



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO



FACULTAD DE INGENIERÍA CIVIL, DE SISTEMAS Y DE ARQUITECTURA

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

TESIS

Numia: Aplicación móvil basada en deep learning para el aprendizaje de números.

PARA OBTENER EL TÍTULO PROFESIONAL DE:

Ingeniero(a) de Sistemas

Autores:

Bach. Gloria Noelia Loro Ayala
Bach. Jonatan Arturo Mayanga Muñoz

Asesor:

Dr. Ing. Juan Elías Villegas Cubas

3 de agosto, 2026

Lambayeque, Perú



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO



FACULTAD DE INGENIERÍA CIVIL, DE SISTEMAS Y DE ARQUITECTURA

ESCUELA PROFESIONAL DE SISTEMAS

TESIS

Numia: Aplicación móvil basada en deep learning para el aprendizaje de números.

PARA OBTENER EL TÍTULO PROFESIONAL DE INGENIERO(A) DE
SISTEMAS

APROBADO POR EL SIGUIENTE JURADO:

Dr. Ing. Ernesto Karlo
Celi Arevalo
Presidente

Msc. Ing. Omar Wilton
Saavedra Salazar
Secretario

Msc. Ing. Robertos Carlos
Arteaga Lora
Vocal



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO



FACULTAD DE INGENIERÍA CIVIL, DE SISTEMAS Y DE ARQUITECTURA

ESCUELA PROFESIONAL DE SISTEMAS

TESIS

Numia: Aplicación móvil basada en deep learning para el aprendizaje de números.

PARA OBTENER EL TÍTULO PROFESIONAL DE INGENIERO(A) DE SISTEMAS

DR. ING. JUAN ELÍAS VILLEGAS CUBAS
ASESOR

BACH. GLORIA NOELIA LORO AYALA
BACH. EN INGENIERÍA DE SISTEMAS

BACH. JONATAN ARTURO MAYANGA MUÑOZ
BACH. EN INGENIERÍA DE SISTEMAS



UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
FACULTAD DE INGENIERÍA CIVIL DE SISTEMAS Y DE ARQUITECTURA
UNIDAD DE INVESTIGACIÓN



ACTA DE SUSTENTACIÓN
N° 012-2026-UI-FICSA



Siendo las 12:00 pm horas del día 03 de agosto del 2026, se reunieron de manera presencial los miembros de jurado de la tesis titulada: **"NUMIA: APLICACIÓN MÓVIL BASADA EN DEEP LEARNING PARA EL APRENDIZAJE DE NÚMEROS"**, con código N° IS_V_2024_033, designado por Resolución Decanal Virtual N° 67-2025-UNPRG-FICSA con la finalidad de Evaluar y Calificar la sustentación de la tesis antes mencionada, conformado por los siguientes docentes:

DR. ING. ERNESTO KARLO CELI AREVALO
MSC. ING. OMAR WILTON SAAVEDRA SALAZAR
MSC. ING. ROBERTO CARLOS ARTEAGA LORA

PRESIDENTE
SECRETARIO
VOCAL

Asesorado por DR. ING. JUAN ELIAS VILLEGAS CUBAS.

El acto de sustentación fue autorizado por OFICIO VIRTUAL N° 173-2026-UIFICSA, la tesis fue presentada y sustentada por los Bachilleres: **GLORIA NOELIA LORO AYALA Y JONATAN ARTURO MAYANGA MUÑOZ**, tuvo una duración de 65..... minutos. Después de la sustentación, y absueltas las preguntas y observaciones de los miembros del jurado; se procedió a la calificación respectiva:

	NUMERO	LETRAS	CALIFICATIVO
GLORIA NOELIA LORO AYALA	18	DIECIOCHO	MUY BUENO
JONATAN ARTURO MAYANGA MUÑOZ	18	DIECIOCHO	MUY BUENO

Por lo que quedan APTOS para obtener el Título Profesional de INGENIERO (A) DE SISTEMAS de acuerdo con la Ley Universitaria 30220 y la normatividad vigente de la Facultad de Ingeniería Civil De Sistemas y de Arquitectura de la Universidad Nacional Pedro Ruiz Gallo.

Siendo las 13:30..... Se dio por concluido el presente acto académico, dándose conformidad al presente acto, con la firma de los miembros del jurado.

DR. ING. ERNESTO KARLO CELI AREVALO
PRESIDENTE

MSC. ING. OMAR WILTON SAAVEDRA SALAZAR
SECRETARIO

MSC. ING. ROBERTO CARLOS ARTEAGA LORA
VOCAL

DR. ING. JUAN ELIAS VILLEGAS CUBAS
ASESOR

JHFF/bcm





CONSTANCIA DE VERIFICACIÓN DE ORIGINALIDAD

Yo, Dr. Ing. Juan Elias Villegas Cubas, **asesor de tesis de los bachilleres:**

- Loro Ayala Gloria Noelia
- Mayanga Muñoz Jonatan Arturo

TITULADA:

“NUMIA: APLICACIÓN MÓVIL BASADA EN DEEP LEARNING PARA EL APRENDIZAJE DE NÚMEROS”

Luego de la revisión exhaustiva del documento constato que la misma tiene un índice de similitud de 10% verificable en el reporte de similitud del programa TURNITIN.

El suscrito analizó dicho reporte y concluyó que cada una de las coincidencias detectadas NO CONSTITUYEN PLAGIO. A mi leal saber y entender la tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad Nacional Pedro Ruiz Gallo.

Se expide la presente según lo dispuesto en la Resolución N° 659-2020-R, de fecha 8 de setiembre de 2020 formativa para la obtención de Grados y Títulos de la UNPRG:

Lambayeque, 14 de mayo del 2026

ATENTAMENTE,

Dr. Ing. Juan Elias Villegas Cubas
DNI. 80103991
Asesor

Se adjunta:
Recibo digital de Turnitin
Revisión de informe en Turnitin

NUMIA: APLICACIÓN MÓVIL BASADA EN DEEP LEARNING PARA EL APRENDIZAJE DE NÚMEROS

INFORME DE ORIGINALIDAD

10%

INDICE DE SIMILITUD

8%

FUENTES DE INTERNET

4%

PUBLICACIONES

4%

TRABAJOS DEL
ESTUDIANTE

FUENTES PRIMARIAS

1

cloud.google.com

Fuente de Internet

<1 %

2

Submitted to Universidad Internacional de la Rioja

Trabajo del estudiante

<1 %

3

alicia.concytec.gob.pe

Fuente de Internet

<1 %

4

repositorio.unjbg.edu.pe

Fuente de Internet

<1 %

5

www.aucal.edu

Fuente de Internet

<1 %

6

www.revista.ister.edu.ec

Fuente de Internet

<1 %

7

repositorio.ucv.edu.pe

Fuente de Internet

<1 %

8

funes.uniandes.edu.co

Fuente de Internet

<1 %



Dr. Ing. Juan Elías Villegas Cubas
Asesor
80103991

9	hdl.handle.net Fuente de Internet	<1 %
10	www.coursehero.com Fuente de Internet	<1 %
11	www.slideshare.net Fuente de Internet	<1 %
12	www.goconqr.com Fuente de Internet	<1 %
13	repositorio.unsaac.edu.pe Fuente de Internet	<1 %
14	Submitted to Universidad de Guayaquil Trabajo del estudiante	<1 %
15	repositorio.uct.edu.pe Fuente de Internet	<1 %
16	repositorio.upt.edu.pe Fuente de Internet	<1 %
17	Submitted to Ilerna Online Blackboard Trabajo del estudiante	<1 %
18	repositorio.uta.edu.ec Fuente de Internet	<1 %
19	dspace.unach.edu.ec Fuente de Internet	<1 %
20	repositorio.unemi.edu.ec Fuente de Internet	<1 %



21 www.fastcheck.cl <1 %
Fuente de Internet

22 Submitted to Universidad Cesar Vallejo <1 %
Trabajo del estudiante

23 repositorio.upla.edu.pe <1 %
Fuente de Internet

24 Submitted to Universidad Católica San Pablo <1 %
Trabajo del estudiante

25 Submitted to Universidad EAN <1 %
Trabajo del estudiante

26 repositorio-api.eespli.edu.pe <1 %
Fuente de Internet

27 revistalatam.redilat.org <1 %
Fuente de Internet

28 Rodríguez López, Jorge Leonardo. "Predicción de Esperanza de Vida Considerando el Covid-19, Bajo Algoritmos de Machine Learning Basados en Datos Epidemiológicos y Patológicos.", Universidad Distrital Francisco José de Caldas (Colombia) <1 %
Publicación

29 Submitted to Universidad Nacional de Cajamarca <1 %
Trabajo del estudiante

Submitted to Universitat Oberta de Catalunya



Dr. Ing. Juan Elías Villegas Cubas
Asesor
80103991

30

Trabajo del estudiante

<1 %

31

repositorio.uss.edu.pe

Fuente de Internet

<1 %

32

Submitted to Unidades Tecnológicas de Santander

Trabajo del estudiante

<1 %

33

Iparraguirre Ramos, Cesar Mariano.
"Resolución de problemas para mejorar el
trabajo autónomo de los estudiantes del
primer grado de secundaria de una
institución educativa, Chimbote, 2023",
Universidad Católica los Ángeles de Chimbote
(Peru)

Publicación

<1 %

34

Submitted to Universidad Nacional del Centro del Peru

Trabajo del estudiante

<1 %

35

repositorio.unjfsc.edu.pe

Fuente de Internet

<1 %

36

repositorio.epnewman.edu.pe

Fuente de Internet

<1 %

37

Leodan Stalin Tapia González, Victoria Marilyn
Salazar Piña. "Ejercicios funcionales: efecto
sobre la fuerza y la resistencia en estudiantes

<1 %



Dr. Ing. Juan Elías Villegas Cubas

Asesor
80103991



Recibo digital

Este recibo confirma que su trabajo ha sido recibido por **Turnitin**. A continuación podrá ver la información del recibo con respecto a su entrega.

La primera página de tus entregas se muestra abajo.

Autor de la entrega: Loro Ayala Gloria Noelia Mayanga Muñoz Jonatan Arturo
Título del ejercicio: Quick Submit
Título de la entrega: NUMIA: APLICACIÓN MÓVIL BASADA EN DEEP LEARNING PARA...
Nombre del archivo: INFORME_FINAL_DE_TESIS_LORO_MAYANGA_agosto_2026.pdf
Tamaño del archivo: 7.59M
Total páginas: 189
Total de palabras: 36,594
Total de caracteres: 200,651
Fecha de entrega: 14-may-2026 11:40p. m. (UTC-0500)
Identificador de la entrega: 2961676929

	UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO FACULTAD DE INGENIERÍA CIVIL, DE SISTEMAS Y DE ARQUITECTURA UNIDAD DE INVESTIGACIÓN	
"NUMIA: APLICACIÓN MÓVIL BASADA EN DEEP LEARNING PARA EL APRENDIZAJE DE NÚMEROS." TESIS Presentada para obtener el Título Profesional de Ingeniero(a) de Sistemas		
Autores Loro Ayala Gloria Noelia Mayanga Muñoz Jonatan Arturo		
Asesor Dr. Juan Elías Villegas Cubas		
Líneas de Investigación Ingenierías y Tecnologías		
Lugar de Ejecución Chiclayo, Lambayeque		
Duración estimada del proyecto 18 meses		
1		


Dr. Ing. Juan Elías Villegas
CubasAsesor
80103991

DEDICATORIA

A mi madre Gloria, por ser el pilar más importante en mi vida y por demostrarme siempre su cariño y apoyo incondicional. A mi padre Fausto por ser un ángel en mi camino. A mis hermanos Liliana, Eder y Milagros por su constante respaldo y compañía. A mis sobrinos, cuya alegría y ocurrencias iluminaron mis días y me regalaron siempre una sonrisa sincera. A mi equipo querido, quienes con su amistad y esfuerzo hicieron de cada ciclo una experiencia enriquecedora y memorable.

