



Validación y prueba piloto de herramientas para medir la integración del celular desde la cognición extendida

Validation and pilot testing of tools to measure the integration of mobile phones in extended cognition

Fecha de recepción:
01 de abril 2025

Fecha de aprobación:
14 de octubre 2025



<https://hdl.handle.net/20.500.14492/32611>

Luis Fernando Castañeda Fuentes

México

Universidad Nacional Rosario Castellanos

castaneda fuentesluis@gmail.com

 <https://orcid.org/0000-0002-1611-5334>

Karen Fernanda Amezcua Kosterlitz

México

Universidad Nacional Autónoma de México

karen.amezcua@gmail.com

 <https://orcid.org/0000-0003-0299-200X>

Resumen

Este estudio se centró en el desarrollo, validación y prueba piloto de dos instrumentos para evaluar la integración del teléfono celular como extensión de la cognición en estudiantes y docentes de secundaria en la Ciudad de México. El objetivo principal de este artículo fue presentar el proceso de desarrollo y validación de estos instrumentos y analizar la relación preliminar entre el uso del teléfono celular en el contexto escolar, la extensión de la cognición y el rendimiento académico, a través de la aplicación de los instrumentos desarrollados, en el marco de una investigación doctoral más amplia que evalúa el efecto de la restricción de su uso. La metodología empleada fue cuantitativa, utilizando cuestionarios autoadministrados que fueron validados a través del juicio de expertos, el coeficiente Alfa de Cronbach y el coeficiente V de Aiken, lo que garantizó su confiabilidad. Tras la prueba piloto, los resultados indicaron que la validación por expertos, junto con los coeficientes de Alfa de Cronbach y de constructo, permitió realizar ajustes importantes en los instrumentos para mejorar su claridad, suficiencia, coherencia y relevancia. Se concluye que los instrumentos son apropiados y aplicables en el contexto educativo de secundaria, proporcionando una herramienta válida para investigar el uso del teléfono celular como extensión cognitiva, como recurso pedagógico y su impacto en el rendimiento académico de los estudiantes.

Palabras clave: aprendizaje, cuestionario, educación básica, proceso de aprendizaje, rendimiento escolar, teléfono móvil.

Abstract

This study focused on the development, validation, and pilot testing of two instruments designed to assess the integration of the cell phone as an extension of cognition in middle school students and teachers in Mexico City. The main objective of this article was to present the development and validation process of these instruments and to analyze the preliminary relationship between cell phone use in the school context, cognitive extension, and academic performance, through the application of the developed tools, within the framework of a broader doctoral research project that evaluates the effect of restricting its use. The methodology employed was quantitative, using self-administered questionnaires validated through expert judgment, Cronbach's Alpha coefficient, and Aiken's V coefficient, which ensured their reliability. Following the pilot test, the results indicated that validation by experts, along with Cronbach's Alpha and construct coefficients, allowed for important adjustments to the instruments to improve their sufficiency, clarity, coherence, and relevance. It is concluded that the instruments are appropriate and applicable in the secondary education context, providing a valid tool to investigate the use of the cell phone as a cognitive extension, as a pedagogical resource, and its impact on students' academic performance.

Keywords: basic education, learning, learning process, mobile phone, questionnaire, school performance.

1. Introducción

El avance tecnológico ha experimentado un crecimiento vertiginoso en las últimas décadas, transformando de manera profunda la manera en que las personas se relacionan entre sí y con su entorno. En este contexto, los teléfonos móviles han adquirido un papel central en la vida cotidiana, ya que facilitan la comunicación, el acceso a la información y la realización de diversas actividades en diferentes ámbitos, incluyendo el educativo. Estos dispositivos proporcionan recursos que potencian la productividad, el aprendizaje y la comunicación social, lo que ha generado un debate en torno a su influencia en el ámbito educativo (Astorga Aguilar et al., 2020).

Este estudio se basa en la teoría de la cognición extendida, propuesta por Clark y Chalmers (1998), la cual sostiene que la mente humana no se limita al cerebro, sino que se extiende al entorno a través del uso de herramientas y dispositivos externos. Desde esta perspectiva, los teléfonos móviles no solo facilitan los procesos cognitivos, sino que también pueden integrarse activamente en funciones como la memoria, la atención y la resolución de problemas, favoreciendo así el procesamiento y la adquisición de conocimientos. En este marco, el entorno juega un papel fundamental en la configuración de la cognición, permitiendo que el celular opere como una extensión del sistema cognitivo del estudiante (Clark y Chalmers, 1998).

Diversos estudios han analizado la incorporación de los teléfonos móviles en la educación secundaria, resaltando tanto sus ventajas como los retos asociados a su uso. Por ejemplo, una investigación realizada por Nikolopoulou (2020) en Grecia, examinó la percepción docente respecto al empleo de estos dispositivos en el aula, evidenciando que, aunque pueden mejorar los procesos de enseñanza-aprendizaje, su implementación enfrenta obstáculos como la insuficiente infraestructura y la falta de normativas claras. De manera similar, el estudio de Martínez Domínguez (2020) identificó la

desigualdad en el acceso a la conectividad como una limitante para el uso pedagógico de los teléfonos móviles en el contexto educativo mexicano.

Asimismo, un estudio llevado a cabo en Colombia, Hernández Silva (2017) identificó que los estudiantes recurren al uso de sus teléfonos en «espacios de fuga», es decir, en momentos y lugares donde pueden evadir las restricciones impuestas por las instituciones educativas en torno a estos dispositivos. En la misma línea, Matamala Riquelme (2016) analizó los efectos de las tecnologías digitales en el aprendizaje en España, concluyendo que, aunque facilitan la gestión de la información, también pueden promover la dispersión de la atención y una comprensión superficial del conocimiento.

En el contexto de la educación superior, Silva Calpa y Martínez Delgado (2017) reconocieron la dualidad inherente de los dispositivos móviles, al postular su potencial como instrumentos tecnopedagógicos, a la vez que advirtieron sobre su capacidad para interrumpir el aprendizaje si no se integran adecuadamente en el proceso didáctico. En un estudio con adolescentes en los Países Bajos, Nijssen et al. (2018) indagaron la externalización de procesos cognitivos y sociales facilitada por los teléfonos celulares, sugiriendo que, bajo un uso controlado y estratégico, estos dispositivos pueden funcionar como extensiones de la cognición humana.

