will be higher.

4.2 Implications of the Study for Teaching Practice

Regarding teaching, the study reveals the importance of continuous teacher training on improving the usefulness of mobile devices, especially for the formative assessments. Work on the usefulness of mobile devices through new assessment techniques and new tools is already a reality in continuous training in many countries (Jin et al., 2021), trying to encourage a change in teaching towards the inclusion of technology in classroom practice.

Moreover, the models developed have confirmed self-efficacy in the use of mobile devices as a relevant factor in both usefulness and intention to use, in line with the studies by Nikou and Economides (2017a, 2017b). Therefore, teaching practice should promote training and the development of specific policies focused not only on the usefulness or the basics of assessment, but also make teachers effective in the use of mobile devices for their teaching-learning processes and their assessments.

Finally, it has also been found that ease of use is only a significant factor in the relationship with the perceived usefulness of formative assessment (model B). This fact can be explained by the regular and standardised use of mobile devices in modern society (Smith, 2021), especially among future teachers, being the highest dimension of the study in the descriptive analysis carried out.

4.3 Implications of the Study for Initial Teacher Training

The research confirms the need for training based on technological and pedagogical innovation rather than training in the use of mobile devices, a less relevant construct for these assessment models. This innovation should be focused on working on self-efficacy in the use of devices and the perceived usefulness of teachers in their use in assessment, assuming that the training in their use has been completed (Domingo-Coscollola et al., 2020).

This new approach of training in assessment through mobile devices concerning self-efficacy and technological usefulness will be of great relevance in order to have future teachers with a high intention of use, who are updated in assessment techniques and the use of mobile devices for this purpose and who have an

intention of use following the new educational scenarios, achieving the transition of education towards the new educational reality (Souabi et al., 2021). These specific changes in the training of future teachers require, ultimately, the establishment of new training policies in higher education institutions and the design of new, updated study plans, with technologically competent teaching staff and with innovation as the sole focus of training, reaffirming the need for a new education mediated by technology (Canales-García et al., 2020; Skulmowski & Rey, 2020).

5 CONCLUSIONS

Technology adoption in education has been addressed in numerous models, mostly in the field of e-learning and in studies related to methodology, but there is little research focused on the technological acceptance of mobile devices in e-assessment (Alrfooh & Lakulu, 2020).

This study has not only explored this area of the teaching process in depth but has also included an exhaustive analysis based on the theoretical currents on the conception of assessment and the study of its usefulness, which is based on the evaluation of three different technology acceptance models in future teachers. This research contributes to the advancement of knowledge from the following perspectives:

- (1) First, this research supports the differences between summative and formative assessments from the teacher's perspective and the importance of considering them specifically as two different usefulness of mobile devices for assessment in technology adoption models, explaining a higher percentage of the intention of use than the TAM model in its generic conception.
- (2) Second, a greater relevance of formative assessment compared to summative assessment is observed in the specific adoption models (MB, MC). Perceived usefulness of formative assessment explains a higher intention to use in the model with two constructs (MB), and more items are representative of formative usefulness than of summative usefulness in the model with the formative construct of perceived usefulness of assessment (MC).
- (3) Third, the results suggest the high predictive power and functioning of self-efficacy in the use of mobile devices on the perceived usefulness of assessment, being an appropriate construct in all three proposed models. However, ease of use is a factor of lesser relevance and explanatory power, possibly a consequence of the current digital competence of the teachers.
- (4) Finally, the initial training of future teachers and active teachers training should focus on the development of formative assessment work, the effective use of mobile devices in teaching, and their usefulness in assessment processes, to increase their intention to use mobile technology in assessment.

This study has some limitations. From a methodological perspective, the results may have been compromised by the access to the sample. Despite trying to ensure heterogeneity by applying the questionnaire to all students enrolled in the master's degree, all participants share the training within the same academic programme. Future studies could widen the participation of students from other institutions and geographical areas, leading to more generalisable results. Future research could also already consider the specific usefulness of summative and formative assessments

as the main usefulness of the model, extend the study and analyse its relationship with other related antecedents.

AUTHORS' CONTRIBUTION

- Conceptualization, data curation, formal analysis, methodology, resources, investigation resources, writing – original draft, writing – review & editing: A.O-L.
- Conceptualization, data curation, formal analysis, methodology, supervision, writing – review & editing: J-C.S-P.
- Conceptualization, formal analysis, investigation, resources, supervision, writing – review & editing: S.O-M.
- All authors have read and agreed to the published version of the manuscript.

REFERENCES

- Abd-Karim, R., Abu, A., Adnan, A., & Suhandoko, A. (2018). *The Use of Mobile Technology in Promoting Education 4.0 for Higher Education*. 2, 34-39. <https://doi.org/10.26666/emp.ajive.2018.3.6>
- Ahmad, B., & Bhat, G. J. (2019). Formative and Summative Evaluation Techniques for Improvement of Learning Process. *European Journal of Business & Social Sciences*, 7(5), 776-785.
- Ajms, E. (2015). Structure Equation Modeling Basic Assumptions and Concepts: A Novices Guide. *Asian Journal of Management Sciences*, 3(1), Article 1. <https://www.ajmsjournal.com/index.php/ajms/article/view/70>
- Al-Emran, M., Arpaci, I., & Salloum, S. A. (2020). An empirical examination of continuous intention to use m-learning: An integrated model. *Education and Information Technologies*, 25(4), 2899-2918. <https://doi.org/10.1007/s10639-019-10094-2>
- Al-Gasawneh, J. A., Al Khoja, B., Al-Qeed, M. A., Nusairat, N. M., Hammouri, Q., & Anuar, M. M. (2022). *Mobile-customer relationship management and its effect on post-purchase behavior: The moderating of perceived ease of use and perceived usefulness*. <https://digitallibrary.aau.ac.ae/handle/123456789/672>
- Aljawarneh, S. A. (2020). Reviewing and exploring innovative ubiquitous learning tools in higher education. *Journal of Computing in Higher Education*, 32(1), 57-73. <https://doi.org/10.1007/s12528-019-09207-0>
- Ally, M., & Prieto-Blázquez, J. (2014). What is the future of mobile learning in education? *RUSC. Universities and Knowledge Society Journal*, 11(1), Article 1. <https://doi.org/10.7238/rusc.v11i1.2033>
- Alrfooh, A., & Lakulu, M. (2020). A systematic review of mobile-based assessment acceptance studies from 2009 to 2019. *Journal of Theoretical and Applied Information Technology*, 97(20), 1-25.
- Alshurideh, M., Al Kurdi, B., Salloum, S. A., Arpaci, I., & Al-Emran, M. (2020). Predicting the actual use of m-learning systems: A comparative approach using PLS-SEM and machine learning algorithms. *Interactive Learning Environments*, 0(0), 1-15. <https://doi.org/10.1080/10494820.2020.1826982>
- Anderson, A. A. (1996). Predictors of computer anxiety and performance in information systems. *Computers in Human Behavior*, 12(1), 61-77. [https://doi.org/10.1016/0747-5632\(95\)00019-4](https://doi.org/10.1016/0747-5632(95)00019-4)
- Anisimova, T. I., Sabirova, F. M., & Shatunova, O. V. (2020). Formation of Design and Research Competencies in Future Teachers in the Framework of STEAM Education. *International Journal of Emerging Technologies in Learning (IJET)*, 15(02), Article 02. <https://doi.org/10.3991/ijet.v15i02.11537>
- Baier, F., & Kunter, M. (2020). Construction and validation of a test to assess (pre-service) teachers' technological pedagogical knowledge (TPK). *Studies in Educational Evaluation*, 67, 100936. <https://doi.org/10.1016/j.stueduc.2020.100936>
- Bayaga, A., & kyobe, M. (2021). PLS-SEM technique and phases of analysis – implications for information systems' exploratory design researchers. *2021 Conference on Information Communications Technology and Society (ICTAS)*, 46-51. <https://doi.org/10.1109/ICTAS50802.2021.9395029>
- Bernacki, M. L., Greene, J. A., & Crompton, H. (2020). Mobile technology, learning, and achievement: Advances in understanding and measuring the role of mobile technology in education. *Contemporary Educational Psychology*, 60, 101827. <https://doi.org/10.1016/j.cedpsych.2019.101827>
- Bizzo, E. (2021). Aceptación y la adopción del e-learning en los países en desarrollo: Una revisión de la literatura. *Ensaio: Avaliação e Políticas Públicas em Educação*, 30, 458-483.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003, enero 1). *Formative and summative assessment: Can they serve learning together?* Annual meeting of the American Educational Research Association, Chicago.
- Black, P. J. (1993). Formative and Summative Assessment by Teachers. *Studies in Science Education*, 21(1), 49-97. <https://doi.org/10.1080/03057269308560014>
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Bollen, K. A. (2011). Evaluating Effect, Composite, and Causal Indicators in Structural Equation Models. *MIS Quarterly*, 35(2), 359-372. <https://doi.org/10.2307/23044047>
- Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3), 345-370. <https://doi.org/10.1007/BF02294361>
- Buchholtz, N. F., Krosanke, N., Orschulik, A. B., & Vorhölter, K. (2018). Combining and integrating formative and summative assessment in mathematics teacher education. *ZDM*, 50(4), 715-728. <https://doi.org/10.1007/s11858-018-0948-y>
- Burbules, N. C., Fan, G., & Repp, P. (2020). Five trends of education and technology in a sustainable future. *Geography and Sustainability*, 1(2), 93-97. <https://doi.org/10.1016/j.geosus.2020.05.001>
- Canales-García, A., Fernández-Valverde, M., & Ulate-Solis, G. (2020). Aprender y enseñar con recursos TIC: Experiencias innovadoras en la formación docente universitaria. *Ensayos Pedagógicos*, 15(1), 235-248.
- Castañeda-Vázquez, C., Espejo-Garcés, T., Zurita-Ortega, F., & Fernández-Revelles, A. (2019). La formación de los futuros docentes a través de la gamificación, tic y evaluación continua. *SPORT TK-Revista EuroAmericana de Ciencias del Deporte*, 8(2), Article 2. <https://doi.org/10.6018/sportk.391751>
- Ciobanu, R.-C. (2022). M-learning and E-learning Educational Solutions Impact in the COVID-19 Pandemic. *Informatica Economica*, 26(3), 64-73.
- Clark, R. M., Kaw, A. K., & Braga Gomes, R. (2022). Adaptive learning: Helpful to the flipped classroom in the online environment of COVID? *Computer Applications in Engineering Education*, 30(2), 517-531. <https://doi.org/10.1002/cae.22470>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2.ª ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Cosí, A., Voltas, N., Lázaro-Cantabrana, J. L., Morales, P., Calvo, M., Molina, S., & Quiroga, M. Á. (2020). Formative assessment at university through digital technology tools. *Profesorado, Revista de Currículum y Formación Del Profesorado*, 24(1), Article 1. <https://doi.org/10.30827/profesorado.v24i1.9314>
- Criollo, S., Guerrero-Arias, A., Jaramillo-Alcázar, Á., & Luján-Mora, S. (2021). Mobile Learning Technologies for Education: Benefits and Pending Issues. *Applied Sciences*, 11(9), Article 9. <https://doi.org/10.3390/app11094111>
- Cruz-Benito, J., Sánchez-Prieto, J. C., Therón, R., & García-Peñalvo, F. J. (2019). Measuring Students' Acceptance to AI-Driven Assessment in eLearning: Proposing a First TAM-Based Research Model. In P. Zaphiris & A. Ioannou (Eds.), *Learning and Collaboration Technologies. Designing Learning Experiences* (pp. 15-25). Springer International Publishing. https://doi.org/10.1007/978-3-030-21814-0_2
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511802843>
- Dixon, D. D., & Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory Into Practice*, 55(2), 153-159. <https://doi.org/10.1080/00405841.2016.1148989>
- Dolin, J., Black, P., Harlen, W., & Tiberghien, A. (2018). Exploring Relations Between Formative and Summative Assessment. In J. Dolin and R. Evans (Eds.), *Transforming Assessment: Through an Interplay Between Practice, Research and Policy* (pp. 53-80). Springer International Publishing. https://doi.org/10.1007/978-3-319-63248-3_3
- Domingo-Coscollola, M., Bosco-Paniagua, A., Carrasco-Segovia, S., & Sánchez-Valero, J.-A. (2020). Fomentando la competencia digital docente en la universidad: Percepción de estudiantes y docentes. *Revista de Investigación Educativa*, 38(1), Article 1. <https://doi.org/10.6018/rie.340551>
- Engard, N. C. (2009). LimeSurvey. *Public Services Quarterly*, 5(4), 272-273. <https://doi.org/10.1080/15228950903288728>
- Evans, C., & Robertson, W. (2020). The four phases of the digital natives debate. *Human Behavior and Emerging Technologies*, 2(3), 269-277. <https://doi.org/10.1002/hbe2.196>
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention And Behavior: An introduction to theory and research*. Addison-Wesley.
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- García-Aretio, L. (2019). Necesidad de una educación digital en un mundo digital. *RIED. Revista iberoamericana de educación a distancia*. <https://doi.org/10.5944/ried.22.2.23911>
- Geisser, S. (1974). A Predictive Approach to the Random Effect Model. *Biometrika*, 61(1), 101-107. <https://doi.org/10.2307/2334290>

- Gómez-Galán, J. (2020). Media Education in the ICT Era: Theoretical Structure for Innovative Teaching Styles. *Information, 11*(5), Article 5. <https://doi.org/10.3390/info11050276>
- Gómez-Ruiz, M. Á., Vázquez-Recio, R., López-Gil, M., & Ruiz-Romero, A. (2022). La pesadilla de la evaluación: Análisis de los sueños de estudiantes universitarios. *Revista Iberoamericana de Evaluación Educativa, 15*(1), Article 1. <https://doi.org/10.15366/riece2022.15.1.008>
- Guardia, J. J., Del Olmo, J. L., Roa, I., & Berlanga, V. (2019). Innovation in the teaching-learning process: The case of Kahoot! *On the Horizon, 27*(1), 35-45. <https://doi.org/10.1108/OTH-11-2018-0035>
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate Data Analysis*. Prentice-Hall.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Evaluation of Formative Measurement Models. En J. F. Hair Jr., G. T. M. Hult, C. M. Ringle, M. Sarstedt, N. P. Danks, & S. Ray (Eds.), *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook* (pp. 91-113). Springer International Publishing. https://doi.org/10.1007/978-3-030-80519-7_5
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice, 19*(2), 139-152. <https://doi.org/10.2753/MTP1069-6679190202>
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2013). Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance. *Long Range Planning, 46*(1). <https://papers.ssrn.com/abstract=2233795>
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2-24. <https://doi.org/10.1108/EBR-11-2018-0203>
- Hair, J., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. <https://doi.org/10.1007/978-3-030-80519-7>
- Hao, S., Dennen, V. P., & Mei, L. (2017). Influential factors for mobile learning acceptance among Chinese users. *Educational Technology Research and Development, 65*(1), 101-123. <https://doi.org/10.1007/s11423-016-9465-2>
- Harchay, A., Berquiga, A., Cheniti-Belcadi, L., & Braham, R. (2019). Student Perception of Mobile Self-assessment: An Evaluation of the Technology Acceptance Model. *Interaction Design and Architecture(s), 109-124*. <https://doi.org/10.55612/s-5002-041-008>
- Harlen, W., & James, M. (1997). Assessment and Learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice, 4*(3), 365-379. <https://doi.org/10.1080/096959497004304>
- Hébert, C., Jensen, J., & Terzopoulos, T. (2021). "Access to technology is the major challenge": Teacher perspectives on barriers to DGBL in K-12 classrooms. *E-Learning and Digital Media, 18*, 2042753021995315. <https://doi.org/10.1177/2042753021995315>
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115-135. <https://doi.org/10.1007/s11747-014-0403-8>
- Horvat, L., Balen, J., & Martinović, G. (2012). Proposal of mLearning system for written exams. *Proceedings ELMAR-2012, 345-348*.
- Hossain, S. F. A., Shan, X., Nurunnabi, M., Tushar, H., Mohsin, A. K. M., & Ahsan, F. T. (2021). Opportunities and Challenges of M-Learning During the COVID-19 Pandemic: A Mixed Methodology Approach. In *E-Collaboration Technologies and Strategies for Competitive Advantage Amid Challenging Times* (pp. 210-227). IGI Global. <https://doi.org/10.4018/978-1-7998-7764-6.ch007>
- Hu, P. J.-H., Clark, T. H. K., & Ma, W. W. (2003). Examining technology acceptance by school teachers: A longitudinal study. *Information & Management, 41*(2), 227-241. [https://doi.org/10.1016/S0378-7206\(03\)00050-8](https://doi.org/10.1016/S0378-7206(03)00050-8)
- Jin, Y., Lin, C.-L., Zhao, Q., Yu, S.-W., & Su, Y.-S. (2021). A Study on Traditional Teaching Method Transferring to E-Learning Under the Covid-19 Pandemic: From Chinese Students' Perspectives. *Frontiers in Psychology, 12*, 632787. <https://doi.org/10.3389/fpsyg.2021.632787>
- Kimmons, R., Clark, B., & Lim, M. (2017). Understanding web activity patterns among teachers, students and teacher candidates. *Journal of Computer Assisted Learning, 33*(6), 588-596. <https://doi.org/10.1111/jcal.12202>
- Knight, P. T. (2002). Summative Assessment in Higher Education: Practices in disarray. *Studies in Higher Education, 27*(3), 275-286. <https://doi.org/10.1080/03075070220000662>
- Lau, A. M. S. (2016). 'Formative good, summative bad?' – A review of the dichotomy in assessment literature. *Journal of Further and Higher Education, 40*(4), 509-525. <https://doi.org/10.1080/0309877X.2014.984600>
- MacLellan, E. (2001). Assessment for Learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education, 26*(4), 307-318. <https://doi.org/10.1080/02602930120063466>
- Marín-Díaz, V., Sampedro, B. E., Aznar, I., & Trujillo, J. M. (2022). Perceptions on the use of mixed reality in mobile environments in secondary education. *Education + Training, 65*(2), 312-323. <https://doi.org/10.1108/ET-06-2022-0248>
- Matas, A. (2018). Diseño del formato de escalas tipo Likert: Un estado de la cuestión. *Revista electrónica de investigación educativa, 20*(1), 38-47.
- Mejía-Pérez, O. (2012). De la evaluación tradicional a una nueva evaluación basada en competencias. *Revista Electrónica Educare, 16*(1), Article 1. <https://doi.org/10.15359/ree.16-1.3>
- Moccozet, L., Benkacem, O., Berisha, E., Trindade, R. T., & Bürgi, P.-Y. (2019). A versatile and flexible e-assessment framework towards more authentic summative examinations in higher-education. *International Journal of Continuing Engineering Education and Life Long Learning, 29*(3), 211-229. <https://doi.org/10.1504/IJCEELL.2019.101032>
- Moore, G. C., & Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research, 2*(3), 192-222.
- Moreno-Guerrero, A.-J., Rodríguez-Jiménez, C., Gómez-García, G., & Ramos Navas-Parejo, M. (2020). Educational Innovation in Higher Education: Use of Role Playing and Educational Video in Future Teachers' Training. *Sustainability, 12*(6), Article 6. <https://doi.org/10.3390/su12062558>
- Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education, 9*(3), e3292. <https://doi.org/10.1002/rev3.3292>
- Mutambara, D., & Bayaga, A. (2021). Learners' and teachers' acceptance of mobile learning: An exploratory study in a developing country. *International Journal of Learning Technology, 16*(2). <https://doi.org/10.1504/IJLT.2021.117763>
- Nikou, S. A., & Economides, A. A. (2021). A Framework for Mobile-Assisted Formative Assessment to Promote Students' Self-Determination. *Future Internet, 13*(5). <https://doi.org/10.3390/fi13050116>
- Nikou, S., & Economides, A. (2016). An Outdoor Mobile-Based Assessment Activity: Measuring Students' Motivation and Acceptance. *International Journal of Interactive Mobile Technologies (IJIM), 10*, 11-17. <https://doi.org/10.3991/ijim.v10i4.5541>
- Nikou, S., & Economides, A. (2017a). Mobile-Based Assessment: Integrating acceptance and motivational factors into a combined model of Self-Determination Theory and Technology Acceptance. *Computers in Human Behavior, 68*, 83-95. <https://doi.org/10.1016/j.chb.2016.11.020>
- Nikou, S., & Economides, A. (2017b). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education, Query date: 2022-03-22 16:32:53*. <https://www.sciencedirect.com/science/article/pii/S0360131517300283>
- Olimov, S. S. (2021). The innovation process is a priority in the development of pedagogical sciences. *European Journal of Research Development and Sustainability, 2*(3), Article 3.
- Patton, M. Q. (1996). A World Larger than Formative and Summative. *Evaluation Practice, 17*(2), 131-144. <https://doi.org/10.1177/109821409601700205>
- Pikkarainen, T., Pikkarainen, K., Karjaluoto, H., & Pahlila, S. (2004). Consumer acceptance of online banking: An extension of the Technology Acceptance Model. *Internet Research, 14*(3), 224-235. <https://doi.org/10.1108/10662240410542652>
- Rahmawati, R. N. (2019). Self-Efficacy and Use of E-learning: A Theoretical Review Technology Acceptance Model (TAM). *American Journal of Humanities and Social Sciences Research (AJHSSR), 3*(5), 41-55.
- Ramayah, T., Hwa, C., Chuah, F., Ting, H., & Memon, M. (2017). PLS-SEM using SmartPLS 3.0: Chapter 8: Assessment of Formative Measurement Models. En *Partial least squares structural equation modeling (PLS-SEM) using smartPLS 3.0: An Updated and Practical Guide to Statistical Analysis*. Pearson.
- Reisoğlu, İ., & Çebi, A. (2020). How can the digital competences of pre-service teachers be developed? Examining a case study through the lens of DigComp and DigCompEdu. *Computers & Education, 156*, 103940. <https://doi.org/10.1016/j.compedu.2020.103940>
- Rothmann, S. (2015). A structural model of technology acceptance. *South African Journal of Industrial Psychology, 41*.
- Sánchez-Prieto, J. C., Olmos-Migueláñez, S., & García-Peñalvo, F. J. (2017a). MLearning and pre-service teachers: An assessment of the behavioral intention using an expanded TAM model. *Computers in Human Behavior, 72*, 644-654. <https://doi.org/10.1016/j.chb.2016.09.061>
- Sánchez-Prieto, J. C., Olmos-Migueláñez, S., & García-Peñalvo, F. J. (2017b). ¿Utilizarán los futuros docentes las tecnologías móviles? Validación de una propuesta de modelo TAM extendido. *Revista de Educación a Distancia (RED), 52*, Article 52. <https://revistas.um.es/red/article/view/282191>
- Sánchez-Prieto, J., Hernández-García, Á., García-Peñalvo, F., Chaparro-Peláez, J., & Olmos, S. (2019). Break the Walls! Second-Order Barriers and the Acceptance of mLearning by First-Year Pre-Service Teachers. *Computers in Human Behavior, 95*. <https://doi.org/10.1016/j.chb.2019.01.019>

- Sar, A., & Misra, S. N. (2020). A study on policies and implementation of information and communication technology (ICT) in educational systems. *Materials Today*, 8. <https://doi.org/10.1016/j.matpr.2020.10.507>
- Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2016). Estimation issues with PLS and CBSEM: Where the bias lies! *Journal of Business Research*, 69(10), 3998-4010. <https://doi.org/10.1016/j.jbusres.2016.06.007>
- Scriven, M. (1967). The Methodology of Evaluation. En R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of Curriculum Evaluation* (pp. 39-83). Rand McNally.
- Scriven, M. (1991). Beyond Formative and Summative Evaluation. *Teachers College Record*, 92(6), 19-64. <https://doi.org/10.1177/016146819109200603>
- Sharma, P., & Kim, K. (2012). *Model Selection in Information Systems Research Using Partial Least Squares Based Structural Equation Modeling*. International Conference on Interaction Sciences. <https://www.semanticscholar.org/paper/Model-Selection-in-Information-Systems-Research-Sharma-Kim/cfde34aa3bd19983b07dc16fe2801cdd377b05d7>
- Sharma, P., Liengaard, B., Jr, H., Sarstedt, M., & Ringle, C. (2022). Predictive model assessment and selection in composite-based modeling using PLS-SEM: Extensions and guidelines for using CVPAT. *European Journal of Marketing*. <https://doi.org/10.1108/EJM-08-2020-0636>
- Shepard, L. (2006). La evaluación en el aula. In R. Brennan (Ed.). En *Educational Measurement* (4 Edition, pp. 623-646). Praeger Westport.
- Simonetto, A. (2012). Formative and reflective models: State of the art. *Electronic Journal of Applied Statistical Analysis*, 5(3), Article 3. <https://doi.org/10.1285/i20705948v5n3p452>
- Skulmowski, A., & Rey, G. D. (2020). COVID-19 as an accelerator for digitalization at a German university: Establishing hybrid campuses in times of crisis. *Human Behavior and Emerging Technologies*, 2(3), 212-216. <https://doi.org/10.1002/hbe2.201>
- Smith, C. A. (2021). Development and Integration of Freely Available Technology into Online STEM Courses to Create a Proctored Environment During Exams. *Journal of Higher Education Theory and Practice*, 4. <https://papers.iafor.org/submission59360/>
- Souabi, S., Retbi, A., Idrissi, M. K., & Bennani, S. (2021). Towards an Evolution of E-Learning Recommendation Systems: From 2000 to Nowadays. *International Journal of Emerging Technologies in Learning (IJET)*, 16(06), Article 06. <https://doi.org/10.3991/ijet.v16i06.18159>
- Stone, M. (1974). Cross-Validatory Choice and Assessment of Statistical Predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2), 111-133. <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x>
- Sun, Y., Li, N., Hao, J. L., Di Sarno, L., & Wang, L. (2022). Post-COVID-19 Development of Transnational Education in China: Challenges and Opportunities. *Education Sciences*, 12(6), Article 6. <https://doi.org/10.3390/educsci12060416>
- Terán-Guerrero, F. N. (2019). Acceptance of university students in the use of Moodle e-learning systems from the perspective of the TAM model. *UNEMI*, 12(29), 63-76.
- Thorsteinsson, G., & Niculescu, A. (2013). Examining teachers' mindset and responsibilities in using ICT. *Studies in Informatics and Control*, 22(2), 315-322. <https://doi.org/10.24846/v22i3y201308>
- Tyler, R. (1950). *Basic principle of curriculum and instruction*. Chicago University.
- Valverde-Berrocoso, J., Fernández-Sánchez, M. R., Dominguez, F. I. R., & Sosadiaz, M. J. (2021). The educational integration of digital technologies preCovid-19: Lessons for teacher education. *PLOS ONE*, 16(8), e0256283. <https://doi.org/10.1371/journal.pone.0256283>
- Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39, 273-315. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>
- Venkatesh, V., & Davis, F. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 184-204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Vieira, H., & Ribeiro, C. P. (2018). Implementing Flipped Classroom in History: The reactions of eighth grade students in a Portuguese school. *Yesterday and Today*, 19, 35-49. <https://doi.org/10.17159/2223-0386/2018/n18a3>
- Vilches, A., & Gil, D. (2010). Máster de formación inicial del profesorado de enseñanza secundaria. Algunos análisis y propuestas. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 661-666.
- Wang, R., Chen, L., & Solheim, I. (2020). Modeling dyslexic students' motivation for enhanced learning in E-learning systems. *The ACM Transactions on Interactive Intelligent Systems*, 2. <https://doi.org/10.1145/3341197>
- Wan-Sulaiman, W. N. A., & Mustafa, S. E. (2020). Usability Elements in Digital Textbook Development: A Systematic Review. *Publishing Research Quarterly*, 36(1), 74-101. <https://doi.org/10.1007/s12109-019-09675-3>

Anexo 5

Publicación en capítulo de libro: Editorial Tirant Lo Blanch

VERSIÓN PREPRINT

Ortiz-López, A.; Olmos-Migueláñez, S.; Sánchez-Prieto, J.C. (En prensa). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En E. Chiner (Ed.), *Nuevas tendencias interdisciplinares en Educación y Conocimiento*. Tirant Lo Blanch

INDICADORES:

Editorial Q1 Ranking SPI: ICEE 15



Universitat d'Alacant
Universidad de Alicante

La abajo firmante, Profa. Dra. Esther Chiner, como editora del libro “Nuevas tendencias interdisciplinarias en Educación y Conocimiento”, que será editado por la Editorial Tirant lo Blanch en 2023,

HACE CONSTAR QUE:

Ha sido aceptado el siguiente capítulo de libro para su publicación:

***Título del capítulo:* ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica**

***Autores (por orden):* Ortiz-López, Alberto, Olmos-Migueláñez, Susana & Sánchez-Prieto, José Carlos**

Título del libro: *Nuevas tendencias interdisciplinarias en Educación y Conocimiento*

Editorial: Tirant lo Blanch

Año de publicación: 2023

Y para que así conste y a los efectos oportunos, se expide el presente escrito, en lugar y fecha indicados,

Alicante, 16 de octubre de 2023

Profa. Dra. Esther Chiner

¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica

Alberto Ortiz-López, Susana Olmos-Migueláñez, José Carlos Sánchez-Prieto

Universidad de Salamanca (España)

Abstract: In the current digital transition in which educational systems are immersed, the evaluation of teaching must be a factor of quality, adaptation, and change, open to the inclusion of mobile technologies; without forgetting the teacher, a key element to achieve an effective and real inclusion, so it is essential to address their acceptance of the use of mobile devices in assessment. This research, starting from the identification of the constructs addressed in the most relevant models of technological acceptance in the field of assessment mediated by mobile devices and theoretical reflection, proposes a new research model based on the Technological Acceptance Model (TAM). This new model has undergone a content validation process conducted by a group of six national and international experts. The resulting model includes a total of eleven constructs and 18 hypotheses that justify their relationships and evaluates the final intention of educational use in a total of 50 items. Future research could consider using the model, already validated, to extend validity and reliability studies and to check the level of prediction of the new dimensions proposed.