Gloria Noelia Loro Ayala

A mis padres, que creyeron en mí incluso en los momentos en que yo mismo dudé. Su confianza fue la brújula cuando no tenía certezas. A mi familia, por entender cada vez que no pude estar, y por seguir ahí cuando regresé. A los docentes que en algún momento de mi formación me enseñaron que la tecnología solo tiene sentido cuando sirve a las personas. Y a NumIA: porque empezó como una idea y terminó siendo la razón por la que hoy me paro ante un jurado con la certeza de que hicimos algo que importa.

Jonatan Arturo Mayanga Muñoz

AGRADECIMIENTO

A Dios, por darnos la fortaleza para no soltar este proyecto en los momentos en que todo parecía más grande que nosotros, y por poner en nuestro camino a las personas correctas en el momento justo.

A nuestro asesor, el Dr. Juan Elías Villegas Cubas, por su orientación constante, su paciencia ante nuestras dudas y por exigirnos con la firmeza justa para que este trabajo alcanzara el rigor que merecía. Su guía fue determinante en cada etapa de esta investigación.

Al profesor Matias Carranza Fernández, por ser el vínculo esencial que facilitó nuestra comunicación con la institución y nos brindó la oportunidad de llevar nuestra investigación a la práctica.

A la Institución Educativa Fe y Alegría N.º 28 de Chiclayo, a su directiva y a las docentes de primer grado, por abrirnos las puertas de sus aulas con confianza y disposición, permitiéndonos trabajar con sus estudiantes en un ambiente de colaboración genuina. Sin ese respaldo institucional, esta investigación no habría sido posible.

A los niños y niñas de primer grado que participaron en este estudio, porque con cada trazo en la pantalla, con cada sonrisa al ver el check verde y con cada intento después de equivocarse, nos recordaron por qué elegimos este camino. Ellos son la razón de ser de NumIA.

A la Facultad de Ingeniería Civil, de Sistemas y de Arquitectura de la Universidad Nacional Pedro Ruiz Gallo, por la formación recibida a lo largo de estos años, y a cada docente que en algún momento de nuestra carrera sembró en nosotros la convicción de que la tecnología tiene sentido cuando resuelve problemas reales de personas reales.

A nuestras familias, por sostener lo que no se ve: las madrugadas, las ausencias, la incertidumbre. Por creer en este proyecto incluso cuando nosotros mismos no teníamos todas las respuestas. Este logro también les pertenece.

Y a quienes, de una u otra forma, contribuyeron con una idea, una corrección, una palabra de aliento o simplemente con su presencia durante este proceso: gracias, porque cada aporte, por pequeño que parezca, dejó huella en este trabajo.

RESUMEN

La escritura correcta de los números arábigos constituye una base esencial para el desarrollo del pensamiento matemático en los primeros grados de educación primaria; sin embargo, la realidad educativa peruana evidencia deficiencias preocupantes en este aspecto, agravadas por la limitada retroalimentación formativa que los docentes pueden ofrecer de manera individualizada en contextos de aulas numerosas. Frente a este escenario, la presente investigación tuvo como objetivo general desarrollar e implementar una aplicación educativa basada en Deep Learning, denominada NumIA, para mejorar el trazado correcto de los números arábigos en estudiantes de primer grado de primaria de la Institución Educativa Fe y Alegría N.º 28, ubicada en Chiclayo, Lambayeque. El estudio se tipificó como cuantitativo, aplicativo, de nivel explicativo y con un diseño cuasi-experimental de tipo pretest-posttest con un solo grupo intacto, conformado por 30 estudiantes seleccionados mediante muestreo no probabilístico intencional, distribuidos equitativamente en tres secciones (10 por sección). La aplicación NumIA integró un modelo de inteligencia artificial basado en Redes Siamesas capaz de evaluar automáticamente la similitud entre el trazo del estudiante y patrones de referencia predefinidos, generando retroalimentación inmediata y personalizada. La intervención se ejecutó durante cinco semanas con tres sesiones semanales de 30 minutos, bajo supervisión docente e investigadora permanente. Los resultados del pretest arrojaron una media global de 72.4%, donde solo el 23% de los estudiantes alcanzaba el nivel de Logro Destacado (AD), mientras que en el posttest la media global ascendió a 95.7%, representando una ganancia de +23.2 puntos porcentuales, con el 90% de los estudiantes ubicados en nivel AD y el 10% restante en nivel A, desapareciendo por completo los niveles B y C. La prueba t de Student para muestras relacionadas arrojó resultados estadísticamente significativos en las tres secciones y en el análisis global ($t = 7.33$, $p < 0.001$), y los tamaños del efecto calculados mediante la d de Cohen superaron el umbral de efecto grande en todos los análisis (d global = 1.34). La Precisión, dimensión más deficitaria en el diagnóstico inicial (71.8%), registró la mayor ganancia en el posttest, evidenciando la eficacia específica del módulo de trazado guiado de NumIA sobre la forma estructural del dígito. En conclusión, NumIA demostró ser una herramienta eficaz para mejorar el trazado correcto de los números arábigos, produciendo una mejora consistente, estadísticamente verificada y de magnitud práctica muy grande, según los criterios establecidos por la literatura especializada.

Palabras clave: Deep Learning, Redes Siamesas, aplicación móvil educativa, trazado numérico, retroalimentación formativa, primer grado de primaria.

ABSTRACT

The correct writing of Arabic numerals constitutes an essential foundation for the development of mathematical thinking in the early grades of primary education; however, the Peruvian educational reality evidences concerning deficiencies in this area, aggravated by the limited formative feedback that teachers can offer individually in contexts of large classrooms. Faced with this scenario, the present research aimed to develop and implement an educational application based on Deep Learning, called NumIA, to improve the correct tracing of Arabic numerals in first-grade primary students at the Fe y Alegría N.º 28 Educational Institution, located in Chiclayo, Lambayeque. The study was classified as quantitative, applied, explanatory in level, and with a quasi-experimental pretest-posttest design with a single intact group, consisting of 30 students selected through intentional non-probabilistic sampling, equally distributed across three sections (10 per section). The NumIA application integrated an artificial intelligence model based on Siamese Networks capable of automatically evaluating the similarity between the student's stroke and predefined reference patterns, generating immediate and personalized feedback. The intervention was carried out over five weeks with three weekly sessions of 30 minutes each, under permanent teacher and researcher supervision. Pretest results yielded a global mean of 72.4%, with only 23% of students reaching the Outstanding Achievement level (AD), while in the posttest the global mean rose to 95.7%, representing a gain of +23.2 percentage points, with 90% of students at the AD level and the remaining 10% at level A, with levels B and C disappearing completely. The paired-samples t-test yielded statistically significant results across all three sections and in the global analysis ($t = 7.33$, $p < 0.001$), and effect sizes calculated using Cohen's d exceeded the large effect threshold in all analyses (global $d = 1.34$). Precision, the most deficient dimension in the initial assessment (71.8%), recorded the greatest gain in the posttest, evidencing the specific efficacy of NumIA's guided tracing module on the structural form of the digit. In conclusion, NumIA proved to be an effective tool for improving the correct tracing of Arabic numerals, producing a consistent, statistically verified improvement of very large practical magnitude, according to the criteria established by the specialized literature.

Keywords: Deep Learning, Siamese Networks, educational mobile application, numerical tracing, formative feedback, first grade of primary school.

INDICE

1.	INTRODUCCIÓN	10
1.1.	Justificación e importancia de la investigación	11
1.2.	Objetivo general.....	12
1.3.	Objetivos específicos	12
1.4.	Estructura del informe.....	13
2.	DISEÑO TEÓRICO.....	14
2.1.	Antecedentes	14
2.1.1.	Antecedentes Internacionales.....	14
2.1.2.	Antecedentes Nacionales	16
2.2.	Bases Teóricas	18
2.2.1.	Aprendizaje de Niños.....	18
2.2.2.	Teorías de Aprendizaje de Números.....	18
2.2.3.	Aprendizaje del Trazado de Números.....	19
2.2.4.	Teoría del Feedback Formativo	19
2.2.5.	Escritura Numérica Árabe.....	20
2.2.6.	Errores en la Escritura Numérica Árabe	20
2.2.7.	Evaluación en el Trazado de Números	21
2.2.8.	Teoría del Aprendizaje Multimedia	23
2.2.9.	Aplicativo Móvil.....	23
2.2.10.	App Lúdica/Educativa	24
2.2.11.	Rol del Color, Música y Repetición en la Creación de Apps	24
2.2.12.	Deep Learning en la Educación	25
2.2.13.	Redes Siamesas	26
2.2.14.	Medición de la Aplicación Móvil	27
2.3.	Operacionalización o categorización de variables	29
3.	DISEÑO METODOLÓGICO.....	31
3.1.	Tipo de Investigación.....	31
3.2.	Diseño de Contrastación de Hipótesis / Procedimiento a Seguir.....	33
3.3.	Población y Muestra	36
3.3.1.	Población.....	36
3.3.2.	Muestra	37

3.4.	Técnicas, Instrumentos, Equipos y Materiales	38
3.5.	Aspectos Éticos de la Investigación.....	41
4.	RESULTADOS.....	43
4.1.	Búsqueda de estrategias pedagógicas para la enseñanza del trazo de números.....	43
4.1.1.	Estrategias Pedagógicas Tradicionales Identificadas.....	43
4.1.2.	Caracterización de la Retroalimentación Docente	46
4.1.3.	Elementos Efectivos del Aprendizaje Multimedia	48
4.1.4.	Síntesis de Hallazgos del Resultado	50
4.2.	Diseño y desarrollo de la aplicación móvil "NumIA".	50
4.2.1.	Diseño del modelo de deep learning (redes siamesas).....	51
4.2.2.	Diseño de la Base de Datos.....	66
4.2.3.	Diseño y desarrollo de la aplicación móvil.....	68
4.2.4.	Síntesis de Hallazgos del Resultado	83
4.3.	Resultados de pretest e implementación de aplicación NumIA.....	85
4.3.1.	Diagnóstico Previo a la Intervención	85
4.3.2.	Implementación de la aplicación NumIA	94
4.3.3.	Usabilidad de la aplicación	98
4.3.4.	Síntesis de Hallazgos del Resultado	99
4.4.	Aplicación del postest y análisis de resultados	100
4.4.1.	Aplicación Postest.....	100
4.4.2.	Análisis Estadístico Inferencial.....	112
4.4.3.	Registros Automáticos de la Aplicación NumIA.....	123
4.4.4.	Síntesis de Hallazgos del Resultado	135
5.	Discusión de resultados.....	137
5.1.	Sobre la mejora en el trazado numérico y su contraste con antecedentes	137
5.2.	Sobre el efecto de la práctica estructurada y la frecuencia de uso	139
5.3.	Sobre los fundamentos teóricos del trazado numérico	140
5.4.	Sobre el marco teórico del aprendizaje digital y la retroalimentación.....	141
5.5.	Sobre la magnitud del efecto y su comparabilidad	142
6.	Conclusiones	144
7.	Recomendaciones	146
8.	Bibliografía	150
9.	ANEXOS	155