A partir de estas investigaciones previas, este estudio tiene como objetivo diseñar y validar dos instrumentos que permitan evaluar el impacto del uso del teléfono celular en los procesos cognitivos de los estudiantes de secundaria en la Ciudad de México y el uso que hacen en el profesorado del teléfono celular como herramienta pedagógica. Para ello, se proponen dos herramientas que garanticen validez y confiabilidad en la medición de la relación entre el uso de estos dispositivos en el aula y su incidencia en el aprendizaje y la participación estudiantil. La consistencia interna de los instrumentos se verificó mediante el coeficiente Alfa de Cronbach, mientras que su validez fue confirmada a través de la opinión de un grupo de expertos en educación y psicometría.

Este trabajo pretende aportar al campo de la cognición extendida y la tecnología educativa a través del diseño de un instrumento de medición que genere evidencia empírica sobre el papel del teléfono celular como extensión cognitiva en el ámbito escolar. Se espera que los resultados obtenidos proporcionen información valiosa para el diseño de estrategias pedagógicas que aprovechen las ventajas de estos dispositivos en el aprendizaje de los estudiantes.

2. Metodología

Se realizó un estudio cuantitativo con el propósito de evaluar la validez de contenido y de constructo de dos instrumentos diseñados para medir el nivel de integración del teléfono celular como extensión de la cognición en estudiantes de secundaria en la Ciudad de México. El objetivo de este artículo fue detallar el proceso de desarrollar y validar dos instrumentos dirigidos al estudiantado y plantilla docente de educación secundaria, con el fin de realizar un análisis descriptivo y correlacional inicial, en el marco del objetivo general de la investigación más amplia que busca analizar la relación entre el uso del teléfono celular en el contexto escolar, la extensión de la cognición y el rendimiento académico, mediante la aplicación de dichos instrumentos.

En el diseño de investigación se utilizó el cuestionario, aplicado a una muestra de tres expertos. Para la recopilación de datos, se diseñaron dos cuestionarios autoadministrados: uno dirigido al estudiantado y otro al cuerpo docente. Ambos cuestionarios fueron desarrollados específicamente para este estudio, siguiendo un proceso de validación por juicio de expertos, en el que se calculó el coeficiente de concordancia y la consistencia interna. La validez de contenido se determinó a través de un panel de expertos (en el cuestionario del estudiantado), lo cual es especialmente relevante cuando existen limitaciones en las observaciones experimentales.

El juicio de expertos se refiere a la evaluación de personas con experiencia y reconocimiento en el tema en cuestión, quienes pueden aportar información, evidencia, juicios y valoraciones informadas (Escobar Pérez y Cuervo Martínez, 2008). Según Hernández Sampieri et al. (2014), la validez de constructo se define como la capacidad de un instrumento para representar y medir de manera efectiva un concepto teórico específico. Este tipo de validación implica determinar qué tan adecuado es el conjunto de preguntas del instrumento para representar el espectro completo de comportamientos relevantes según el objetivo de la medición. Los elementos de este conjunto pueden ser denominados como reactivos o ítems.

2. 1 Instrumento de recolección de datos para el estudiantado

Para la elaboración del instrumento para la población estudiantil, se realizó una revisión de la literatura relacionada con el uso del teléfono celular en el ámbito educativo y el concepto de cognición extendida. A partir de esto, se generó una lista de indicadores que sirvieron como base para la creación de las preguntas y el desarrollo de los ítems. En este proceso, se consideraron tres variables principales: el grado en que el celular facilita las tareas cognitivas, la frecuencia y los patrones de uso del celular por parte del estudiantado y la cognición extendida en el estudiantado a través del uso del teléfono celular. Estas variables se utilizaron para construir los reactivos, siguiendo las ideas planteadas por Martínez Ramírez (2019). Como señala el autor, cualquier instrumento resulta de una cuidadosa integración entre paradigma, epistemología, perspectiva teórica, metodología y las técnicas empleadas para la recolección y análisis de datos.

El instrumento para el estudiantado fue validado a través de un juicio de expertos, en el que participaron tres profesionales especializados en el área. Aunque no existe un consenso definitivo sobre la cantidad óptima de expertos requeridos para la validación de instrumentos de investigación, diversos autores han propuesto criterios clave para su selección. Cabrero y Llorente (2013) sostienen que el número de expertos depende de la facilidad de acceso a estos y de la disponibilidad de especialistas con conocimientos en la temática de estudio. Escobar Pérez y Cuervo-Martínez (2008) enfatizan que la cantidad de jueces en un proceso de validación debe considerar tanto el nivel de especialización como la diversidad del conocimiento aportado. De manera similar, Maldonado Suárez y Santoyo-Téllez (2024) sugieren que la conformación del panel de expertos debe atender a factores como la facilidad de comunicación, la disponibilidad de tiempo, la alta cualificación de los panelistas y la definición de un perfil específico. En este sentido, Hernández Nieto (2002, 2011) plantea que la validación por juicio de expertos puede llevarse a cabo con un panel de entre tres y cinco jueces, permitiendo un análisis cuantitativo mediante el coeficiente de validez de contenido.

Esta consideración resulta especialmente relevante en este estudio, ya que se empleó un panel de tres jueces para la validación. La selección de estos expertos se basó en criterios como su trayectoria

en el campo, el reconocimiento dentro de la comunidad académica, su experiencia en validaciones o investigaciones similares y su disponibilidad para participar activamente en el proceso. Estos elementos garantizaron una evaluación rigurosa y fundamentada, contribuyendo a la optimización del instrumento validado.

Para realizar esta validación, se diseñó una herramienta virtual con el fin de facilitar la recopilación de datos. Esta herramienta consistió en una plantilla creada en hojas de cálculo de Google, siguiendo las directrices establecidas por Escobar Pérez y Cuervo Martínez (2008), así como por Galicia Alarcón et al. (2017).