Keywords: acceptance, m-learning, assessment, technology acceptance model, teachers.

1. INTRODUCCIÓN

La irrupción de las tecnologías de la información y la comunicación en educación ha supuesto un cambio radical en la totalidad de sus niveles (Gros-Salvat, 2018), revolucionando los procesos de enseñanza-aprendizaje gracias a los rápidos avances tecnológicos experimentados, los cuales posibilitan a los docentes contar con una infinidad de herramientas, medios y técnicas para plantear nuevos escenarios de aprendizaje (Colás-Bravo et al., 2018). Esta introducción tecnológica en educación ha permitido el desarrollo de enseñanzas en red (Souabi et al., 2021) y la virtualización de enseñanzas presenciales a campus digitales empleando dispositivos como ordenadores, tabletas o teléfonos móviles para su conexión y acceso (Delgado y Vélez, 2021). Entre estos avances destaca el exponencial desarrollo de los dispositivos móviles, siendo un elemento presente en prácticamente la totalidad de los alumnos que pasan hoy por el sistema educativo, con una presencia incluso superior a los ordenadores portátiles (Sánchez-Prieto et al., 2019), dadas sus posibilidades de inmediatez, la situatividad, la ubicuidad, la conveniencia y la contextualidad (Gabriel-Pedro et al., 2018).

En una sociedad inmersa en un proceso de transformación digital a todos los niveles y en todos los ámbitos (García-Peñalvo, 2005; Uriarte y Acevedo, 2018), se podría deducir que se espera que los docentes utilicen dispositivos móviles en procesos tan importantes como los evaluativos (Arancibia-Muñoz et al., 2017); actuando como docentes innovadores (Córica y García-Aretio, 2018) y teniendo esto una clara repercusión también en la concepción del alumnado, quienes demandan una mayor actualización académica a

los tiempos y a los medios disponibles, adaptando tanto el proceso de enseñanza-aprendizaje como los procesos de evaluación de las materias, uno de los componentes más rigurosos de la docencia (Fondo, 2019).

Investigaciones previas han abordado los factores condicionantes de la aceptación tecnológica en docentes en diferentes campos relativos al eLearning, siendo el más estudiado la aceptación de dispositivos móviles en enseñanza (Sánchez-Prieto et al., 2017). Este estudio trata de abordar un campo más específico, la aceptación de la evaluación a través de dichos dispositivos; sobre lo que se encuentran un menor número de estudios (Nikou y Economides, 2016), profundizando en las variables que predicen la intención de uso de tecnologías móviles en procesos evaluativos en el aula, dando respuesta a los objetivos y preguntas de investigación presentadas en la siguiente sección.

2. OBJETIVO Y PREGUNTAS DE INVESTIGACIÓN

El objetivo del estudio es la validación de un nuevo modelo, basado en el modelo TAM, para evaluar la aceptación tecnológica de dispositivos móviles para evaluación en docentes. Además, se plantean las siguientes preguntas de investigación:

- Q1. ¿Qué variables pueden ayudar a predecir la intención de uso de tecnologías móviles en evaluación por parte de los docentes?
- Q2. ¿Qué factores influyen y condicionan la aceptación tecnológica, por parte de docentes, de los dispositivos móviles en procesos de evaluación?

Las respuestas a las preguntas de investigación planteadas pretenden describir los factores determinantes en la aceptación docente de evaluaciones a través de dispositivos móviles, para lo que se propone la validación de un nuevo modelo de aceptación tecnológica. La sección tercera plantea la introducción a los modelos actuales. La cuarta sección es la reflexión teórica, las variables, las hipótesis y el modelo de investigación propuesto. En la sección quinta, se presenta el análisis de los datos y los procesos de validación junto con sus resultados; en la sección sexta se discuten y analizan los resultados e implicaciones de estos y, finalmente, se resumen las contribuciones del estudio.

3. MODELOS DE ACEPTACIÓN: TAM Y MODELOS POSTERIORES

El modelo de aceptación tecnológica (*Technology Acceptance Model, TAM*) (Davis, 1989) trata de predecir la intención conductual de uso a través de la actitud y el efecto mediador de la intención conductual. El modelo TAM mide la intención conductual de uso en base a dos factores principales, la utilidad percibida y la facilidad de uso percibida (Figura 1), y sienta las bases de los modelos y estudios de aceptación posteriores.

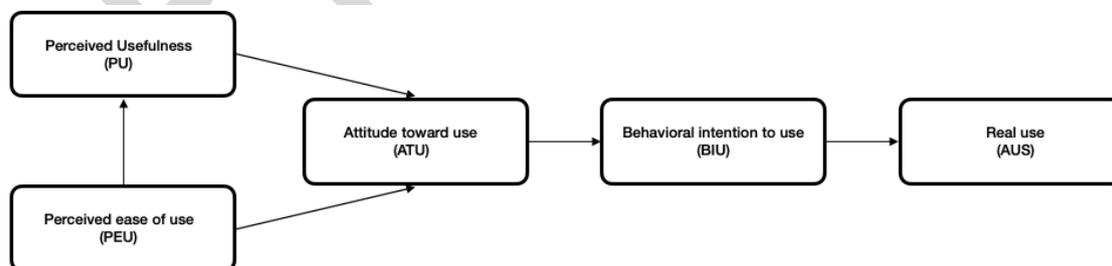


Figura 1. Modelo TAM (Davis, 1989)

El modelo TAM evoluciona y surgen el modelo TAM 2 (Venkatesh y Davis, 2000) que atiende a factores externos que influyen en el uso de la tecnología, y la última versión, TAM 3, que incluye el efecto de la confianza y el riesgo percibido de los sistemas (Venkatesh y Bala, 2008). En el campo educativo, el modelo TAM también ha sido

desarrollado y es en la última década en la que el modelo está experimentando su mayor presencia, ya que las TIC están irrumpiendo en el sistema educativo a todos los niveles y la aceptación se ha convertido en uno de los principales temas de estudio tanto en enseñanzas presenciales como en la modalidad eLearning (Bizzo, 2021; Pino-Varela, 2022).

Los modelos derivados han ocupado ámbitos específicos del conocimiento como la aceptación tecnológica de dispositivos móviles (modelo MAM-PET) (Sánchez-Prieto, 2018) o, en estadios más iniciales, la aceptación de éstos en evaluaciones, como el modelo MBAAM (Nikou y Economides, 2017a). Este último campo incipiente del conocimiento será el que centrará el desarrollo de un nuevo modelo de aceptación, que parte de las teorías y modelos aquí descritos y que tratará de explicar nuevos factores condicionantes sobre la aceptación de dispositivos móviles en evaluación por parte del colectivo docente.

4. MÉTODO: PROPUESTA Y VALIDACIÓN DE UN NUEVO MODELO

El modelo propuesto parte del modelo TAM (Davis, 1989), y amplía su foco con otros constructos de distintas teorías y variables relativas a los procesos de evaluación: intención de uso (BIU), facilidad de uso percibido (PEOU), utilidad percibida (PU), norma subjetiva (SN), resistencia al cambio (RC), autonomía (AUT), autoeficacia móvil (MSE) e innovación personal (PI); las variables más relevantes en los modelos de aceptación centrados en evaluación a través de dispositivos móviles.

Tras la realización de una revisión sobre dichos modelos, se ha determinado la necesidad de plantear una nueva propuesta que combine otras variables y que plantee la introducción de constructos propios que, según la literatura, podrían influir en esta aceptación. Para ello, se incorporan otros modelos no vinculados a la e-evaluación pero relevantes en sus estudios de aceptación de la norma subjetiva (SN, adaptada de TAM 3) (Venkatesh y Bala, 2008) y la resistencia al cambio (RC, adaptada de MAM-PET) (Sánchez-Prieto, 2018), así como constructos derivados de la reflexión teórica: la ansiedad ante la evaluación (ASAN), la ventaja relativa de la evaluación formativa (RAFA) y la ventaja relativa de la evaluación sumativa (RASA).

Por tanto, una vez seleccionados los constructos que formarán parte del modelo inicial, se han formulado las hipótesis relacionales que justifican el modelo (Figura 2).

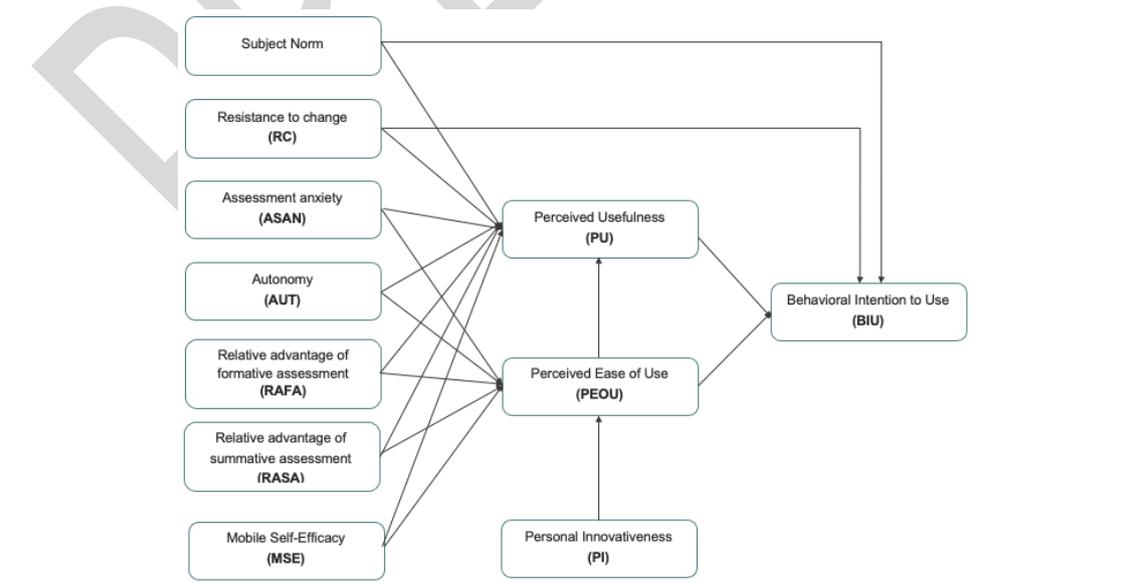


Figura 2. Modelo inicial propuesto para la evaluación de la aceptación

BIU	6,48	,352	6	6,25	,639	6	5,20	,630	6
PEOU	5,98	,457	6	6,02	,362	6	5,11	,774	6
PU	6,66	,264	6	6,55	,366	6	5,92	,276	6
SN	6,25	,356	6	6,26	,426	6	6,24	,822	6
RC	5,63	,217	6	6,70	,137	6	6,44	,416	6
ASAN	6,17	,750	6	6,23	,710	6	5,684	,807	6
AUT	4,86	,125	6	4,93	,309	6	4,82	,629	6
RAFA	6,63	,165	6	6,49	,204	6	6,28	,360	6
RASA	6,72	,134	6	6,63	,195	6	6,35	,264	6
MSE	6,33	,441	6	6,17	,454	6	6,08	,739	6
PI	6,44	,520	6	6,35	,565	6	6,41	,383	6

Respecto a la validez de las dimensiones y a la validez global del instrumento, éstas también resultaron positivas (puntuaciones medias comprendidas entre 5,83 y 6,67). El único valor por debajo de lo permitido se encuentra nuevamente en la autonomía (5,33), y confirma la necesidad de reestructurar dicha dimensión. Los resultados pueden ser consultados en el enlace previo.

Seguidamente, los expertos valoraron la validez general del instrumento, determinando si el mismo se adecúa al objetivo del estudio y al objeto de medida, y otorgando una alta validez global de medida al mismo (Tabla 3).

Tabla 3. Validez global del instrumento

	Adecuación al objetivo			Validez global de medida		
	AVG	SDT	N	AVG	SDT	N
Instrumento	6,333	,816	6	6,085	,367	6

La validez ha sido evaluada también mediante el coeficiente de validez de contenido (CVC) (Hernández-Nieto, 2002), uno de los más recomendados en estos procesos de validación (Pedrosa et al., 2013), que recomienda mantener ítems con CVC superior a 0,80 y modificarlos en valores comprendidos entre 0,70-0,80 (Balbinotti, 2005). De los ítems, nueve contaban con un CVC comprendido entre 0,70 y 0,80 (modificados) y tres han obtenido una puntuación menor (eliminados). Los CVC pueden consultarse en: <http://bit.ly/3YhEbwX>

Las dimensiones muestran un CVC alto (valores entre 0,815 y 0,938), salvo la dimensión autonomía (0,697). El instrumento obtiene un CVC medio de 0,867; por lo que se puede afirmar que cumple con la validez de contenido tras la validación realizada. Respecto a la evaluación cualitativa, centrados los comentarios en la séptima dimensión (autonomía), los revisores han propuesto su reformulación partiendo de los trabajos de Reeve (Reeve y Cheon, 2021; Su y Reeve, 2011).

5. RESULTADOS: DESCRIPCIÓN DEL MODELO VALIDADO

Una vez validado el modelo por jueces expertos y realizadas las modificaciones, se desarrolló la propuesta final con los constructos e hipótesis (Figura 3). El modelo (50 ítems) puede consultarse en el siguiente enlace: <https://bit.ly/3MBRwwc>

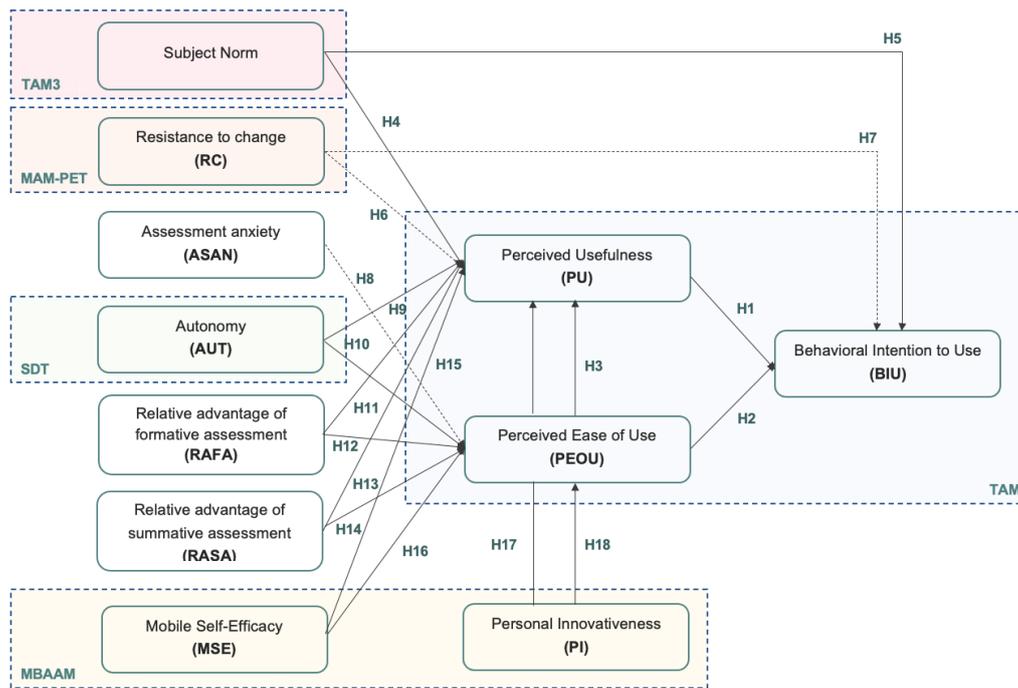


Figura 3. Modelo final

Para la comprensión de las relaciones planteadas y de los constructos que conforman el modelo, se plantea a continuación la justificación teórica de las hipótesis establecidas.

5.1. Intención de uso, utilidad percibida y facilidad de uso percibido

El modelo TAM (Davis, 1989) mide la intención conductual de uso en base a la utilidad percibida y la facilidad de uso percibida. La utilidad percibida (PU) plantea que todo aquel docente que comprenda las posibilidades y utilidades de un sistema asumirá una mayor usabilidad del mismo (Davis, 1989). Por tanto, este constructo tendrá una influencia positiva sobre la intención de uso, por lo que:

- **H1.** La utilidad percibida predice de forma positiva la intención de los profesores de utilizar dispositivos móviles para la e-evaluación.

Por otro lado, Davis (1989) plantea la facilidad de uso que los docentes perciben sobre estas herramientas (PEOU) como un factor determinante para su uso. Es decir, si el uso de herramientas de e-evaluación disminuyera su carga de trabajo estos se posicionarían más a favor de su uso. Sánchez-Prieto et al. (2019) defienden además que, en las primeras fases de adopción de herramientas tecnológicas, la facilidad de uso percibida puede llegar a ser una barrera interna. Por tanto, se propone que:

- **H2.** La facilidad de uso percibida predice positivamente la intención de los profesores de utilizar dispositivos móviles para la e-evaluación.
- **H3.** La facilidad de uso percibida por los profesores en procesos de e-evaluación predice positivamente la utilidad percibida.

5.2. Norma subjetiva

La decisión de los docentes de utilizar herramientas de e-evaluación en sus procesos de enseñanza-aprendizaje no sólo depende de un proceso interno en el que tanto la práctica como la experiencia actúan como factores modeladores e influyentes, sino que también se han de contemplar factores externos que pueden condicionar la intención de uso final de la tecnología (Teo, 2015). Estas presiones se categorizan en tres niveles jerárquicos adaptados al caso de la enseñanza: el nivel superior (equipos directivos), el nivel de

iguales (equipo docente) y el nivel inferior (alumnado); y modelan la intención de uso (Teo, 2010), la facilidad y la utilidad percibida (Teo, 2019). Por tanto, se propone que:

- **H4.** La norma subjetiva predice negativamente la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
- **H5.** La norma subjetiva predice negativamente la intención de los profesores de utilizar dispositivos móviles para la e-evaluación.

5.3. Resistencia al cambio

La resistencia al cambio es definida como la dificultad para cambios en una rutina establecida o situaciones estresantes ante una expectativa de cambio (Sánchez-Prieto, 2005). Es decir, es un factor condicionante negativo de la intención de uso (Al-Somali et al., 2009). Se entiende como resistencia al cambio la negativa docente al empleo de e-evaluaciones en su docencia, manteniendo procesos evaluativos tradicionales y desechando el empleo de los dispositivos móviles en evaluación (Karahanna et al., 2006; Sánchez-Prieto et al., 2019). Las hipótesis de este constructo se basan en las investigaciones de Battacherjee y Hikmet (2007) y Guo et al. (2012):

- **H6.** La resistencia al cambio predice negativamente la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
- **H7.** La resistencia al cambio predice negativamente la intención de los profesores de utilizar dispositivos móviles para la e-evaluación.

5.4. Ansiedad ante la evaluación

Estudios de aceptación recientes han abordado la ansiedad en relación con la tecnología, los dispositivos tecnológicos, o las gestiones en línea (Belanger et al., 2002; Leso y Peck, 1992; Saade y Kira, 2007; Yu et al., 2009); pero es un constructo poco explorado en los estudios de aceptación de la ansiedad ante procesos evaluativos.

Esta ansiedad hacia la evaluación es entendida como la tendencia a manifestar conductas estresadas en situaciones evaluativas (Torrano-Martínez et al., 2017). El estudio de la ansiedad se ha centrado en el colectivo de los estudiantes (Frías-Navarro et al., 2020), pero también es importante atender a cómo afrontan los docentes situaciones evaluativas en nuevos entornos de enseñanza. Para ello, tomando como partida la ansiedad propuesta por Venkatesh y Bala (2008), se plantea aquí una adaptación de los ítems para la evaluación de la ansiedad ante la evaluación asumiendo que la ansiedad tiene un efecto negativo sobre la facilidad de uso percibido, por lo que:

- **H8.** La ansiedad ante la evaluación tendrá un efecto negativo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

5.5. Autonomía

El uso de dispositivos móviles en procesos evaluativos dota de una mayor autonomía al estudiante y al docente en estos procesos y posibilita un mayor control personal sobre el momento de la evaluación, experimentando un sentido de la propiedad (Jones y Issroff, 2007). La autonomía permite también una mayor independencia de la evaluación, lo que deriva en que se perciban estos procesos como útiles y fáciles (Reeve y Cheon, 2021; Su y Reeve, 2011; Cheon et al., 2012), influyendo de forma positiva en la utilidad de uso y en la facilidad percibida según la Teoría de la Autodeterminación (Deci y Ryan, 1985); por lo que se propone que:

- **H9.** La autonomía tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

- **H10.** La autonomía tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

5.6. Ventajas relativas de las evaluaciones formativa y sumativa

La ventaja relativa (grado en el que una innovación se percibe como mejor que la idea que sustituye) (Rogers, 1962) es un constructo que emerge en estudios de aceptación en la IDT Refinada (Moore y Benbasat, 1991). La ventaja relativa ha sido ya justificada como factor influyente en la intención de uso, influyendo de forma positiva tanto en la utilidad percibida como en la facilidad de uso (Henderson et al., 2012). En este caso, se distinguirá si los profesores diferencian entre ventaja relativa para la evaluación formativa y ventajas relativas para la evaluación sumativa (Scriven, 1967) tratando de comprender la influencia de ambos tipos de evaluación en la intención de uso. Por tanto, las hipótesis establecidas para las dos modalidades evaluativas son:

- **H11.** La ventaja relativa de la evaluación formativa tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
- **H12.** La ventaja relativa de la evaluación formativa tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
- **H13.** La ventaja relativa de la evaluación sumativa tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
- **H14.** La ventaja relativa de la evaluación sumativa tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

5.7. Autoeficacia en el uso de dispositivos móviles

La autoeficacia móvil es definida como la percepción del individuo sobre su capacidad para utilizar los dispositivos móviles para realizar determinadas tareas (Nikou y Economides, 2017b). Esta variable es considerada relevante para la explicación de la intención en su variante de “autoeficacia en el uso tecnológico”, y son Nikou y Economides (2017b) quienes la introducen en el campo de la evaluación a través de dispositivos móviles. Por tanto, se podrá deducir que cuanto mayor sea el nivel de habilidad con el uso de dispositivos mayor será la facilidad de uso y la utilidad percibida en procesos evaluativos (Pikkarainen et al., 2004; Wang et al., 2020). En base a estas premisas, se propone que:

- **H15.** La autoeficacia móvil tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.
- **H16.** La autoeficacia móvil tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

5.8. Innovación personal

La innovación personal es definida como la disposición de un individuo al empleo de una nueva tecnología (Agarwal y Prasad, 1998). Además, esta capacidad para innovar actúa como un factor importante en la determinación de la aceptación tecnológica (Turan et al., 2015; Yi et al., 2006), y se considera alta cuando una persona adopta una innovación de forma temprana (Xu y Gupta, 2009). Esta innovación aumentará la intención de uso de herramientas en los procesos de enseñanza; aumentando tanto la utilidad percibida como la facilidad de uso percibida (Nikou y Economides, 2015). Por tanto, se plantea que:

- **H17.** La innovación personal tendrá un efecto positivo en la utilidad percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

- **H18.** La innovación personal tendrá un efecto positivo en la facilidad de uso percibida por los docentes en el uso de dispositivos móviles para la e-evaluación.

6. DISCUSIÓN Y CONCLUSIONES

La evaluación a través de dispositivos móviles afronta un nuevo reto en las modalidades de enseñanza virtual y en las metodologías que hoy se implementan en las aulas, lo que sitúa al docente en el centro del proceso (Riveros et al., 2021). Por tanto, es importante asumir que la aceptación docente tiene que actuar como elemento mediador entre las prácticas de evaluación y la tecnología, y es fundamental establecer políticas de formación, capacitación y preparación docente que conduzcan a una aceptación efectiva en la práctica (Cabero-Almenara et al., 2018; Jiménez-Martínez et al., 2022).

Esta investigación ha constatado la existencia de publicaciones centradas en los estudios de aceptación tecnológica en ordenadores en el ámbito educativo, pero se han encontrado escasos resultados centrados en la evaluación a través de dispositivos móviles, pese a ser un campo con interés creciente (Fombona-Cadavieco y Rodil-Pérez, 2018; Villavicencio y Paredes, 2020; Williams-Muller et al., 2020). En la actualidad solamente se encuentran escasos modelos que han tratado de abordar esta temática, especialmente los desarrollados por Nikou y Economides (2015, 2017a, 2017b, 2019), quienes basan sus modelos en el modelo TAM y lo complementan con constructos provenientes de diferentes modelos.

De la literatura consultada sobre esta modalidad de evaluación, y del análisis teórico de los modelos de aceptación más relevantes, se han determinado nuevas dimensiones que pueden ser de especial relevancia para el estudio de esta aceptación, por lo que la propuesta que aquí se realiza consiste en un nuevo modelo diseñado para estudiar el nivel de aceptación tecnológica de los docentes al evaluar con dispositivos móviles

La propuesta que aquí se realiza toma cinco constructos nuevos que no probados en el campo del MBA, y el modelo se sustenta a partir del modelo TAM y se amplía con otros modelos y teorías; además de tres constructos propios no derivados de otros modelos. Estos constructos no explorados son la ansiedad ante los procesos de evaluación, entendiendo que un docente que experimente estados de ansiedad tendrá un nivel de predisposición menor, y la ventaja relativa de la evaluación. Este último diferencia entre evaluación formativa y sumativa, y no solo trata de determinar si una modalidad u otra favorece más la aceptación, si no que permite la comparación futura entre las dos modalidades en el empleo de dispositivos móviles en cada una.

El resultado es un modelo teórico que aúna los constructos más relevantes para explicar la aceptación docente a través de once variables (dimensiones) que se relacionan mediante 18 hipótesis y 50 ítems. Este modelo ha sido aquí sometido a un proceso de validación profunda por seis expertos nacionales e internacionales, asegurando la validez de su contenido.

6.1. Limitaciones y prospectiva

En cuanto a las limitaciones de la investigación, destaca el posible sesgo derivado de la representatividad de los jueces elegidos (Robles y Del Carmen, 2015). Para su minimización, se ha establecido el reparto de jueces en dos expertos en evaluación, dos expertos en sistemas e-Learning y m-Learning y dos expertos en el campo de la metodología.

Esta propuesta establece una nueva línea de investigación en el campo de la aceptación tecnológica de la evaluación a través de dispositivos móviles, y el modelo aquí presentado debe ser aplicado en un entorno empírico real, permitiendo la realización de un estudio

de validez estadística con el fin de asegurar y ampliar la propuesta. Otra línea posible es el estudio específico de las nuevas dimensiones que aquí se han presentado y justificado como factores relevantes, para describir así si pueden actuar como un nuevo determinante mediador en la aceptación; con el objetivo último de establecer todos los factores que condicionan la aceptación docente del empleo de dispositivos móviles en evaluación, para poder contar así con un modelo ampliamente validado que permita una transición efectiva y real a la nueva educación mediada por la tecnología.