INDICE DE TABLAS

Tabla 1 Operacionalización de la variable Independiente	30
Tabla 2 Operacionalización de la variable Dependiente.....	31
Tabla 3 Distribución de la muestra por sección.....	37
Tabla 4 Técnicas de Recolección de Datos.....	38
Tabla 5 Instrumentos.....	39
Tabla 6 Resultados de validación de contenido por juicio de expertos (V de Aiken)	40
Tabla 7 Equipos y Materiales.....	40
Tabla 8 Arquitectura CNN de la Red Siamesa de NumIA	59
Tabla 9 Métricas de evaluación del modelo de Redes Siamesas de NumIA (n = 10,000 pares)	65
Tabla 10 Matriz de confusión del modelo de Redes Siamesas por dígito (n = 1,000 por clase)	65
Tabla 11 Rubrica de trazado	86
Tabla 12 Estadísticos descriptivos de los subcriterios de la rúbrica pretest (n = 30)	88
Tabla 13 Niveles de logro y umbrales de clasificación	89
Tabla 14 Resultado de la evaluación pretest.....	91
Tabla 15 Análisis global del diagnóstico	93
Tabla 16 Resultados del System Usability Scale (SUS) aplicado a docentes y estudiantes...	98
Tabla 17 Resultado de la evaluación post test	102
Tabla 18 Comparación individual – sección A.....	105
Tabla 19 Comparación individual – sección B.....	107
Tabla 20 Comparación individual – sección C	108
Tabla 21 Estadísticos descriptivos comparativos pretest–postest.....	109
Tabla 22 Distribución de niveles pretest vs. postest por sección.....	111
Tabla 23 Resultados de Shapiro-Wilk	113
Tabla 24 Cálculo de diferencias individuales y estadísticos Sección 1ºA.....	115
Tabla 25 Cálculo de diferencias individuales y estadísticos Sección 1ºB.....	117
Tabla 26 Cálculo de diferencias individuales y estadísticos Sección 1ºC.....	119
Tabla 27 Resultados consolidados de la prueba t de Student pareada	121
Tabla 28 Verificación por componente de la hipótesis de investigación.....	122
Tabla 29 Intentos de práctica registrados en NumIA por semana de intervención y sección	124

Tabla 30 Mensajes de retroalimentación emitidos por NumIA por estudiante (n = 30).....	126
Tabla 31 Tiempo prom. Por intento emitidos por NumIA por estudiante (n = 30)	128
Tabla 32 Frecuencia de práctica por sección	131
Tabla 33 <i>Frecuencia de práctica por estudiante registrada en NumIA (n = 30)</i>	132

INDICE DE FIGURAS

Figura 1 Ejemplo de arquitectura de red siamesa	51
Figura 2 Ejemplo de multitrazo	54
Figura 3 Generación de guías.	57
Figura 4 Arquitectura del modelo.	61
Figura 5 Entrenamiento del modelo.....	62
Figura 6 Pruebas del modelo.	64
Figura 7 Diagrama de base de datos.	67
Figura 8 Estructura de archivos en flutter.....	71
Figura 9 Estructura de archivos en python.	72
Figura 10 Conexión de base de datos.	73
Figura 11 Pantalla de logeo de usuario.	76
Figura 12 Pantalla de Inicio.	77
Figura 13 Pantalla de capitulos.	78
Figura 14 Pantalla de preguntas de conteo y escritura.....	79
Figura 15 Pantalla de preguntas con recurso multimedia.	80
Figura 16 Pantalla de retroalimentación positiva.	80
Figura 17 Pantalla de retroalimentación con corrección.	81
Figura 18 Pantalla de personalización de errores.	82
Figura 19 Pantalla de perfil de usuario.	82
Figura 20 Flujo de usuario.	83

1. INTRODUCCIÓN

La escritura correcta de números arábigos constituye una base esencial para el desarrollo del pensamiento matemático en grados posteriores de los estudiantes de educación primaria. Según el Currículo Nacional de Educación Básica del Perú (MINEDU, 2016), al finalizar el primer grado, los estudiantes deben ser capaces de reconocer, leer y escribir números naturales del 0 al 100 en formato arábigo. Sin embargo, la realidad educativa peruana evidencia deficiencias preocupantes en este aspecto. La Evaluación Nacional de Logros de Aprendizaje (ENLA, 2023) reveló que únicamente el 11.2% de los estudiantes de segundo grado alcanzaron el nivel satisfactorio en Matemática, lo que indica un rezago que se arrastra desde primer grado.

Esta problemática se acentúa debido a errores comunes en el trazado numérico, tales como la inversión de números (escritura espejo), discontinuidad en los trazos y orientación incorrecta. Según Cabezas (2023), aproximadamente el 38% de niños de primer grado invierten números y letras al escribir, lo cual, de no corregirse oportunamente, puede convertirse en un obstáculo para el aprendizaje matemático posterior. Villarroel et al. (2013) enfatiza que el primer grado es una etapa decisiva para consolidar patrones motores correctos en la escritura de números, haciendo crucial la intervención temprana.

Un factor agravante de esta situación es la limitada retroalimentación formativa en el entorno escolar peruano. Llanos Masciotti & Tapia (2021) señalan que la interacción docente-estudiante suele limitarse a una participación reducida, donde la retroalimentación es meramente correctiva: el docente solo indica si algo está bien o mal, sin ofrecer orientaciones específicas ni correcciones inmediatas que permitan al estudiante comprender y mejorar sus errores en el momento oportuno.

Las consecuencias de esta problemática son amplias y preocupantes. A nivel individual, los niños que no desarrollan adecuadamente la habilidad de trazar números enfrentan

dificultades para realizar operaciones matemáticas básicas, lo que se traduce en bajo rendimiento académico, frustración, pérdida de autoestima y desinterés por las matemáticas (Carreira, 2013).

Frente a este escenario, surge la necesidad de investigar e implementar soluciones tecnológicas como herramientas educativas que proporcionen retroalimentación inmediata, personalizada y formativa con la finalidad de reforzar el aprendizaje del trazado de números en estudiantes de primer grado, todo ello dentro de la línea de investigación de Ingenierías y Tecnologías, específicamente en la aplicación de soluciones tecnológicas innovadoras para resolver problemas educativos reales mediante inteligencia artificial.

1.1. Justificación e importancia de la investigación

Justificación teórica: El presente estudio aporta al campo científico ya que explora la aplicación del Deep Learning, particularmente en las redes siamesas, en el contexto educativo peruano. Ya que estas arquitecturas han demostrado eficacia en verificación de firmas y reconocimiento de escritura manuscrita (Dey et al., 2017), su aplicación específica para la enseñanza del trazado de números en niños de primer grado constituye un área poco explorada que amplía el entendimiento sobre cómo las tecnologías de inteligencia artificial pueden adaptarse a contextos pedagógicos específicos.

Justificación práctica: La investigación tiene un aporte práctico porque se desarrolla un modelo de Deep learning, el cual se usa como base para la construcción de un aplicativo móvil educativo para ser usada como herramienta tecnológica accesible en aulas, proporcionando retroalimentación inmediata y personalizada que los docentes, a veces no pueden ofrecer de manera individual, especialmente en aulas numerosas donde la atención individualizada constituye un desafío estructural (Llanos Masciotti & Tapia, 2021).

Justificación metodológica: Esta investigación se justifica porque aporta una metodología que usa instrumentos de medición específicos para evaluar el trazado numérico en tres dimensiones: precisión, secuencia y orientación. Para la medición operacional, la dimensión secuencia se desagrega en dos subcriterios diferenciados secuencia de orden y secuencia de fluidez, lo que permite capturar tanto el aspecto procedimental como la continuidad del trazo. Estos instrumentos, junto con el modelo de Deep Learning implementado, ofrece un marco de referencia metodológico para investigaciones futuras en áreas relacionadas.

Justificación social: Esta investigación busca generar un impacto social positivo al mejorar las competencias matemáticas básicas de los estudiantes de educación primaria. Debido a que las dificultades en matemáticas pueden perpetuar desigualdades educativas y limitar oportunidades futuras, como evidencia el hecho de que solo el 11.2% de los estudiantes de segundo grado alcanzaron el nivel satisfactorio en Matemática según la ENLA (2023).

Con base a lo descrito anteriormente, la pregunta de investigación que guía el estudio es: **¿De qué manera influye el uso de la aplicación móvil NumIA en la mejora del trazado correcto de los números arábigos en estudiantes de primer grado de primaria?**

Para dar respuesta a esta interrogante, se plantea lo siguiente:

1.2. Objetivo general

Desarrollar e implementar una aplicación educativa basada en deep learning para mejorar el trazado correcto de los números arábigos en estudiantes de primer grado de primaria.

1.3. Objetivos específicos

- a.** Identificar y analizar las estrategias pedagógicas utilizadas por docentes de primer grado para enseñar el trazo de números, especialmente con el apoyo de tecnologías digitales.

- b. Diseñar y desarrollar la aplicación móvil "NumIA", incorporando una interfaz gráfica amigable y módulos interactivos de práctica de trazos numéricos evaluados mediante un modelo de inteligencia artificial basado en Redes Siamesas.
- c. Evaluar el nivel inicial del trazado numérico mediante una prueba diagnóstica (pretest) e implementar la aplicación NumIA en sesiones de práctica estructurada durante cinco semanas con estudiantes de primer grado de una institución educativa pública.
- d. Analizar el impacto de la intervención mediante una evaluación final (postest) y registros automáticos de la aplicación, comparando los resultados con el diagnóstico inicial para determinar las mejoras logradas en precisión, secuencia y orientación del trazado.

1.4. Estructura del informe

Introducción. Presenta una visión general del estudio, incluyendo la descripción de la problemática abordada, el objetivo general y los objetivos específicos a cumplir, por último, se menciona la justificación de la investigación.

Diseño Teórico. Explora una revisión de los antecedentes de investigación relevantes y de similitud en desarrollo de herramientas educativas, a nivel internacional y nacional. Profundiza las bases teóricas que sustentan el estudio, abarcando teorías sobre el aprendizaje de números, el feedback formativo, el aprendizaje multimedia y los fundamentos de las Redes siamesas.

Metodología. Detalla el tipo de investigación cuantitativa, el diseño cuasi-experimental utilizado, la descripción de la población y muestra estudiadas, así como las técnicas e instrumentos de recolección de datos usados. Además, se especifican los aspectos éticos considerados en la investigación.

Resultados. Describe los hallazgos obtenidos, así como el desarrollo de cada uno de los objetivos específicos planteados.

Conclusiones y recomendaciones: Se discuten los resultados comparando con los antecedentes seleccionados.

2. DISEÑO TEÓRICO

2.1. Antecedentes

El aprendizaje correcto del trazado de los números arábigos en los primeros años de escolaridad es fundamental para el desarrollo de competencias matemáticas y de motricidad fina en los niños. En este sentido, distintas investigaciones a nivel internacional, nacional han abordado estrategias y herramientas tecnológicas, como aplicaciones educativas, para fortalecer habilidades de escritura numérica en estudiantes de educación primaria.