En el proceso de recolección de datos, se solicitó información personal de los expertos, que incluyó su nombre completo, formación académica, áreas de experiencia profesional, cargo actual e institución de adscripción. A cada experto seleccionado se le envió una invitación por correo electrónico, detallando su inclusión en el proceso de evaluación de un instrumento diseñado para medir el grado de cognición extendida entre los estudiantes de secundaria que utilizan teléfonos celulares en el aula.

La invitación contenía información sobre los objetivos de la investigación, los fines específicos del juicio de expertos y de la prueba, así como detalles sobre la población objetivo y el tiempo estimado para completar la evaluación.

En otra sección de la herramienta, se presentaron los indicadores a través de los cuales los expertos debían evaluar cada ítem según categorías previamente definidas. Las categorías de evaluación fueron: claridad, coherencia, relevancia y suficiencia, basadas en las propuestas de Escobar Pérez y Cuervo Martínez (2008). Para la categoría «suficiencia», la evaluación no se realizó a nivel de ítem individual, sino que se aplicó a grupos de ítems que correspondían a una variable específica. En cuanto a la estructura del cuestionario, cada ítem fue evaluado por los expertos mediante una escala ordinal.

La escala empleada permitió clasificar las respuestas según niveles previamente definidos, asignando valores que oscilan entre «1. No cumple con el criterio» y «4. Alto nivel». Cada nivel de la escala se respaldó por indicadores específicos que guiaron la evaluación. Por ejemplo, la calificación «No cumple con el criterio» (1), en la variable de «Coherencia», se sustentó en el indicador que señalaba que el ítem no presenta una relación lógica con la variable. En contraste, la calificación «Alto nivel» (4) se apoyó en el indicador que señalaba que el ítem está completamente alineado con la variable evaluada. Estos indicadores proporcionaron un marco claro y objetivo para valorar cada ítem, lo que fortaleció la coherencia y la confiabilidad del proceso de evaluación realizado por los expertos.

El instrumento para la evaluación de expertos dirigido al estudiantado, se organizó en torno a tres variables clave:

1. El grado en que el teléfono celular facilita tareas cognitivas.
2. La frecuencia y patrones de uso del teléfono celular.
3. El uso del teléfono celular como extensión de la mente.

Con un total de 40 ítems, se garantizó que cada variable tuviera suficientes preguntas para evaluar de manera exhaustiva sus diferentes aspectos. Además, se asignaron diferentes niveles de importancia a las variables, lo que se reflejó en la cantidad de ítems por variable, logrando un equilibrio en la ponderación de la prueba.

Cada pregunta del cuestionario estuvo vinculada a una escala de Likert, con valores que van del 1 al 4: el 1 representaba «no cumple con el criterio», el 2 correspondía a un «bajo nivel», el 3 indicaba un «nivel moderado», y el 4 reflejaba un «alto nivel». Este enfoque proporcionó la base para una evaluación precisa de las respuestas y la obtención de resultados, permitiendo a los expertos asignar calificaciones de acuerdo con la escala establecida. Además, se incluyó una sección destinada a observaciones generales sobre cada ítem, así como un espacio para señalar si alguna dimensión del constructo estaba ausente o no fue evaluada.

Este documento fue enviado de manera individual a cada experto validador por correo electrónico, lo que facilitó su revisión y evaluación dentro del plazo estipulado, garantizando así una evaluación precisa y efectiva por parte de cada uno de los expertos.

Dentro de este proceso de validación por juicio de expertos, las respuestas obtenidas fueron analizadas mediante el coeficiente de concordancia “V” de Aiken (Aiken, 1980; 1985). Este método fue seleccionado debido a su facilidad de cálculo y su fundamentación lógica, basada en la opinión de los expertos sobre un material evaluativo. El coeficiente “V” es una técnica empleada para cuantificar la validez de contenido o la relevancia de un ítem con respecto a un dominio específico, contando con la participación de N jueces. Su valor oscila entre 0.00 y 1.00, donde 1.00 indica un acuerdo perfecto entre los jueces sobre la validez del contenido evaluado. La interpretación del coeficiente se basa en dos aspectos esenciales: su magnitud, que refleja el grado de acuerdo entre los jueces, y su significancia estadística, la cual se obtiene al comparar el valor del coeficiente con los valores críticos establecidos en las tablas de Aiken (Aiken, 1985; Merino Soto y Livia Segovia, 2009).

La fórmula para calcular el coeficiente es la siguiente:

$$V = \frac{\bar{X} - l}{k}$$

En la ecuación del coeficiente “V” de Aiken, V representa el coeficiente en sí, $\bar{X}$ es la media de las calificaciones dadas por los jueces, l es la calificación mínima posible en una escala Likert, y k es la diferencia entre la calificación máxima y mínima que los jueces pueden asignar en el instrumento (Penfield y Giacobbi, 2004). En este estudio, la escala Likert empleada incluye las siguientes opciones de calificación:

- No cumple con el criterio
- Bajo nivel
- Nivel moderado
- Alto nivel

Por lo tanto, en este caso, el valor de “k” se calcula como 4-1, lo que resulta en “k” =3.

Para establecer los niveles de confianza, se utilizó el Método Score, el cual se emplea para cuantificar la relevancia de los ítems en el contenido del cuestionario. Este método se basa en el enfoque de Penfield y Giacobbi (2004), quienes utilizaron el coeficiente “V” de Aiken como una proporción para calcular el intervalo de confianza en un nivel específico. En otras palabras, el valor de “V” se empleó como una medida para estimar la confiabilidad del instrumento, basándose en los resultados obtenidos a través de este coeficiente. Dado que se trata de una escala ordinal, esta medida es adecuada para el estudio, ya que permite reflejar con precisión y significado el grado de acuerdo entre los evaluadores.

La ecuación para el límite inferior del intervalo es:

$$L = \frac{2nkV + z^2 - z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)}$$

Y la ecuación para el límite superior es:

$$U = \frac{2nkV + z^2 + z\sqrt{4nkV(1-V) + z^2}}{2(nk + z^2)}$$

Los valores utilizados para calcular los límites inferior y superior del intervalo de confianza son los siguientes: L representa el límite inferior, U el límite superior y Z corresponde al valor en la distribución normal estándar según el nivel de confianza seleccionado. Se optó por un nivel de confianza del 95 %, lo que determina que Z es igual a 1.96. Con base en este cálculo mediante el Método Score, se decidió eliminar tres ítems que resultaban poco claros, se repetían o no medían adecuadamente la variable correspondiente.