REFERENCIAS

- Agarwal, R., y Prasad, J. (1998). A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research*, 9(2), 204-215. <https://doi.org/10.1287/isre.9.2.204>
- Al-Somali, S. A., Gholami, R., y Clegg, B. (2009). An investigation into the acceptance of online banking in Saudi Arabia. *Technovation*, 29(2), 130-141.
- Arancibia-Muñoz, M. L., Halal-Orfali, C., y Romero-Alonso, R. (2017). Valoración y barreras en la integración del e-portafolio en el proceso de práctica inicial por parte de docentes y estudiantes de Educación Superior. *Pixel-Bit*, 51, 151-163. <https://doi.org/10.12795/pixelbit.2017.i51.10>
- Balbinotti, M. A. A. (2005). Para se avaliar o que se espera: Reflexões acerca da validade dos testes psicológicos. *Aletheia*, 21, 43-52.
- Belanger, F., Hiller, J. S., y Smith, W. J. (2002). Trustworthiness in electronic commerce: The role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, 11(3), 245-270. [https://doi.org/10.1016/S0963-8687\(02\)00018-5](https://doi.org/10.1016/S0963-8687(02)00018-5)
- Bhattacharjee, A., y Hikmet, N. (2007). Physicians' resistance toward healthcare information technologies: A dual-Factor model. *Abstracts Proceedings*, 141. <https://doi.org/10.1109/HICSS.2007.437>
- Bizzo, E. (2021). Aceptación y a la adopción del e-learning en los países en desarrollo: Una revisión de la literatura. *Ensaio: Avaliação e Políticas Públicas em Educação*, 30, 458-483.
- Cabero-Almenara, J., Gallego, O., Puentes, Á., y Jiménez, T. (2018). La "aceptación de la tecnología de la formación virtual" y su relación con la capacitación docente en formación virtual. *EDMETIC*, 7(1), 225-241.
- Cheon, J., Lee, S., Crooks, S. M., y Song, J. (2012). An investigation of mobile learning readiness in higher education based on the theory of planned behavior. *Computers & Education*, 59(3), 1054-1064. <https://doi.org/10.1016/j.compedu.2012.04.015>
- Colás-Bravo, M. P. C., de Pablos, J., y Ballesta, J. (2018). Incidencia de las TIC en la enseñanza en el sistema educativo español: Una revisión de la investigación. *Revista de Educación a Distancia (RED)*, 56, Article 56. <https://revistas.um.es/red/article/view/321471>
- Córica, J. L., y García-Aretio, L. G. (2018). Teachers' reluctance to ICT in Argentina. A qualitative approach. *Educación Superior*, 25, Article 25. <https://revistavipi.uapa.edu.do/index.php/edusup/article/view/164>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340.

<https://doi.org/10.2307/249008>

- Deci, E. L., y Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US. <https://doi.org/10.1007/978-1-4899-2271-7>
- Delgado, J. E., y Vélez, J. M. (2021). La plataforma MOODLE: Caracterización, aplicaciones y beneficios para las competencias docentes. *Revista Cognosis*, 6(4), Article 4. <https://doi.org/10.33936/cognosis.v6i4.3046>
- Escobar-Pérez, J., y Cuervo, A. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medicina*, 6, 27-36.
- Fombona-Cadavieco, J., y Rodil-Pérez, F. J. (2018). Niveles de uso y aceptación de los dispositivos móviles en el aula. *Pixel-Bit. Revista de Medios y Educación*, 52, Article 52. <https://doi.org/10.12795/pixelbit.2018.i52.02>
- Fondo, M. del C. (2019). Seis competencias docentes clave para el siglo XXI. *marcoELE. Revista de Didáctica Español Lengua Extranjera*, 29. <https://www.redalyc.org/journal/921/92159587007/>
- Gabriel-Pedro, L. F. M., de Oliveira-Barbosa, C. M. M., y Neves-Santos, C. M. das. (2018). A critical review of mobile learning integration in formal educational contexts. *International Journal of Educational Technology in Higher Education*, 15(1), 10. <https://doi.org/10.1186/s41239-018-0091-4>
- García-Llorente, H. J., Martínez-Abad, F., y Rodríguez-Conde, M. J. (2019). Validación de un instrumento de evaluación de competencias informacionales autopercibidas en educación secundaria obligatoria. *Anales de Documentación*, 22(1), Article 1. <https://doi.org/10.6018/analesdoc.22.1.305641>
- García-Peñalvo, F. J. (2005). Estado actual de los sistemas e-learning. *Education in the knowledge society (EKS)*, 6(2), 1.
- Gros-Salvat, B. (2018). La evolución del e-learning: Del aula virtual a la red. *RIED. Revista iberoamericana de educación a distancia*, 21(2), 69-82. <https://doi.org/10.5944/ried.21.2.20577>
- Guo, X., Sun, Y., Wang, N., Peng, J., y Yan, Z. (2012). The dark side of elderly acceptance of preventive mobile health services in China. *Electronic Markets*, 23(1). <https://doi.org/10.1007/s12525-012-0112-4>
- Hasani, I. B., Chroqui, R., Okar, C., y Talea, M. (2017). Literature review: All about IDT and TAM. *EST Berrechid*. 1st edition of a scientific day: performance management.
- Henderson, D., Sheetz, S. D., y Trinkle, B. S. (2012). The determinants of inter-organizational and internal in-house adoption of XBRL: A structural equation model. *International Journal of Accounting Information Systems*, 13(2), 109-140.
- Hernández-Nieto, R. A. (2002). *Contributions to Statistical Analysis*. Universidad de Los Andes.
- Jiménez-Martínez, K. A., Zamudio-Rodríguez, B. R., y Martínez-Moreno, M. (2022). Evaluación de la aceptación de las herramientas digitales para la enseñanza a través del modelo de aceptación tecnológica. *Revista Internacional de Desarrollo Regional Sustentable*, 6(1-2), Article 1-2.
- Jones, A., y Issroff, K. (2007). Motivation and mobile devices: Exploring the role of appropriation and coping strategies. *ALT-J: Research in Learning Technology*, 15(3). <https://doi.org/10.1080/09687760701673675>

- Karahanna, E., Agarwal, R., y Angst, C. M. (2006). Reconceptualizing compatibility beliefs in technology acceptance research. *MIS Quarterly*, 30(4), 781-804.
- Leso, T., y Peck, K. L. (1992). Computer anxiety and different types of computer courses. *Journal of Educational Computing Research*, 8(4), 469-478.
- Martínez-Abad, F. (2013). Evaluación y formación en competencias informacionales en la educación secundaria obligatoria. Universidad de Salamanca.
- Matas, A. (2018). Diseño del formato de escalas tipo Likert: Un estado de la cuestión. *Revista electrónica de investigación educativa*, 20(1), 38-47.
- Moore, G. C., y Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research*, 2(3), 192-222.
- Nikou, S., y Economides, A. (2016). An outdoor mobile-based assessment activity: Measuring students' motivation and acceptance. *International Journal of Interactive Mobile Technologies (iJIM)*, 10, 11-17.
- Nikou, S., y Economides, A. (2017a). Mobile-based assessment: Integrating acceptance and motivational factors into a combined model of Self-Determination Theory and Technology Acceptance. *Computers in Human Behavior*, 68, 83-95. <https://doi.org/10.1016/j.chb.2016.11.020>
- Nikou, S., y Economides, A. (2017b). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 109, 56-73. <https://doi.org/10.1016/j.compedu.2017.02.005>
- Nikou, S., y Economides, A. (2015). *The role of Personal Innovativeness and Previous Experience in explaining and predicting Mobile-based Assessment adoption*. E-LEARN 2015 - AACE World Conference on E-Learning, Association for the Advancement of Computing in Education, Hawaii.
- Pascual-Soler, M., Broseta-Aguilar, E., Monterde-i-Bort, H., Soriano-Pastor, J., Fornés-Vives, J., Fuentes-Durán, M. del C., García-Pérez, F., Dolz-Serra, L., y D'Ocón-Gimenez, A. (2020). Escala breve de ansiedad ante la evaluación académica (EBAEA-3): Brief scale of anxiety of the academic evaluation (EBAEA-3). *ENSAYOS. Revista de la Facultad de Educación de Albacete*, 35(1), Article 1.
- Pedrosa, I., Suárez-Álvarez, J., y García-Cueto, E. (2013). Evidencias sobre la validez de contenido: Avances teóricos y métodos para su estimación. *Acción Psicológica*, 10(2), 3-18. <https://doi.org/10.5944/ap.10.2.11820>
- Pikkarainen, T., Pikkarainen, K., Karjaluoto, H., y Pahlila, S. (2004). Consumer acceptance of online banking: An extension of the Technology Acceptance Model. *Internet Research*, 14(3), 224-235. <https://doi.org/10.1108/10662240410542652>
- Pino-Varela, J. J. (2022). Validación del Modelo de Aceptación Tecnológica (TAM) para medir la competencia digital en estudiantes de educación primaria. *EDMETIC*, 11(1), <https://doi.org/10.21071/edmetic.v11i1.13508>
- Reeve, J., y Cheon, S. H. (2021). Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist*, 56(1), 54-77. <https://doi.org/10.1080/00461520.2020.1862657>
- Riveros, J. M. G., Pimentel, J. F. F., Meza, L. C. F., y Solís, A. R. M. (2021). Evaluación formativa: Un reto para el docente en la educación a distancia. *Delectus*, 4(2),

Article 2. <https://doi.org/10.36996/delectus.v4i2.130>

- Robles, P. R., y Del Carmen, M. (2015). La validación por juicio de expertos: Dos investigaciones cualitativas en Lingüística aplicada. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 18, Article 18.
- Rogers, E. M. (1962). *Diffusion of innovations*. Free Press of Glencoe.
- Saade, R., y Kira, D. (2007). Mediating the impact of technology usage on perceived ease of use by anxiety. *Computers & Education*, 49(4), 1189-1204. <https://doi.org/10.1016/j.compedu.2006.01.009>
- Sánchez-Prieto, J. C. (2018). Diseño de un modelo de adopción tecnológica para evaluar la aceptación de tecnologías móviles en el profesorado de primaria [Tesis Doctoral]. Universidad de Salamanca.
- Sánchez-Prieto, J. C., Olmos-Migueláñez, S., y García-Peñalvo, F. J. (2017). M-Learning and pre-service teachers: An assessment of the behavioral intention using an expanded TAM model. *Computers in Human Behavior*, 72, 644-654.
- Sánchez-Prieto, J., Hernández-García, Á., García-Peñalvo, F., Chaparro-Peláez, J., y Olmos, S. (2019). Break the Walls! Second-Order Barriers and the Acceptance of mLearning by First-Year Pre-Service Teachers. *Computers in Human Behavior*, 95.
- Sánchez-Prieto, J., Olmos-Migueláñez, S., y García-Peñalvo, F. (2015). Informal tools in formal contexts: Development of a model to assess the acceptance of mobile technologies among teachers. *Computers in Human Behavior*, 55.
- Scriven, M. (1967). The Methodology of Evaluation. En R. W. Tyler, R. M. Gagne, y M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Rand McNally.
- Souabi, S., Retbi, A., Idrissi, M. K., y Bennani, S. (2021). Towards an evolution of e-Learning recommendation systems: From 2000 to nowadays. *International Journal of Emerging Technologies in Learning (IJET)*, 16(06).
- Su, Y.-L., y Reeve, J. (2011). A Meta-analysis of the Effectiveness of Intervention Programs Designed to Support Autonomy. *Educational Psychology Review*, 23(1), 159-188. <https://doi.org/10.1007/s10648-010-9142-7>
- Teo, T. (2019). The impact of subjective norm and facilitating conditions on pre-service teachers' attitude toward computer use: A structural equation modeling of an extended Technology Acceptance Model. *Journal of Educational Computing Research*, 40(1), 89-109. <https://doi.org/10.2190/EC.40.1.d>
- Teo, T. (2010). Examining the influence of subjective norm and facilitating conditions on the intention to use technology among pre-service teachers: A structural equation modeling of an extended technology acceptance model. *Asia Pacific Education Review*, 11(2), 253-262. <https://doi.org/10.1007/S12564-009-9066-4>
- Teo, T. (2015). Comparing pre-service and in-service teachers' acceptance of technology: Assessment of measurement invariance and latent mean differences. *Computers & Education*, 83. <https://doi.org/10.1016/j.compedu.2014.11.015>
- Torrano-Martínez, R., Ortigosa-Quiles, J. M., Riquelme-Marín, A., y López-Pina, J. A. (2017). Evaluación de la ansiedad ante los exámenes en estudiantes de Educación Secundaria Obligatoria. *Revista de Psicología Clínica con Niños y Adolescentes*, 4(2), 103-110.
- Turan, A., Tunç, A. Ö., y Zehir, C. (2015). A theoretical model proposal: Personal

innovativeness and user involvement as antecedents of Unified Theory of Acceptance and Use of Technology. *Procedia - Social and Behavioral Sciences*, 210, 43-51. <https://doi.org/10.1016/j.sbspro.2015.11.327>

- Uriarte, L. M., y Acevedo, M. (2018). Sociedad Red y transformación digital: Hacia una evolución de la consciencia de las organizaciones. *Economía industrial*, 407, 35-49.
- Venkatesh, V., y Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39, 273-315.
- Venkatesh, V., y Davis, F. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46(2), 186-204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Villavicencio, A. M. V., y Paredes, E. V. P. (2020). Relación del uso del teléfono celular y los niveles de atención en el proceso de enseñanza – aprendizaje. *Encuentros*, 18(1), 11-22.
- Wang, R., Chen, L., y Solheim, I. (2020). Modeling dyslexic students' motivation for enhanced learning in E-learning systems. *The ACM Transactions on Interactive Intelligent Systems*, 2. <https://doi.org/10.1145/3341197>
- Williams-Muller, E., Castro-Galviz, M. Y., Cruz-López, M. D. la, Pinillas, L. N., García, L. R., y Carvalho, J. L. T. (2020). Uso de los teléfonos móviles en el aula de educación primaria. *Revista Electrónica de Conocimientos, Saberes y Prácticas*, 3(2). <https://doi.org/10.5377/recsp.v3i2.10688>
- Xu, H., y Gupta, S. (2009). The effects of privacy concerns and personal innovativeness on potential and experienced customers' adoption of location-based services. *Electronic Markets*, 19(2-3), 137-149.
- Yi, M., Jackson, J., Park, J., y Probst, J. (2006). Understanding information technology acceptance by individual professionals: Toward an integrative view. *Information & Management*, 43(3), 350-363.
- Yu, J.-C., Kuo, L.-H., Chen, L.-M., Yang, H.-J., Yang, H.-H., y Hu, W.-C. (2009). Assessing and managing mobile technostress. *WSEAS Transactions on Communications*, 8, 416-425.

Anexos 8A-8E
Comunicaciones a Congreso
derivadas de la Tesis Doctoral

Anexo 8A

Comunicación a Congreso Internacional XI Conference on Virtual Campus (2021)

VERSIÓN POST-PRINT

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2021). eLearning, assessment and Higher Education in times of pandemic: A mapping of the literature. En A. García-Holgado, F.J. García-Peñalvo, C.S. González-González, A. Infante-Moro & J.C. Infante-Moro (Eds.), *2021 XI International Conference on Virtual Campus (JICV)* (pp. 1-4). IEEExplore.
DOI: 10.1109/JICV53222.2021.9600437.

Enlace a la publicación completa:

<https://ieeexplore.ieee.org/document/9600437/metrics#metrics>

eLearning, assessment and Higher Education in times of pandemic: A mapping of the literature

Alberto Ortiz-López
University of Salamanca
Salamanca, Spain
aortiz@usal.es

Susana Olmos-Migueláñez
GRIAL Research Group, Educational
Research Institute
University of Salamanca
Salamanca, Spain
solmos@usal.es

José Carlos Sánchez-Prieto
GRIAL Research Group, Educational
Research Institute
University of Salamanca
Salamanca, Spain
josecarlos.sp@usal.es

Abstract—At this time of pandemic in which society is immersed, the abandonment of presentiality and the educational transition to the digital environment have been the main line of action followed in educational institutions in Higher Education. In this transition, the evaluation must be regarded as a guarantee for the quality, continuity, effectiveness and efficiency of educational systems and their eLearning teachings. This research proposes a mapping of the literature with the aim of discovering the conditions of research on evaluation in eLearning during the pandemic in Higher Education. For this purpose, it is based on objective criteria for selecting publications, and the phases followed throughout all the process. The results show the growing trend in research and the necessary continuity in the study, and the opening of new lines that place evaluation at the centre of this new educational reality.

Keywords— *e-Learning, Evaluation, Higher Education, Mapping Study, COVID-19*

I. INTRODUCTION

The coronavirus crisis has led to a change in the policies and initiatives of university institutions. Since the middle of March of last year, they have been forced to implement significant changes and adaptations aimed at non-presence and e-Learning teaching. Not only did this rapid change suppose the effort of all the agents involved (teachers, students and university personnel) in a quick transition of the contents, the methodology, the classroom sessions, the technical support and the evaluation for the end of the academic year [1]; but also the opening of a new future scenario in which eLearning is already positioned as the centre of university teaching through totally online or hybrid methodologies. These have been the protagonists of the 2020/2021 academic year in almost every university centre and will also be so in the future, which will bring the consolidation of this teaching modality.

This effort carried out by the institutions and their professionals mean that the support, resources and training are infinitely higher than at pre-pandemic levels nowadays. These resources should not be forgotten and should be the focus of the future progress of Higher Education in eLearning [2]. Thus, it can be said that we are facing a new paradigm in Higher Education that is supported by technology, safe attendance and those changes that the pandemic has caused at all levels, including education.

In this new context, evaluation has been one of the main concerns of teachers, students, and society, in order to assure the continuity of university education. The search for new evaluation methods, adaptation to online evaluation and the

concern to implement a valid and reliable evaluation is today the line of advance and consolidation of eLearning and of the professionals who develop it. They have followed the recommendations of the universities as far as evaluation is concerned, and in some way, they have had to find new different models and options for their subjects [3].

Consequently, it can be stated that eLearning is no longer a utopia in our educational system, otherwise it is the reality of millions of university students and teachers who try to look ahead with this new paradigm as the centre of their academic activity. In this realm, the evaluation should continue ensuring its quality in an unprecedented scenario, full of an infinity of possibilities and advances. In other words, the evaluation has to maintain its essence, its principles of usefulness, viability, legitimacy and precision in the post-COVID era [4].

Therefore, given the lack of studies that bring together what has been published on evaluation in eLearning in Higher Education during the pandemic, this research tries to fill the gap in the literature through a mapping focused on different aspects. This will take into account the number of publications that have addressed the subject, the place where they was carried out, the journals that lead this field, the authors who support it, the field of knowledge and the way in which this evaluation had been studied; trying to make an objective synthesis for the continuity of the studies and the opening of new lines of research.

Below the research objectives and questions are presented. The third section focuses on the methodology, phases, criteria and the complete review process. After this, there are the results obtained from the analysis of the publications that try to answer the questions. Finally, the fifth section includes the implications of the results, the limitations of the study and the new lines of research.

II. OBJECTIVE AND RESEARCH QUESTIONS

A. Objective

This research aims to know the current state of research on evaluation in eLearning in Higher Education during the COVID-19 crisis to check the trend of these studies, their origin and typology, areas and future lines.

B. Research questions

The questions posed for the research are consistent with the methodology followed. These questions, which the research tries to answer, are the following:

- M1: How many studies have been published on evaluation in e-Learning in Higher Education during the years that comprise the COVID pandemic?
- M2: Who are the active authors in eLearning assessment in Higher Education during the pandemic?
- M3: Which countries are leading the studies on evaluation in eLearning in Higher Education during the pandemic?
- M4: Which journals publish the progress and studies related to evaluation in eLearning in Higher Education during the pandemic?
- M5: In which areas of knowledge has an eLearning evaluation in Higher Education been studied during the pandemic?
- M6: What type of study has been developed to evaluate eLearning in Higher Education during the pandemic?

III. METHODOLOGY

The methodology consisted in carrying out a mapping of the literature to know the status of the investigations. This is defined as a methodological process that "offers a careful selection of works about the production on the subject and enables interested researchers and readers to draw precise routes of inquiry" [5, p. 1]. In conclusion, it tries to structure a knowledge area [6].

The outlined review process follows the structure proposed by Kitchenham and Charters [7] and is organized into three stages: review design (topic, search, criteria), review process (selection, study and acceptance/exclusion) and report of the review (publication of results). The research search string was: ("e-Learning" OR "eLearning" OR "Distance Education" OR "Educación a Distancia") AND ("COVID" OR "COVID19" OR "COVID-19" OR "Pandemic" OR "Pandemia") AND ("Higher Education" OR "University" OR "Educación Superior" OR "Universidad").

On June 18, 2021, the search was carried out in Scopus and Web of Science, and the resulting publications were evaluated with six total inclusion/exclusion criteria that determined their linkage. In an attempt to establish a preliminary comparison between all the publications related to eLearning in Higher Education during the pandemic and those only focused on evaluation, a two-phase investigation is proposed.

A. Phase 1: Selection of studies on e-Learning in Higher Education during the pandemic.

After entering the search string in the repositories, 385 publications were obtained, which, after eliminating duplicates, remained at 375 (97 from Scopus and 288 from Web of Science). The criteria in this phase, which aims to extract all the articles on evaluation in eLearning in Higher Education in the pandemic, are the following:

- 1) Any publication that is not a scientific article is not included nor published in World of Science and Scopus.
- 2) Articles that do not deal with e-Learning in Higher Education have been excluded.
- 3) Articles that do not address the situation, changes and research on e-Learning during the pandemic (COVID) have been excluded.

4) Articles that are written in languages other than Spanish or English have been excluded.

5) Articles that do not deal with e-Learning in a broader sense (b-Learning, m-Learning) or that make reference to open courses (MOOC) have been excluded.

After applying the criteria, 216 articles were accepted (acceptance rate: 60%). Therefore, it can be deduced from this that during the pandemic, there were a total of 216 publications that addressed eLearning during the COVID crisis, treating it from a myriad of different perspectives (evaluation, perception, roles, platforms, methodologies, content, access...).

B. Phase 2: Selection of studies on evaluation in e-Learning in Higher Education during the pandemic.

In the second phase of the research, an additional inclusion/exclusion criterion was applied focused on the current research topic. This criterion is:

6) Articles that do not present evaluation as the central theme of the study (methods, procedures, perceptions and development) have been excluded.

After applying this new criterion, a total of 18 articles were selected for addressing the evaluation in Higher Education during the pandemic (acceptance rate: 4.8%). The final database is available in: <https://bit.ly/3rqPCDR>. The process followed is detailed below:

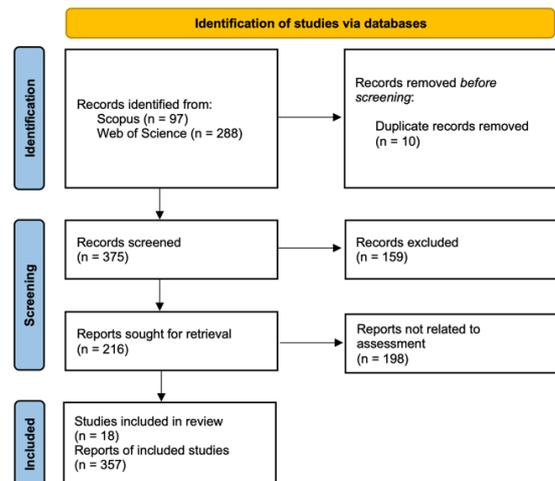


Fig. 1. PRISMA Flow Diagram of this research

Therefore, it can be stated that during the pandemic, a total of 216 articles were published that addressed eLearning in Higher Education, but only 18 of them did so from the perspective of the evaluation. In other words, only 8.33% of the publications that address eLearning in a pandemic refer directly to evaluation.

IV. RESULTS

The first question posed (M1) for the study focuses on knowing in which year of the pandemic a more significant number of publications have been developed (considering that the current year is ongoing and only half of it has elapsed). In the results, it is observed how 2021 is the one that already leads the publications with respect to the previous year (eleven compared to seven), so the growing trend of the studies can be deduced.

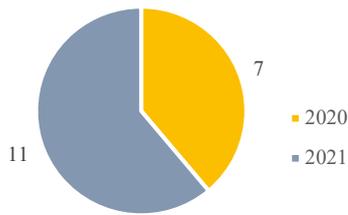


Fig. 2. Evolution of publications during the pandemic

Secondly, it is essential to know the authors who lead the publications on evaluation in this area (M2) since they focus on the advancement of research and the generation of new theories. A total of 55 authors are extracted. Of these, we can highlight García-Peñalvo, with four publications, Corell-Almuzara with three and Abella-García and Grande-de-Prado with two publications.

TABLE I. AUTHORS OF THE ANALYZED STUDIES

Nº	Author
4	García-Peñalvo, F.J.
3	Corell-Almuzara, A.
2	Abella-García, V.; Grande-De-Prado, M.
1	Ahmadi, S.; Al Abdulmonem, W.; Al-Azzam, N.; Alharthi, N.S.; Arbabisarjou, A.; Cárdenas Cabello, F.; Chertakova, E.M.; Chiyón, I.; Churi, P.; Dastyar, N.; De La Fuente, J.M.; Del Castillo-Olivares, A.; Del Castillo-Olivares, J.M.; Diez-Gutiérrez, E.J.; El Sadik, A.; Elmosaad, Y.M.; Elsalem, L.; Eltayeb, L.B.; Elzainy, A.; Espinoza, K.G.; Feijoo, J.C.M.; García-Alberti, M.; García-Holgado, A.; Gudiño Paredes, S.; Idnani, D.; Jacques, S.; Jain, Y.; Jasso Peña, F.D.J.; Jum'ah, A.A.; Junus, K.; Kubadia, A.; Lequeu, T.; Luna Nemecio, J.; Lyusev, V.N.; Martha, A.S.D.; Mokievskaya, N.E.; Nazarenko, A.V.; Obeidat, N.; Olivares, S.T.; Osipova, N.V.; Ouahabi, A.; Sánchez-Prieto, J.C.; Salari, N.; Santoso, H.B.; Shafei, S.; Suárez F.; Suhartanto, H.; Toledano, R.M.; Vázquez Ingelmo, A.; Vazquez, A.M.; Waggiallah, H.A.

Regarding the place where these studies were carried out (M3), which is also an indicator of those countries that concentrate the research, Spain stands out with seven of the eighteen publications; followed by Saudi Arabia with three (this, to a certain extent, may be due to the fact that Spanish is, together with English, part of the research search chain).

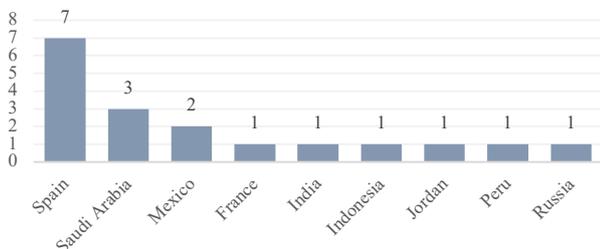


Fig. 3. Countries according to the number of publications

Fourthly (M4), a reliable source for consulting the new literature on the subject are scientific journals. In this case, the Spanish magazine Virtual Campus leads this field with three

related publications. The rest of the data is presented in the following table.

TABLE II. NUMBER OF ARTICLES PER JOURNAL

Nº	Journal
3	Campus Virtuales
2	Education Sciences, International Journal of Engineering Pedagogy
1	Anal. of Medicine and Surgery, Bordón Revista de Pedagogía, Distance Education, Education in the Knowledge Society, Eduweb-Revista de Tecnología de Información y Comunicación en Educación, International Journal of Pharmaceutical and Phytopharmacological Research, Journal of Taibah University Medical Sciences, Pakistan Journal of Medical & Health Sciences, Revista Española de Educación Comparada, Revista Iberoamericana de Educación a Distancia, Universidad y Sociedad

Fifthly (M5) the areas of knowledge in which this evaluation has been studied in eLearning during the pandemic are presented. Here the fields of engineering, social sciences and health sciences stand out as those with the highest number of publications.

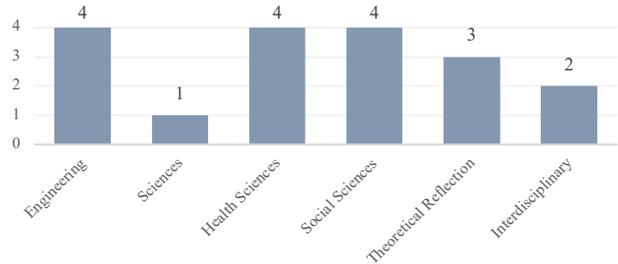


Fig. 4. Scope of knowledge of the analyzed studies

Finally, it is also paramount to pay attention to how each article presents the evaluation study (M6). In general terms, there are empirical articles on satisfaction with evaluation through eLearning, publications on tools and methodologies, and reviews on evaluation recommendations, as it is shown below.