2.1.1. Antecedentes Internacionales

Zachry et al. (2020) realizaron un estudio piloto en Estados Unidos con el objetivo de evaluar si el uso de aplicaciones en tabletas iPad podía mejorar la legibilidad de la escritura a mano en alumnos de 1° y 2° grado, comparándolo con métodos tradicionales de terapia ocupacional. Los niños de primer grado frecuentemente presentan caligrafía poco legible, afectando su desempeño académico.

El estudio empleó un diseño experimental pretest-postest con asignación aleatoria, comparando tres grupos: (1) práctica con aplicaciones iPad, (2) combinación de apps iPad con enfoque multisensorial tradicional, y (3) enfoque multisensorial tradicional únicamente. Los resultados, mediante análisis cualitativo ciego de muestras de escritura, mostraron mejoras observables en la formación de letras en el grupo que utilizó únicamente aplicaciones iPad, comparado con los otros dos grupos. Este estudio fundamenta en enfoques terapéuticos para mejorar la escritura a mano, incluyendo práctica terapéutica tradicional y enfoques sensoriomotores, partiendo de la

premisa que las intervenciones multisensoriales y tecnológicas pueden influir positivamente en las habilidades de escritura infantil.

Niklas et al. (2025) desarrollaron un estudio experimental en Alemania para analizar si el uso de aplicaciones educativas de alta calidad desde el hogar podía mejorar las competencias matemáticas y lingüísticas de niños en edad preescolar (4 a 6 años). El problema abordado fue la baja preparación escolar en estas áreas, especialmente en entornos desfavorecidos, influyendo directamente en el desempeño posterior en primaria. El estudio se sustentó en teorías del desarrollo cognitivo infantil y el enfoque del aprendizaje temprano mediado por tecnología. Se diseñó una intervención de seis meses con 500 niños, distribuidos aleatoriamente en cuatro grupos: (1) intervención en lectoescritura, (2) intervención en matemáticas, (3) grupo control con tableta sin contenido educativo, y (4) grupo control sin tableta. Las aplicaciones incluían contenidos interactivos, retroalimentación inmediata, reforzadores positivos y adaptabilidad al ritmo del niño. En matemáticas se utilizaron las pruebas MARKO-S y WVT; para lenguaje, AWST-R, PPVT y SETK 3-5. Los resultados mostraron mejoras estadísticamente significativas en las competencias evaluadas, con el uso de aplicaciones explicando entre 1% y 2% de la varianza en el rendimiento. Lo notable fue que estas mejoras se lograron con sesiones muy breves: los niños usaban las aplicaciones entre 2.5 y 4.5 minutos diarios en promedio, demostrando que intervenciones de baja duración pueden tener impacto real cuando las herramientas están bien diseñadas.

Bang et al. (2023) abordaron el problema del bajo rendimiento académico en matemática en primeros años de primaria, donde el 40% de niños alcanzan la competencia esperada según NAEP 2019. Propusieron el uso de "My Math Academy", aplicación que ofrece contenido personalizado y evaluaciones adaptativas integradas para mejorar el aprendizaje y motivación de los estudiantes. La investigación se basó en teorías del desarrollo y cognición

que destacan la importancia de habilidades matemáticas tempranas como predictores del éxito académico futuro y preparación para carreras STEM. El estudio fue experimental con grupos de tratamiento y control, participando estudiantes de kindergarten y primer grado (N = 505 en tratamiento, 481 en control). Los resultados indicaron que estudiantes usando "My Math Academy" lograron mejoras significativas en matemáticas.

2.1.2. Antecedentes Nacionales

Núñez (2021) identificó un rendimiento deficiente en nociones numéricas de niños de 5 años (educación inicial), en su investigación halló que los métodos tradicionales no lograban un aprendizaje sólido de conceptos numéricos básicos. El estudio buscó introducir un enfoque innovador mediante una aplicación móvil con Realidad Aumentada (RA) para enseñar números, frente al problema de falta de atención y aprendizaje memorístico en el aula inicial. Se sustentó en teorías de aprendizaje lúdico y multisensorial en la primera infancia, donde la RA permite a los niños interactuar con contenido virtual sobre el mundo real, alineado con el constructivismo de Piaget (aprendizaje mediante experiencia activa). También se apoyó en el enfoque de Rencoret para la enseñanza de matemáticas en inicial, enfatizando la construcción gradual de nociones numéricas. El estudio fue de tipo aplicado, cuantitativo, con diseño cuasi-experimental, trabajando con un grupo de niños de 5 años usando la app con RA en sesiones controladas, comparando capacidades matemáticas antes y después. La app incorporó juegos de RA para que los niños manipulen números y objetos virtuales, reforzando conceptos como conteo, correspondencia uno a uno y reconocimiento de dígitos. Los resultados demostraron mejora en capacidades matemáticas fundamentales en edad preescolar, específicamente en aprendizaje de números naturales. Los niños aumentaron su habilidad para identificar y nombrar números del 1 al 10, contarlos en orden, asociar cantidades con cifras y resolver situaciones sencillas de conteo. También desarrollaron mejor comprensión de conceptos como tamaños, formas y relaciones espaciales gracias a la interactividad de la RA.

Estañiz (2020) realizó un estudio en Andahuaylas, ciudad donde no suelen incorporar las TIC en educación, generando bajo interés y atención hacia las matemáticas en los niños. En primer grado de primaria es fundamental lograr conocimiento básico de números, identificarlos y contarlos (unidades, decenas) para posteriormente realizar operaciones básicas. El estudio se basó en enfoques transversales del Currículo Nacional de Educación Básica respecto al uso de entornos TIC en el aprendizaje, enfatizando constructivismo y aprendizaje activo: la app combina juego y contenido en quechua (lengua materna local), factor importante para facilitar la asimilación de conceptos numéricos básicos.

También consideró la teoría de aprendizaje significativo, integrando el nuevo recurso tecnológico al contexto rural de los niños. El autor aplicó un enfoque cuantitativo con diseño experimental (pretest-posttest), desarrollando una aplicación móvil llamada Yupay Yachay ("Aprende a contar" en quechua) para reforzar matemáticas básicas en alumnos de 1° de primaria. La app se alineó con el currículo oficial de primaria. Para evaluar resultados se aplicaron pruebas de rendimiento antes y después de su uso, demostrando mejora significativa en el aprendizaje matemático. En el pretest los promedios eran bajos, y tras usar la app hubo incrementos notables en reconocimiento de números, conteo de objetos y realización de operaciones básicas (sumas y restas). El estudio concluyó que el uso de Yupay Yachay mejoró significativamente el aprendizaje de matemáticas en comparación con la enseñanza tradicional.

Estos antecedentes demuestran que las aplicaciones móviles educativas, especialmente aquellas que integran tecnologías como Realidad Aumentada, gamificación y personalización adaptativa, pueden mejorar significativamente el aprendizaje de números y habilidades matemáticas básicas en niños de educación inicial y primaria. La retroalimentación inmediata, los entornos lúdicos y la adaptación al contexto cultural del estudiante emergen como factores clave del éxito de estas intervenciones.

2.2. Bases Teóricas

2.2.1. Aprendizaje de Niños

El aprendizaje de los niños se define como el proceso mediante el cual adquieren conocimientos, habilidades, valores y actitudes a través de sus experiencias, la interacción con el entorno y la mediación de los adultos, viéndose influenciado por factores biológicos, emocionales, sociales y culturales. Piaget & Inhelder (1997) establecen que el aprendizaje en la infancia ocurre a través de la asimilación y acomodación, mediante la interacción activa del niño con su ambiente. Recoge información, pero también construye su propio saber a partir de la manipulación, la exploración y el juego. Por su parte Vergel (2014) retoma la perspectiva de Vygotski para destacar que el aprendizaje infantil está profundamente mediado por la interacción social y el lenguaje, señalando la importancia de la mediación de adultos y pares en el desarrollo de los procesos psicológicos superiores.

2.2.2. Teorías de Aprendizaje de Números

El aprendizaje de los números en los niños constituye un proceso cognitivo complejo que involucra múltiples habilidades como el conteo, la identificación de cantidades, la comparación numérica y la representación simbólica. Este proceso ha sido estudiado por diversas teorías que consideran tanto el desarrollo madurativo como la influencia del entorno en la construcción del conocimiento matemático.

Gelman & Gallistel (1978) proponen que la comprensión numérica en los niños pequeños se inicia a partir de una serie de principios, entre los que destacan la correspondencia uno a uno, el principio de orden estable, la cardinalidad y la abstracción. Estos principios permiten que el niño no solo conozca la serie numérica, sino que represente el número con su cantidad correspondiente.

El aprendizaje de los números no puede entenderse sin considerar el sentido numérico, definido como la comprensión intuitiva de las cantidades, la capacidad de estimar, comparar y

operar con números en contextos cotidianos (National Research Council, 2009). Esta capacidad constituye un pilar clave para el desarrollo de competencias matemáticas más avanzadas y es enriquecida por experiencias de aprendizaje significativas en edades tempranas.

2.2.3. Aprendizaje del Trazado de Números

El trazado de números implica aprender a escribir dígitos numéricos en su forma convencional, ayudando al niño a asociar el símbolo escrito con la cantidad que representa. Zhou & Wang (2004) afirman que la habilidad de los niños para representar números por escrito se relaciona estrechamente con su comprensión de la cardinalidad. A medida que el niño aprende a trazar correctamente cada dígito, va consolidando la comprensión de que cada símbolo numeral tiene un significado numérico específico, fortaleciendo así la conexión entre la forma escrita del número y lo que representa.

Sinvani et al. (2023) mencionan que el aprendizaje de la escritura de números es un proceso complejo que integra habilidades cognitivas, perceptuales y motoras finas. Para trazar un dígito de forma convencional, el niño debe emplear la percepción visoespacial (para captar la forma y orientación del número), la discriminación de tamaño y proporción (para respetar las dimensiones apropiadas del trazo) y la discriminación de orientación (para no invertir o rotar incorrectamente el símbolo). Estos procesos perceptuales se combinan con la coordinación visomotora y el control motor fino de los dedos y la muñeca para ejecutar los trazos exactos.

2.2.4. Teoría del Feedback Formativo

El feedback formativo o retroalimentación formativa se define como la provisión de información al estudiante con el fin de cerrar la brecha entre su desempeño actual y el nivel esperado, ayudándolo a mejorar su aprendizaje mediante orientación clara, oportuna y concreta.

Hattie & Timperley (2007) plantean que el feedback formativo cumple una función reguladora del aprendizaje, al proporcionar información clara y específica respecto al nivel de desempeño del estudiante en relación con los criterios establecidos. Es un proceso continuo y progresivo, por lo que la información formativa debe ser entregada durante la ejecución de la tarea, y no solo al final.

En el contexto de la educación, el feedback formativo debe adaptarse a las características del desarrollo cognitivo de los niños, utilizando lenguaje accesible, ejemplos concretos y apoyos visuales o lúdicos que favorezcan la comprensión de los errores y la exploración de nuevas formas de resolver problemas (Hattie & Timperley, 2007).