Tabla 1. Resultados de la validación de contenido con coeficiente V de Aiken

Ítem	Criterio	Media por ítem	V de Aiken	Intervalos de confianza límite inferior al 95 %	Intervalos de confianza límite superior al 95 %
Ítem 1	Suficiencia	1,00	1,00	0,70	1,00
	Coherencia		1,00	0,70	1,00
	Relevancia		1,00	0,70	1,00
	Claridad		1,00	0,70	1,00
Ítem 2	Suficiencia	0,92	1,00	0,70	1,00
	Coherencia		0,89	0,56	0,98
	Relevancia		0,89	0,56	0,98
	Claridad		0,89	0,56	0,98

Ítem 3	Suficiencia	0,92	1,00	0,70	1,00
	Coherencia		0,89	0,56	0,98
	Relevancia		0,89	0,56	0,98
	Claridad		0,89	0,56	0,98
Ítem 4	Suficiencia	1,00	1,00	0,70	1,00
	Coherencia		1,00	0,70	1,00
	Relevancia		1,00	0,70	1,00
	Claridad		1,00	0,70	1,00
Ítem 5	Suficiencia	0,86	1,00	0,70	1,00
	Coherencia		1,00	0,70	1,00
	Relevancia		0,78	0,45	0,78
	Claridad		0,67	0,35	0,88

Nota. La tabla muestra los resultados del juicio de expertos aplicando el coeficiente V de Aiken a los 5 primeros ítems del cuestionario al estudiantado. El color rojo indica un área de mejora en la claridad del ítem 5.

Adicionalmente, se evaluó la confiabilidad del instrumento utilizando el coeficiente Alfa de Cronbach. Este coeficiente se emplea para determinar el grado de correlación entre los ítems de un instrumento. El coeficiente Alfa de Cronbach (α) es la medida de consistencia interna más utilizada en psicometría. Determina la homogeneidad o coherencia de los ítems al reflejar el promedio de las correlaciones entre ellos, asegurando que un conjunto de ítems mide de manera efectiva un factor común o constructo.

La fórmula matemática para su cálculo es:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k \sigma_i^2}{\sigma_T^2} \right)$$

Donde:

k = Número de ítems de la escala.

$\sum_{i=1}^k \sigma_i^2$ = Suma de las varianzas de los puntajes de cada ítem.

σ_T^2 = Varianza total de los puntajes observados de la escala.

La medición refleja el promedio de las correlaciones entre los ítems que conforman el instrumento, e indica en qué medida cada ítem contribuye a medir un concepto o factor específico. Generalmente, un valor elevado de Alfa de Cronbach sugiere que los ítems están midiendo de manera efectiva un factor común (Celina Oviedo y Campo Arias, 2005, p. 575). Según los mismos autores, el coeficiente Alfa de Cronbach debe ser al menos 0.70 para considerar que la consistencia interna de la escala es aceptable. Un valor inferior a este umbral indica una baja consistencia interna, mientras que un valor cercano a 0.90 es ideal, ya que sugiere una medición confiable sin redundancia excesiva entre los ítems.

El proceso de validación incluyó una evaluación detallada de los ítems en relación con criterios de claridad, coherencia, relevancia y suficiencia, utilizando una escala ordinal de tipo Likert. Para validar el contenido, se aplicaron el coeficiente de concordancia V de Aiken y el Método Score, mientras que la confiabilidad del instrumento se evaluó mediante el cálculo del coeficiente Alfa de Cronbach, obteniendo un valor de 0.778.

Tabla 2. Confiabilidad del instrumento: Alfa de Cronbach

Alfa de Cronbach	Número de ítems
0.778	37

Nota. Elaboración propia con información de encuesta piloto aplicada al estudiantado.

2.2 Muestra

La selección de participantes se planificó mediante un muestreo probabilístico, específicamente Muestreo Aleatorio Simple. Este diseño teórico se adoptó para garantizar que cada individuo de la población tuviera la misma probabilidad de ser incluido, permitiendo la representatividad y reduciendo el sesgo de la recolección de datos (Hernández Sampieri et al., 2014)

El tamaño mínimo de la muestra para la población estudiantil (N=65) se calculó utilizando la fórmula para poblaciones finitas, con un nivel de confianza del 90 % (Z=1.645) y un margen de error del 6 % (c=0.06).

$$n = \frac{N \cdot Z^2 \cdot p(1 - p)}{c^2(N - 1) + Z^2 \cdot p(1 - p)}$$

Donde:

n=Tamaño de la muestra requerido.

N=Tamaño de la población total (65 estudiantes)

Z=Nivel de confianza (1.646 para 90 %)

p=Proporción esperada (0,06).

c=Margen de error (0,06).

El tamaño muestral mínimo requerido se estableció en n=49, siendo la muestra final obtenida de 51 estudiantes, superando el mínimo y fortaleciendo la robustez estadística del estudio. A pesar del diseño teórico probabilístico, la recolección de datos se realizó mediante la invitación a la totalidad de la población estudiantil y docente de la institución seleccionada, lo que derivó en una muestra por conveniencia con participación voluntaria.

En el caso del estudiantado, la población de tercer grado, que en conjunto sumaba 65 estudiantes, fue invitada a participar en la prueba piloto, de los cuales 51 respondieron (78 %). En lo que respecta al personal docente, la población estuvo conformada por los 16 profesores que laboraban en la institución educativa al momento del estudio. Se invitó a todos ellos a participar voluntariamente en la investigación, respondiendo al cuestionario diseñado para su perfil. De este total, 11 docentes aceptaron y completaron el instrumento, lo que representa el 69 % de la población docente. La inclusión de este grupo buscó incorporar la perspectiva del profesorado en relación con las dimensiones evaluadas, favoreciendo un análisis más integral de los resultados. La participación se llevó a cabo bajo condiciones de voluntariedad, garantizando la confidencialidad de las respuestas y la protección de los datos personales.