TABLE III. TYOLOGY OF THE ANALYZED STUDIES

Article	Main topic
[8]	Analysis of six engineering subjects in Peru and Spain: Evaluation strategies and reflections.
[9]	Assessment exercise at the University of La Laguna (18 professors from the social branch and 24 from the scientific one)
[10]	Synthesis of recommendations and published instruments.
[11]	Mixed study to understand the impact of supervised online exams.
[12]	Theoretical study of the effect of the pandemic on universities and future challenges.
[13]	Development of the remote evaluation. Observation of students of Social Sciences.
[14]	Online evaluation case in the subject of Degree in Computer Engineering.
[15]	Procedures followed for online evaluation in Faculty of Medicine.
[16]	Instrument for the evaluation of online learning.
[17]	Evaluation of experiences with respect to electronic test preferences and academic dishonesty.

Article	Main topic
[18]	Guide analysis of recommendations for online evaluation.
[19]	Students' satisfaction questionnaire in relation to the evaluation through virtual platforms.
[20]	Online exams and assessment acceptance study.
[21]	Evaluation procedures in eLearning.
[22]	Students' perception of remote exams.
[23]	Survey of attitude towards assessment in remote conditions (attitude, control, tools and forms).
[24]	Guidelines and recommendations for online assessment.
[25]	Online assessment possibilities with students.

V. CONCLUSIONS

All things considered, we can conclude by reflecting on the importance of the increasing trend of publications observed after the mapping indicates the effort and continuity that evaluation research is taking.

At the other end of the scale, it is also deduced that Spain stands out in the study on evaluation in eLearning during the pandemic, proposing studies that address satisfaction, the proposals made for this purpose and presenting the results and the new lines to be followed after the pandemic. It is also where the authors with the most significant number of publications are found (García-Peñalvo and Corell-Almuzara).

The main limitation of the present study has been the temporary shortage given the recent events discussed, so continuity in research is required to study future evolution and trends in this new educational scenario that opens in Higher Education.

Notwithstanding, publications on evaluation in eLearning during the pandemic only represent a low percentage (8.33%) of the total publications made on eLearning in Higher Education. Therefore, we need a more significant effort and the opening of new lines of research that address the study of new possibilities, tools, platforms and resources for the reaffirmation of evaluation as an element of quality in this new paradigm, a paradigm in which eLearning is already positioned as the centre of university education in the new post-COVID age.

REFERENCES

- [1] L.A. Gazca-Herrera, "Implicaciones del coronavirus COVID-19 en los procesos de enseñanza en la Educación Superior", *RIDE: Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, vol. 11, no. 02, doi: 10.23913/ride.v11i21.753
- [2] C.T. Zambrana, "Las consecuencias del Coronavirus en la Educación Superior", en *Cuarentena Multidisciplinar*, vol. I, U.R. Qusipe et al., Fondo Editorial de la Universidad Nacional Autónoma de Huanta, 2020, pp. 212-228.
- [3] D.O. Ortigoza, J.R. Rodríguez and A.M. Inchaurredo, "Educación superior y la COVID-19: Adaptación metodológica y evaluación online en dos universidades de Barcelona", *Revista Digital de Investigación en Docencia Universitaria*, vol. 15, no. 1, Jun. 2021
- [4] J.L. Abreu, "Tiempos de Coronavirus: La Educación en Línea como respuesta a la crisis", *Daena: International Journal of Good Conscience*, vol. 15, no. 1, pp. 1-15, May. 2020.
- [5] C. Navarro and M.S. Ramírez-Montoya, "Mapeo sistemático de la literatura sobre evaluación docente (2013-2017)", *Educação e Pesquisa*, vol. 44, pp. 1-22, 2018.
- [6] G. Tebes, D. Peppino, P. Becker and L. Olsina, "Proceso para revisión sistemática de la literatura y mapeo sistemático", *Electronic Journal of SADIO (EJS)*, vol. 19, no. 02, pp. 94-118, 2020.
- [7] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering", *Keele University and Durham University Joint Report*, p. 44.
- [8] M. García-Alberti, F. Suárez, I. Chiyón and J.C. Mosquera, "Challenges and Experiences of Online Evaluation in Courses of Civil Engineering during the Lockdown Learning Due to the COVID-19 Pandemic", *Education Sciences*, vol. 11, no. 1, pp. 59-78., Feb. 2021, doi: 10.3390/educsci11020059
- [9] J.M. Castillo-Olivares and A. Castillo-Olivares, "El impacto de la COVID-19 en el profesorado de educación superior y sus concepciones sobre la evaluación", *Campus Virtuales*, vol. 10, no. 1, pp. 89-10, Jan. 2021.
- [10] M. Grande de Prado, F.J. García-Peñalvo, A. Corell-Almuzara and V. Abella-García, "Evaluación en Educación Superior durante la pandemia de la COVID-19", *Campus Virtuales*, vol. 10, no. 1, pp. 49-58, Jan. 2021.
- [11] S. Gudiño-Paredes, F.J. Jasso-Peña and J.J. de la Fuente, "Remote proctored exams: Integrity assurance in online education?", *Distance Education*, vol. 42, no. 2, pp. 200-2018, Apr. 2021, doi: 10.1080/01587919.2021.1910495
- [12] F.J. García-Peñalvo and A. Corell-Almuzara, "La COVID-19: ¿enzima de la transformación digital de la docencia o reflejo de una crisis metodológica y competencial en la educación superior?", *Campus Virtuales*, vol. 9, no. 2, Oct. 2020.
- [13] F. Cárdenas-Cabello and J.M. Luna-Nemecio, "Evaluación en línea ante la pandemia por COVID-19: retos y oportunidades para las universidades mexicanas", *Universidad y Sociedad*, vol. 12, no. 6, Dec. 2020.
- [14] F.J. García-Peñalvo, A. García-Holgado, A. Vázquez-Ingelmo and J.C. Sánchez-Prieto, "Planning, Communication and Active Methodologies: Online Assessment of the Software Engineering Subject during the COVID-19 Crisis", *RIED. Revista Iberoamericana de Educación a Distancia*, vol. 24, no. 2, Jun. 2021, doi: 10.5944/ried.24.2.27689.
- [15] N. Salari, S. Ahmadidarrehisima, A. Arbabisarjou, N. Dastyar and S. Shafei, "Evaluation of learning of students of jiroft university of medical sciences through e-learning during the COVID-19 pandemic", *Pakistan Journal of Medical and Health Sciences*, vol. 15, no. 4, 2021
- [16] A.S.D. Martha, K. Junus, H.B. Santoso and H. Suhartanto, "Assessing Undergraduate Students' e-Learning Competencies: A Case Study of Higher Education Context in Indonesia", *Education Sciences*, vol. 11, no. 4, pp. 189-207, Apr. 2021.
- [17] L. Elsalem, N. Al-Azzam, A. Jum'ah and N. Obeidat, "Remote E-exams during Covid-19 pandemic: A cross-sectional study of students' preferences and academic dishonesty in faculties of medical sciences", *Annals of Medicine and Surgery*, vol. 62, pp.326-33, Feb.2021
- [18] E.J. Díez-Gutiérrez and K.G. Espinoza, "Higher Education and online assessment in the time of Coronavirus. What do students think?", *Bordón: Revista de Pedagogía*, vol. 73, no. 1, pp. 39-57, 2021.
- [19] S. Tirado-Olivares, A.M. Vázquez and R.M. Toledano, "La Docencia Virtual o e-Learning como Solución a la Enseñanza de la Física y Química de los Futuros Maestros en tiempos de COVID-19", *Revista Española de Educación Comparada*, no. 38, pp 190-210, 2021
- [20] D. Idnani, A. Kubadia, Y. Jain and P.P. Churi, "Experience of Conducting Online Test During COVID-19 Lockdown: A Case Study of NMIMS University", *International Journal of Engineering Pedagogy*, vol. 11, no. 1, 2021.
- [21] A. Elzainy, A. El-Sadik and W. Al Abdulmonem, "Experience of e-learning and online assessment during the COVID-19 pandemic at the College of Medicine, Qassim University", *Journal of Taibah University Medical Sciences*, vol. 15, pp. 456-462, Dec. 2020
- [22] L.B. Eltayeb, N.S. Alharthi, Y.M. Elmosaad and H.A. Waggiallah, "Students' perception on E. Learning and Remote Exams during COVID 19 Outbreak 2020", *International Journal of Pharmaceutical and Phytopharmacological Research*, vol. 15, no. 5, pp. 142-148, Oct. 2020.
- [23] V.N. Lyusev, A.V. Nazarenko, N.V. Osipova, N.E. Mokievskaya and E.M. Chertakova, "Monitoring and evaluation of the educational results of University students in a pandemic", *Revista de Tecnología de Información y Comunicación en Educación*, vol. 14, no. 2, pp. 301-312, Dec. 2020.
- [24] F.J. García-Peñalvo, A. Corell-Almuzara, V. Abella-García and M. Grande, "Online Assessment in Higher Education in the Time of COVID-19", *Education in the Knowledge Society*, vol. 21, May. 2020
- [25] S. Jacques, A. Ouahabi and T. Lequeu, "Remote Knowledge Acquisition and Assessment During the COVID-19 Pandemic", *International Journal of Engineering Pedagogy*, vol. 10, no. 6, 2020.

Anexo 8B

Comunicación a Congreso Internacional Technological Ecosystems for Enhancing Multiculturality (2022)

VERSIÓN POST-PRINT

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2023). TAM, Assessment, and Mobile Technologies: A Study of Teachers' Acceptance in Higher Education. En F. J. García-Peñalvo & A. García-Holgado (Eds.), *Proceedings TEEM 2022: Tenth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 1263-1271). Springer Nature. https://doi.org/10.1007/978-981-99-0942-1_133

Enlace a la publicación completa :

https://link.springer.com/chapter/10.1007/978-981-99-0942-1_133

TAM, assessment, and mobile technologies: A study of teachers' acceptance in Higher Education

Alberto Ortiz-López*

Instituto Universitario de Ciencias de la Educación, GRIAL, Universidad de Salamanca, aortiz@usal.es, 0000-0003-3164-5680

Susana Olmos-Migueláñez

Instituto Universitario de Ciencias de la Educación, GRIAL, Universidad de Salamanca, solmos@usal.es, 0000-0002-0816-4179

José Carlos Sánchez-Prieto

Instituto Universitario de Ciencias de la Educación, GRIAL, Universidad de Salamanca, josecarlos.sp@usal.es, 0000-0002-8917-9814

Assessment processes and the digital transition are one of the main challenges for educational institutions in roadmaps and future development plans, in an education that tends to digitize teaching at a dizzying pace. In this rapid transition, it is essential to pay attention to the importance of knowing teachers' acceptance of the changes that are implemented, as it is they who will ultimately decide the methodologies followed in a classroom. Therefore, it can be deduced that, without correct acceptance, the challenges will not be met. This doctoral thesis is carrying out the design of a model that, based on the most relevant acceptance models, can predict, and study the factors that condition the use of mobile devices in assessment by teachers. This knowledge will make it possible to address the specific factors that will influence future teacher training policies, both in basic training stages and in continuing education, with the aim of facilitating the digital transition, removing the barriers that teachers observe and moving towards a new era in education, mediated by technology and with assessment at the center of the process.

Keywords: Assessment, Technology, Acceptance, teachers, higher education

1. Context and motivation that drives the dissertation research

Online learning is currently positioned as the objective and the next challenge in today's educational systems [1]. Evaluation in these contexts is fundamental to guarantee the quality of the educational processes, which is why it must be a central theme in the development of e-Learning. A relevant factor in these processes, responsible for the development of these teachings, is the teacher [2]. The aim of this research is the development and validation of an acceptance model that, based on the TAM model [3], can predict and evaluate the factors that condition the intention of these teachers to use e-evaluation tools, thus advancing in the continuous improvement of evaluation tools and the way in which teachers conceive them, all with the aim of guaranteeing a reliable, valid and accepted evaluation within the framework of online education in our educational systems.

The present thesis attempts to continue a field in which universities are making recent efforts to develop, especially since the pandemic, digital education, or e-Learning. These efforts are what motivate the development of an investigation on this field, to continue and follow the effort made in the digital transition; respecting at all times the principles of evaluation and always bearing in mind the importance of teachers in the face of these changes, since their acceptance or not is fundamental for an effective and efficient digital adaptation.

2. State of art

These acceptance models [3], has been widely validated in teaching practice, and this Doctoral Thesis seeks to develop its own model which, based on it, focuses on teachers' acceptance of assessment through mobile devices (Mobile-Based Assessment), with the aim also of finding out the factors that condition the intention to use these tools in the educational community.

In this acceptance, the most studied model, according to García-Peñalvo et al. [4] is the Technology Acceptance Model (TAM), a model derived from the Theory of Reasoned Action [5] which aims to explain the factors that lead an individual to accept or reject a given information system through two factors: perceived usefulness (defined as the "degree to which a person believes that the use of a particular system will improve his or her performance" [3, p. 320], and perceived ease of use (understood as the "degree to which a person believes that using a particular system will be effortless" [3, p. 320]).

Subsequently, other models adapted to technological acceptance have been generated from theories such as IDT (Innovation Diffusion Theory), TTF (Task-Technology Fit), UTAUT (Unified Theory of Acceptance and Use of Technology), models based on TRA (Theory of Reasoned Action) or the TPB model (Theory of Planned Behaviour) [6].

In addition, relevant models already addressed in the eLearning field have been taken as the SDT (Self-Determination Theory), the Decomposed TPB, the combination between TAM and the TPB and, finally, the only model focused on the mobile based assessment, the MBAAM [7].

Based on this premise, we propose the development of our own model of acceptance that allows us to analyse acceptance and how this conditions the use of different assessment tools, thus making a new contribution to the field of e-assessment that allows the assessment of the adoption of the tools that will be at the centre of the teaching assessment processes and also allows us to establish a link between software developers and online assessment tools and educational practice, all with the aim of achieving an improvement in the quality of assessment. This line of attention to e-assessment as the central object of development of this research is based on the current situation that social systems, and specifically educational systems, have been going through in the last decade and which has been accentuated after the COVID-19 pandemic, which has placed e-Learning and online teaching at the centre of teaching and teaching-learning processes [8].

Current research is scarce about the central theme of the thesis, the acceptance of e-assessment through mobile devices, with only a total of 15 publications focusing on these models (detailed in section six of this document). Therefore, it can be deduced that this is an incipient field of knowledge. Parallel studies on the subject have been extracted from more developed studies in other fields (generic acceptance studies) or focused on other digital media, but there are still very few studies on mobile devices. Therefore, the aim of this research to fill a gap in the literature and promote the digital transition of teachers and, therefore, of the University.

3. Hypothesis and problem statement

This research is based on a clear approach, which seeks to know the factors that condition the acceptance of the use of e-assessment tools among university teachers, to improve the teaching-learning processes and have more effective evaluation processes.

The study contemplates the contextual characteristics related to the teachers that will serve to differentiate the different states of acceptance: their type of teaching, their teaching, their field of knowledge, their teaching experience... In other words, the aim is to know which dimensions influence which cases of teaching.

Furthermore, what is fundamental for this research is also to check whether the dimensions of acceptance that work best in other models analyzed in the reviews also have an influence in the Spanish university context (the first study of these characteristics carried out in our country), and to analyze the possible influence of other dimensions on acceptance.

4. Research objectives/goals

The main objectives of the thesis are as follows:

- To develop a technology acceptance model based on TAM to assess the factors that condition the intention to use e-evaluation tools in the educational community.
- To understand the factors that influence the process of technological adoption in the educational community.

On the other hand, a series of specific objectives are also presented:

- To study the level of technological acceptance of e-learning assessment tools in the educational community.
- To find out the state of research in the field of e-evaluation through mobile devices.
- To determine the conditioning factors in the acceptance of e-assessment technologies by teachers.

These objectives are set out with the intention of expanding an incipient field of knowledge, the acceptance of assessment processes in mobile environments.

5. Research approach and methods, including relevant rationale

The work methodology proposed for the development of the thesis is based on a quantitative approach to its development, one of the main lines followed in model acceptance and validation studies [9].

Firstly, with the aim of establishing and delimiting the current research carried out in the field of e-assessment, an initial systematic review of the literature will be carried out to present the most widely used lines of research, as well as the platforms and trends followed in these studies. After this, a second, more specific review will be carried out, focusing on the acceptance of e-assessment on mobile devices, which will position the field in which the current thesis is being developed. These reviews will follow the structure proposed by Kitchenham and Charters [10] and the recommendations proposed by García-Peñalvo [11].

Subsequently, the current models of acceptance in the field and the trends obtained from the review will be studied, to proceed to the proposal of a new model and relational hypotheses adjusted to the objectives that will explore and expand the field of research. This model will be subjected to validation processes by expert judges and intermediate studies, which will determine the accuracy and the changes that will have to be made to the final model.

Ultimately, the final acceptance model will be proposed and evaluated in teaching practice, studying the acceptance based on the constructs marked. This, to study the

status and factors that influence its acceptance, which may lead to policies and programmes aimed at strengthening these factors in order to improve the adoption and digital transition of education.

It should also be noted that this investigation will respect, always, the principles of research ethics (British Educational Research Association [12]) and the different educational and data protection regulations in force.

For the development of the Thesis, different objectives are also established with respect to the publication and dissemination of results. These objectives include publication in international conferences (TEEM, ECER, WERA; among others) and the publication of results in journals indexed in Web of Science and Scopus. These publications will revolve around the systematic reviews carried out, the proposal and validation of the model and the publications of the results, and are organised as follows:

- 2022: Submission and presentation of the research plan proposed here at the TEEM 2022 Congress (Technological Ecosystems for Enhancing Multiculturality) in the track relating to the Doctoral Programme.
- 2023: Publication of the systematic literature reviews in journals indexed in both Web of Science and Scopus (journals focused on e-Learning and mobile devices). In addition, participation in international conferences (ECER, WERA), where the designs of the initial model can be published, is planned for the current academic year.
- 2024: Publication and dissemination of the model, intermediate results, and final validation results both in indexed journals and in conferences related to technology, evaluation and teaching.

6. Results to date

So far, the main results achieved are focused on the development of a generic e-evaluation mapping, an SLR on e-evaluation acceptance and, after revisions, also on the construction of a model to assess acceptance in university teachers (current phase). The results achieved in each of the phases are summarized below.

- Phase 1: Initial mapping: The initial mapping was carried out in December 2021 and started with a broad search string and a total of 1136 results to answer 10 mapping questions. After seven exclusion criteria were applied to limit the publications, a total of 159 publications were accepted (acceptance rate of 13.99%). The results show that the USA is the country with the highest number of studies, that S.A. Nikou and A.A. Economides are the authors with the most publications, that the International Journal of Emerging Technologies in Learning is the journal with the most publications, that Higher Education students are the most explored field, that quantitative studies are the predominant ones, that questionnaires are the most used technique and that studies of results in LMS systems are

the most developed. Furthermore, the need to focus the study on teachers at other educational levels, on other technologies that go beyond the LMS, and on the extension of the studies to acceptance, effectiveness and quality are highlighted as future lines following this mapping.

- Phase 2: SLR acceptance of e-assessment on mobile devices: The second phase comprised a more comprehensive review, guided by the first mapping of the literature. To this end, 11 search strings were designed that related evaluation through mobile devices (central field of the Thesis) with 11 of the most studied acceptance models; and a search was carried out in 9 representative databases of recognized prestige. The results were 572 and, after applying 5 exclusion criteria and 10 quality criteria, a total of 15 publications were selected as relevant to answer the 10 research questions posed. The results focus on an upward trend of studies, Nikou and Economides as the most relevant authors, Greece as the predominant country, acceptance studies in Higher Education as the most addressed, TAM as the most used model and PLS as the chosen analysis technique. In addition, a total of 29 different constructs measured in the 15 investigations were analyzed and a mean explained variance of 59.18%. 59,18%.
- Phase 3: Model building (Figure 1): This is the current phase of the Thesis. At this moment, we are building our own model to evaluate teacher acceptance, which is based on TAM and is based on the 15 publications extracted in the previous review. In addition, our own models have been proposed to address this acceptance. This model, after its construction and hypothesis design, will undergo a process of validation by expert judges and intermediate studies prior to the construction of the final model.

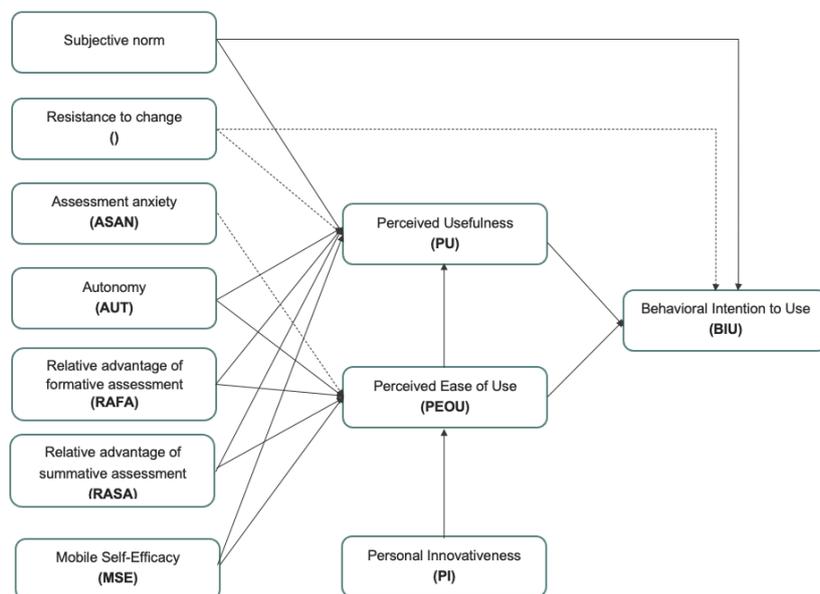


Figure 1. Model in construction

7. Dissertation status

The methodology proposed for the Doctoral Thesis establishes different phases for the in-depth study of the field of e-assessment and for the proposal and validation of the model. These phases are:

- F1. Systematic reviews of the literature: This phase comprises a first systematic review of the literature to ascertain the state of research on e-assessment tools, and to determine which can form part of the new model. On the other hand, a second review on the state of research on TAM and e-assessment is also proposed, trying to find out what has been researched to establish the starting point for the Doctoral Thesis.
- F2. Proposal and validation of the model: This second phase comprises the complete process of design and validation of a new model based on TAM for e-assessment, with which to address the factors that condition the intention to use e-assessment tools in the educational community.
- F3. Intermediate studies on the validated model: In this phase, once the model has been validated, inter- and intra-dimensional studies will be carried out to address the functioning of each item and dimension in the total explanation of the model.
- F4. Final phase of the model: Final development of the TAM-based technology acceptance model to assess the factors conditioning the intention to use e-assessment tools.
- F5. Study of the acceptance and application of the new model in educational practice.

The following image shows the expected design for the development of the Thesis, organized in three academic years, distributed quarterly, for the five phases of its development (Figure 2).

Phases	Year 1				Year 2				Year 3			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
F1) Systematic review 1: e-evaluation tools	■	■										
F1) Systematic review 2: TAM and e-assessment			■	■								
F2) Model design and proposal					■							
F2) Model validation						■	■					
F3) Intermediate studies								■	■	■		
F4) Final phase of the model									■	■	■	
F5) Acceptance study and application of the final model											■	■

Figure 2. Timetable established for the Thesis

8. Current and expected contributions

For the moment, the main contributions achieved focus on what is referred to in the sixth section, on having a marked and complete delimitation of everything written about the acceptance of e-assessment processes through mobile devices, having an incipient field with only 15 publications in a university system that tends towards the digital and towards assessment in these contexts.

What this thesis seeks to develop is a model specifically adapted to measure the factors that condition teachers' acceptance of mobile technologies as e-assessment tools, a subject that has not yet been explored in the literature [13]. Knowing these conditioning factors opens a new avenue for the development and advancement of these tools and techniques, as it will allow them to be adapted to the needs of teachers and to those dimensions that they consider most relevant to teaching, their application, and their usefulness in assessment practice.

On the other hand, it is hoped that the fulfilment of the objectives set for the development of the Thesis will provide a list of conditioning factors on the use and acceptance of e-assessment and the specific tools, thus being able to establish a determining model in the evaluation of the factors of teaching acceptance.

This Thesis is a continuation of the line of evaluation marked within the Interaction and eLearning Research Group (GRIAL) [14,15] and the University Institute of Education Sciences (IUCE); it is also annexed to related lines of research also worked on in the GRIAL group on technological acceptance and new e-Learning contexts in the field of education. The Thesis presented here is part of the line of Educational Evaluation and Guidance of the Doctoral Programme "Training in the Knowledge Society".

For its development, the Doctoral Thesis "Design of a technology adoption model to assess the acceptance of mobile technologies among primary school teachers." [16] will be taken as references, focusing on this case the study on the adoption of e-evaluation tools in mobile technologies; and the Thesis "Mobile-based Assessment: An exploration of Motivation and Acceptance factors" [17].

References

1. Castañeda LJ (2019) Debates on Technology and Education: Contemporary Pathways and Pending Conversations. RIED. Revista iberoamericana de educación a distancia, 22(1): 29-39. doi.org/10.5944/ried.22.1.23020
2. Cascante N, Villanueva L (2020) Teacher training in university didactics in the pandemic: Between pedagogical reflection and instrumentalisation. InterCambios. Dilemas y transiciones de la Educación, 7(2).
3. Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3): 319-340.