2.2.5. Escritura Numérica Árabe

Según la RAE, el sistema arábigo (o sistema de numeración indo-arábigo) cuenta con diez signos llamados dígitos (0, 1, 2, 3, 4, 5, 6, 7, 8, 9), cuyo valor se determina por la posición que ocupan dentro del número representado. Es el sistema más extendido y utilizado actualmente, permitiendo generar con pocos signos una serie ilimitada de números.

Detalles relevantes incluyen:

Posición y valor: El valor de cada dígito depende de su posición dentro del número. Por ejemplo, en el número 217, el 2 representa 200, el 1 representa 10 y el 7 representa 7.

Dirección de escritura: Los números arábigos se escriben de izquierda a derecha, al igual que la escritura occidental.

2.2.6. Errores en la Escritura Numérica Árabe

Los errores en la escritura numérica arábigo son dificultades específicas que presentan los niños al representar los números del sistema decimal (0 al 9), ya sea en su forma, orden, orientación o valor posicional. Dos de los errores más comunes en la etapa preescolar son la inversión de números y la discontinuidad en el trazo.

Inversión de números: Consiste en escribir los dígitos en una dirección incorrecta o reflejada. Se observa comúnmente en números como el 3, 5, 6, 7 y 9, propensos a escribirse en forma de espejo. Zapateiro-Segura et al. (2018) indican que estas inversiones están vinculadas al desarrollo de la orientación espacial y no deben considerarse patológicas en los primeros años, formando parte del proceso de adquisición de la lateralidad y del control visomotor.

Discontinuidad en el trazo: Ocurre cuando el niño representa un número utilizando múltiples partes o fragmentos en lugar de un solo trazo continuo. Un ejemplo típico es la escritura del número "8" mediante dos círculos superpuestos, en vez de un solo trazo en forma de lazo. Este tipo de error puede estar asociado a una coordinación motriz fina insuficientemente desarrollada o a una percepción inadecuada del número como unidad visual (Chang & Yu, 2022).

2.2.7. Evaluación en el Trazado de Números

La evaluación del trazado numérico se refiere a medir cómo el estudiante escribe las cifras numéricas de forma correcta y legible. Villarroel et al. (2013) señala que dominar el trazado no implica solo reproducir la forma gráfica de cada dígito, sino también comprender que los símbolos numéricos tienen significado. Lograr la correcta escritura de números arábigos supone dos aprendizajes básicos: (1) establecer rutinas motoras específicas para cada dígito y (2) reconocer que la notación numérica es simbólica, donde cada símbolo posee un valor cuantitativo particular (Hederich-Martínez et al., 2016).

Al evaluar el trazado numérico, la literatura especializada identifica tres dimensiones de desempeño que permiten un análisis integral y sistemático del trazo:

- La precisión hace referencia a la claridad y reconocibilidad del número, la forma debe ser simétrica, acorde al modelo convencional y libre de distorsiones significativas Calero-Brito et al. (2024). Feder & Majnemer (2007) documentan que la precisión de la forma del símbolo es el criterio central en la evaluación clínica de

la escritura infantil, ya que permite distinguir si las dificultades del niño corresponden a errores de representación visual del dígito o a errores de ejecución motriz

- La secuencia del trazo evalúa que el estudiante siga el orden apropiado de los segmentos y la dirección correcta al formar cada número. Esta dimensión se desagrega en dos subcriterios evaluables de forma diferenciada: la secuencia de orden, que evalúa el aspecto procedimental inicio, dirección y orden lógico de los segmentos, y la secuencia de fluidez, que evalúa la continuidad y automatización del movimiento trazo fluido, sin interrupciones ni fragmentaciones. Chang & Yu (2022), en un estudio computarizado de evaluación grafomotora con niños de primer a tercer grado de primaria, identifican el orden del trazo y la fluidez del movimiento como dimensiones directamente asociadas al desempeño académico en escritura y al nivel de automatización grafomotora alcanzado, confirmando que ambos aspectos pueden fallar de forma independiente y requieren medición separada.
- La orientación adecuada implica que la cifra no aparezca invertida ni rotada incorrectamente. Zapateiro-Segura et al. (2018) señalan que las inversiones están vinculadas al desarrollo de la orientación espacial y constituyen un indicador del control visomotor del estudiante.

Para evaluar estas dimensiones de manera integral y sistemática, es recomendable el uso de rúbricas analíticas que describan indicadores y niveles de logro para cada criterio (Ministerio de Educación, 2022). Este tipo de instrumento permite identificar con precisión qué aspecto específico del trazado requiere intervención, lo que es especialmente relevante en contextos donde la retroalimentación individualizada es limitada.

2.2.8. Teoría del Aprendizaje Multimedia

Mayer (2010) define el aprendizaje multimedia como el proceso mediante el cual los estudiantes construyen representaciones mentales a partir de palabras e imágenes.

Esta teoría se basa en la idea de que existen tres tipos de almacenaje en la memoria (memoria sensorial, memoria de trabajo y memoria a largo plazo) y que los individuos poseen dos canales separados para procesar información: uno para material verbal y otro para material visual. Cada canal solo puede procesar una pequeña cantidad de información a la vez, pudiendo darse soporte procesando un contenido presentado de dos formas diferentes y complementarias.

Esta teoría resalta tres procesos cognitivos importantes para el aprendizaje significativo:

- **Selección de palabras e imágenes:** El estudiante se concentra en percibir las palabras e imágenes relevantes del material presentado.
- **Organización:** El alumno organiza mentalmente la información en dos modelos: verbales lógicos y representaciones visuales.
- **Integración:** Estos dos tipos de representaciones se integran entre sí y se asimilan con el conocimiento previo.

Mayer propone principios de diseño instruccional como el principio de contigüidad (presentar texto e imagen juntos), el principio de coherencia (eliminar información irrelevante) y el principio de modalidad (usar narración en lugar de texto escrito junto a la imagen).

2.2.9. Aplicativo Móvil

Gómez & Mesías (2021) **definen el aplicativo móvil, también conocido como aplicación móvil o app**, como un programa informático desarrollado para ejecutarse en dispositivos móviles, como teléfonos inteligentes o tabletas, con un propósito específico. Romero et al. (2024) señalan que estas aplicaciones están desarrolladas específicamente para aprovechar las características de los dispositivos móviles, permitiendo realizar muchas tareas

y funciones. Su alcance es amplio, utilizándose en diferentes contextos empresariales, educativos, de entretenimiento y demás ámbitos.

2.2.10. App Lúdica/Educativa

Las apps lúdicas son aplicativos móviles que tienen el propósito de enseñar a los usuarios por medio de juegos. Se definen como herramientas tecnológicas desarrolladas para promover el aprendizaje mediante experiencias de juego, incorporando principios del aprendizaje basado en el juego, utilizando estrategias como recompensas, niveles, retroalimentación inmediata y avatares para involucrar al usuario en un proceso activo y motivador.

Hassinger-Das et al. (2017) indican que este tipo de aplicaciones son efectivas en la educación infantil porque logran integrar el entretenimiento con objetivos pedagógicos, haciendo más flexible la comprensión de conceptos abstractos y el desarrollo de habilidades cognitivas y motoras. El componente lúdico no solo aumenta la motivación, sino que permite al niño repetir tareas sin percibirlo como una obligación, facilitando así el proceso de adquisición de habilidades como la escritura de números.

2.2.11. Rol del Color, Música y Repetición en la Creación de Apps

Según Ortiz (2006), los colores llamativos o fuertes actúan como estímulos didácticos que favorecen la atención de los estudiantes y refuerzan los procesos de aprendizaje. El color puede funcionar como una clave mnemotécnica, facilitando la retención de contenidos. La inclusión adecuada de la música en entornos digitales de aprendizaje también desempeña un papel relevante, siendo una herramienta efectiva para captar y mantener el interés de los estudiantes durante el proceso de aprendizaje (Dzib Moo et al., 2023).

En complemento a estos elementos, la repetición de contenidos se identifica como una estrategia didáctica eficaz. Vargas-Murillo (2020) menciona que repetir ejercicios y contenidos favorece la asimilación progresiva del conocimiento, el desarrollo de la comprensión y la

retención de la información. Robert (2001) sostiene que la repetición no solo mejora la retención de información, sino que también permite al estudiante construir significados de forma progresiva y sostenible.

2.2.12. Deep Learning en la Educación

El Deep Learning en la educación se refiere al uso de algoritmos de aprendizaje profundo para simular el complejo poder de toma de decisiones del cerebro humano. Según IBM, el aprendizaje profundo permite a los sistemas simular este poder de toma de decisiones, resultando en aplicaciones que mejoran la automatización y realizan tareas analíticas sin intervención humana.

El Deep Learning es útil para analizar y procesar grandes cantidades de datos educativos, con el fin de mejorar la enseñanza y el aprendizaje. Estos algoritmos, basados en redes neuronales artificiales de múltiples capas, son capaces de aprender automáticamente patrones complejos a partir de datos como escritura, voz, imágenes o respuestas de los estudiantes.

En el sector educativo, el Deep Learning se emplea para desarrollar sistemas inteligentes que identifican grafías manuales, evalúan la pronunciación, adaptan contenidos según el rendimiento del estudiante e incluso predicen dificultades de aprendizaje. Tales herramientas permiten individualizar la enseñanza, automatizar la retroalimentación y mejorar la eficacia del aprendizaje en tiempo real.

Silva-Peñafiel et al. (2024) señalan que la inteligencia artificial, y sobre todo su rama de Deep Learning, está transformando la educación mediante la personalización de la enseñanza y la mejora de los procesos de evaluación, permitiendo que con el uso de algoritmos se personalice el material educativo, adaptándose a las carencias y al modo de aprendizaje de cada estudiante.

2.2.13. Redes Siamesas

Una Red Siamesa es una arquitectura de deep learning compuesta por dos (o más) subredes idénticas que comparten parámetros (pesos y arquitectura), lo cual garantiza que ambas aprendan de manera sincrónica las mismas características. Cada subred procesa una entrada independiente y genera un vector de características o embedding. Posteriormente, una función de similitud como la distancia Euclídea o la similitud del coseno compara estos embeddings para determinar el grado de semejanza entre las entradas (Bromley et al., 1993)

De Rosa & Papa (2018) explican que este tipo de redes consisten en un par de redes neuronales que comparten sus pesos y tienen como objetivo calcular funciones de similitud. En esencia, su objetivo principal es identificar si un par de datos es diferente. El objetivo principal de esta arquitectura es que las entradas pertenecientes a la misma clase presenten embeddings cercanos en el espacio latente, mientras que entradas de clases distintas se ubican alejadas, haciéndolas útiles en tareas de verificación y comparación más que de clasificación tradicional.

- **Origen y evolución:** El modelo fue introducido por Bromley et al. (1993) en el contexto de la verificación de firmas manuscritas, convirtiéndose en un referente pionero en el ámbito de la biometría. A diferencia de las redes de clasificación convencionales, la Red Siamesa se centra en aprender una función de distancia discriminativa capaz de generalizar incluso a ejemplos no vistos durante el entrenamiento.
- **Áreas de aplicación:** Desde su propuesta inicial, las redes siamesas se han aplicado exitosamente en diversos campos:
- **Verificación de firmas offline:** El modelo SigNet Dey et al. (2017) demostró que las redes siamesas pueden detectar falsificaciones sin requerir entrenamiento específico por usuario, superando a métodos tradicionales de extracción manual de características.