2.3 Prueba piloto con el estudiantado

Una vez construido y validado el cuestionario, se realizó la prueba piloto con la muestra de 51 estudiantes de tercer grado de secundaria, con edades comprendidas entre 13 y 16 años, de una escuela pública del Estado de México, inscritos en el ciclo escolar 2023-2024. Tras analizar los resultados obtenidos, se realizaron ajustes en la redacción y claridad de algunas preguntas. Además, se calculó el coeficiente Alfa de Cronbach, el cual arrojó un valor de 0.858, lo que, según los parámetros establecidos por Tuapanta Dacto et al. (2017), se clasifica como una confiabilidad muy buena. A continuación, se presenta la tabla con los niveles de confiabilidad basados en el valor del Alfa de Cronbach:

Tabla 3. Niveles de confiabilidad según Alfa de Cronbach

Índice	Nivel de confiabilidad	Valor de Alfa de Cronbach
1	Excelente	0.9 - 1
2	Muy bueno	0.7 - 0.9
3	Bueno	0.5 - 0.7
4	Regular	0.3 - 0.5
5	Deficiente	0 - 0.3

Nota. Elaboración propia con datos de la investigación y con base en los autores Tuapanta Dacto et al. (2018).

Por consiguiente, los ítems del cuestionario evidencian una adecuada consistencia interna, lo que respalda su aplicabilidad en otros contextos y unidades de análisis. A continuación, se presentan en la siguiente tabla los resultados obtenidos del coeficiente Alfa de Cronbach en la prueba piloto aplicada al estudiantado:

Tabla 4. Confiabilidad de la prueba piloto con Alfa de Cronbach

Alfa de Cronbach	Número de ítems
0.858	37

Nota. Elaboración propia con información de la encuesta piloto aplicada al estudiantado.

2.4 Instrumento de recolección de datos para la plantilla docente

El instrumento diseñado para el personal docente, compuesto por 45 ítems, fue sometido a un proceso de validación de contenido mediante el coeficiente V de Aiken y el Método Score, mientras que su consistencia interna se evaluó a través del coeficiente Alfa de Cronbach. El cuestionario inicial se estructuró en torno a cuatro variables principales:

- Frecuencia de uso del teléfono celular por parte del estudiantado durante las clases.
- Actividades pedagógicas mediadas por el uso del teléfono celular.
- Percepción del personal docente sobre el impacto del teléfono celular como herramienta educativa en el aprendizaje.
- Percepción del personal docente sobre el teléfono celular como facilitador de tareas cognitivas.
- Esta organización permitió analizar tanto las prácticas docentes como las percepciones sobre la integración del teléfono celular en el entorno educativo.

2.5 Prueba piloto a la plantilla docente

Se realizó una prueba con 11 docentes de la misma institución en la que se realizó la prueba piloto con el estudiantado. Tras analizar sus respuestas y observaciones verbales, se efectuaron ajustes en la redacción de algunas preguntas, así como en el formato y en las opciones de respuesta. El coeficiente Alfa de Cronbach alcanzó un valor de 0.871, lo cual, conforme a los parámetros establecidos, indica una muy buena confiabilidad. Este resultado demuestra que los ítems del instrumento presentan una alta consistencia interna, lo que valida su uso en otros contextos. A continuación, se presentan los resultados obtenidos:

Tabla 5. Resultado de confiabilidad de la prueba piloto al cuerpo docente con Alfa de Cronbach

Alfa de Cronbach	Número de ítems
0.871	45

Nota. Elaboración propia con información de la prueba piloto aplicada al cuerpo docente.

2.6 Consideraciones éticas

El desarrollo del estudio se realizó conforme a los principios éticos que rigen la investigación en el ámbito educativo. Se obtuvo el consentimiento informado de todo el personal docente participante, así como el asentimiento informado del estudiantado, garantizando en todo momento la protección de sus datos personales y la confidencialidad de la información recopilada. En el caso del alumnado menor de edad, se contó además con el consentimiento informado de madres, padres o tutores legales. Los datos obtenidos fueron utilizados exclusivamente con fines de investigación para la elaboración del presente trabajo y se resguardaron conforme a las disposiciones aplicables. Asimismo, se gestionó y obtuvo la

autorización formal de las instituciones educativas involucradas para la aplicación de los instrumentos y la recolección de la información.

3. Resultados

3.1 Validación de contenido por juicio de expertos

La validez de contenido del instrumento fue evaluada por un panel de tres personas expertas en el área, quienes analizaron los ítems del cuestionario en función de su claridad, coherencia, relevancia y suficiencia. Para ello, cada experta o experto valoró los 40 ítems del cuestionario utilizando una escala ordinal de 1 a 4, donde 1 representa «no cumple con el criterio» y 4 «alto nivel».

La organización de las respuestas se realizó de manera sistemática, asignando cada respuesta a su respectivo ítem y dimensión dentro del cuestionario. Posteriormente, se realizó una revisión individual de cada ítem, prestando especial atención a las observaciones proporcionadas por el panel. Durante este proceso, se evaluó la coherencia de las respuestas en relación con el contenido y los objetivos específicos del instrumento.

El análisis de consensos y discrepancias permitió identificar patrones comunes y áreas de acuerdo entre las personas expertas, así como diferencias significativas en sus valoraciones. Esta revisión detallada contribuyó a un refinamiento del cuestionario, asegurando su pertinencia y precisión.

3.2 Resultados de la evaluación por categoría

Los resultados del juicio de expertos indicaron que la mayoría de los ítems alcanzaron puntuaciones promedio entre 3 y 4 en las dimensiones de claridad y coherencia, lo que sugiere que fueron considerados comprensibles y alineados con los constructos que pretendían medir. En la categoría de suficiencia, los ítems obtuvieron una puntuación promedio de 3.8, lo que evidencia su adecuación para evaluar las dimensiones propuestas en el instrumento. No obstante, se recomendó mejorar la redacción de algunos ítems con el fin de hacer el lenguaje más accesible para el estudiantado.

3.3 Cálculo del coeficiente de concordancia: V de Aiken

Para cuantificar la validez de contenido y el grado de concordancia entre las personas evaluadoras, se utilizó el coeficiente V de Aiken. El análisis arrojó un valor promedio de 0.85 en la dimensión de claridad, 0.83 en coherencia y 0.87 en relevancia. Estos valores reflejan un alto nivel de acuerdo entre las personas expertas, ya que se aproximan al valor ideal de 1.00, lo que indica una validación sólida del contenido del instrumento.