4. García-Peñalvo F, Olmos-Migueláñez S, & Sánchez-Prieto J (2015) Las tecnologías móviles en educación primaria: Estudio sobre la actitud de los futuros docentes. *Série-Estudos - Periódico do Programa de Pós-Graduação em Educação da UCDB*, 40: 35-55.
5. Fishbein M, & Ajzen I (1975) *Belief, Attitude, Intention And Behavior: An introduction to theory and research*. Addison-Wesley.
6. Nikou S, Economides A (2014) Acceptance of Mobile-Based Assessment from the Perspective of Self-Determination Theory of Motivation. *IEEE, International Conference of Advanced Learning Technologies: ICALT 2014*. 10.1109/ICALT.2014.136
7. Nikou S, Economides, A (2017) Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 23.
8. Grande-de-Prado M, García-Peñalvo F J, Almuzara AC, & Abella-García V (2021) Assessment in Higher Education during the COVID-19 Pandemic. *Campus Virtuales*, 10(1): 49-58.
9. Monje CA (2011) *Quantitative and Qualitative Research Methodology: A Teaching Guide*.
10. Kitchenham B, & Charters S (2007) *Guidelines for performing Systematic Literature Reviews in Software Engineering (Keele University and Durhan University Joint Report) [EBSE Technical Report]*. Keele University and Durhan University.
11. García-Peñalvo FJ (2014). Training in the knowledge society, a doctoral programme with an interdisciplinary perspective. *Education in the Knowledge Society*, 15(1): 4-9. doi.org/10.14201/eks.11641
12. British Educational Research Association [BERA] (2019). *Ethical Guidance for Educational Research (4.a ed.)* (L. Rivera Otero and R. Casado-Muñoz, Trads.), Londres. <https://www.bera.ac.uk/publication/guia-etica-para-la-investigacion-educativa>
13. Nikou S, Economides A (2019) A comparative study between a computer-based and a mobile-based assessment: Usability and user experience. *Interactive Technology and Smart Education*. doi.org/10.1108/ITSE-01-2019-0003
14. García-Peñalvo, FJ, Rodríguez-Conde, MJ, Therón R, García-Holgado A, Martínez-Abad F, & Benito-Santos A (2019a) GRIAL Group. *IE Communications. Iberoamerican Journal of Educational Informatics* (30): 33-48.
15. Grupo GRIAL (2019). *Scientific output of the GRIAL Group from 2011 to 2019 (GRIAL-TR-2019-010)*.
16. Sánchez-Prieto JC (2018) *Design of a technology adoption model to evaluate the acceptance of mobile technologies in primary school teachers [Doctoral Thesis]*. University of Salamanca.
17. Nikou S (2017) *Mobile-based assessment: An exploration of acceptance and motivation factors [Tesis Doctoral, University of Macedonia]*. <https://dspace.lib.uom.gr/handle/2159/21275>

Anexo 8C

Comunicación a Congreso Internacional Congreso Internacional: Educación y Conocimiento (2023)

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2023). Evaluando a través de dispositivos móviles: Validación de un nuevo modelo para determinar la aceptación tecnológica en docentes. En R. Roig-Vila, F Manuel-Sirignano & J.M. Antolí-Martinez (Eds.), *Actas del II Congreso Internacional: Educación y Conocimiento* (p. 403). Octaedro Editorial. ISBN: 9788419690975

Evaluando a través de dispositivos móviles: Validación de un nuevo modelo para determinar la aceptación tecnológica en docentes

Alberto Ortiz-López, Susana Olmos-Migueláñez, José Carlos Sánchez-Prieto

Universidad de Salamanca

Los dispositivos móviles tratan de romper hoy la barrera que separa los centros educativos de las nuevas metodologías de enseñanza, donde estos dispositivos se vuelven fundamentales en la mediación profesor-alumno. En este cambio que la educación está experimentando, la evaluación de las enseñanzas ha de ser un factor de calidad, de adaptación y cambio abierto a la inclusión de las tecnologías móviles en sus procesos, ya que estas actúan hoy como soporte principal y como vía de desarrollo metodológico en la docencia del siglo XXI. Pese a las ventajas ya contrastadas de la tecnología en el ámbito educativo, es todavía difícil encontrar aulas que empleen el dispositivo móvil de forma continuada y con finalidades evaluativas, estando ante una resistencia de múltiples causas que provoca que las experiencias que se desarrollan sean implementaciones esporádicas. Entre estas fuentes destaca la resistencia de los docentes a su introducción, y son ellos además en quienes recae la decisión última de su implementación y desarrollo. Por tanto, se puede afirmar que todo aquel docente no conforme con la idea de emplear dispositivos móviles en su evaluación empleará una metodología alternativa o tradicional con la que sí muestre conformidad. Con el objetivo de desarrollar y validar un nuevo modelo para el estudio de la aceptación tecnológica en docentes sobre la inclusión de dispositivos móviles en evaluación, tomando como referencia el análisis de los principales modelos disponibles y sus constructos, y partiendo de la reflexión teórica sobre la evaluación mediada por la tecnología, se ha diseñado un nuevo modelo específico basado en el Modelo de Aceptación Tecnológica (TAM) que contempla, para determinar la intención de uso, la utilidad percibida, la facilidad percibida, la norma subjetiva, la resistencia al cambio, la ansiedad ante la evaluación, la autonomía, las ventajas relativas de la evaluación formativa y sumativa y la autoeficacia en el uso de dispositivos móviles. Seguidamente, este nuevo modelo ha sido sometido a un proceso de validación de contenido realizado por un grupo de seis expertos nacionales e internacionales en materia tecnológica y evaluativa. Los expertos han sometido al modelo a tres fases de revisión: la evaluación individual de los ítems propuestos (evaluando coherencia, relevancia y claridad), la valoración global del modelo (adecuación al tema, validez y relevancia de los constructos) y una evaluación cualitativa final. Del análisis de resultados extraídos tras el proceso de validación se destaca el alto acuerdo de los evaluadores en las dimensiones medidas, contando con puntuaciones elevadas en prácticamente la totalidad del modelo (ítems y dimensiones). Las principales modificaciones se han centrado en aspectos de claridad (redacción) y la coherencia de varios ítems con su dimensión. Tras las modificaciones, el modelo resultante validado cuenta con un total de once constructos y 18 hipótesis que justifican las relaciones entre ellos, y evalúa la aceptación docente en un total de 50 ítems para describir la intención final de uso de dispositivos móviles en procesos de evaluación. En conclusión, la propuesta que aquí se realiza toma cinco constructos nuevos que no han sido aún probados en este campo y trata de determinar, de forma específica, los factores condicionantes en la inclusión tecnológica en docentes, permitiendo así su conocimiento y la posibilidad de un desarrollo institucional y formativo para la inclusión efectiva de la tecnología en los procesos de evaluación en educación. Investigaciones futuras pueden partir del modelo, ya validado, para ampliar los estudios de validez y fiabilidad y comprobar el nivel de predicción de las nuevas dimensiones propuestas, así como el contraste de las hipótesis establecidas.

Anexo 8D

Comunicación a Congreso Internacional Congreso Internacional Aprendizaje, Innovación y Cooperación (2023)

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2023). Utilidad percibida por docentes de un dispositivo móvil para evaluación formativa y sumativa: Un estudio comparativo. En M.L. Sein-Echaluce Lacleta, A. Fidalgo-Blanco & F.J. García-Peñalvo (Eds.), *Innovación educativa en los tiempos de la inteligencia artificial. Actas del VII Congreso Internacional sobre Aprendizaje, Innovación y Cooperación. CINAIC 2023* (pp. 396-299). Servicio de Publicaciones: Universidad de Zaragoza. <https://doi.org/10.26754/uz.978-84-18321-92-4>

Utilidad percibida por docentes de un dispositivo móvil para evaluación formativa y sumativa: Un estudio comparativo

Teachers' perceived usefulness of mobile devices for formative and summative assessment: A comparative study.

Alberto Ortiz-López¹, Susana Olmos-Migueláñez², José Carlos Sánchez-Prieto³
aortiz@usal.es, solmos@usal.es, josecarlos.sp@usal.es

¹ Instituto Universitario de Ciencias de la Educación (IUCE)
Universidad de Salamanca
Salamanca, España

² Instituto Universitario de Ciencias de la Educación (IUCE)
Universidad de Salamanca
Salamanca, España

³ Instituto Universitario de Ciencias de la Educación (IUCE)
Universidad de Salamanca
Salamanca, España

Resumen- La evaluación en los entornos virtuales es un reto para los futuros docentes, quienes se enfrentan a una nueva realidad educativa mediada por la tecnología. Por tanto, conocer los factores que condicionan la aceptación o rechazo de la tecnología para evaluación, y la existencia de diferencias entre la utilidad que los docentes le conciben para evaluación sumativa y para evaluación formativa, todo ello con el objetivo de contribuir al avance de la educación. Esta comunicación presenta los resultados alcanzados en un estudio realizado en 262 estudiantes del máster universitario en Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas de la Universidad de Salamanca mediante la aplicación de un cuestionario de 18 ítems. Los resultados muestran un alto nivel de aceptación tecnológica y manifiestan diferencias entre la utilidad de los dispositivos en función de la modalidad de evaluación a realizar.

Palabras clave: Aceptación, Tecnología, Evaluación, Docentes

Abstract- Assessment is a challenge for future teachers, who face a new educational reality mediated by technology. Therefore, this paper aims to know the factors that condition the acceptance or rejection of technology for assessment, and the existence of differences between the usefulness that teachers perceive for summative and formative assessment, all with the aim of contributing to the development of the educational field. This paper presents the results achieved in a study carried out on 262 students of the master's degree in Teacher of Compulsory Secondary Education and Baccalaureate, Vocational Training and Language Teaching of the University of Salamanca, through the application of an 18-item questionnaire. The results show a high level of technological acceptance and show differences between the usefulness of the devices depending on the assessment modality performed.

Keywords: Acceptance, Technology, Assessment, Teachers

1. INTRODUCCIÓN

La educación se encuentra hoy inmersa en un profundo cambio impulsado, en parte, por la revolución tecnológica que atraviesa la sociedad; replanteando la forma en la que el estudiante aprende y el docente enseña (Kim & Park, 2020). La tecnología posibilita hoy un cambio metodológico, didáctico y evaluativo nunca antes visto en educación mediante la

implementación de nuevos dispositivos, nuevas herramientas y nuevos recursos para el acceso a la información, la secuenciación didáctica y la evaluación final (Area-Moreira & Adell-Segura, 2021).

Es en este último aspecto, la evaluación, en el que la comunicación actual centra su foco. La evaluación es hoy una de las partes más importantes del proceso de enseñanza-aprendizaje en la certificación del rendimiento, de la aptitud del estudiante y de la calidad de la docencia (Mellado-Moreno et al., 2021); por lo que atender a su implementación y adaptación en los nuevos contextos de enseñanza debe ser acción prioritaria para el docente (Ibarra-Sáiz et al., 2020).

Por tanto, es importante abordar desde la comunidad investigadora la transformación que los docentes experimentan hoy para conocer, de forma objetiva, la aceptación que experimentan ante cambios de tal magnitud. Conocer los factores que conducen a un docente a la introducción de cambios (tecnológicos y evaluativos, en este caso) es necesario para el desarrollo de los futuros planes de formación y las políticas formativas en los niveles básicos (Grados), avanzados (Másteres) y en la formación continua del profesorado.

En la actualidad, la determinación de la aceptación se desarrolla principalmente mediante los modelos de adopción tecnológica, siendo el modelo TAM (Davis, 1989) el más empleado. Este modelo, basado en la Teoría de la Acción Razonada (Fishbein & Ajzen, 1975) y en la Teoría del Comportamiento Planeado (Ajzen, 1985), determina la intención de uso tecnológico a través del estudio de la actitud frente al uso (A), la utilidad percibida (PU) y la facilidad de uso percibido (PEOU) (Figura 1).

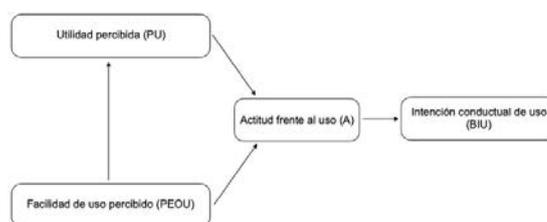


Figura 1. Modelo TAM (Davis, 1989)

La investigación que se presenta es parte de un estudio sobre la aceptación tecnológica de dispositivos móviles para procesos de evaluación en futuros docentes, estudiantes del máster universitario en Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas de la Universidad de Salamanca .

Dado que revisiones previas no han constatado la existencia estudios en el campo que determinen las diferencias entre la evaluación formativa y sumativa, ni que profundicen en sus utilidades diferenciadas, a continuación, se propone el estudio de la utilidad de forma específica para el campo de la evaluación mediada por dispositivos móviles.

La sección segunda presenta el contexto y la descripción de la investigación, la tercera sección los resultados alcanzados y, finalmente, la cuarta sección presenta las conclusiones del estudio.

2. CONTEXTO Y DESCRIPCIÓN

A continuación, se presenta un modelo para determinar la aceptación tecnológica en futuros docentes, que parte de la propuesta de Davis (1989) y desarrolla el constructo utilidad percibida para las evaluaciones formativas y las evaluaciones sumativas.

La hipótesis que se plantean para la investigación son la existencia de relaciones entre la intención conductual de uso y las tres utilidades propuestas (1) y la existencia de diferencias entre la utilidad percibida para la evaluación formativa y la utilidad percibida para la evaluación sumativa (2).

A. Metodología: Modelo de investigación propuesto

El modelo empleado para el análisis es el Modelo de Aceptación Tecnológica ya descrito en la introducción (TAM), y reformula el constructo de la utilidad percibida para las dos modalidades específicas de la evaluación (sumativa y formativa) en base a la diferenciación realizada por Scriven (1991).

1) Variables del modelo TAM

El modelo TAM (Davis, 1989) propone los constructos intención conductual de uso, utilidad percibida y facilidad de uso percibido para la determinación de la aceptación. En la propuesta original, el modelo plantea también la actitud como elemento mediador sobre la intención final, pero, dado su bajo poder explicativo, ha sido eliminado de la investigación en consonancia con la tendencia seguida en la literatura (Davis & Venkatesh, 1996). Además, dado que el objetivo de esta investigación es la diferenciación entre las utilidades, se ha eliminado del modelo la facilidad de uso para abordar, de forma directa y única, la utilidad percibida.

Por ello, partiendo de estos constructos se plantea la siguiente hipótesis:

- La utilidad percibida predice positivamente la intención conductual de usar dispositivos móviles para procesos de evaluación en futuros docentes (H1).

2) Modelo TAM para el campo de la evaluación

Seguidamente, se ha reformulado el constructo de la utilidad percibida para la adaptación al campo de la evaluación. Dado que el objetivo de la investigación es el planteamiento de un estudio comparativo, se ha adaptado el modelo TAM a través de dos constructos para la evaluación formativa y sumativa.

En base a esta construcción, se plantean por tanto las siguientes hipótesis:

- La utilidad percibida de los dispositivos móviles para procesos de evaluación formativa predice positivamente la intención conductual de usar dispositivos móviles para procesos de evaluación en futuros docentes (H2).
- La utilidad percibida de los dispositivos móviles para procesos de evaluación sumativa predice positivamente la intención conductual de usar dispositivos móviles para procesos de evaluación en futuros docentes (H3).

Por tanto, el modelo final planteado es un modelo triple que plantea la relación de las utilidades con la intención conductual (Figura 2).

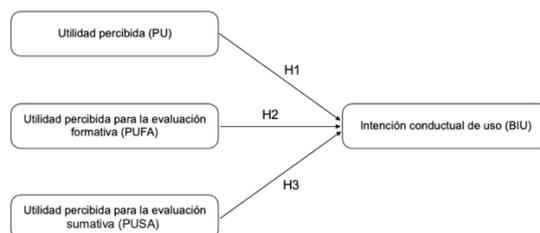


Figura 2. Modelo de investigación propuesto

3) Población y muestra

Poblacionalmente, el estudio se centra en los estudiantes del Máster en Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas; la formación habilitante para el ejercicio de la docencia posterior al Grado Universitario (N= 292).

La muestra resultante es de 262 participantes (89,7%).. En cuanto a su distribución, ésta se encuentra conformada por un 58,4% de mujeres, un 41,2% de hombres y un participante que marcó la opción “Otro” en la variable género.

4) Instrumento

El instrumento propuesto ha sido un cuestionario con escala tipo Likert 1-7, aplicado a través de la aplicación LimeSurvey e informado favorablemente para su aplicación por el Comité de Ética de la Universidad de Salamanca. El cuestionario se encuentra conformado por las variables de identificación (género y grado académico cursado previo al Máster) y los ítems del modelo, un total de 18.

Estos ítems se agrupan en cuatro dimensiones. En primer lugar, la intención conductual de uso (BIU) y la utilidad percibida (PU) han sido tomados del modelo TAM:

- **Intención conductual de uso (BIU):** *Pretendo utilizar dispositivos móviles para evaluar en mi futura docencia (BIU_01), En caso de que tuviera acceso a dispositivos móviles para evaluar, predigo que los utilizaría (BIU_02), Asumiendo que dispusiese de dispositivos móviles en el aula, mi intención sería utilizarlos para evaluar (BIU_03), Estoy dispuesto a utilizar dispositivos móviles para evaluar en mi futura función docente (BIU_04).*
- **Utilidad percibida (PU):** *Emplear dispositivos móviles en evaluación mejoraría mi productividad (PU_01), El uso de dispositivos móviles en el proceso de evaluación me hará más efectivo (PU_02), Los móviles en el proceso de evaluación facilitarán mi futuro trabajo docente*

(PU_03), *El uso de dispositivos móviles en el proceso de evaluación es útil en la docencia* (PU_04).

Seguidamente, se presenta la adaptación realizada de los ítems para evaluación formativa y sumativa, que parte también del modelo TAM y de la literatura sobre evaluación (Black, 2008; Dixson & Worrell, 2016).

- **Utilidad percibida para evaluación formativa (PUFA):** La evaluación basada en el uso dispositivo móvil es útil para evaluaciones formativas (orientada a la mejora) (PUFA_01), *Utilizar dispositivos móviles me permitiría realizar evaluaciones formativas de una forma más eficaz* (PUFA_02), *Emplear móviles en evaluación formativa mejorará la calidad de mi docencia* (PUFA_03), *El uso de dispositivos móviles en evaluaciones formativas hará mi trabajo más fácil* (PUFA_04), *Las ventajas de usar dispositivos móviles en evaluaciones formativas supera ampliamente las desventajas* (PUFA_05).
- **Utilidad percibida para evaluación sumativa (PUSA):** La evaluación basada en el uso dispositivo móvil es útil para evaluaciones sumativas (finales) (PUSA_01), *Utilizar dispositivos móviles me permitiría realizar evaluaciones sumativas de una forma más eficaz* (PUSA_02), *Emplear móviles en evaluación sumativa mejorará la calidad de mi docencia* (PUSA_03), *El uso de dispositivos móviles en evaluaciones sumativas hará mi trabajo más fácil* (PUSA_04), *Las ventajas de usar dispositivos móviles en evaluaciones sumativas supera ampliamente las desventajas* (PUSA_05).

Para la determinación de la consistencia interna del instrumento de medida se ha realizado a través del coeficiente α de Cronbach, que refleja una alta consistencia interna ($\alpha=0,944$).

3. RESULTADOS

En primer término, se presentan los resultados relativos a la aceptación tecnológica evaluada en futuros docentes a través del cuestionario descrito. Los descriptivos de la muestra (media y desviación típica) y la distribución en sus respuestas, puede ser consultada en la Tabla 1.

Tabla 1. Descriptivos de la muestra para los ítems

ITEM	M	SD	% de respuesta (Escala Likert 1-7)						
			1	2	3	4	5	6	7
BIU_01	4.0	1.71	9,1	11,5	14,9	19,1	24,1	13,7	7,6
BIU_02	4.31	1.77	6,9	14,5	10,7	15,6	24,4	16,8	11,1
BIU_03	4.19	1.81	9,2	14,1	11,5	16,4	19,8	19,8	9,2
BIU_04	4.46	1.79	8,8	8,0	11,8	16,4	22,9	18,7	13,4
PU_01	4.05	1.75	10,3	11,8	14,1	20,2	20,6	14,9	8,0
PU_02	4.00	1.75	11,8	9,2	16,8	21,0	19,5	13,4	8,4
PU_03	4.23	1.74	8,0	11,5	15,6	15,3	22,5	17,9	9,2
PU_04	4.61	1.64	4,2	9,2	9,2	21,8	24,8	15,6	15,3
PUFA_01	4.65	1.49	3,8	5,0	10,7	24,8	25,2	20,2	10,3
PUFA_02	4.42	1.69	5,7	9,2	14,1	22,5	17,2	19,5	11,8
PUFA_03	3.72	1.71	12,6	16,4	12,6	24,4	17,6	11,5	5,0
PUFA_04	4.17	1.65	7,3	10,7	14,9	24,0	18,3	17,9	6,9
PUFA_05	4.00	1.56	7,3	11,1	16,4	28,6	16,8	15,6	4,2
PUSA_01	4.08	1.86	10,7	13,4	14,9	18,7	15,6	14,5	12,2
PUSA_02	4.16	1.80	9,5	9,9	17,6	20,2	16,4	13,7	12,6
PUSA_03	3.44	1.73	17,2	18,7	12,2	25,2	13,7	8,0	5,0
PUSA_04	4.13	1.82	10,7	11,5	14,1	19,5	16,8	17,6	9,9
PUSA_05	3.77	1.67	11,8	13,4	16,4	24,0	18,3	10,7	5,3

Elaboración propia

Los resultados demuestran una alta aceptación en futuros docentes, observándose que prácticamente la totalidad de las medias para los ítems se encuentran en valores superiores a 4 en la escala medida, tanto para los ítems del modelo TAM (intención de uso y utilidad percibida), como para las dimensiones utilidad percibida de los dispositivos móviles para la evaluación sumativa y la utilidad percibida de éstos para la evaluación. A continuación, se presentan los resultados de los ítems para las dimensiones, agrupados en los cuatro constructos que aborda el estudio (Tabla 2).

Tabla 2. Descriptivos de la muestra para las dimensiones

	M	SD	As	Court	Min	Max	N
BIU	4,27	1,65	-,337	-,764	1,00	7,00	262
PU	4,23	1,53	-,222	-,669	1,00	7,00	262
PUFA	4,19	1,34	-,233	-,400	1,00	7,00	262
PUSA	3,91	1,49	-,149	-,663	1,00	7,00	262

Elaboración propia

De forma global, las valoraciones superan los 4 puntos de media en tres de las cuatro dimensiones propuestas, encontrándose la utilidad percibida de la tecnología para evaluaciones sumativas (PUSA) como principal excepción tomando un valor inferior a 4.

En primer término, se aborda la relación entre las dimensiones planteadas mediante el coeficiente de correlación correspondiente, para comprobar así la relación entre las intenciones de uso con las utilidades propuestas (Tabla 3).

Tabla 3. Correlación entre las dimensiones

	Pearson	Sig.	N
BIU-PU	,686	<,001	262
BIU-PUFA	,730	<,001	262
BIU-PUSA	,620	<,001	262

Elaboración propia

Los resultados del análisis correlacional muestran que las tres utilidades medidas tienen relación con la intención de uso tecnológico, tanto la utilidad percibida de la tecnología de forma genérica como la utilidad percibida para evaluaciones formativa y sumativas.

Justificada su relación, y con el objetivo de ampliar el modelo TAM hacia el campo específico de la evaluación mediada por la tecnología, se propone un estudio comparativo entre las dimensiones PUFA y PUSA en la muestra, para determinar si los participantes otorgan puntuaciones significativamente mayores a una u otra dimensión. Por tanto, se procede a la realización de un contraste de hipótesis.

Dicho contraste parte de la evaluación de la distribución muestral a través de la prueba de Kolmogorov-Smirnov y Shapiro-Wilk (Tabla 4).

Tabla 4. Contraste de normalidad

Item	Kolmogorov-Smirnov			Shapiro-Wik		
	Est.	Df.	Sig.	Est.	Df.	Sig.
PU	,062	262	,018	,985	262	,007
PUFA	,050	262	,002	,979	262	<,001
PUSA	,074	262	,001	,977	262	<,001

Elaboración propia

Los resultados de la prueba de normalidad reflejan que la muestra no sigue una distribución normal, por lo que se

continúa el contraste de hipótesis a través de la realización de pruebas no paramétricas que determinen la existencia (o no) de diferencias en las dos nuevas utilidades diseñadas.

Para dicho contraste, se plantea la prueba de los rangos con signos de Wilcoxon (prueba no paramétrica). Los resultados de la prueba se presentan por pares de ítems, comprobando sus diferencias en las puntuaciones de los sujetos (Tabla 5).

Tabla 5. W de Wilcoxon para los ítems

	W	Z	Sig.
PUFA_01 – PUSA_01	12415,500	4,245	<,001
PUFA_02 – PUSA_02	7678,000	2,426	0,013
PUFA_03 – PUSA_03	7584,000	2,916	0,003
PUFA_04 – PUSA_04	6737,000	0,089	0,928
PUFA_05 – PUSA_05	6594,000	2,569	0,008

Elaboración propia

Por último, se presentan los resultados de la prueba para el constructo general, resultado del cálculo de la media de los cinco ítems de cada constructo (Tabla 6).

Tabla 6. W de Wilcoxon para las dimensiones

	W	Z	Sig.
PUFA – PUSA	17395,000	3,250	0,001

Elaboración propia

El análisis de la prueba muestra diferencias significativas en 4 de los 5 ítems de cada dimensión (a excepción de PUFA_04), por lo que se puede afirmar que los futuros docentes difieren significativamente sobre la utilidad que confieren a los dispositivos móviles en función del tipo de evaluación a realizar. Analizando las medias presentadas en la Tabla 1, se observa cómo las diferencias son significativamente más altas para los ítems de la utilidad formativa, por lo que la utilidad sumativa obtiene puntuaciones significativamente menores. En una representación más visual, se observan las diferencias por pares de ítems en la siguiente figura (Figura 3).

Figura 3. Representación gráfica de las medias de los pares de ítems

En la siguiente sección se presentan las principales conclusiones alcanzadas tras la investigación.

4. CONCLUSIONES

Los resultados alcanzados tras la investigación ponen de manifiesto, en primer lugar, el elevado nivel de aceptación tecnológica manifestado por futuros docentes, estudiantes del Máster de Profesor en Educación Secundaria en el curso académico 2022/2023. Estos altos niveles de aceptación son indispensables para el avance de la educación y la inclusión

tecnológica, ya que garantizan la actitud positiva del profesorado ante los cambios que las nuevas realidades educativas demandan.

Por otro lado, la investigación ha realizado un esfuerzo por ampliar el modelo TAM mediante la propuesta de dos dimensiones nuevas no abordadas aún en la literatura, la utilidad percibida por los docentes de los dispositivos móviles para evaluaciones formativas (PUFA) y la utilidad percibida por los docentes para el uso de dispositivos móviles en evaluaciones sumativas (PUSA).

Los resultados reflejan diferencias significativas entre las utilidades diferenciadas en función de la modalidad evaluativa; por lo que se puede afirmar que los docentes conciben más útil el uso de la tecnología para evaluaciones formativas que para evaluaciones sumativas, lo que abre una nueva vía en la investigación de los modelos de adopción tecnológica por separado y el trabajo diferenciado en la formación entre evaluación sumativa y formativa.

REFERENCIAS

- Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. En J. Kuhl & J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11-39). Springer.
- Area-Moreira, M., & Adell-Segura, J. (2021). *Tecnologías Digitales y Cambio Educativo. Una Aproximación Crítica*.
- Black, P. J. (1993). Formative and Summative Assessment by Teachers. *Studies in Science Education*, 21, 49-97.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.
- Davis, F.D., Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: three experiments. *International Journal of Human-Computer Studies*, 45, 19-45
- Dixson, D. D., & Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory Into Practice*, 55(2), 153-159.
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention And Behavior: An introduction to theory and research*. Addison-Wesley.
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., Boud, D., Rotsaert, T., Brown, S., Salinas-Salazar, M. L., & Rodríguez-Gómez, H. M. (2020). El futuro de la evaluación en la Educación Superior. *Revista electrónica de investigación y evaluación educativa*.
- Kim, J., & Park, C.-Y. (2020). Education, skill training, and lifelong learning in the era of technological revolution: A review. *Asian-Pacific Economic Literature*, 34(2), 3-19.
- Mellado-Moreno, P. C., Sánchez-Antolín, P., Blanco-García, M., Mellado-Moreno, P. C., Sánchez-Antolín, P., & Blanco-García, M. (2021). Tendencias de la evaluación formativa y sumativa del alumnado en Web of Sciences. *ALTERIDAD.Revista de Educación*, 16(2), 170-183.
- Scriven, M. (1991). Beyond Formative and Summative Evaluation. *Teachers College Record*, 92(6), 19-64.

Anexo 8E

Comunicación a Congreso Internacional Technological Ecosystems for Enhancing Multiculturality (2023)

VERSIÓN PREPRINT

Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (En prensa). Behavioral intention to use mobile devices for assessment in pre-service teachers: A gender or an initial training issue?. En *Technological Ecosystems for Enhancing Multiculturality, TEEM 2023*. Springer.

Behavioral intention to use mobile devices for assessment in pre-service teachers: A gender or an initial training issue?

Alberto Ortiz-López¹[0000-0003-3164-5680], Susana Olmos Migueláñez²[0000-0002-0816-4179] and José Carlos Sánchez-Prieto³[0000-0002-8917-9814]

¹ Instituto Universitario de Ciencias de la Educación, GRIAL, Universidad de Salamanca

² Instituto Universitario de Ciencias de la Educación, GRIAL, Universidad de Salamanca

³ Instituto Universitario de Ciencias de la Educación, GRIAL, Universidad de Salamanca

aortiz@usal.es

Abstract. The technological acceptance of teachers in the face of new changes and advances in the educational system is one of the main challenges of research, and the assessment field is one of those involved in this transformation. Understanding the factors that condition teachers' technological adoption is a crucial task for the advance of training policies in initial, basic, and continuous training to ensure the progress of education towards the new digital realities. This paper presents the results of a descriptive study about the intention to use mobile devices for assessment processes in master's degree students of the University of Salamanca and seeks to know if gender and previous training condition its technological acceptance. The sample is composed of 262 students who answered a 24-item questionnaire, based on the TAM model and extended with the constructs autonomy and mobile self-efficacy. The results show a high level of acceptance among the participants, and significant differences according to the area of knowledge of the degree studied. As for gender, no significant differences in technological acceptance have been detected.

Keywords: Acceptance, Assessment, Technology Acceptance Model, Pre-service Teachers

1 Introduction

The implementation of technology in the educational environment generates, at this time, a huge impact on classrooms, redefining the role of all agents involved in the teaching-learning process [1]. The introduction of electronic resources such as computers, tablets and mobile devices has expanded the range of resources and tools available to teachers, who are forced to redefine their role in education assuming new challenges [2]. Therefore, it should be a fundamental task for the research to know the attitude that teachers have in the face of the rapid changes they experience, especially in the face of the inclusion and irruption of technology in the teaching [3,4]. Knowing the factors that determine their attitude is a basic requirement to predict the success or limitations of technological implementations [5], enabling their development from the initial, basic, or continuous training of teachers and future teachers.