- **Reconocimiento de escritura manuscrita:** Variantes modernas que integran redes siamesas con arquitecturas Transformer han permitido realizar verificación few-shot en contextos de datos limitados (Majumder et al., 2024).

Justificación para NumIA: La elección de una red siamesa se fundamenta en:

- **Naturaleza del problema:** No se busca solo clasificar un dígito entre 0–9, sino verificar si el trazo cumple con las características correctas de escritura.
- **Eficiencia de datos:** Al trabajar con pocas muestras de referencia, se evita entrenar un clasificador complejo que requiera miles de ejemplos por clase.
- **Robustez a la variabilidad:** Los niños difieren en tamaño, orientación, presión y estilo de escritura. Al comparar contra un embedding normalizado, la red siamesa aprende a abstraer esas diferencias superficiales.
- **Escalabilidad:** Una vez entrenado el backbone, se pueden agregar nuevas variantes de trazos sin necesidad de reentrenar el modelo completo.

2.2.14. Medición de la Aplicación Móvil

La evaluación de una aplicación móvil educativa es esencial para garantizar su calidad tanto pedagógica como técnica. Dada la gran cantidad de apps destinadas al aprendizaje, se vuelve indispensable verificar su pertinencia educativa (Pinilla & Badilla, 2020).

Lee & Kim (2015) señalan que la evaluación de aplicaciones para niños debe considerar cuatro dimensiones fundamentales: contenido, diseño, funcionalidad y calidad técnica. Estas categorías incluyen detalles importantes como:

Usabilidad de la aplicación: Se refiere a qué tan fácil de usar resulta una app para los usuarios.

La norma internacional ISO 9241-11 define la usabilidad como "el grado en que un sistema puede ser utilizado por usuarios específicos para lograr objetivos específicos con efectividad, eficiencia y satisfacción en un contexto de uso determinado" (ISO, 2018). Una revisión sistemática destaca que los atributos más valorados en la usabilidad de aplicaciones educativas

móviles son la eficiencia, la eficacia, la facilidad de aprendizaje y la satisfacción del usuario. (Mkpojiogu et al., 2018).

Tiempo de respuesta y rendimiento: El tiempo de respuesta de una aplicación es la rapidez con la que el sistema procesa las acciones del usuario, factor técnico clave que incide directamente en la experiencia de uso. En el caso de las aplicaciones infantiles, incluso breves retrasos pueden generar frustración o desinterés, afectando la continuidad del aprendizaje (Mkpojiogu et al., 2018).

Por lo tanto, una aplicación educativa debe garantizar tiempos de respuesta mínimos para preservar la fluidez de la interacción. Esto se considera al evaluar si la app funciona bien técnicamente, las cuales son revisadas con pruebas de rapidez de rendimiento y análisis de registros (logs), permitiendo identificar latencias y detectar fallos en la respuesta del sistema. La optimización de estos tiempos contribuye no solo a mejorar la usabilidad, sino también a favorecer el compromiso y la eficacia del proceso de aprendizaje (Lee & Kim, 2015).

Calidad del trazo escrito: Hace referencia al nivel de precisión, claridad y uniformidad al escribir letras o números. Según Skar et al. (2022), la fluidez de escritura entendida como la rapidez y precisión con que los estudiantes producen letras correctamente predice de forma significativa la calidad del texto escrito en los primeros grados de primaria. Por su parte, (Hamdi et al., 2020) describen un sistema digital de evaluación de escritura infantil que analiza el trazo en tiempo real considerando criterios como la forma, la dirección, el orden de los segmentos y la cinemática, permitiendo detectar dificultades tempranas y diseñar intervenciones personalizadas.

Interacción y retroalimentación: Resultan ser elementos esenciales dentro de las apps didácticas diseñadas para niños, ya que determinan en gran medida su nivel de compromiso y motivación. Liu & Correia (2021) afirman que la retroalimentación inmediata y las actividades

variadas contribuyen a experiencias educativas más significativas y a un mayor nivel de autonomía del estudiante.

Frecuencia de uso y engagement: La frecuencia de uso de una aplicación educativa es un indicador clave de su valor pedagógico y atractivo. Gómez Vera (2019) destaca que el uso de apps móviles puede mejorar la motivación del estudiante y su involucramiento en el aprendizaje, reflejándose en un mayor tiempo de dedicación a las actividades propuestas.

Instrumentos de evaluación:

- **Rúbricas de evaluación:** Guías de evaluación con criterios e indicadores concretos que permiten calificar cualitativa o cuantitativamente distintos aspectos de la aplicación.
- **Encuestas y cuestionarios:** Técnica de recolección de información que facilita la obtención de datos de modo rápido y eficaz (Casas Anguita et al., 2003).
- **Observación directa:** Forma sistematizada y lógica para el registro visual y verificable de lo que ocurre en el mundo real (Campos & Lule, 2012).

2.3. Operacionalización o categorización de variables

Variable Independiente: Uso de la aplicación NumIA

Definición conceptual: Es la interacción del estudiante con la aplicación móvil educativa basada en Deep Learning, diseñada para proporcionar ejercicios interactivos de trazado numérico con retroalimentación inmediata y personalizada.

Definición operativa: Se medirá a través de registros automáticos generados por la aplicación que cuantifican: (1) el tipo y frecuencia de retroalimentación proporcionada al estudiante, (2) el tiempo de dedicación a cada actividad, y (3) el número de preguntas respondidas por semana.

Tabla 1
Operacionalización de la variable Independiente

Variable	Dimensión	Indicador	Escala de Medición
Uso de la aplicación NumIA (X)	Retroalimentación	Proporción de mensajes correctivos frente a mensajes afirmativos (refuerzo positivo)	Razón (%)
	Interactividad	Promedio de tiempo que el usuario dedica a cada actividad o ejercicio en la app	Razón (minutos)
	Frecuencia de práctica	Número de preguntas respondidas por semana.	Razón (cantidad)

Nota. Elaboración Propia.

Variable Dependiente: Trazado correcto del número

Definición conceptual: El trazado correcto del número se define como la habilidad del estudiante para escribir dígitos numéricos arábigos (0-9) de manera legible, siguiendo la forma convencional, la secuencia apropiada de trazos y la orientación correcta, sin inversiones ni rotaciones.

Definición operativa: Se medirá mediante una rúbrica de evaluación aplicada en pretest y posttest que evalúa tres dimensiones conceptuales: (1) orientación adecuada del trazo, (2) secuencia del trazo desagregada operacionalmente en secuencia de orden y secuencia de fluidez para mayor precisión diagnóstica, y (3) precisión de la forma del numeral. Cada subdimensión se califica en una escala ordinal de tres niveles: Nivel 0 (No logrado: 0 puntos), Nivel 1 (En proceso: 1 punto), Nivel 2 (Logrado: 2 puntos), generando un puntaje máximo de 8 puntos por ítem.

Tabla 2
Operacionalización de la variable Dependiente

Variable	Dimensión	Indicador	Escala de Medición
Trazado correcto del número (Y)	Precisión	Porcentaje de dígitos trazados correctamente en cuanto a su forma estructural (semejanza con el modelo estándar)	Ordinal (Niveles 0-2)
	Secuencia de orden (So)	Número de dígitos en los que el estudiante respeta el orden lógico y progresivo de los segmentos del trazo (inicio, dirección y orden de líneas)	Ordinal (Niveles 0-2)
	Secuencia de fluidez (Sf)	Número de dígitos trazados de forma fluida y continua, sin interrupciones ni fragmentaciones del trazo	Ordinal (Niveles 0-2)
	Orientación	Porcentaje de dígitos escritos con orientación correcta (no invertidos ni rotados)	Ordinal (Niveles 0-2)

Nota. Elaboración Propia.

3. DISEÑO METODOLÓGICO

3.1. Tipo de Investigación

La presente investigación se tipifica como cuantitativa, aplicada, de nivel explicativo y con un diseño cuasi-experimental de tipo pretest-posttest con un solo grupo intacto.

Enfoque Cuantitativo

El estudio adopta un enfoque cuantitativo, ya que se centra en la medición objetiva del efecto que tiene el uso de la aplicación educativa sobre el trazado correcto del número en estudiantes de primer grado. El enfoque cuantitativo permite recoger datos numéricos, analizarlos estadísticamente y contrastar hipótesis con base en evidencia empírica. Según Hernández (2014), el enfoque cuantitativo utiliza la recolección de datos para probar hipótesis, con base en la medición numérica y el análisis estadístico, para establecer patrones de comportamiento y probar teorías. En este sentido, se busca determinar si existen diferencias significativas entre el desempeño de los estudiantes antes y después de la intervención con la app, lo que exige un análisis estadístico riguroso de los resultados.

Tipo de Investigación: Aplicativo

La investigación es de tipo aplicativo puesto que tiene como finalidad proponer e implementar una solución tecnológica concreta, la aplicación NumIA, para mejorar un problema identificado en el entorno educativo: las dificultades de los niños de 1° grado para realizar correctamente el trazado de los números. De acuerdo con Hernández (2014), la investigación aplicada se realiza con el propósito de resolver problemas prácticos inmediatos y busca generar conocimientos útiles para la acción.

Nivel de Investigación: Explicativo

La investigación explicativa es aquella que tiene relación causal, no solo persigue describir o acercarse a un problema, sino que intenta precisar las causas del mismo (Guevara Alban et al., 2020). El nivel de la investigación es explicativo porque no solo describe ni relaciona variables, sino que busca comprobar el efecto causal del uso de la app NumIA en el trazado correcto del número. En otras palabras, pretende identificar si el uso de la aplicación produce mejoras significativas en la variable dependiente.

Diseño de Investigación: Cuasi-experimental

Se ha optado por un diseño cuasi-experimental de tipo pretest-posttest con un solo grupo intacto, debido a las condiciones propias del contexto educativo (no se puede asignar aleatoriamente a los estudiantes a un grupo control y uno experimental). En este diseño, se aplica una prueba antes y después de la intervención al mismo grupo, con el fin de comparar los resultados y medir el impacto del uso de la app. Para Tamayo (2004), el diseño cuasi-experimental es apropiado cuando no se tiene un control total de las variables, pero se puede aplicar una intervención y observar sus efectos en un grupo específico.

3.2. Diseño de Contrastación de Hipótesis / Procedimiento a Seguir

El procedimiento de investigación se estructuró en seis pasos sistemáticos:

Paso 1: Selección y Caracterización de la Muestra

La muestra estuvo compuesta por estudiantes de primer grado de primaria de la Institución Educativa Pública Fe y Alegría 28 en Chiclayo, Lambayeque, seleccionada mediante un muestreo no probabilístico intencional. Se consideraron criterios específicos para garantizar la homogeneidad y relevancia del grupo:

Edad: niños entre 6 y 7 años

Acceso a un dispositivo móvil, una tableta de la institución

Estas características son fundamentales para asegurar que los participantes cumplan con los requisitos mínimos de la intervención tecnológica, tal como lo recomienda Hernández (2014), quien indica que el muestreo intencional es válido cuando se pretende estudiar grupos particulares en contextos específicos.

Paso 2: Aplicación del Pretest (Evaluación Inicial)

Antes de implementar la intervención con la aplicación móvil NumIA, se realizó una medición inicial (pretest) del trazado correcto de números arábigos a través de una rúbrica estandarizada.