La interpretación estadística del coeficiente V de Aiken, con un nivel de confianza del 95 % ($Z = 1.96$), mostró intervalos de confianza entre 0.80 y 0.89 en todas las dimensiones, lo que respalda la consistencia y pertinencia de las evaluaciones realizadas. Estos resultados sugieren que los ítems cumplen con los criterios teóricos necesarios para medir el constructo de cognición extendida mediante el uso del teléfono celular.

3. 4 **Análisis de confiabilidad: Alfa de Cronbach**

Para evaluar la consistencia interna del instrumento, se calculó el coeficiente Alfa de Cronbach a partir de los datos recolectados. El coeficiente general obtenido fue de 0.92, lo que indica un alto nivel de confiabilidad en la medición de las dimensiones del cuestionario. De manera específica, las subvariables también mostraron valores elevados de Alfa de Cronbach: 0.90 para el grado en que el celular facilita tareas cognitivas, 0.89 para el uso del celular como extensión de la mente y 0.88 para la frecuencia y patrones de uso del dispositivo.

Estos valores sugieren que el instrumento presenta una estructura interna sólida y que los ítems seleccionados son consistentes en la medición de cada variable, lo que refuerza su idoneidad para futuras aplicaciones en contextos educativos.

3. 5 **Observaciones adicionales de las personas expertas**

La retroalimentación proporcionada por las personas expertas, junto con los hallazgos derivados de la prueba piloto, evidenció la necesidad de realizar ajustes en la formulación de ciertos ítems del instrumento, con el propósito de incrementar su claridad y garantizar la coherencia con los objetivos planteados en cada dimensión de análisis. Asimismo, se recomendó la utilización de un lenguaje más accesible para el estudiantado, favoreciendo así una mejor comprensión de las preguntas.

En este proceso, también se identificó la conveniencia de eliminar ítems redundantes o que funcionaban como subcategorías de otros, con el fin de optimizar la estructura y la pertinencia del instrumento. Finalmente, se sugirió la incorporación de escalas de valoración alternativas, orientadas a obtener datos más precisos y detallados en determinadas respuestas. Estas observaciones fueron integradas al análisis final del instrumento, fortaleciendo su estructura y claridad.

3. 6 **Organización y análisis de las respuestas de las personas expertas**

La recopilación y organización de respuestas se realizó de manera sistemática, asignando cada respuesta a su ítem y dimensión correspondiente. Se identificaron patrones de consenso y discrepancia entre las opiniones del panel, con un alto nivel de acuerdo en las dimensiones de suficiencia y coherencia. Sin embargo, el consenso en claridad fue moderado, por lo que se recomendó revisar la redacción de las preguntas, reformular algunas, eliminar aquellas redundantes, reubicar ciertos ítems en dimensiones más apropiadas y corregir errores ortográficos que pudieran afectar la comprensión.

3. 7 **Coeficiente V de Aiken y resultados por ítems**

El análisis reveló un consenso alto en la dimensión de suficiencia en las cuatro categorías evaluadas¹, lo que indica que la cantidad de ítems en cada dimensión es adecuada para medir las variables propuestas. Asimismo, los indicadores de coherencia y relevancia reflejan un nivel elevado de acuerdo entre las personas expertas. No obstante, el indicador de claridad obtuvo un nivel moderado de consenso, lo que

¹ Ver tabla 4.

sugiere la necesidad de ajustes adicionales en la redacción de preguntas y opciones de respuesta. Entre las sugerencias realizadas, se incluyó la reformulación de ciertos ítems, la eliminación de preguntas redundantes y la reubicación de algunas dentro de dimensiones más apropiadas para su análisis.

3. 8 Puntuaciones y promedio general de las personas expertas por dimensión

La siguiente tabla resume el nivel de consenso en las dimensiones de suficiencia y coherencia, y destaca las áreas de mejora en claridad para optimizar el instrumento.

Tabla 6. *Puntuaciones y promedio general de expertos por variable*

Experto	Dimensión	Suficiencia	Coherencia	Relevancia	Claridad	Promedio
Exp. 1	Dim. 1	100 %	100 %	100 %	100 %	98.3 %
	Dim. 2	100 %	93.3 %	93.3 %	93.3 %	
	Dim. 3	100 %	98.3 %	96.6 %	93.3 %	
Exp. 2	Dim. 1	100 %	82.5 %	85 %	82.5 %	83.5 %
	Dim. 2	100 %	95 %	91.6 %	83.3 %	
	Dim. 3	10 0%	93.3 %	88.3 %	71.6 %	
Exp. 3	Dim. 1	100 %	92.5 %	92.5 %	92.5 %	95.1 %
	Dim. 2	100 %	91.6 %	91.6 %	91.6 %	
	Dim. 3	100 %	93.3 %	93.3 %	93.3 %	
Promedio						92.3 %

Nota. Elaboración propia con las puntuaciones de expertos por dimensión del instrumento, según suficiencia, coherencia, relevancia y claridad. Valores superiores al 80 % indican alta validez de contenido.

El primer experto otorgó puntuaciones elevadas a las tres dimensiones, evidenciando un alto grado de conformidad y satisfacción. Por su parte, el segundo experto, aunque mostró plena concordancia en la dimensión de suficiencia, otorgó calificaciones más bajas en la dimensión de coherencia y claridad, señalando aspectos susceptibles de mejora. En contraste, el tercer experto mantuvo una evaluación consistente, asignando puntuaciones elevadas en todas las dimensiones y en los cuatro indicadores, lo que refleja un respaldo generalizado al instrumento.

El promedio general obtenido al calcular la sumatoria de promedios de cada persona experta fue de 92.3 %, lo que indica un alto grado de validez del instrumento.