Technology in education offers several opportunities to enhance the teaching-learning process, with many advantages focused on its flexibility and the individualization of teaching; but it also poses challenges that require constant updating and adaptation by teachers, and there may be resistance in its implementation [6,7].

The present study attempts to determine the current acceptance of the intention to use mobile devices in the assessment process through technology adoption models, the most common tool used for the study of acceptance. Among these models, the TAM model proposed by Davis [8] stands out as the most widely applied, based on the Theory of Planned Behavior [9] and the Theory of Reasoned Action [10]. The TAM model consists of four constructs (Figure 1) that determine the final intention of technological use (BI), all based on attitude towards its use (A), perceived usefulness (PU) and perceived ease of use (PE).

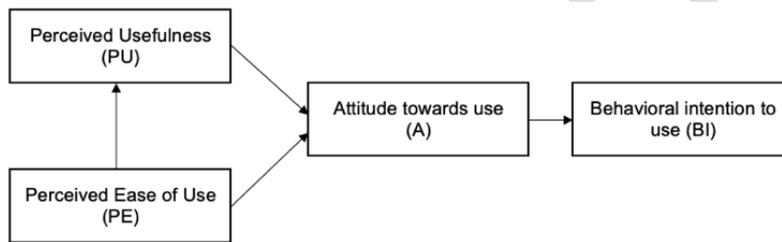


Fig. 1. TAM model proposed by Davis (1989).

The TAM model has been applied, developed, and validated in many fields, especially those related to online learning and the use of technology [11-15]. The present research is framed within a broad study on the technological acceptance of mobile devices for assessment processes in pre-service teachers, students of the Master of Teacher of Compulsory Secondary Education and Baccalaureate, Vocational Training and Language Teaching at the University of Salamanca.

The second section presents the methodology followed in the research, the third section the results achieved and, finally, the fourth section presents the conclusions of the study.

2 Methodology

The model presented here attempts to determine the intention to use technology, as well as several factors that may influence this intention, as a function of gender and area of knowledge of the graduates of the Master's Degree in Teaching at the University of Salamanca. To this end, the following hypotheses are proposed: gender influences the technological acceptance of mobile devices in evaluation processes (1) and the area of knowledge influences the technological acceptance of mobile devices in evaluation processes (2). The objective of the research is, therefore, to determine the level of technological acceptance, understood as teachers' predisposition to the use of mobile technology for evaluation processes in the classroom in their future teaching practice.

The methodology presented below summarizes the proposed research model, the variables addressed in the study, the population and sample used, and the explanation of the instrument developed.

2.1 Proposed research model

The research model proposed in this study is based on the TAM model [8] and is extended with two more constructs already addressed in this study line, [16] and mobile self-efficacy [17]. The model is presented below in terms of the constructs and hypotheses proposed.

TAM model variables. The proposed model is based on the initial assumptions proposed by Davis [9], including the constructs: perceived usefulness, perceived ease of use and behavioral intention to use. In the original proposal, the construct attitude towards use, eliminated in numerous studies due to its low explanatory and mediating power over intention is included. Therefore, based on these three constructs and in relation to the subject of the study, the following hypotheses are proposed:

1. Perceived ease of use positively predicts behavioral intention to use mobile devices for assessment processes in prospective teachers (H1).
2. Perceived usefulness positively predicts behavioral intention to use mobile devices for assessment processes in prospective teachers (H2).
3. Perceived ease of use positively predicts perceived usefulness of mobile devices for assessment processes in prospective teachers (H3).

Autonomy. The second construct proposed is the autonomy, understood as the perception of control and independence that technology grants to the teacher in evaluation processes [18]. This independence is the subsequent result of a greater perception of usefulness and ease in the implementation of evaluation processes by the teacher, so it can be affirmed that this construct will have a positive influence on those described in the TAM model [19]. The proposed construct has been raised based on Self-Determination Theory [20], and the following hypotheses are put forward:

4. Autonomy positively predicts teachers' perceived usefulness of using mobile devices in assessment processes (H4).
5. Autonomy positively predicts teachers' perceived ease of use of mobile devices in assessment processes (H5).

Mobile self-efficacy. The last construct addressed in this study is the mobile self-efficacy, defined as the perception that the teacher has about himself/herself when carrying out certain activities, in this case the assessment via mobile devices. The construct has been proposed for this field by Nikou & Economides [17], affirming that the greater the self-efficacy that a teacher perceives, the greater the perceived usefulness of the tool and the greater the ease with which he/she conceives its use for assessments [21].

6. Mobile self-efficacy positively predicts teachers' perceived usefulness for assessment processes (H6).
7. Mobile self-efficacy positively predicts teachers' perceived ease of use for evaluation processes (H7).

Therefore, the final proposed model is summarized in a total of five constructs and seven hypotheses to determine intention to use (Figure 2).

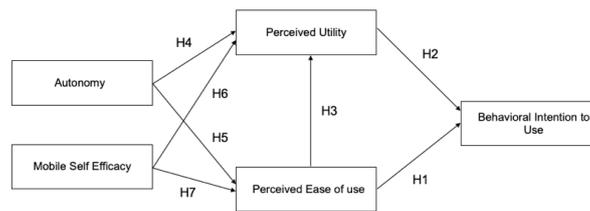


Fig. 2. Proposed model.

2.2 Variables

This study includes three different variables: exogenous, endogenous and variables used for the identification of the sample, which are described below.

- Exogenous: Perceived usefulness, perceived ease, autonomy, mobile self-efficacy.
- Endogenous: Behavioral intention to use.
- Identification variables: Gender, university degree taken.

2.3 Population and sample

The study population is composed of students of the Master's Degree in Teacher of Compulsory Secondary Education and Baccalaureate, Vocational Training and Language Teaching; the qualifying training for the exercise of teaching after the University Degree (N= 292). After electronic distribution through the LimeSurvey application [22], a total of 262 responses were obtained (89.7% of the population). The sample is composed of 58.4% women, 41.2% men and one participant who marked the option "Other" in the gender variable. As for training prior to the master's degree, the degrees were grouped into their respective branches, with 51.1% of subjects corresponding to the branch of Arts and Humanities, 28.2% to the branch of sciences (grouping the Science, Health Sciences and Engineering branches) and 16.7% corresponding to the branch of Social Sciences. The remaining 4% corresponds to missing values.

2.4 Instrument

The instrument designed consists of two sections, one corresponding to the identification variables (gender, University Degree) and the other relating to the items of the

proposed dimensions, which have been evaluated on a Likert-type scale (1-7). The questionnaire and its application have been favorably evaluated by the Research Ethics Committee of the University of Salamanca. The items of the model according to the dimension to which they belong are available at: <https://docs.google.com/document/d/1IRAUFfMEigo71Rt5sF3tDVcw9jEbP3xDfu7Vf1zAadU/edit?usp=sharing>

The items of the TAM model have been adapted from the proposals of the original model and the evolution of its versions [9,23,24]. The items of the autonomy construct have been adapted from Self-Determination Theory, specifying in its formulation the field of e-assessment [25]. Finally, the items belonging to the mobile self-efficacy construct have been adapted from the proposals of Nikou & Economides [17].

Prior to the presentation of results, the internal consistency of the instrument was evaluated through Cronbach's α coefficient, which reflects a high internal consistency ($\alpha=0.948$).

3 Results

To evaluate the intention to use mobile devices in assessment processes in future teachers, and to analyze their differences according to gender and the area of knowledge in which they feel their initial training, the results of the descriptive analysis performed for the constructs are presented. The results show, in the first place, a good attitude of the participants towards technological acceptance evaluated through their positive response to the proposed items, with practically all the averages being above 4 (on a Likert scale 1-7). The only two items below 4 are in the autonomy dimension (items AUT_01 and AUT_04), being above 3.9 in both cases. These results indicate that both self-efficacy, autonomy and TAM model values are high for the sample. Once the descriptive results have been presented, the differences according to gender and the area of knowledge of the initial degree taken by the students are tested.

The first variable presented for analysis is gender, taking in the study the values: male (108), female (153) and other (1). For the statistical analysis, the subject who indicated the gender "other" was not considered, given the differences in size between the groups and the almost null representativeness of the third group. The means of the items according to gender are presented below (Table 1).

Table 1. Descriptive by gender

Item	Male			Female		
	Average	S.D.	N	Average	S.D.	N
BIU_01	4,10	1,84	108	4,09	1,62	153
BIU_02	4,36	1,91	108	4,27	1,68	153
BIU_03	4,17	1,96	108	4,22	1,72	153
BIU_04	4,39	1,93	108	4,52	1,69	153
PEOU_01	4,24	2,02	108	3,90	1,86	153
PEOU_02	4,83	1,79	108	4,67	1,74	153
PEOU_03	5,77	1,43	108	5,49	1,40	153
PEOU_04	4,77	1,59	108	4,75	1,51	153

PU_01	4,12	1,79	108	4,02	1,73	153
PU_02	4,17	1,72	108	3,90	1,78	153
PU_03	4,31	1,74	108	4,18	1,77	153
PU_04	4,59	1,64	108	4,65	1,67	153
AUT_01	4,01	1,58	108	3,95	1,60	153
AUT_02	4,46	1,85	108	4,39	1,71	153
AUT_03	4,82	1,61	108	4,90	1,61	153
AUT_04	3,96	1,59	108	3,95	1,59	153
AUT_05	4,94	1,55	108	5,04	1,70	153
AUT_06	3,94	1,98	108	4,41	1,78	153
MSE_01	1,67	1,70	108	4,80	1,52	153
MSE_02	4,81	1,73	108	4,75	1,54	153
MSE_03	5,15	1,48	108	4,81	1,64	153
MSE_04	5,21	1,50	108	5,11	1,69	153
MSE_05	5,81	1,39	108	5,86	1,39	153

An initial analysis shows that the means do not seem to reflect big differences between the genders. To test the significance of the differences, a hypothesis test was performed. First, Kolmogorov-Smirnov and Shapiro Wilk normality tests (Table 2) were applied to determine the appropriate method for subsequent hypothesis testing. The results show rejection of normality ($p < 0.05$), so non-parametric techniques were applied.

Table 2. Normality tests

Item	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	Df.	Sig.	Statistic	Df.	Sig.
BIU_01	,157	262	<.001	0.942	1.706	<.001
BIU_02	,175	262	<.001	0.930	1.770	<.001
BIU_03	,159	262	<.001	0.928	1.815	<.001
BIU_04	,168	262	<.001	0.926	1.791	<.001
PEOU_01	,134	262	<.001	0.926	1.926	<.001
PEOU_02	,169	262	<.001	0.917	1.756	<.001
PEOU_03	,239	262	<.001	0.846	1.412	<.001
PEOU_04	,153	262	<.001	0.933	1.539	<.001
PU_01	,140	262	<.001	0.940	1.753	<.001
PU_02	,126	262	<.001	0.941	1.759	<.001
PU_03	,166	262	<.001	0.936	1.749	<.001
PU_04	,149	262	<.001	0.934	1.647	<.001
AUT_01	,138	262	<.001	0.949	1.591	<.001
AUT_02	,130	262	<.001	0.935	1.774	<.001
AUT_03	,168	262	<.001	0.915	1.605	<.001
AUT_04	,154	262	<.001	0.945	1.586	<.001
AUT_05	,178	262	<.001	0.895	1.639	<.001
AUT_06	,142	262	<.001	0.935	1.751	<.001
MSE_01	,178	262	<.001	0.926	1.587	<.001
MSE_02	,210	262	<.001	0.908	1.617	<.001
MSE_03	,185	262	<.001	0.908	1.582	<.001
MSE_04	,188	262	<.001	0.885	1.604	<.001
MSE_05	,221	262	<.001	0.794	1.389	<.001

Therefore, assuming a non-normal distribution in the sample and categorizing the gender variable into two categories, the Mann-Whitney U test (Table 3) was applied for the contrast. The results indicate that significant differences are only found in 2 items of the 24 analyzed (PEOU_03 and AUT_06), so it cannot be affirmed that gender is an

influential variable in the technological acceptance of mobile devices for evaluation (n.s. 0.05).

Table 3. Mann-Whitney U Test

Item	U. M-W	W Wilcoxon	Z	Sig.(bilateral)
BIU_01	8149,000	19930,000	-,191	,849
BIU_02	7939,000	19720,000	-,546	,585
BIU_03	8196,500	14082,500	-,110	,912
BIU_04	8101,000	13987,000	-,272	,786
PEOU_01	7453,000	19234,000	-1,362	,173
PEOU_02	7746,500	19527,500	-,872	,383
PEOU_03	7079,000	18860,000	-2,041	,041
PEOU_04	8118,500	19899,500	-,243	,808
PU_01	7995,000	19776,000	-,451	,652
PU_02	7611,500	19392,500	-1,098	,272
PU_03	7932,000	19713,000	-,557	,577
PU_04	7993,500	13879,500	-,455	,649
AUT_01	8161,000	19942,000	-,171	,864
AUT_02	8008,500	19789,500	-,428	,669
AUT_03	8007,000	13893,000	-,433	,665
AUT_04	8156,000	14042,000	-,180	,857
AUT_05	7683,500	13569,500	-,984	,325
AUT_06	6832,500	12718,500	-2,416	,016
MSE_01	7996,000	13882,000	-,452	,651
MSE_02	7892,000	19673,000	-,630	,529
MSE_03	7371,000	19152,000	-1,516	,130
MSE_04	8174,000	19955,000	-,150	,881
MSE_05	8074,000	13960,000	-,330	,742

The second factor addressed in the study is the area of knowledge of the degree studied. These fields have been categorized into three groups: Arts and Humanities (1), Sciences (including Sciences, Health Sciences and Engineering) (2) and Social Sciences (3). For the contrast of the second factor, the same procedure is followed as above, presenting the means according to the branch (Table 4) and showing few differences, more observable in lower means in Arts and Humanities than in Sciences and Social Sciences.

Table 4. Average according to the field of knowledge of the Degree

Item	Arte y Humanidades			Ciencias			Ciencias Sociales		
	Aver.	S.D.	N	Aver.	S.D.	N	Aver.	S.D.	N
BIU_01	3,86	1,60	134	4,34	1,61	74	4,50	1,89	44
BIU_02	4,14	1,72	134	4,49	1,63	74	4,61	2,00	44
BIU_03	4,04	1,78	134	4,36	1,71	74	4,48	1,95	44
BIU_04	4,32	1,78	134	4,66	1,70	74	4,66	1,88	44
PEOU_01	3,72	1,96	134	4,14	1,96	74	4,68	1,60	44
PEOU_02	4,61	1,81	134	4,89	1,68	74	1,95	1,67	44
PEOU_03	5,46	1,47	134	5,91	1,04	74	5,66	1,49	44
PEOU_04	4,52	1,63	134	5,04	1,30	74	5,09	1,44	44
PU_01	3,81	1,81	134	4,14	1,66	74	4,82	1,45	44
PU_02	3,74	1,82	134	4,08	1,65	74	4,86	1,47	44
PU_03	3,99	1,85	134	4,22	1,64	74	4,93	1,44	44
PU_04	4,46	1,70	134	4,69	1,58	74	5,02	1,53	44
AUT_01	3,72	1,61	134	4,00	1,41	74	4,75	1,33	44
AUT_02	4,16	1,84	134	4,58	1,65	74	5,09	1,46	44

AUT_03	4,73	1,67	134	5,01	1,58	74	5,14	1,34	44
AUT_04	3,80	1,59	134	4,00	1,52	74	4,43	1,57	44
AUT_05	4,89	1,65	134	5,15	1,68	74	5,16	1,52	44
AUT_06	4,01	1,80	134	4,46	1,72	74	4,39	1,67	44
MSE_01	4,55	1,59	134	4,86	1,63	74	5,02	1,34	44
MSE_02	4,52	1,64	134	5,08	1,44	74	5,14	1,55	44
MSE_03	4,67	1,71	134	5,27	1,30	74	5,27	1,39	44
MSE_04	4,96	1,77	134	5,46	1,22	74	5,30	1,50	44
MSE_05	5,65	1,58	134	6,09	1,08	74	5,95	1,22	44

To test the significance of the means, assuming a distribution that does not conform to the normal one, and with a polytomous grouping variable, the Kruskal-Wallis H test was used (Table 5).

Table 5. Kruskal-Wallis H Test

Item	H de Kruskal-Wallis	Gl.	Sig.
BIU_01	6,859	2	,032
BIU_02	3,720	2	,156
BIU_03	2,890	2	,236
BIU_04	2,753	2	,252
PEOU_01	8,500	2	,014
PEOU_02	1,728	2	,421
PEOU_03	3,631	2	,163
PEOU_04	7,147	2	,028
PU_01	10,212	2	,006
PU_02	13,550	2	,001
PU_03	8,971	2	,011
PU_04	4,108	2	,128
AUT_01	14,173	2	<,001
AUT_02	9,635	2	,008
AUT_03	2,692	2	,260
AUT_04	4,611	2	,100
AUT_05	2,140	2	,343
AUT_06	3,631	2	,163
MSE_01	4,686	2	,096
MSE_02	9,289	2	,010
MSE_03	7,099	2	,029
MSE_04	2,869	2	,238
MSE_05	2,934	2	,231

The table shows significant differences in 10 of the 24 items proposed (41.7% of the items). The differences are found mainly in the perceived usefulness dimension (significant differences in PU_01, PU_02, PU_03), the same situation as in the significant items of the autonomy (AUT_01, AUT_02) and self-efficacy (MSE_02, MSE_03). In addition, there are significant differences in items BIU_01 and PEOU_01).

In all cases, these differences reflect that the highest scores are found in graduates in the Social Sciences area, significantly higher than for graduates in Sciences, who in turn obtain significantly higher scores than graduates in Arts and Humanities, who give the lowest scores in the study; a situation that results in a greater predisposition to evaluation through mobile devices in subjects from Sciences and greater resistance in subjects from Arts and Humanities.

4 Conclusions

The results have shown a high behavioral intention to use mobile devices for assessment processes in future teachers, considering the high mean values for all items, especially for those that explicitly assess the behavioral intention to use technology (BIU). What the results suggest is the high importance that future teachers place on the inclusion of technology in education (and specifically in assessment processes), considering not only its usefulness and ease of use, but also the autonomy it confers on the teacher and the importance of self-efficacy in its use.) These high values in the behavioral intention to use mobile devices follow the trend of previous research conducted in the field [26].

Secondly, the first hypothesis test has not shown significant differences in terms of gender, a factor that may be due to the high knowledge and preparation that the population currently has (especially the younger ones, future teachers in this case) about technology, which means that its use and implementation does not depend on and is not related to variables such as gender [27].

Finally, the second hypothesis has shown partial differences among the participants according to the area of knowledge to which their degree belongs. These differences may be based on the different theoretical-practical conceptions of the initial training, linked to the theoretical in some and more related to the practical in others. This hypothesis opens a new avenue in the research to know explicitly what factors affect, to what extent they do so and how to work on their development to favor an adequate relationship between future teachers and technology.

References

1. Gómez-Vera, K.: El desafío de las nuevas tecnologías: El uso del aula virtual y su influencia en el rendimiento académico. *ReHuSo: Revista de Ciencias Humanísticas y Sociales*, 4(3), 48-56 (2019).
2. Guerrero, I., & Kalman, J.: La inserción de la tecnología en el aula: Estabilidad y procesos instituyentes en la práctica docente. *Revista Brasileira de Educação*, 15, 213-229 (2010).
3. Aguaded, I., & Tirado-Morueta, R.: Influencias de las creencias del profesorado sobre el uso de la tecnología en el aula. *Revista de Educación (Madrid)*, 230-255 (2014).
4. Escofet, A., Gros, B., López, M., & Marimon-Martí, M.: Percepción del profesorado sobre la integración de la tecnología en el espacio escolar. *RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa* (2019).
5. Sánchez-Prieto, J., Olmos-Migueláñez, S., & García-Peñalvo, F.: Informal tools in formal contexts: Development of a model to assess the acceptance of mobile technologies among teachers. *Computers in Human Behavior*, 55 (2015).
6. Gómez, V. L., & Álvarez, G: Tecnologías digitales en la escuela primaria: Las perspectivas de los docentes sobre su inclusión y la enseñanza en las aulas. *Virtualidad, Educación y Ciencia*, 11(20), 9-26 (2020).
7. Granados-Maguiño, M. A., Romero, S., Rengifo Lozano, R. A., & Garcia Mendocilla, G. F.: Tecnología en el proceso educativo: Nuevos escenarios. *Revista Venezolana de Gerencia: RVG*, 25(92), 1809-1823 (2020).

8. Davis, F. D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340 (1989).
9. Ajzen, I.: From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl & J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11-39). Springer (1985).
10. Fishbein, M., & Ajzen, I. : *Belief, Attitude, Intention And Behavior: An introduction to theory and research*. Addison-Wesley (1975).
11. Abdirahman, A. A., Jabar, M. A., & Hashi, A. O.: Attitude towards intention to use Mobile-Based Teaching Assessment in relation to Technology Acceptance Model. *International Journal of Advances in Electronics and Computer Science*, 5(2), 35-40 (2018).
12. Abdirahman, A. A., Jabar, M. A., Hashi, A. O., Elmi, M. A., & Rodriguez, O. E. R. Attitude Towards Intention to Use Mobile-Based Teaching Assessment Based on TAM and a Novel Extended TAM Model. In F. Saeed, F. Mohammed, & F. Ghaleb (Eds.), *Advances on Intelligent Informatics and Computing*, 583-593 (2022).
13. Alshurideh, M., Al Kurdi, B., Alhamad, A. Q., Salloum, S., Alkurdi, S., Dehghan, A., Abuhashesh, M., & Masa'deh, R.: Factors Affecting the Use of Smart Mobile Examination Platforms by Universities' Postgraduate Students during the COVID 19 Pandemic: An Empirical Study. *Informatics*, 8(32) (2021).
14. Harchay, A., Berguiga, A., Cheniti-Belcadhi, L., & Braham, R.: Student Perception of Mobile Self-assessment: An Evaluation of the Technology Acceptance Model. *Interaction Design and Architecture(s)*, 109-124 (2019).
15. Nikou, S., & Economides, A.: Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 109, 56-73 (2017).
16. Reeve, J., & Cheon, S. H.: Autonomy-supportive teaching: Its malleability, benefits, and potential to improve educational practice. *Educational Psychologist*, 56(1), 54-77 (2021).
17. Nikou, S., & Economides, A.: Mobile-Based Assessment: Integrating acceptance and motivational factors into a combined model of Self-Determination Theory and Technology Acceptance. *Computers in Human Behavior*, 68, 83-95 (2017).
18. Jones, A., & Issroff, K.: Motivation and mobile devices: Exploring the role of appropriation and coping strategies. *ALT-J: Research in Learning Technology*, 15(3) (2007).
19. Su, Y.-L., & Reeve, J.: A Meta-analysis of the Effectiveness of Intervention Programs Designed to Support Autonomy. *Educational Psychology Review*, 23(1), 159-188 (2011).
20. Deci, E. L., & Ryan, R. M.: Self-determination theory. In *Handbook of theories of social psychology*, 416-436), Sage Publications (2012).
21. Anderson, A.: Predictors of computer anxiety and performance in information systems. *Computers in Human Behavior*, 12(1), 61-77 (1996).
22. Engard, N. C.: LimeSurvey. *Public Services Quarterly*, 5(4), 272-273 (2009).
23. Venkatesh, V., & Bala, H.: Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39, 273-315 (2008).
24. Venkatesh, V., & Davis, F.: A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46, 186-204 (2000).
25. Deci, E. L., & Ryan, R. M.: Self-determination theory. In *Handbook of theories of social psychology*, 416-436). Sage Publications (2012).
26. Al-Emran, M., Mezhyuev, V., & Kamaludin, A.: Technology Acceptance Model in M-learning context: A systematic review. *Computers & Education*, 21 (2018).
27. Ramírez-Correa, P.; Rondán-Cataluña, F.J. & Arenas-Gaitán, J.: Influencia del género en la percepción and adopción de e-learning: Estudio exploratorio en una universidad chilena. *Journal of Technology Management and Innovation*, 5, 129-141 (2010).

Anexos A-C
Anexos de la Tesis Doctoral

Anexo A

Informe favorable del Comité de Ética en la Investigación de
la Universidad de Salamanca



El Comité de Ética de la Investigación de la Universidad de Salamanca, en su reunión ordinaria celebrada el 20 de octubre de 2022, ha considerado las circunstancias que concurren en el proyecto de investigación titulado *“Estudio de los factores que condicionan la aceptación y la intención de uso docente de dispositivos móviles en e-evaluación”*, que tiene como investigador/a principal a el/la Dr/a. Susana Olmos Miguelañez

A la vista de la documentación presentada, este Comité ha acordado **informar favorablemente** el proyecto de investigación con nº de registro 903, ya que cumple los requisitos éticos requeridos para su ejecución.

Y para que así conste lo firmo en Salamanca a 24 de octubre de 2022

MUÑOZ DE LA
PASCUA LUIS JOSE
- DNI 31238752Z

Firmado digitalmente por
MUÑOZ DE LA PASCUA LUIS
JOSE - DNI 31238752Z
Fecha: 2022.10.24 11:08:24
+02'00'

Fdo.: Luis Muñoz de la Pascua
Secretario del CBE

GUTIERREZ
RODILLA
BERTA MARÍA
- 07843572C

Firmado digitalmente
por GUTIERREZ
RODILLA BERTA
MARIA - 07843572C
Fecha: 2022.10.24
09:42:15 +02'00'

Fdo.: Berta Gutiérrez Rodilla
Presidente del CBE

Anexo B

Ítems del modelo F/S-MAAM

Modelos F/S-MAAM: Listado de ítems

Constructo	Ítem
Intención conductual de uso de dispositivos móviles en evaluaciones formativas (BIUF)	
BIUF_01	Pretendo utilizar dispositivos móviles móviles para la realización de evaluaciones continuas.
BIUF_02	Mi intención es utilizar tecnologías móviles para actividades que lleven un feedback al estudiante.
BIUF_03	Estoy dispuesto a utilizar dispositivos móviles para evaluaciones de un trabajo.
BIUF_04	Usaría dispositivos móviles para realizar evaluaciones orientadas a la mejora del proceso de enseñanza-aprendizaje.
Intención conductual de uso de dispositivos móviles en evaluaciones sumativas (BIUS)	
BIUS_01	Estoy dispuesto a utilizar dispositivos móviles móviles para la realización de pruebas objetivas teóricas.
BIUS_02	Pretendo utilizar tecnologías móviles para evaluaciones finales de una asignatura.
BIUS_03	Mi intención es utilizar dispositivos móviles para evaluaciones orientadas a una calificación.
BIUS_04	Usaría dispositivos móviles para realizar una evaluación final de unas prácticas.
Utilidad percibida de los dispositivos móviles en evaluaciones formativas (PUFA)	
PUFA_01	La evaluación basada en el dispositivo móvil es útil para evaluaciones formativas (orientada a la mejora).
PUFA_02	El uso de dispositivos móviles me permite realizar las evaluaciones formativas más eficaces.
PUFA_03	El uso de dispositivos móviles en evaluación formativa mejora la calidad de mi docencia.
PUFA_04	Las ventajas de usar dispositivos móviles en evaluaciones formativas supera ampliamente las desventajas.
Utilidad percibida de los dispositivos móviles en evaluaciones sumativas (PUSA)	
PUSA_01	La evaluación basada en el dispositivo móvil es útil para evaluaciones sumativas (finales).
PUSA_02	El uso de dispositivos móviles me permite realizar las evaluaciones sumativas más eficaces.
PUSA_03	El uso de dispositivos móviles en evaluación sumativa mejora la calidad de mi docencia.
PUSA_04	Las ventajas de usar dispositivos móviles en evaluaciones sumativas supera ampliamente las desventajas.
Facilidad de uso percibido de los dispositivos móviles en evaluación (PEOU)	
PEOU_01	Encuentro fácil que las tecnologías móviles hagan lo que yo quiero que hagan.
PEOU_02	Para mí sería fácil ser hábil en el uso de evaluaciones a través del móvil.
PEOU_03	Mi interacción con dispositivos móviles a la hora de evaluar es clara y comprensible.
PEOU_04	Las evaluaciones mediadas por dispositivos móviles son fáciles de usar.
Norma subjetiva (SN)	
SN_01	Mis compañeros piensan que debería utilizar dispositivos móviles para la evaluación.
SN_02	Mis alumnos piensan que debería utilizar dispositivos móviles para la evaluación.
SN_03	Las autoridades educativas demandan el uso de tecnologías móviles para evaluar en educación.
SN_05	Las personas importantes para mí creerán que debo utilizar dispositivos móviles para evaluar.
SN_05	La sociedad espera que los profesores empleen dispositivos móviles en evaluación.
Resistencia al cambio (CR)	
RC1	Me resisto a que las tecnologías móviles cambien la manera en que se desarrolla mi función docente.
RC2	Me opongo a que las tecnologías móviles cambien el sentido de la evaluación.
RC3	Me niego a que las tecnologías móviles cambien la interacción profesor-alumno.
RC4	Es contraproducente que las evaluaciones dependan de factores tecnológicos.
RC5	Los dispositivos móviles me suponen un cambio difícil de asumir para evaluar.
Ansiedad ante el uso de dispositivos móviles en procesos de evaluación (ASAN)	
ASAN_01	Los dispositivos móviles no me dan ningún miedo.
ASAN_02	Trabajar con un dispositivo móvil me pone nervioso.
ASAN_03	Los dispositivos móviles me hacen sentir incómodo.
ASAN_04	Los dispositivos móviles me inquietan.
Autonomía (AUT)	
AUT_01	Me sentiré cómodo si tengo que evaluar con dispositivos móviles.
AUT_02	Evaluar mediante dispositivos móviles me ofrecerá opciones y elecciones interesantes.
AUT_03	El uso de dispositivos móviles para evaluar me permitirá realizar evaluaciones de forma alternativa.
Autoeficacia en el uso de dispositivos móviles (MSE)	
MSE_01	Sería capaz de integrar un dispositivo móvil en mis procesos de evaluación de forma efectiva.
MSE_02	Sabría desarrollar un proceso de evaluación utilizando un dispositivo móvil.
MSE_03	Enstaría capacitado para ofrecer retroalimentación digital a una evaluación realizada a través de un dispositivo móvil.