Este pretest permitió obtener un diagnóstico inicial que sirve como referencia para evaluar posteriormente los cambios generados por la intervención educativa (Tamayo, 2004)

El pretest consistió en una prueba de desempeño diseñada para recopilar información base sobre la habilidad de trazo numérico de cada niño. Los estudiantes escribieron en papel números representativos, los cuales fueron evaluados mediante una rúbrica estandarizada considerando tres criterios: (a) precisión en la forma del numeral, (b) secuencia del trazo evaluada mediante dos subcriterios diferenciados: secuencia de orden (So.) y secuencia de fluidez (Sf.), y (c) orientación/colocación adecuada. Cada subcriterio recibió una puntuación de 0 a 2 puntos según niveles de logro predefinidos, con un puntaje máximo de 8 puntos por ítem, proporcionando medidas cuantitativas iniciales por estudiante y por criterio evaluado.

Paso 3: Implementación de la Intervención

La intervención de la aplicación fue en sesiones estructuradas donde los estudiantes utilizaron la aplicación NumIA durante un período de cinco semanas. Cada sesión tuvo una duración aproximada de 30 minutos y se realizó tres veces por semana, bajo la supervisión directa del docente responsable e investigadores del proyecto. La aplicación NumIA incluyó actividades interactivas basadas en deep learning que proporcionaron retroalimentación inmediata sobre el trazo numérico realizado por los estudiantes.

Este procedimiento sigue las recomendaciones metodológicas planteadas por estudios similares, como el de Sáez-López et al. (2019), donde se destacó la importancia de sesiones regulares, supervisadas y con retroalimentación constante para garantizar la efectividad del aprendizaje mediado por tecnología.

En estas sesiones, los estudiantes utilizaron dispositivos táctiles (tabletas) para practicar la escritura de números con un lápiz digital, siguiendo las indicaciones de la aplicación. NumIA integró algoritmos de reconocimiento de patrones basados en Deep Learning para analizar en tiempo real los trazos realizados en la pantalla. La aplicación proporcionó retroalimentación

automática e inmediata: cada vez que el niño dibujó un número, el sistema verificó el trazo indicando si fue correcto mediante un icono de check y mensaje aprobatorio, o si hubo algún error en la formación, secuencia u orientación.

Adicionalmente, la app incorporó ejercicios lúdicos y elementos de gamificación para mantener alta la motivación de los niños, incluyendo minijuegos, medallas virtuales por progreso y niveles de dificultad graduados. Estas características lúdicas buscaron que la práctica repetitiva fuera atractiva y se generara un ambiente de aprendizaje activo.

Paso 4: Aplicación del Postest (Evaluación Final)

Finalizada la intervención con la app NumIA, se realizó una evaluación final (postest) utilizando el mismo instrumento aplicado en la evaluación inicial, garantizando así la comparabilidad de los resultados, el cual permite determinar objetivamente si el uso de la aplicación tuvo un impacto significativo sobre el desempeño de los estudiantes en el trazado numérico.

El uso del mismo instrumento en ambas mediciones es una estrategia válida en diseños cuasi-experimentales de tipo pretest–postest, ya que permite comparar los efectos de una intervención en un mismo grupo de participantes, evaluando posibles mejoras de forma directa y controlada (Alessandri et al., 2017).

Paso 5: Análisis Estadístico

Se aplicó la prueba t de Student para muestras relacionadas (t pareada) con el propósito de comparar los resultados obtenidos en el pretest y postest del mismo grupo de estudiantes. Esta prueba estadística paramétrica es adecuada cuando se busca determinar si existe una diferencia significativa entre dos mediciones aplicadas al mismo grupo de forma consecutiva.

Previamente, se verificó el cumplimiento del supuesto de normalidad de las diferencias mediante la prueba de Shapiro-Wilk, recomendada para muestras pequeñas ($n < 50$). Si el valor de significancia (p) es igual o superior a 0.05, se consideró que los datos se distribuyen normalmente, pudiendo aplicar la prueba t pareada con validez estadística. Esta metodología fue aplicada en la tesis de Estañiz (2020), quien señala: "Comprobamos la normalidad de los datos con la prueba Shapiro Wilk para muestras pequeñas menores de 50 personas. Si $p \geq 0.05$, los datos corresponden a una distribución normal."

Paso 6: Discusión de Resultados

La discusión de los resultados se orientó a comparar los hallazgos del presente estudio con investigaciones previas, así como con los postulados teóricos sobre el aprendizaje del trazado numérico y el uso de tecnologías educativas. En caso de encontrar diferencias significativas tras la aplicación de la prueba t , se pudo afirmar que el uso de la aplicación NumIA favoreció el desarrollo del trazado correcto del número, cumpliendo así con los objetivos y confirmando la hipótesis general del estudio.

3.3. Población y Muestra

3.3.1. Población

La población del presente estudio estuvo conformada por todos los estudiantes matriculados en el primer grado de educación primaria de la Institución Educativa Fe y Alegría N.º 28, ubicada en el distrito de Chiclayo, región Lambayeque, durante el año escolar 2025. Este grupo representó un total de 90 estudiantes, distribuidos en tres aulas de aproximadamente 30 estudiantes cada una.

3.3.2. Muestra

La muestra seleccionada para la investigación fue no probabilística, por conveniencia, compuesta por 30 estudiantes, quienes participaron del proceso experimental. Esta selección respondió a criterios de accesibilidad y viabilidad operativa, permitiendo realizar el trabajo de campo en un solo grupo intacto, minimizando las interferencias que podrían alterar el desarrollo natural de las clases.

Tabla 3

Distribución de la muestra por sección

Sección	Estudiantes matriculados	Seleccionados para la muestra
1ºA	30	10
1ºB	30	10
1ºC	30	10
Total	90	30

Nota. Elaboración Propia.

La distribución equitativa de 10 estudiantes por sección garantiza la comparabilidad entre grupos y proporciona potencia estadística suficiente para detectar efectos de magnitud mediana o grande en las pruebas de contraste de hipótesis, según los criterios de Cohen (1988) para diseños de medidas repetidas con $\alpha = 0.05$.

Los 10 estudiantes seleccionados por sección no fueron elegidos de forma arbitraria, sino a partir del criterio de representatividad de la distribución real de niveles de desempeño del aula. Así, la muestra de cada sección incluyó estudiantes ubicados en los cuatro niveles de logro del Currículo Nacional En Inicio (C), En Proceso (B), Logro Esperado (A) y Logro Destacado (AD) en proporciones que reflejan la heterogeneidad natural de cada grupo, lo que fortalece la validez interna de los resultados.

Criterios de Inclusión:

- Estar matriculado/a en primer grado de primaria en la institución educativa participante

- Tener 6 o 7 años de edad (rango etario correspondiente al primer grado regular)
- Contar con el consentimiento del docente encargado en la sección, además del asentimiento del niño o niña
- Asistencia regular a clases y disponibilidad para usar la aplicación NumIA durante el período de la intervención

Criterios de Exclusión:

- Inasistencia en las evaluaciones clave (ausencia en el pretest o posttest) o abandono de la intervención antes de su finalización
- Presentar alguna condición que impida el uso adecuado de la aplicación que pueda interferir con la participación efectiva y la medición del aprendizaje de números

3.4. Técnicas, Instrumentos, Equipos y Materiales

La investigación adoptó un enfoque cuantitativo con un diseño cuasi-experimental de único grupo, bajo un esquema pretest–posttest, llevado a cabo en una institución educativa pública de Lambayeque. A continuación, se describen las técnicas de recolección de datos a emplear, los instrumentos, los equipos y materiales.

Tabla 4
Técnicas de Recolección de Datos

Técnica	Descripción	Propósito
Evaluación de Desempeño	Pruebas escritas (pretest-posttest) de trazado numérico	Medir dominio cuantitativo del trazado antes/después de la intervención
Registro Automático	Datos generados por la app (tiempo de uso, progreso, errores)	Obtener métricas objetivas del desempeño individual

Nota. Elaboración Propia.

Tabla 5
Instrumentos

Instrumento	Componentes Evaluados	Escala
Rúbrica de Trazado Numérico	Orientación, Secuencia de orden (So.), Secuencia de fluidez (Sf.), Precisión	Ordinal de 4 subcriterios (0-2 puntos por subcriterio; máximo 8 puntos por ítem)
Pruebas de Desempeño (Pretest-Postest)	Escritura de números (0-100), Identificación	Puntaje numérico
System Usability Scale (SUS)	Facilidad de uso, eficiencia, satisfacción del usuario	Likert 1–5 (5 ítems)

Nota. Elaboración Propia.

La validez de contenido de la rúbrica de trazado numérico fue establecida mediante juicio de expertos, siguiendo el procedimiento propuesto por Aiken (1985). Se consultó a tres especialistas con experiencia en educación primaria, quienes valoraron cada uno de los cuatro criterios del instrumento orientación, secuencia de orden, secuencia de fluidez y precisión en términos de claridad, pertinencia y relevancia, utilizando una escala del 1 al 4. Los coeficientes V de Aiken obtenidos oscilaron entre 0.815 y 0.963, superando en todos los casos el umbral de aceptabilidad de 0.80 establecido por la literatura especializada, lo que confirma la validez de contenido del instrumento.

Tabla 6*Resultados de validación de contenido por juicio de expertos (V de Aiken)*

Criterio	Experto 1	Experto 2	Experto 3	V de Aiken	Decisión
Orientación	4.00	4.00	3.67	0.963	Válido
Secuencia de orden	4.00	4.00	3.67	0.963	Válido
Secuencia de fluidez	4.00	3.67	4.00	0.963	Válido
Precisión	4.00	3.00	3.33	0.815	Válido
GLOBAL				0.926	Válido

Nota. V de Aiken calculado con escala 1–4 en dimensiones: Claridad, Pertinencia y Relevancia (Aiken, 1985). $V \geq 0.80$ indica validez aceptable. Elaboración propia.

La validez de contenido de la rúbrica de trazado numérico fue establecida mediante juicio de expertos, siguiendo el procedimiento propuesto por Aiken (1985). Se consultó a tres especialistas con experiencia en educación primaria, de 20, 12 y 4 años de experiencia en dicho nivel educativo. Cada especialista valoró los cuatro criterios del instrumento orientación, secuencia de orden, secuencia de fluidez y precisión en términos de claridad, pertinencia y relevancia, utilizando una escala del 1 al 4. Los coeficientes V de Aiken obtenidos oscilaron entre **0.815 y 0.963**, superando en todos los casos el umbral de aceptabilidad de 0.80, con un coeficiente global de **V = 0.926**, lo que confirma la validez de contenido del instrumento.

Tabla 7*Equipos y Materiales*

Categoría	Especificaciones/Componentes	Propósito
Dispositivos Móviles	Tablets/smartphones (Android 10+, 8", RAM 4GB)	Ejecutar la app NumIA en sesiones
Software	App NumIA (Flutter, reconocimiento de trazos con Deep Learning mediante Redes Siamesas)	Intervención educativa y retroalimentación automática

Materiales	Cuadernillos de evaluación, guías de uso,	Evaluación inicial/final y
Impresos	formularios de juicio de expertos, cuestionario SUS impresos	validación de instrumentos
Equipos	Proyector, computadora portátil	Instrucción grupal
Adicionales		

Nota. Elaboración Propia.