La interpretación de los datos se realizó conforme a la escala propuesta por Celina Oviedo y Campo-Arias (2005), quienes establecen que un valor mínimo aceptable para Alfa de Cronbach es 0.70, ya que valores inferiores sugieren una consistencia interna baja. Por otro lado, valores superiores

a 0.90 pueden indicar redundancia o duplicación en los ítems del cuestionario. En este estudio, los coeficientes obtenidos confirman la confiabilidad del instrumento, asegurando su aplicabilidad en investigaciones futuras sobre el uso del teléfono celular en contextos educativos.

4. Discusión

Los resultados obtenidos del análisis preliminar del instrumento desarrollado muestran áreas clave que deben ser ajustadas para mejorar la validez y la claridad del mismo. Se identificaron varias sugerencias útiles que contribuirán a optimizar la efectividad del instrumento en la medición de las dimensiones de interés. Entre las recomendaciones más destacadas se encuentra la necesidad de ajustar la redacción de algunas preguntas, para asegurar que cada una refleje claramente el objetivo que se explora en cada dimensión. Esta modificación contribuyó a mejorar la alineación de las preguntas con las dimensiones correspondientes y garantizó que las respuestas sean precisas y claras.

Además, se sugirió la incorporación de escalas de valoración distintas en algunas respuestas, lo que permitiría obtener datos más precisos y significativos. El empleo de diferentes escalas también podría enriquecer el análisis de los resultados, proporcionando una visión más detallada y compleja de las percepciones de los participantes. Por otro lado, se destacó la necesidad de eliminar preguntas redundantes, aquellas que exploran objetivos similares o que se constituyen como subcategorías de otras preguntas. Esta recomendación tiene como fin evitar la repetición de contenido y mejorar la economía del instrumento, concentrándose en los aspectos esenciales de la medición.

Otro aspecto importante subrayado en el análisis es la precisión en la elección de palabras y verbos utilizados en las preguntas. Es fundamental que las preguntas sean formuladas de manera precisa para evitar malentendidos que puedan generar respuestas erróneas o imprecisas por parte de los participantes. De igual forma, se identificaron algunos errores ortográficos y de puntuación, como la omisión de signos de interrogación, que deben corregirse para garantizar una presentación profesional y clara del instrumento. Estos detalles, aunque menores, tienen un impacto significativo en la percepción de los participantes sobre la calidad y seriedad del instrumento.

En cuanto a los resultados cuantitativos obtenidos con el coeficiente V de Aiken, estos permiten valorar la validez del contenido del instrumento a través de la calificación proporcionada por los expertos. El uso de este coeficiente, combinado con el método de Penfield y Giacobbi (2004), permitió calcular los intervalos de confianza para cada ítem, lo cual resultó ser una herramienta útil para identificar los ítems que no cumplían con el criterio de validez establecido. Los intervalos de confianza del 95 % señalaron cuáles ítems requerían ajustes en sus formulaciones o en la dimensión que pretendían medir, lo que permitió una retroalimentación precisa sobre los aspectos problemáticos del instrumento.

En términos de consistencia interna, el Alfa de Cronbach de 0.778 indica que el instrumento posee una confiabilidad adecuada. Este valor refleja que las dimensiones evaluadas están interrelacionadas de manera coherente y que el instrumento es capaz de medir los constructos de manera fiable. La combinación de un coeficiente de V de Aiken adecuado dentro de los intervalos de confianza y un Alfa de Cronbach satisfactorio confirma que el instrumento es apropiado para su aplicación en la población seleccionada.

Los hallazgos sugieren que el instrumento es, en general, adecuado para medir las dimensiones previstas, pero su efectividad y precisión pueden mejorarse mediante ajustes en la redacción de las preguntas, la eliminación de redundancias y la corrección de pequeños detalles ortográficos. Las sugerencias aportadas por los expertos y los resultados cuantitativos respaldan estos cambios y son cruciales para garantizar que el instrumento sea una herramienta válida, confiable y eficiente en la recolección de datos.

5. Conclusiones

El proceso de validación del instrumento mediante juicio de expertos ha demostrado ser fundamental para evaluar su calidad y adecuación. Este proceso permitió identificar consensos y discrepancias entre los expertos, así como calificaciones tanto por dimensión como por ítem. De este modo, se lograron señalar áreas clave que requieren ajustes o mejoras. Las observaciones de los expertos no solo proporcionaron una evaluación cuantitativa, sino que también proporcionaron sugerencias concretas para optimizar el instrumento, tales como la revisión de la redacción, la mejora en la claridad de las preguntas, la eliminación de ítems redundantes y la mejora de la estructura general del instrumento.

Se alcanzó el objetivo de desarrollar dos instrumentos válidos y confiables para medir la integración del teléfono celular como extensión cognitiva en estudiantes de secundaria. La validez de contenido, corroborada a través del juicio de expertos, mostró puntuaciones superiores a 0.80 en términos de claridad, coherencia y relevancia de los ítems. Además, el coeficiente Alfa de Cronbach, con un valor de 0.92, refleja una alta consistencia interna, lo que refuerza la fiabilidad del cuestionario en la medición de las dimensiones que pretende evaluar.

Aunque el instrumento muestra altos niveles de validez y confiabilidad, persisten interrogantes que deben ser investigados más a fondo, tales como el impacto del uso del celular en los procesos cognitivos de los estudiantes y su posible efecto distractor en el aula. Se recomienda que futuras investigaciones se enfoquen en explorar estos aspectos, así como en identificar estrategias efectivas para mitigar los efectos negativos del uso del celular en el entorno educativo, tal como lo sugieren estudios previos (Hernández Silva, 2017).

La implementación del instrumento en investigaciones futuras puede contribuir significativamente a comprender cómo los estudiantes integran el celular en sus procesos de pensamiento, resolución de problemas y aprendizaje diario. Con una herramienta validada y confiable, se podrá obtener información precisa que profundice el análisis de la cognición extendida y su impacto en el desarrollo educativo, además de facilitar el diseño de estrategias pedagógicas que reconozcan el celular como un recurso valioso dentro del proceso de enseñanza-aprendizaje.

Agradecimientos

Extendemos nuestro sincero agradecimiento a la dirección escolar, el personal docente y administrativo de la institución educativa, Secundaria Oficial 1126 «Carlos Fuentes Macías», por las facilidades otorgadas para la realización de este trabajo. De igual manera agradecemos al alumnado que participó con entusiasmo y a los padres, madres y tutores por su valiosa cooperación y consentimiento informado.