Modelos F/S-MAAM: Listado de ítems

Constructo	Ítem
------------	------

Innovación personal (PI)

- | | |
|-------|---|
| PI_01 | Me gustaría experimentar con nuevas tecnologías para evaluar. |
| PI_02 | Las nuevas tecnologías tienen que ser parte importante en el sistema educativo. |
| PI_03 | Innovar en evaluación es importante para el avance y actualización del sistema educativo. |
| PI_04 | La innovación en herramientas de evaluación es una de mis principales preocupaciones. |
| PI_05 | Lideraría un proyecto de implementación de nuevas tecnologías para evaluar. |

Rechazo ante la incertidumbre (UBR)

- | | |
|--------|--|
| UBR_01 | Las normas o los reglamentos específicos son importantes para m. |
| UBR_02 | Los requisitos detallados son importantes para mi. |
| UBR_03 | Tener instrucciones detalladas es una parte importante para mi. |
| UBR_04 | Los precedentes/antecedentes operativos estandarizados me ayudan a seguir un ejemplo |
| UBR_05 | Lo mejor es seguir al pie de la letra los requisitos, instrucciones y procedimientos |

Compatibilidad con el estilo de trabajo (COMP)

- | | |
|---------|---|
| COMP_01 | Utilizar tecnologías móviles para evaluar en mi docencia sería compatible con mi estilo de trabajo. |
| COMP_02 | Utilizar tecnologías móviles para evaluar en mi docencia sería compatible con mi forma de trabajar. |
| COMP_03 | Utilizar tecnologías móviles para evaluar a mis alumnos encajaría con mi estilo de vida. |
-

Anexo / Appendix C

Extended abstract of the Ph.D. Dissertation



University Institute of Educational Sciences (IUCE)
Doctorate Program in Education in the Knowledge Society

DOCTORAL DISSERTATION

**Design of a technology adoption model for
mobile devices in formative and summative
evaluation**

Author: Alberto Ortiz López

Directors: Susana Olmos Migueláñez y José Carlos Sánchez Prieto

November 2023



VNIVERSIDAD D SALAMANCA

UNIVERSIDAD DE SALAMANCA

Extended abstract of the PhD Dissertation**Design of a technology adoption model for mobile devices in formative and summative assessment**

Alberto Ortiz-López

Keywords: Technology adoption, Technology Acceptance Model, assessment, mobile devices, teachers, quality, structural equation models.

In the extensive journey of educational assessment, its terminological approach has undergone a continuous transformation. From a historical, social, and ideological perspective, the assessment of learning in any formal educational context is a primary concern for educators, students, and educational institutions (Black & Wiliam, 1998; Hattie & Timperley, 2007; Ibarra-Sáiz & Rodríguez-Gómez, 2020; Perrenoud, 1991; Scriven, 1967; Shepard, 2000).

Originally, assessment was synonymous with measuring students' learning outcomes and performance (Rodríguez-Neira et al., 2001). Today, the notion of assessment encompasses not only academic assessment but also all the interconnected elements in the complex teaching-learning process (James, 2006), evolving into conceptions such as assessment for learning (Barrientos-Hernán et al., 2020), assessment as learning (Ibarra-Sáiz & Rodríguez-Gómez, 2019), sustainable assessment (Boud, 2000; Boud & Soler, 2016), or assessment for empowerment (Fetterman et al., 2015; Fetterman & Wandersman, 2004), among others.

Educational assessment is defined as a "process of collecting and analyzing relevant information to discover any facet of the educational reality and formulate a judgment about its adequacy to a pre-established pattern or criterion which constitutes the basis for decision-making" (De la Orden, 1982, p. 2). This process focuses on determining the achievement of objectives (Anderson, 2005) and supporting students with feedback in the teaching-learning process (Morris et al., 2021). Among the different perspectives conceptualizing authentic assessment, Gulikers et al. (2004) present five dimensions: the assessment task, the physical context, the social context, the assessment outcome, and the established criteria for issuing value judgment.

Therefore, assessing involves being part of a systematic information collection process (planned, sequenced, and with control over agents, techniques, and instruments) linked to issuing a value judgment based on assessment criteria and aimed at decision-making and improving educational practice (Tejada, 1997). In this regard, Pérez-Juste and García-Ramos (1989, p. 23) present a definition that seeks to contextualize assessment as a whole:

“el acto de valorar una realidad, formando parte de un proceso cuyos momentos previos son los de fijación de las características de la realidad a valorar, y de recogida de información sobre las mismas, y cuyas etapas posteriores son la información y la toma de decisiones”

[the act of assessing a reality as part of a process whose preceding stages involve establishing the features of the reality to be assessed and gathering information about them, with subsequent stages encompassing information analysis and decision-making].

Within the educational sphere and guided by standards of utility, feasibility, ethics, and honesty (Joint Committee on Standards for Educational Evaluation, 1988), evaluation assumes various functions that confer a holistic and intrinsic character to the entire educational process (Stufflebeam & Shinkfield, 1995). Variably described in the literature, these functions are sorted by authors such as Hernández (1998) or Posner (1998) into four main functions: diagnostic, instructive, educational, and self-formative. Other authors, such as García-Ramos (1989), expand on these by adding predictive, guiding, and control functions. These functions assigned to evaluation construct it as a broad process focused on various aspects and stages related to teaching and the educational domain. It also results from a modeling process in the interaction between educational agents and the evaluation.

Regarding assessment modalities in education, the research delves into formative and summative assessment as fundamental components of its development. Scriven (1967) delineates both modalities and asserts that summative assessment takes place after the teaching-learning process and is oriented towards decision-making regarding certification or the attainment of degrees, while formative assessment focuses on the process itself to guide improvement during the developmental phase.

The first of these, summative assessment, closely related to early models of educational assessment (quantitative), prioritizes the assessment of the product over the process and directs measurement towards the degree of goal achievement or final certification (Harlen & James, 1997; Knight, 2002). This assessment encompasses all techniques aimed at certification, such as exams, tests, or final assessment exercises, among others (Kibble, 2017).

Derived from discrepancies with the rigid summative assessment models (Sacristán-Díaz et al., 2010), assessment expands its scope to become synonymous with a comparison- and improvement-

oriented appraisal (González-Monteaudo, 2001): formative assessment. Bell and Cowie (2001, p. 536) define this approach as "the process used by teachers and students to recognize and respond to student learning to enhance that learning during the learning." This assessment stands out for the significance of feedback, an aspect addressed from a dual perspective according to Miras and Solé (1990), as it impacts both the student (aid in their learning at a given moment) and the teacher (information about the state of the learning process and stimulation of metacognition) (Morris et al., 2021).

The research on the development of the described assessment modalities delineates three trends regarding the teaching function and its relationship with the development of these assessments. The first trend assumes the non-differentiation between formative and summative assessment, considering assessment as an inherent whole in the teaching practice. The second trend acknowledges the coexistence of both modalities but emphasizes its independence in practice. Finally, the third trend advocates for their necessary complementarity for an adequate assessment process.

On the one hand, the first trend assumes educational assessment as a single set of processes, without distinguishing between the process and the achieved result, and understands its utility in terms of effectiveness, performance, and goal responsiveness (Mejía-Pérez, 2012). This trend aligns with the classical conception of assessment (Shepard, 2006; Tyler, 1949).

On the other hand, the second trend assumes the independence of both modalities and seeks to differentiate and separate their effects on students and teaching (Patton, 1996). It proposes formative assessments on one side and summative assessments on the other (Black & Wiliam, 1998). This proposition aligns with one of the most established and classical views on assessment, one that distinguishes between their effects and assumes that a combined and unitary model can be a challenge for teachers who intend to integrate different techniques and purposes (MacLellan, 2001).

Finally, a third perspective acknowledges the duality of assessment in its formative and summative nature. This implies that educators conceive the assessment process as the indispensable combination of both (Ahmad & Bhat, 2019; Buchholtz et al., 2018), as their purposes are intricately related and complementary to achieving a comprehensive assessment (Black et al., 2003). This occurs in a process mediated by feedback and guidance to reach milestones and a final certification (Dixon & Worrell, 2016; Dolin et al., 2018).

Despite summative assessments being a mandatory requirement in the current educational system (given the necessary certification of student performance), educators must make decisions about the assessment process, including when, how, and why to assess and for what purpose.

Currently, these evaluative modalities are compromised by the rapid changes and advancements in educational technology, leading to a transition and adaptation of assessment to new environments (Tomas et al., 2015). This process has revalued the assessment of meaningful learning, fostered student participation in assessment, promoted collaboration among students, and encouraged self-assessment and peer assessment (Ibarra-Sáiz et al., 2020) in virtual environments. All of this is supported by technology (García-Peñalvo, 2021; Hernández-Sellés et al., 2018), which has enabled the creation of more flexible, engaging, and user-friendly scenarios for both students and educators, where assessment plays an active, objective, and practical role (Ghouali et al., 2020).

Assessment in virtual environments seeks to counteract the rigidity associated with many classical assessment methods, which often focus on accountability and student grading while overlooking creativity and originality—two essential aspects for student motivation (Pearse-Romera & Ruiz-Cecilia, 2019). Consequently, institutions currently face a comprehensive challenge in overcoming the obstacles posed by this adaptation, enhancing the integration of education and assessment in technology-mediated contexts, and embracing the concept of assessment as synonymous with guidance for learning improvement, synonymous with the enhancement of teaching quality (Jiménez et al., 2021).

In this regard, it is crucial to emphasize the intrinsic connection between quality and assessment in the educational context (Koslowski, 2006). Research findings underscore the critical importance of determining the quality of assessment practices to ensure confidence in the process and support the successful implementation of technology in the new educational paradigm (Connolly et al., 2005).

Therefore, integrating formative and summative assessments with Information and Communication Technologies (ICT) is essential in today's educational landscape. This focus optimizes the teaching-learning process and contributes to greater technological inclusion. Delving into this relationship, the research focuses on using mobile devices for evaluative processes, providing a robust foundation for the future of teaching, methodologies, and assessment (Fombona-Cadavieco & Rodil-Pérez, 2018).

Mobile learning (m-learning) is defined as the implementation of mobile devices in e-learning systems to develop learning experiences (Elkheir & Abdul-Mutalib, 2015). This process allows for innovative teaching based on ubiquity, immediacy, mobility, permanent and open access, and contextualization (Gabriel-Pedro et al., 2018). These advancements lead to a new context that enables the online development of modalities such as self-assessment, peer assessment, or hetero-assessment (depending on the involved agent), as well as formative assessments, performance-based assessments, summative assessments, or competency-based assessments (depending on the

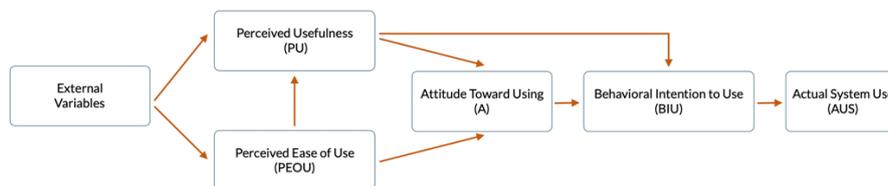
purpose) (Rodríguez-Gómez et al., 2013). Additionally, technology has been a determining factor in situations like the COVID-19 pandemic (2020), generating high expectations (García-Peñalvo et al., 2020).

Despite its advantages, integrating mobile devices in classrooms and their evaluative use is still challenging, facing multifactorial resistance that leads to only sporadic implementations (Álvarez-Herrero, 2018). Furthermore, teacher resistance to their use has been defined as one of the main current limitations (Córica & García-Aretio, 2018), as they are the ultimate decision-makers for their introduction in the classroom. Therefore, it can be asserted that a teacher who does not embrace technological development in their assessments will employ a traditional methodology that does not incorporate mobile technology (Taherdoost, 2019).

The present research focuses on understanding the factors that condition the teachers' technological adoption of these devices, which will enable the design of practical training and teacher updating practices—a crucial task to ensure the introduction of functional changes in educational systems. The literature has not previously described the relationship between assessment modalities and technology acceptance. This dissertation aims to fill a gap that delves into formative and summative assessment and the specific determinants of their adoption among teachers.

Understanding these factors aligns with exploring the theoretical structures of technology adoption models, which analyze and explain the acceptance and use of technologies in various contexts. Among these models, one of the most widely recognized (Hernández-García, 2012) is the Technology Acceptance Model (TAM) (Davis, 1989). The TAM model (Figure 1) has undergone a long process from its formulation to its current representation in acceptance studies. This model is the result of the principles of the Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1975) and the Theory of Planned Behavior (TPB) (Ajzen, 1985). It aims to predict behaviors through the study of behavioral intentions. In other words, it determines whether the subject is likely to perform a specific behavior by studying their attitude and the mediating effect of behavioral intention.

FIGURE 1. Technology Acceptance Model. Adapted from Davis (1989)



In the approach to Davis's model (1989), the behavioral intention to use a technological system is determined by the attitude toward its use as a precursor to this intention and by studying the perceived ease of use and usefulness. For the author, these constructs are defined as follows:

- Perceived Ease of Use: “The degree to which a person believes that using a particular system would be free of effort” (Davis, 1989, p. 320).
- Perceived Usefulness: “The degree to which a person believes that using a particular system would enhance his or her job performance” (Davis, 1989, p. 320).
- Attitude Toward Use: “The person’s general feeling of favorableness or unfavorableness toward some stimulus object” (Fishbein & Ajzen, 1975, p. 216).

Following the TAM model, literature has seen the emergence of models that integrate relevant adoption theories, such as the combination of the TAM and TPB (C-TAM-TPB) (Taylor & Todd, 1995a) or the combination of the TAM and IDT and TPB theories (D-TPB) (Taylor & Todd, 1995b). Additionally, to overcome the limitations of the TAM model, two evolutionary models have surfaced years after Davis's proposal (1989): the TAM2 model (Venkatesh & Davis, 2000) and the TAM3 model (Venkatesh & Bala, 2008).

The second version of the model (TAM2) retains the TAM model constructs, excluding attitude toward use due to its limited moderator effect, as evidenced by previous (Davis & Venkatesh, 1996) and subsequent studies (Hu et al., 2003). The third version (TAM3) maintains the proposed constructs and groups the antecedents of perceived ease of use into two categories: anchoring and adaptation factors (Venkatesh & Bala, 2008).

Despite its limitations, the TAM model is one of the most studied and developed in acceptance studies (King & He, 2006). In the educational field, this model is also the most representative in adoption studies (Scherer et al., 2019).

In the search for reviews that synthesize the relationship between Davis's model (1989) and technology-mediated teaching, Abduyah and Ward (2016) conducted a meta-analysis of 107 adoption models to identify the most used external factors in the context of e-learning adoption. Results showed that self-efficacy, subjective norm, technology anxiety, perceived enjoyment, and experience are the most used external factors besides the TAM model.

Further delving into the relationship between mobile devices and the TAM model, a synthesis of 64 relevant models by Alsharida (2021) stands out. Key findings have confirmed that self-efficacy is the most frequently employed factor in TAM models in m-learning, followed by subjective norm, perceived enjoyment, anxiety in mobile device use, facilitating conditions, social influence, innovation, and satisfaction, respectively. This study also confirms the growing trend of publishing expanded TAM models.

Lastly, regarding the relationship between technological adoption and mobile-based assessment, research focusing on technological adoption in this area still needs to be explored (Alrofou et al., 2019). Prominent among the expanded TAM models for the study of adoption in this

context are the "Mobile-Based Assessment Acceptance Model" (MBAAM) (Nikou & Economides, 2017) and the "Teachers' Acceptance Mobile-Based Assessment" (TAMBA) (Nikou & Economides, 2019).

Therefore, given the nascent state of this field, based on the current relevance of these models, and considering that previous research has delved into factors related to assessment but not the specific characteristics of the process, it can be asserted that the development of studies and models on the adoption of mobile devices in assessment is a future research avenue.

The methodological proposal presented here aims to contribute to the literature on technological adoption models of mobile devices in assessment processes by teachers. Consequently, the following general objective is proposed:

To design a technological adoption model to study the factors influencing teachers' use of mobile devices in assessment.

This overarching objective is broken down into five specific objectives that are related to the studies conducted in the context of this thesis:

- OE1. To analyze the relationship between assessment and quality and its relevance in technology-mediated teaching and learning.
- OE2. To describe the current state of research in e-assessment in recent years, its trends, and future research directions.
- OE3. To identify the main models of technological adoption of mobile devices in assessment processes.
- OE4. To validate an initial proposal for a technological adoption model of mobile devices in assessment.
- OE5. To empirically determine the functioning of the factors included in the initial model.

This document follows the format of a Doctoral Thesis by a compendium of publications. Table 1 shows the relationship between the specific objectives and the publications derived from the Thesis (in article format) (Ortiz-López et al., 2021, 2022, 2023; In press,a; In press,b), some of which are currently in the review phase as of the date of deposit.

TABLE 1. Journal publications based on the specific objectives

S.O.	Derived publication
SO1	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C (2022). Evaluación de la calidad en e-learning en Educación Superior: Una revisión sistemática de la literatura. <i>Education in the Knowledge Society (EKS)</i> , 23. https://doi.org/10.14201/eks.26986 Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2021). Calidad en e-learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. <i>RIED. Revista Iberoamericana de Educación a Distancia</i> , 24(2). https://doi.org/10.5944/ried.24.2.29073

S.O.	Derived publication
SO2	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. <i>Education and Information Technologies</i> . https://doi.org/10.1007/s10639-023-12117-5
SO3	Ortiz-López, A., Olmos-Migueláñez, S., y Sánchez-Prieto, J. C. (Under review). Mobile-based assessment acceptance: A systematic literature review in the educational context
SO4	Ortiz-López, A.; Olmos-Migueláñez, S.; Sánchez-Prieto, J.C. (In press). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. <i>En Eds., Nuevas tendencias interdisciplinarias en Educación y Conocimiento</i> . Tirant Lo Blanch
SO5	Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (In press). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. <i>Journal of New Approaches in Educational Research</i> . Ortiz-López, A.; Sánchez-Prieto, J.C.; Olmos-Migueláñez, S. (Under review). Mobile technology in assessment processes: Comparing its usefulness and antecedents in CB-SEM

The research proposed to address the established objectives has followed three phases, including a literature review (Phase 1), the initial proposal of a model (Phase 2), and the empirical validation of the model's constructs (Phase 3).

Various methodologies and analytical techniques have been used throughout the research process, taking part in the described phases to reach the final model proposal. These include:

- **Mapping and Systematic Literature Reviews:** Mapping and systematic literature reviews are two methodologies focused on compiling published studies for analysis through an objective and structured process (García-Peñalvo, 2017). This method enables the identification, evaluation, and interpretation of the work of researchers, academics, and professionals in a chosen field (Fink, 1998).

Systematic literature reviews have followed the proposal made by Kitchenham and Charters (2007), who outline steps such as establishing research questions (1), designing inclusion and exclusion criteria (2), selecting databases (3), designing the search string (4), selecting publications based on established criteria (5), and extracting information to address research questions (6). Additionally, the reviews submitted adhere to the PRISMA flow diagram (Page et al., 2021).

- **Model Validation Process:** The two content validation processes by expert judges have followed the guidelines proposed by García-Llorente et al. (2019) for their development.

The validations have consisted of three phases. The first phase involved the individual evaluation of the items. For this purpose, experts were asked to assess, based on the recommendations of Escobar-Pérez and Cuervo (2008), their coherence, relevance, and clarity. Secondly, experts were requested to validate the constructs that make up the model. Another scale was included to measure their validity and space for relevant

qualitative considerations. Finally, the process concluded with an overall assessment of the model, evaluating the appropriateness, validity, and relevance of the proposed constructs.

- **Structural Equation Modeling (CB-SEM and PLS-SEM):** These models are data analysis methods that examine complex relationships between constructs and indicators.

The first technique, Covariance-Based Structural Equation Modeling (CB-SEM), focuses on estimating model parameters that minimize the differences between the covariance matrix of the observed sample (calculated before the analysis) and the estimated covariance matrix once the revised theoretical model has been confirmed (Hair et al., 2017).

The second technique, Partial Least Squares Structural Equation Modeling (PLS-SEM), applies partial least squares regression to minimize measurement errors of endogenous constructs by estimating coefficients (model relationships) for prediction purposes (within the sample) and maximizing the variance of constructs (Hair et al., 2021).

This thesis is part of the Ph.D. program in Education in the Knowledge Society (García-Peñalvo, 2014), an interdisciplinary doctoral program at the University of Salamanca, which guides teaching and learning processes toward the Knowledge Society in nine open research lines. Secondly, this doctoral thesis is associated with the University Institute of Educational Sciences (IUCE) at the University of Salamanca and the Research Group on Interaction and e-Learning (GRIAL) (García-Peñalvo et al., 2019)

Subsequently, the results of the research process are presented. These results comprise five studies that address the proposed objectives.

Firstly (**study 1**), a systematic review has been conducted related to models for assessing the quality of Higher Education, a vital component of the evaluation process. This review has revealed 25 models available for quality assessment, focusing on evaluating the system's quality, teaching profile, and student profile as the most described dimensions.

Additionally, the research aims to find a consensus model that combines the currently described indicators and constructs. For this reason, the proposed and validated model for e-learning quality assessment is presented. This model, validated by four evaluation and quality experts, includes four dimensions (institutional quality, teaching quality, learning system quality, and assessment quality), 14 sub-dimensions, and a total of 80 indicators.

Secondly (**study 2**), a literature mapping has been carried out, focusing on e-assessment over the last five years (2017-2021). The mapping analyzed 159 publications after applying seven inclusion criteria in its final phase. The results have shown a growing trend in e-assessment, with a predominance of studies on student satisfaction, with Learning Management Systems (LMS) being the most addressed avenue. This mapping has also identified the main characteristics of the field

and determined technological adoption and the use of mobile devices in technology-mediated education as developing research avenues.

Thirdly (**study 3**), a systematic literature review has been developed, focusing on existing technological adoption models for mobile-mediated assessment. Due to the low number of studies (15 included), the review has shown the initial state of technological acceptance studies of mobile-mediated assessments.

In these studies, the TAM model has positioned itself as the most described model, aligning with Zaineideen et al.'s (2020) assertion that it is the predominant model in education acceptance research, mainly applied to Higher Education students (Sáez-López, 2015). Additionally, it has confirmed the relevance of the MBAAM model (Nikou and Economides, 2017) as the most representative in the field. It has reflected that the validation of a model in teachers is one of the most explorable in the literature.

In the fourth place (**study 4**), a model has been proposed for determining the teachers' technological adoption of mobile devices in assessment processes. Based on the TAM model, the new proposal retains the constructs of behavioral intention to use, perceived usefulness, and perceived ease of use, eliminating the attitude toward the use construct due to its low mediating effect (Hu et al., 2003). The proposal introduces constructs used in other adoption models but not in the assessment field, such as 'subjective norm' (Venkatesh & Bala, 2008) and 'resistance to change.' Additionally, it adapts from the MBAAM model (Mobile-based assessment Acceptance Model) (Nikou & Economides, 2017) the constructs 'self-efficacy in the use of mobile devices' and 'personal innovation.' Finally, it proposes the introduction of the relative advantages of formative and summative assessment as antecedents to the intention of use, a distinction between assessment modalities not provided before in studies in this field.

The model has been validated by a group of six expert evaluators, resulting in a model composed of 11 constructs, 18 relational hypotheses, and 50 items for determining the intention to use mobile devices in assessment by teachers.

Finally (**study 5**), the designed model has been applied to student teachers enrolled in the Master's in Secondary Education and Baccalaureate, Vocational Training, and Language Teaching at the University of Salamanca, a specific training for teachers before reaching a professional position. The sample consisted of 262 students (N= 292 subjects). Although validating the model with practicing teachers would be ideal, the inability to access a large sample for its application led to the decision to use student teachers as a sample for the intermediate study, given their demonstrated similarity in previous studies (Teo, 2015).

Before applying the model, after reflecting on the study design and consulting with experts in technological adoption models, it was determined that the formulation of the constructs 'relative

advantage of summative assessment' and 'relative advantage of formative assessment' referred to perceived utilities of assessment. This realization led to formulating the constructs as 'perceived usefulness of summative assessment' and 'perceived usefulness of formative assessment' and redirecting the model towards a new specific focus on the usefulness of assessment in different modalities (formative and summative).

From this study, two independent analyses have emerged: a double-comparative analysis to validate the functioning of the newly perceived usefulness of mobile devices for assessment (with the introduction of the constructs autonomy and self-efficacy) and a second comparative analysis of three technological adoption models with different formulations of perceived usefulness in each of them (adapted to different teaching conceptions of formative and summative assessments), with self-efficacy in mobile device use as an antecedent.

These analyses employed the structural equation modeling technique based on covariance (CB-SEM) and partial least squares structural equation modeling (PLS-SEM), respectively. Performing an analysis using PLS-SEM is justified by a formative construct in a proposed model, an analysis that is only possible through PLS-SEM (Hair et al., 2021).

The first analysis (CB-SEM) presents two research models, the TAM model and a model with a construct of the perceived usefulness of mobile devices in assessment. After the study, it was confirmed that a usefulness explicitly formulated for the assessment field performs better than a model that maintains a generic usefulness based on the TAM model, more parsimonious and generalizable. Secondly, the study has confirmed the functioning of autonomy and self-efficacy in using mobile devices for assessment as conditioning factors for acceptance, in line with the research of Castiblanco-Jiménez et al. (2020).

The second comparative analysis (PLS-SEM) studies three models framed in different assessment trends described in the introduction and complemented with the construct 'self-efficacy in the use of mobile devices' as an antecedent to perceived usefulness. This analysis has shown that the two newly proposed models that reform perceived usefulness into specific utilities (model 1) or a single construct of mobile device usefulness in assessment (model 2) are more predictive than the original TAM model.

Therefore, in line with the specificity made for perceived utilities of assessment, the proposal is directed towards two versions of a model adapted to each assessment modality (formative/summative). To achieve this, the 'behavioral intention to use' has been reformulated for the final proposal to adapt to formative and summative assessments, presented below."