Agradecimiento a los revisores

La Revista «La Universidad» agradece a los siguientes revisores por su evaluación y sugerencias en este artículo:

MSc. Marta Guadalupe Cienfuego Paniagua

Docente Investigadora, Universidad Técnica Latinoamericana
marta.paniagua@utla.edu.sv

MSc. Melvin Adalberto Cruz Cruz

Especialista informático y Técnico de Planificación
y Evaluación Institucional, Universidad de El Salvador
melvin.cruz@ues.edu.sv

Sus aportes fueron fundamentales para mejorar la calidad y rigor de esta investigación.

6. Referencias

- Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and Psychological Measurement*, 40, 955-959. <https://doi.org/10.1177/001316448004000419>
- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45, 131-142. <https://doi.org/10.1177/0013164485451012>
- Astorga Aguilar, C., León Anchía, R. y Schmidt Fonseca, I. (2020). El teléfono celular como herramienta de apoyo en el proceso de aprendizaje para el estudiante no indígena y el estudiante indígena de la Sección Regional Huetar. *Revista Ensayos Pedagógicos*, 15(2), 155-172. <https://doi.org/10.15359/rep.15-2.7>
- Cabrero Almenara, J. y Llorente Cejudo, M. C. (2013). La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información (TIC). *Eduweb. Revista de Tecnología de Información y Comunicación en Educación*, 7(22), 11-22. <https://dialnet.unirioja.es/servlet/articulo?codigo=4857163>
- Celina Oviedo, H. y Campo Arias, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572-580. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0034-74502005000400009
- Clark, A., y Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19. <https://doi.org/10.1093/analys/58.1.7>
- Escobar Pérez, J., Cuervo Martínez, A. (2008). Validez de contenido y juicio de expertos: Una Aproximación a su utilización. *Avances en Medición*, 6, 27-36.
- https://www.researchgate.net/publication/302438451_Validez_de_contenido_y_juicio_de_expertos_Una_aproximacion_a_su_utilizacion
- Galicia Alarcón L. Balderrama Trápaga, J., y Navarro, E. (2017) Validez de contenido por juicio de expertos: propuesta de una herramienta virtual. *Apertura*, 9, (2), 42-53. <http://dx.doi.org/10.32870/Ap.v9n2.993>
- Hernández-Nieto, R. A. (2002). *Contributions to Statistical Analysis*. Mérida: Universidad de Los Andes.
- Hernández-Nieto, R. (2011). *Instrumentos de recolección de datos en ciencias sociales y ciencias biomédicas*. Universidad de los Andes. https://www.academia.edu/37886946/Instrumentos_de_recoleccion_de_datos_en_ciencias_sociales_y_ciencias_biomedicas_Rafael_Hernandez_Nieto_pdf
- Hernández Sampieri, R., Fernández Collado, C. y Baptista Lucio, M. (2014). *Metodología de la Investigación*. Sexta edición. Mc Graw Hill Education. <https://www.uncuyo.edu.ar/ices/libro-metodologia-de-la-investigacion-6ta-edicion>

- Hernández Silva, D. (2017). Etnografía escolar del celular como un recurso para la construcción de espacios de fuga. *Civilizar Ciencias Sociales y Humanas*, 18(34),167-178. <https://revistas.usergioarboleda.edu.co/index.php/ccsh/article/view/v18n34a11>
- Maldonado Suárez, N. y Santoyo-Téllez, F. (2024). Validez de contenido por juicio de expertos: Integración cuantitativa y cualitativa en la construcción de instrumentos de medición. *REIRE Revista d'Innovació i Recerca en Educació*, 17(2), 1-19. <https://dialnet.unirioja.es/servlet/articulo?codigo=9622062>
- Martínez Domínguez, M. (2020). La desigualdad digital en México: un análisis de las razones para el no acceso y el no uso de internet. *Paakat: Revista de Tecnología y Sociedad*, 10(19). <http://dx.doi.org/10.32870/Pk.a10n19.519>
- Martínez Ramírez, J. L. (2019). El proceso de elaboración y validación de un instrumento de medición documental. *Acción y reflexión educativa*, 44. <https://portal.amelica.org/ameli/jatsRepo/226/226955004/index.html>
- Matamala Riquelme, C. (2016). Uso de las TIC en el hogar: Entre el entretenimiento y el aprendizaje informal. *Estudios Pedagógicos (Valdivia)*, 42(3), 293–311. https://www.scielo.cl/scielo.php?pid=S0718-07052016000400016&script=sci_arttext
- Merino Soto, C. y Livia Segovia, J. (2009). Intervalos de confianza asimétricos para el índice de validez de contenido: Un programa Visual Básica para V de Aiken. *Anales de Psicología*, 25(1), 169-171. <https://doi.org/10.15359/mhs.20-1.3>
- Nijssen, S., Schaap, G. y Verheijen, G. (2018). Has your smartphone replaced your brain? Construction and validation of the Extended Mind Questionnaire (XMQ). *PLoS One*, 13(18). <https://doi.org/10.1371/journal.pone.0202188>
- Nikolopoulou, K. (2020). Secondary education teachers' perceptions of mobile phone and tablet use in classroom: constraints and concerns. *Journal of Computers in Education*, 7, 257-275. <https://doi.org/10.1007/s40692-020-00156-7>
- Penfield, R. D. y Giacobbi, P. R. (2004). Applying a Score Confidence Interval to Aiken's Item Content-Relevance Index. *Measurement in Physical Education and Exercise Science*, 8(4), 213-225. DOI: 10.1207/s15327841mpee0804_3.
- Silva Calpa, A. C. y Martínez Delgado, D.G. (2017). Influencia del Smartphone en los procesos de aprendizaje y enseñanza. *Suma de Negocios*, 8(17), 1-6. <https://doi.org/10.1016/j.sumneg.2017.01.001>
- Tuapanta Dacto, J., Duque Vaca, M. y Mena Reinoso, A. (2017). Alfa de Cronbach para validar un cuestionario de usos de TIC en docentes universitarios. *Revista mkt Descubre*, 10, 37-48. <https://core.ac.uk/download/pdf/234578641.pdf>