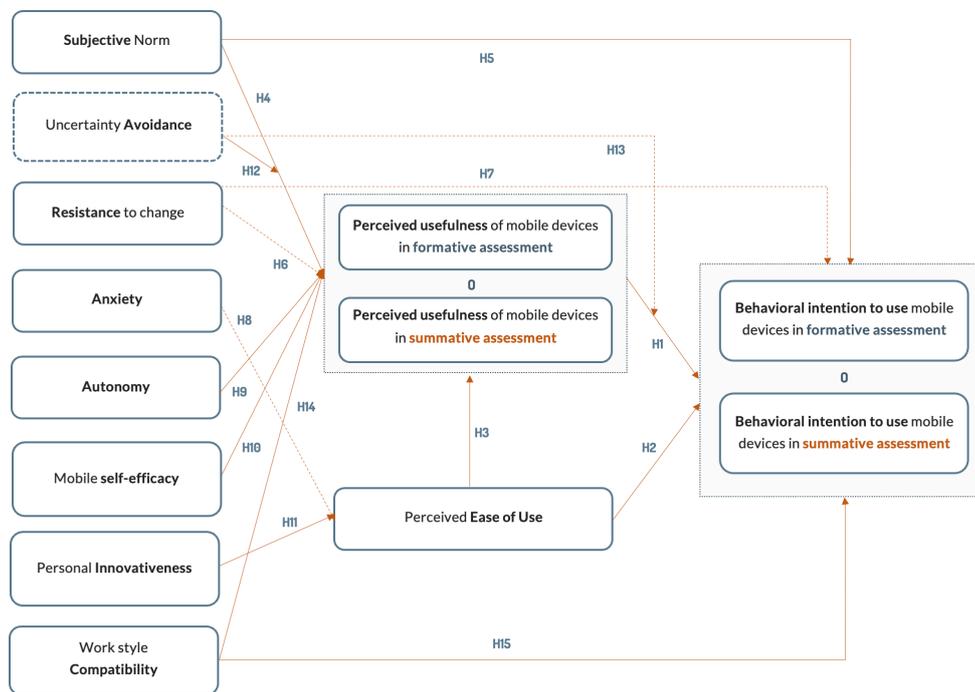
As a result of this process, the final proposal of the F/S-MAAM model (Formative/Summative – Mobile-based Assessment Acceptance Model) has been developed. The first conclusion after conducting the interim study is the confirmation that the usefulness proposed by Davis (1989) does

not fit into a specific model for assessment, explaining a lower percentage of variance than a model with a specific usefulness. Furthermore, it is confirmed that designing two distinct utilities explains, in turn, a higher rate of variance than a single construct of assessment usefulness; therefore, the design should focus on proposing two separate models.

These findings are also reaffirmed when analyzing the fit indicators of the previously evaluated models, resulting in the new proposals being more parsimonious and generalizable. Therefore, a final model is proposed with two versions that differ in perceived usefulness and behavioral intention to use: one for adopting mobile devices in formative assessment processes and another for summative assessment processes.

The constructs that make up the final model proposal (Figure 2) are the intention to use mobile devices in formative (BIUF) or summative (BIUS) assessments, perceived usefulness of mobile devices for formative (PUFA) or summative (PUSA) assessment, perceived ease of use (PEOU), subjective norm (SN), resistance to change (RC), assessment anxiety (ASAN), autonomy in the use of mobile devices for assessment (AUT), self-efficacy in the use of mobile devices (MSE), personal innovation (PI), rejection of uncertainty (UBR), and compatibility with work style (COMP).

FIGURE 2. F/S-MAAM model (Formative/Summative – Mobile-based Assessment Acceptance Model)



The described model (and its versions) is justified by 15 relational hypotheses for the 11 proposed constructs. These hypotheses vary in each model, adapting to formative and summative assessment, and are as follows:

- F/S-MAAM.H1. The perceived usefulness of mobile devices for assessment (formative or summative) positively predicts the teachers' intention to use mobile devices in formative or summative assessments in their teaching practice.
- F/S-MAAM.H2. The perceived ease of use of mobile devices for assessment positively predicts the teachers' intention to use them in formative or summative assessments in their teaching practice.
- F/S-MAAM.H3. The perceived ease of use positively predicts the perceived usefulness of mobile devices for formative or summative assessments in teaching practice.
- F/S-MAAM.H4. The subjective norm has a positive relationship with the perceived usefulness of mobile devices for formative or summative assessments by teachers.
- F/S-MAAM.H5. The subjective norm positively correlates with the teachers' intention to use mobile technologies in assessment processes.
- F/S-MAAM.H6. Resistance to change negatively predicts the perceived usefulness (formative or summative) of mobile devices for teacher assessment.
- F/S-MAAM.H7. Resistance to change negatively predicts the teachers' intention to use mobile devices in formative or summative assessment processes.
- F/S-MAAM.H8. Anxiety about assessment will have a negative effect on the perceived ease of use by teachers of mobile devices in assessment.
- F/S-MAAM.H9. Autonomy will have a positive effect on the perceived usefulness (formative or summative) by teachers in using mobile devices for assessment.
- F/S-MAAM.H10. Mobile self-efficacy in the use of mobile devices will have a positive effect on the perceived usefulness (formative or summative) by teachers.
- F/S-MAAM.H11. Personal innovativeness will have a positive effect on the perceived ease of use by teachers in using mobile devices for assessment.
- F/S-MAAM.H12. Uncertainty avoidance positively moderates the predictive relationship between subjective norm and the perceived usefulness (formative or summative) of mobile devices for assessment.
- F/S-MAAM.H13. Uncertainty avoidance negatively moderates the relationship between the perceived usefulness (formative or summative) of mobile devices for assessment and their behavioral intention to use (formative or summative).
- F/S-MAAM.H14. Perceived compatibility positively influences the perceived usefulness of mobile devices in formative or summative assessments.

F/S-MAAM.H15. Perceived compatibility is positively related to the intention to use mobile devices in formative/summative assessments.

In conclusion, the research has successfully addressed the objectives outlined in its initial proposal. Firstly, it has developed an approach to quality assessment and its intrinsic relationship.

The review identified twelve distinct dimensions influencing quality, with nine overlapping with a previous review (Marciniak & Gairín-Sallán, 2018). Conversely, five dimensions did not align, emphasizing the breadth of quality components and the challenge in their delineation (Fernández, 2005).

Additionally, a model for quality assessment in Higher Education was proposed and validated, enabling a thorough analysis of the main components of quality. The final model, validated by experts in quality and assessment, consists of four dimensions, 14 sub-dimensions, and 80 indicators for evaluating quality in technology-mediated education in Higher Education.

Secondly, the research described the current state of e-assessment through systematic literature mapping. The results indicate a high level of interest in the scientific and research community regarding assessment in digital environments, evidenced by the growing amount of publications confirming the benefits of technological inclusion in assessment processes (Huda et al., 2020). Notably, a significant representation of research focused on student satisfaction in Higher Education with the assessment of Learning Management System platforms, with studies on technological adoption among faculty emerging as an exploratory avenue in the literature.

Thirdly, this doctoral thesis aimed to identify critical models of technological adoption related to mobile-mediated assessment. In this regard, the finalized review contrasts with the previous one by Alrofou et al. (2019), who selected eight models exclusively based on TAM (Davis, 1989) focused on mobile-mediated assessment, all included in this study. Therefore, the present research delved into many adoption models and repositories to reflect almost all publications on the topic.

The systematic review to address the third specific objective highlighted the initial state of studies on the technological adoption of mobile devices in assessment processes (15 studies). The conclusion asserts that this field is in its early stages, keeping future development avenues open to support recent technological and educational changes and advancements in adapting assessment processes to new educational scenarios. Furthermore, understanding the factors influencing teachers' acceptance of these devices is fundamental for effectively implementing technology in teaching and learning processes.

Subsequently, the fourth specific objective proposed to design a new model for the technological acceptance of mobile devices in assessment processes. These devices have seen extensive development in methodological approaches in the classroom (Williams-Muller et al.,

2020), but have been relatively scarce in assessment processes. This objective is justified by proposing a theoretical model validated by experts and two derived models validated after application to future secondary education teachers (fifth objective).

Given the linear format of the research, this model (F/S-MAAM) has been reformulated into two differentiated versions as the final proposal of the research. The evaluation of usage intention, as analyzed in the models presented, has highlighted that the effective incorporation of mobile devices in teaching and assessment requires teachers to perceive their usefulness, consider them effective in their handling, and recognize their efficacy in educational tasks, differentiating between formative and summative assessment processes. This realization directly connects with the need to reconfigure initial and continuous training. Therefore, the contributions of the proposed adoption model not only enrich the research field, but also hold crucial importance in the educational and formative realm for in-service and future education professionals by comprehending the factors influencing the final acceptance of technology and the intention to use it in educational assessment processes.

In this context, the research results offer specific dimensions that condition technological adoption, an essential element for the fundamental training of educational professionals. Understanding how teachers perceive and adopt these technologies, their barriers, and how to overcome them will provide educational institutions with the necessary strategies to adapt and improve their initial training programs. This will effectively address the changing needs in technology-mediated education.

The importance and implications of the proposed model extend beyond the described initial training. Continuous professional development for active teachers is crucial in designing specific programs and actions seeking the acceptance of technological innovations and new pedagogical practices. By considering the factors conditioning the adoption of mobile devices in assessment, a helpful framework will be provided for educators' ongoing training, which will facilitate the integration of technology into their daily practices.

The future research directions derived from the Doctoral Thesis are presented in two categories: those research lines already initiated and future lines of exploration. Among the already initiated research lines, the application of the F/S-MAAM model has been carried out in the degrees of Early Childhood Education, Primary Education, and the Double Degree in Early Childhood Education and Primary Education at the University of Salamanca on its three campuses (Ávila, Salamanca, Zamora) in a study currently in the phase of results analysis. On the other hand, a second open line for the study of adopting and validating F-MAAM and S-MAAM versions in different Portuguese universities and polytechnics is in the data collection phase (a phase initiated during the international research stay).

Finally, future research lines include the application of the proposed quality assessment model, continuity in modeling the perceived usefulness of assessment in adoption studies or modeling perceived ease of use (given discrepancies in this research with hypotheses regarding perceived ease of use in the original TAM model application). Furthermore, future avenues of research include the application of the F/S-MAAM model to in-service teachers, designing two models with differentiated constructs adapted to the peculiarities of each assessment modality (formative, F-MAAM, and summative, S-MAAM), expanding the antecedents of these last two models, and opening a joint international research line with relevant researchers, aiming at expanding adoption studies, and advancing mobile-mediated assessment.

References

- Abdullah, F., & Ward, R. (2016). Developing a General Extended Technology Acceptance Model for E-Learning (GETAMEL) by analysing commonly used external factors. *Computers in Human Behavior*, 56, 238-256. <https://doi.org/10.1016/j.chb.2015.11.036>
- Ahmad, B., & Bhat, G. J. (2019). Formative and Summative Evaluation Techniques for Improvement of Learning Process. *European Journal of Business & Social Sciences*, 7(5), 776-785.
- Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl & J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11-39). Springer. https://doi.org/10.1007/978-3-642-69746-3_2
- Alrofou, A., Lakulu, M. M., & Almaiah, M. (2019). A Systematic Review of Mobile-Based Assessment Acceptance Studies From 2009 to 2019. *Journal of Theoretical and Applied Information Technology*, 97(20), 1-25.
- Alsharida, R., Hammood, M., & Al-Emran, M. (2021). Mobile Learning Adoption: A Systematic Review of the Technology Acceptance Model from 2017 to 2020. *International Journal of Emerging Technologies in Learning (IJET)*, 16(5), 147-162.
- Álvarez-Herrero, J. F. (2018). Aprendizaje móvil fuera del aula en Educación Secundaria. In E. López-Meneses, D. Cobos-Sanchiz, A.H. Martín-Padilla, L. Molina-García y A. Jaén-Martínez (Eds.), *Experiencias pedagógicas e innovación educativa. Aportaciones desde la praxis docente e investigadora* (pp. 937-943). Ediciones Octaedro.
- Anderson, L. W. (2005). Objectives, evaluation, and the improvement of education. *Studies in Educational Evaluation*, 31(2), 102-113. <https://doi.org/10.1016/j.stueduc.2005.05.004>
- Barrientos-Hernán, E. J., López-Pastor, V. M., & Pérez-Brunicardi, D. (2020). Evaluación Auténtica y Evaluación Orientada al Aprendizaje en Educación Superior. Una Revisión en Bases de Datos Internacionales. *Revista Iberoamericana de Evaluación Educativa*, 13(2), 67-83.
- Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85(5), 536-553. <https://doi.org/10.1002/sce.1022.abs>
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Formative and summative assessment: Can they serve learning together?* Annual meeting of the American Educational Research Association, Chicago.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>

- Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400-413. <https://doi.org/10.1080/02602938.2015.1018133>
- Buchholtz, N. F., Krosanke, N., Orschulik, A. B., & Vorhölter, K. (2018). Combining and integrating formative and summative assessment in mathematics teacher education. *ZDM*, 50(4), 715-728. <https://doi.org/10.1007/s11858-018-0948-y>
- Castiblanco-Jiménez, I. A., Cepeda-García, L. C., Grazia-Violante, M., Marcolin, F., & Vezzetti, E. (2020). Commonly Used External TAM Variables in e-Learning, Agriculture and Virtual Reality Applications, *Future Internet*, 13(1), 1-21.
- Connolly, M., Jones, N., & O'shea, J. (2005). Quality assurance and e-learning: Reflections from the front line. *Quality in Higher Education*, 11(1), 59-67. <https://doi.org/10.1080/13538320500077660>
- Córica, J. L., & García-Aretio, L. G. (2018). Teachers' reluctance to ICT in Argentina. A qualitative approach. *Educación Superior*, 25.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: Three experiments. *International Journal of Human-Computer Studies*, 45(1), 19-45. <https://doi.org/10.1006/ijhc.1996.0040>
- De la Orden, A. (1982). *La evaluación educativa*. Docencia.
- Dixson, D. D., & Worrell, F. C. (2016). Formative and Summative Assessment in the Classroom. *Theory Into Practice*, 55(2), 153-159. <https://doi.org/10.1080/00405841.2016.1148989>
- Dolin, J., Black, P., Harlen, W., & Tiberghien, A. (2018). Exploring Relations Between Formative and Summative Assessment. In J. Dolin and R. Evans (Eds.), *Transforming Assessment: Through an Interplay Between Practice, Research and Policy* (pp. 53-80). Springer International Publishing. https://doi.org/10.1007/978-3-319-63248-3_3
- Elkheir, Z., & Abdul-Mutalib, A. (2015). Mobile Learning Applications Designing Concepts and Challenges: Survey. *Research Journal of Applied Sciences, Engineering and Technology*, 10, 438-442. <https://doi.org/10.19026/rjaset.10.2509>
- Escobar-Pérez, J., & Cuervo, A. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 6, 27-36.
- Fernández, V. (2005). Calidad y e-Learning. In F. Martínez (Coord.), *E-aprendizaje en bibliotecología. Perspectivas globales*. Universidad Nacional Autónoma de México.
- Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (2015). *Empowerment Evaluation: Knowledge and Tools for Self-Assessment, Evaluation Capacity Building, and Accountability*. SAGE Publications, Inc. <https://doi.org/10.4135/9781483387079>
- Fetterman, D. M., & Wandersman, A. (2004). *Empowerment Evaluation Principles in Practice*. Guilford Press.
- Fink, A. (1998). *Conducting literature research reviews: From paper to the Internet*. Sage Publications.
- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention And Behavior: An introduction to theory and research*. Addison-Wesley.
- Fombona-Cadavieco, J., & Rodil-Pérez, F. J. (2018). Niveles de uso y aceptación de los dispositivos móviles en el aula. *Pixel-Bit. Revista de Medios y Educación*, 52. <https://doi.org/10.12795/pixelbit.2018.i52.02>
- Gabriel-Pedro, L. F., de Oliveira-Barbosa, C. M. M., & Neves-Santos, C. M. das. (2018). A critical review of mobile learning integration in formal educational contexts. *International Journal of Educational Technology in Higher Education*, 15(1), 10. <https://doi.org/10.1186/s41239-018-0091-4>

- García-Llorente, H. J., Martínez-Abad, F., & Rodríguez-Conde, M. J. (2019). Validación de un instrumento de evaluación de competencias informacionales autopercibidas en educación secundaria obligatoria. *Anales de Documentación*, 22(1). <https://doi.org/10.6018/analesdoc.22.1.305641>
- García-Peñalvo, F. J. (2017). *Revisión sistemática de literatura en los Trabajos de Final de Máster y en las Tesis Doctorales*. <https://knowledgesociety.usal.es/sites/default/files/20170316%20-%20Seminaro%20SLR.pdf>
- García-Peñalvo, F. J. (2021). Transformación digital en las universidades: Implicaciones de la pandemia de la COVID-19. *Education in the knowledge society: EKS*, 22, 1-6. <https://doi.org/10.14201/eks.25465>
- García-Peñalvo, F. J. (2014). Formación en la sociedad del conocimiento, un programa de doctorado con una perspectiva interdisciplinar. *Education in the Knowledge Society (EKS)*, 15(1). <https://doi.org/10.14201/eks.11641>
- García-Peñalvo, F. J., Rivero-Ortega, R., Rodríguez-Conde, M. J., & Rodríguez-García, N. (2020). The institutional decisions to support remote learning and teaching during the COVID-19 pandemic. *2020 X International Conference on Virtual Campus (JICV)*, 1-5. <https://doi.org/10.1109/JICV51605.2020.9375683>
- García-Peñalvo, F. J., Rodríguez-Conde, M. J., Therón, R., García-Holgado, A., Martínez-Abad, F., & Benito, A. (2019). Grupo GRIAL. *IE Comunicaciones: Revista Iberoamericana de Informática Educativa*, 30, 33-48.
- García-Ramos, J. M. (1989). *Bases pedagógicas de la evaluación. Guía práctica para educadores*. Editorial Síntesis.
- Ghouali, K., Benmoussat, S., & Ruiz-Cecilia, R. (2020). E-assessment on the spotlight: Present and future prospects. *ReiDoCrea: Revista electrónica de investigación Docencia Creativa*, 9, 52-62. <https://doi.org/10.30827/Digibug.59151>
- González-Monteaquedo, J. (2001). John Dewey y la pedagogía progresista. In J. Trilla (Coord.), *El legado pedagógico del siglo XX para la escuela del siglo XXI*. Grao.
- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52(3), 67-86. <https://doi.org/10.1007/BF02504676>
- Hair, J.F., Babin, B., & Krey, N. (2017). Covariance-Based Structural Equation Modeling in the Journal of Advertising: Review and Recommendations. *Journal of Advertising*, 46(3), 454-454. <https://doi.org/10.1080/00913367.2017.1329496>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-80519-7>
- Harlen, W., & James, M. (1997). Assessment and Learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365-379. <https://doi.org/10.1080/0969594970040304>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Hernández, P. (1998). *Diseñar y Enseñar*. Narcea.
- Hernández-García, Á. (2012). Desarrollo de un modelo unificado de adopción del comercio electrónico entre empresas y consumidores finales. Aplicación al mercado español [Tesis Doctoral, E.T.S.I. Telecomunicación (UPM)]. <https://oa.upm.es/10281/>
- Hernández-Sellés, N., Muñoz-Carril, P. C., & Sanmamed, M. G. (2018). La e-evaluación en el trabajo colaborativo en entornos virtuales: Análisis de la percepción de los estudiantes. *EduTec. Revista Electrónica de Tecnología Educativa*, 65. <https://doi.org/10.21556/edutec.2018.65.997>

- Hu, P. J.-H., Clark, T. H. K., & Ma, W. W. (2003). Examining technology acceptance by school teachers: A longitudinal study. *Information & Management*, 41(2), 227-241. [https://doi.org/10.1016/S0378-7206\(03\)00050-8](https://doi.org/10.1016/S0378-7206(03)00050-8)
- Huda, S. S. M., Kabir, M., & Siddiq, T. (2020). E-Assessment in Higher Education: Students' Perspective. *International Journal of Education and Development Using Information and Communication Technology*, 16(2), 250-258.
- Ibarra-Sáiz, M., & Rodríguez-Gómez, G. (2019). Una evaluación como aprendizaje. In J. Paricio, A. Fernández y I. Fernández (Eds.), *Cartografía de la buena docencia universitaria. Un marco para el desarrollo del profesorado basado en la investigación* (pp. 175-196). Narcea.
- Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2020). Aprendiendo a Evaluar para Aprender en la Educación Superior. *Revista Iberoamericana de Evaluación Educativa*, 13(1), Article 1.
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., Boud, D., Rotsaert, T., Brown, S., Salinas-Salazar, M. L., & Rodríguez-Gómez, H. M. (2020). El futuro de la evaluación en la Educación Superior. *Revista Electrónica de Investigación y Evaluación Educativa*. <https://doi.org/10.7203/relieve.26.1.17323>
- James, M. (2006). Assessment, Teaching and Theories of Learning. In J. Gardner (Ed.), *Assessment and Learning* (pp. 47-60). Sage Publications. <https://doi.org/10.13140/2.1.5090.8960>
- Jiménez, Y. I., Hernández, J., & Rodríguez, E. (2021). Educación en línea y evaluación del aprendizaje: De lo presencial a lo virtual. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 12(23). <https://doi.org/10.23913/ride.v12i23.1005>
- Joint Committee on Standards for educational evaluation. (1988). *Normas de evaluación para programas, proyectos y material educativo* (Versión original en inglés: Standards for evaluation of educational programs, projects and materials). Editorial Trillas.
- Kibble, J. D. (2017). Best practices in summative assessment. *Advances in Physiology Education*, 41(1), 110-119. <https://doi.org/10.1152/advan.00116.2016>
- King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, 43(6), 740-755. <https://doi.org/10.1016/j.im.2006.05.003>
- Kitchenham, B., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering (EBSE Technical Report Version 2.3; pp. 1-57). Keele University.
- Knight, P. T. (2002). Summative Assessment in Higher Education: Practices in disarray. *Studies in Higher Education*, 27(3), 275-286. <https://doi.org/10.1080/03075070220000662>
- Koslowski, F. A. (2006). Quality and assessment in context: A brief review. *Quality Assurance in Education*, 14(3), 277-288. <https://doi.org/10.1108/09684880610678586>
- MacLellan, E. (2001). Assessment for Learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education*, 26(4), 307-318. <https://doi.org/10.1080/02602930120063466>
- Marciniak, R., & Gairín-Sallán, J. (2018). Dimensiones de evaluación de calidad de educación virtual: Revisión de modelos referentes. *RIED. Revista Iberoamericana de Educación a Distancia*, 21(1). <https://doi.org/10.5944/ried.21.1.16182>
- Mejía-Pérez, O. (2012). De la evaluación tradicional a una nueva evaluación basada en competencias. *Educare*, 16(1). <https://doi.org/10.15359/ree.16-1.3>
- Miras, M., & Solé, I. (1990). La evaluación del aprendizaje y la evaluación del proceso de enseñanza - aprendizaje. In C. Coll, J. Palacios, y A. Marchesi (Eds.), *Desarrollo psicológico y educación, II Psicología de la Educación* (pp. 419-434). Alianza.
- Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education:

- A systematic review. *Review of Education*, 9(3), e3292. <https://doi.org/10.1002/rev3.3292>
- Nikou, S., & Economides, A. (2017). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 109, 56-73. <https://doi.org/10.1016/j.compedu.2017.02.005>
- Nikou, S., & Economides, A. (2019). Factors that influence behavioral intention to use mobile-based assessment: A STEM teachers' perspective. *British Journal of Educational Technology*, 50(2). <https://doi.org/10.1111/bjet.12609>
- Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2021). Calidad en e-Learning: Identificación de sus dimensiones, propuesta y validación de un modelo para su evaluación en Educación Superior. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2). <https://doi.org/10.5944/ried.24.2.29073>
- Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2022). Evaluación de la calidad en e-Learning en Educación Superior: Una revisión sistemática de la literatura. *Education in the Knowledge Society (EKS)*, 23. <https://doi.org/10.14201/eks.26986>
- Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (2023). Toward a new educational reality: A mapping review of the role of e-assessment in the new digital context. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12117-5>
- Ortiz-López, A., Olmos-Migueláñez, S., & Sánchez-Prieto, J. C. (In press). ¿Un móvil para evaluar en el aula? Diseño y validación de un modelo específico de aceptación tecnológica. En *Nuevas tendencias interdisciplinarias en Educación y Conocimiento*. Tirant Lo Blanch.
- Ortiz-López, A., Sánchez-Prieto, J. C., & Olmos-Migueláñez, S. (In press). Perceived Usefulness of Mobile Devices in Assessment: A Comparative Study of Three Technology Acceptance Models Using PLS-SEM. *Journal of New Approaches in Educational Research*.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(71). <https://doi.org/10.1136/bmj.n71>
- Patton, M. Q. (1996). A World Larger than Formative and Summative. *Evaluation Practice*, 17(2), 131-144. <https://doi.org/10.1177/109821409601700205>
- Pearse-Romera, C. R., & Ruiz-Cecilia, R. (2019). A Motivation Case Study of Students Learning English at a Secondary School in Granada, Spain. *The International Journal of Interdisciplinary Educational Studies*, 14(1), 31-45. <https://doi.org/10.18848/2327-011X/CGP/v14i01/31-45>
- Pérez-Juste, R., & García-Ramos, J. M. (1989). *Diagnóstico, evaluación y toma de decisiones*. Rialp.
- Perrenoud, P. (1991). *La evaluación de los alumnos: De la producción de la excelencia a la regulación de los aprendizajes*. Alternativa Pedagógica.
- Posner, G. (1998). *Análisis del curriculum*. McGraw Hill.
- Rodríguez-Gómez, G., Ibarra-Saiz, M., & García-Jiménez, E. (2013). Autoevaluación, evaluación entre iguales y coevaluación: Conceptualización y práctica en las universidades españolas. *Revista de Investigación en Educación*, 2(11), 198-210.
- Rodríguez-Neira, T., Álvarez-Pérez, L., Cadrecha y Caparrós, M. Á., Hernández-García, J., Luengo-García, M. A., Muñoz-Fernández, J., Ordóñez-Alvarez, J. J., & Soler-Vázquez, E. (2001). *La evaluación en el aula*. Ediciones Nobel.
- Sacristán-Díaz, M., Garrido-Vega, P., Alfalla-Luque, R., & González-Zamora, M. del M. (2010). De la evaluación sumativa a la formativa a través de las plataformas de enseñanza virtual. In R. del Pozo (Coord.), *Nuevas formas de docencia en el área económico-empresarial* (pp. 151-167). Edición Digital Atres.

- Sáez-López, J. M. (2015). Actitudes de los docentes respecto a las tic, a partir del desarrollo de una práctica reflexiva. *Escuela Abierta*, 13, 37-54.
- Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Scriven, M. (1967). The Methodology of Evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of Curriculum Evaluation* (pp. 39-83). Rand McNally.
- Shepard, L. (2006). La evaluación en el aula. In R. Brennan (Ed.), *Educational Measurement* (4 Edition, pp. 623-646). Praeger Westport.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Stufflebeam, D., & Shinkfield, A. (1995). *Evaluación sistemática: Guía teórica y práctica*. Centro de Publicaciones del Ministerio de Educación y Ciencia, Ediciones Paidós.
- Taherdoost, H. (2019). Importance of Technology Acceptance Assessment for Successful Implementation and Development of New Technologies, *Global Journal of Engineering Sciences*, 1. <https://doi.org/10.33552/GJES.2019.01.000511>
- Taylor, S., & Todd, P. (1995a). Assessing IT Usage: The Role of Prior Experience. *MIS Quarterly*, 19(4), 561-570. <https://doi.org/10.2307/249633>
- Taylor, S., & Todd, P. (1995b). Understanding information technology usage: A test of competing models. *Information Systems Research*, 6(2), 144-176.
- Tejada, J. (1997). La evaluación. In J. Gairin y A. Ferrández (Eds.), *Planificación y gestión de instituciones de formación* (pp. 243-268). Praxis.
- Tomas, C., Borg, M., & McNeil, J. (2015). E-assessment: Institutional development strategies and the assessment life cycle. *British Journal of Educational Technology*, 46(3), 588-596. <https://doi.org/10.1111/bjet.12153>
- Tyler, R. (1949). *Basic principle of curriculum and instruction*. Chicago University.
- Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39(2), 273-315. <https://doi.org/10.1111/j.1540-5915.2008.00192.x>
- Venkatesh, V., & Davis, F. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46, 186-204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Williams-Muller, E., Castro-Galviz, M. Y., Cruz-López, M. D. la, Pinillas, L. N., García, L. R., & Carvalho, J. L. T. (2020). Uso de los teléfonos móviles en el aula de educación primaria. *Revista Electrónica de Conocimientos, Saberes y Prácticas*, 3(2). <https://doi.org/10.5377/recsp.v3i2.10688>
- Zaineldeen, S., Li, H., Koffi, A., & Mohammed, B. (2020). Technology Acceptance Model' Concepts, Contribution, Limitation, and Adoption in Education. *Universal Journal of Educational Research*, 8(11), 5061-5071. <https://doi.org/10.13189/ujer.2020.081106>



PROGRAMA DE DOCTORADO: FORMACIÓN EN LA SOCIEDAD DEL CONOCIMIENTO