3.5. Aspectos Éticos de la Investigación

La presente investigación se condujo bajo estrictos principios éticos que garantizaron el respeto, la dignidad y el bienestar de los participantes, especialmente considerando que se trabajó con menores de edad. Los aspectos éticos considerados fueron:

✓ Consentimiento Informado

Se solicitó el consentimiento informado a los docentes a cargo de los estudiantes. Además, se obtuvo el asentimiento verbal de los niños participantes, explicándoles de manera comprensible y apropiada para su edad en qué consistiría su participación.

✓ Confidencialidad y Protección de Datos

- Toda la información recopilada fue tratada con estricta confidencialidad
- Los datos personales de los estudiantes fueron codificados y almacenados de manera segura
- Solo los investigadores principales tuvieron acceso a la información identificable.
- Los datos almacenados en la aplicación NumIA cumplieron con normativas de protección de datos personales

✓ Beneficencia y No Maleficencia

- La intervención buscó maximizar los beneficios educativos para los estudiantes, proporcionando una herramienta adicional de aprendizaje
- Se minimizaron los riesgos, asegurando que la aplicación fuera apropiada para la edad y nivel cognitivo de los participantes.
- El uso de la aplicación no interfirió negativamente con el plan de estudios regular ni sobrecargó a los estudiantes.

✓ **Equidad y Justicia**

- Todos los estudiantes del aula tuvieron la misma oportunidad de participar en la intervención
- No se discriminó por razones de género, nivel socioeconómico, rendimiento académico previo u otras características personales.

✓ **Transparencia y Honestidad Científica**

- Los datos fueron recopilados y analizados con rigor metodológico
- Los resultados se reportaron de manera íntegra, sin manipulación ni falsificación
- Se reconocieron las limitaciones del estudio
- Se citaron adecuadamente todas las fuentes utilizadas

✓ **Aprobación Institucional**

Se obtuvieron los permisos necesarios de:

- La dirección de la Institución Educativa Fe y Alegría N.º 28
- Los docentes de cada aula de 1er grado.

✓ **Retorno de Resultados**

Se comprometió a:

- Dejar la aplicación NumIA disponible para uso continuo de la institución educativa.

4. RESULTADOS

4.1. Búsqueda de estrategias pedagógicas para la enseñanza del trazo de números

Para dar cumplimiento al primer objetivo específico: "Identificar y analizar las estrategias pedagógicas utilizadas por docentes de primer grado para enseñar el trazo de números, especialmente con el apoyo de tecnologías digitales", se realizó técnicas de recolección, las cuales fueron:

- Revisión sistemática de literatura pedagógica especializada en enseñanza de escritura numérica
- Análisis documental de investigaciones nacionales e internacionales sobre prácticas docentes
- Sistematización de hallazgos reportados en antecedentes de investigación
- Análisis de manuales y guías pedagógicas del Ministerio de Educación del Perú

4.1.1. Estrategias Pedagógicas Tradicionales Identificadas

4.1.1.1. *Práctica Repetitiva mediante Cuadernos de Caligrafía*

Los docentes utilizan principalmente cuadernos de aprestamiento donde los estudiantes:

- Trazan números siguiendo líneas punteadas prediseñadas
- Realizan ejercicios de "planas" (repetición del mismo número en múltiples renglones)
- Progresan desde trazos grandes hasta más pequeños.

Villarroel et al. (2013) documenta que el primer grado es una etapa decisiva para consolidar patrones motores correctos en la escritura de números. La práctica repetitiva con cuadernos de aprestamiento es reconocida como una de las estrategias más utilizadas en aulas de educación primaria para lograr este objetivo (MINEDU, 2016).

4.1.1.2. *Demostración y Modelado del Docente*

Los docentes realizan demostraciones explícitas del trazado correcto:

- Dibujo del número en pizarra con verbalización del proceso
- Indicación de punto de inicio y dirección del trazo
- Uso de flechas y marcadores de colores para señalar la secuencia

Santillana Educación (n.d.) en su manual de matemáticas para primaria recomienda que los docentes pongan especial atención en el punto de inicio y el sentido del trazo al escribir cada número, como parte de las orientaciones pedagógicas estándar implementadas en instituciones educativas.

4.1.1.3. *Material Manipulativo y Multisensorial*

Los docentes incorporan materiales concretos para el aprendizaje:

- Números en lija o texturas para repasar con el dedo
- Plastilina o arcilla para modelar números
- Palitos, cuerdas o materiales para formar números
- Tarjetas con números en diversos tamaños

Sinvani et al. (2023) documentan que "el aprendizaje de la escritura de números es un proceso complejo que integra habilidades cognitivas, perceptuales y motoras finas", y que los docentes utilizan enfoques multisensoriales que combinan percepción visoespacial, discriminación de tamaño y coordinación visomotora para facilitar este aprendizaje.

4.1.1.4. *Uso de Color como Recurso Didáctico*

Los docentes emplean colores llamativos en materiales educativos:

- Carteles numéricos con colores brillantes
- Códigos de color para diferentes segmentos del número
- Decoración del aula con números coloridos
- Materiales didácticos con diseños atractivos

Ortiz (2006) reporta que en la práctica docente "los colores llamativos o fuertes actúan como estímulos didácticos que favorecen la atención de los estudiantes y refuerzan los procesos de aprendizaje", funcionando además como claves mnemotécnicas que facilitan la retención de contenidos.

4.1.1.5. *Sistema de Evaluación por Criterios*

Los docentes evalúan el trazado numérico considerando tres dimensiones principales: precisión, secuencia y orientación. En la presente investigación, la dimensión secuencia se desagrega en dos subcriterios para mayor granularidad diagnóstica orden del trazo y fluidez del trazo, siguiendo la distinción establecida por Chang & Yu (2022).

Criterio 1: Precisión de la grafía

- Claridad y reconocibilidad del número
- Semejanza con la forma convencional
- Ausencia de distorsiones significativas

Calero-Brito et al. (2024) señalan que "el número debe ser reconocible y estar bien formado, sin distorsiones, lo cual es clave para un buen desempeño académico".

Criterio 2: Secuencia del trazo

- Orden correcto de inicio del trazo
- Dirección apropiada (arriba-abajo, izquierda-derecha)
- Continuidad sin interrupciones innecesarias

Para efectos de medición en la rúbrica, este criterio se operacionaliza en dos subcriterios: la secuencia de orden, que evalúa el aspecto procedimental (inicio, dirección y orden de segmentos), y la secuencia de fluidez, que evalúa la continuidad y automatización del movimiento. Esta distinción responde a que ambos aspectos pueden fallar de forma independiente: un estudiante puede trazar los segmentos en orden correcto, pero con discontinuidades, o viceversa

Hederich-Martínez et al. (2016) documentan que "para lograr la correcta escritura de números arábigos es necesario establecer rutinas motoras específicas para cada dígito", enfatizando que los docentes evalúan el seguimiento de estas rutinas.

Criterio 3: Orientación correcta

- Ausencia de inversiones (números en espejo)
- Posición apropiada sin rotaciones
- Reconocimiento de la forma estándar

Zapateiro-Segura et al. (2018) reportan que "las inversiones están vinculadas al desarrollo de la orientación espacial" y son evaluadas por los docentes como indicadores del dominio grafomotor, aunque "no deben considerarse patológicas en los primeros años, ya que forman parte del proceso de adquisición de la lateralidad".

El Ministerio de Educación (2022) establece en sus orientaciones pedagógicas que "las rúbricas de evaluación describen indicadores y lo que se espera en distintos niveles para cada competencia o criterio a evaluar", siendo este el marco que guía la evaluación docente del trazado numérico.

4.1.2. Caracterización de la Retroalimentación Docente

4.1.2.1. Naturaleza Correctiva de la Retroalimentación

Llanos Masciotti & Tapia (2021) documentan que en el contexto educativo peruano "la interacción docente-estudiante suele limitarse a una participación reducida del alumno, y la

retroalimentación brindada por muchos maestros es meramente correctiva", lo que significa que:

- El docente indica únicamente si algo está bien o mal
- No se ofrecen orientaciones específicas para la mejora
- Las correcciones no son inmediatas durante la ejecución
- El estudiante no comprende qué aspecto específico debe mejorar

4.1.2.2. Limitaciones Contextuales Documentadas

Ratio estudiante-docente:

El Ministerio de Educación (2022) destaca un punto crucial: en entornos donde la cantidad de estudiantes por docente es alta (más de 30 alumnos por profesor), la atención individualizada se convierte en un desafío. La complejidad de ofrecer retroalimentación precisa y oportuna a cada estudiante constituye un desafío estructural en este contexto.

Esta situación presenta algunas implicaciones importantes:

- Retroalimentación en Diferido: La revisión de trabajos, como los cuadernos, debe realizarse fuera del horario de clase.
- Observación Limitada: La oportunidad de guiar el proceso de aprendizaje en tiempo real se reduce.
- Identificación Retrasada: Los errores, una vez consolidados, pueden ser más difíciles de corregir.
- Aprendizaje Independiente: Los estudiantes practican de manera autónoma, con menos oportunidades de recibir guía inmediata.

Marco teórico del feedback formativo:

En contraste, Hattie & Timperley (2007) establecen que el feedback formativo efectivo debe:

1. Ser específico respecto al nivel de desempeño
2. Entregarse durante la ejecución de la tarea

3. Orientar sobre cómo mejorar
4. Regular el proceso de aprendizaje

4.1.3. Elementos Efectivos del Aprendizaje Multimedia

4.1.3.1.1. Teoría del Aprendizaje Multimedia

Hallazgo teórico de Mayer (2010):

Define el aprendizaje multimedia como el proceso mediante el cual "los estudiantes construyen representaciones mentales a partir de palabras e imágenes".

- Tres tipos de almacenaje en memoria: sensorial, trabajo, largo plazo
- Dos canales separados: verbal y visual
- Cada canal procesa pequeñas cantidades a la vez
- Pueden apoyarse mutuamente procesando información complementaria

Procesos cognitivos clave:

1. **Selección:** Percibir palabras e imágenes relevantes
2. **Organización:** Crear modelos verbales lógicos y representaciones visuales
3. **Integración:** Combinar ambas representaciones con conocimiento previo

Aplicación al trazo numérico:

- Presentar número visualmente (imagen)
- Acompañar con instrucción verbal (narración)
- Integrar experiencia kinestésica (dibujar)
- Proporcionar retroalimentación multimodal

4.1.3.2. Rol del Color en el Aprendizaje

Hallazgo de Ortiz (2006)

Documenta que "los colores llamativos o fuertes actúan como estímulos didácticos que favorecen la atención de los estudiantes y refuerzan los procesos de aprendizaje", además de

que "el color puede funcionar como una clave mnemotécnica, facilitando la retención de contenidos".

Aplicación pedagógica identificada:

- Carteles con números en colores vivos
- Códigos de color para segmentos del trazo
- Diferenciación visual entre números similares
- Retroalimentación cromática (verde=correcto, rojo=revisar)

4.1.3.3. *Importancia de la Música*

Hallazgo de Dzib Moo et al. (2023):

Reportan que "la inclusión adecuada de la música en entornos digitales de aprendizaje también desempeña un papel relevante, ya que es una herramienta efectiva para captar y mantener el interés de los estudiantes durante el proceso de aprendizaje".

Tipos de aplicación musical:

- Música de fondo ambiental (concentración)
- Efectos sonoros de retroalimentación
- Narración y pronunciación de números
- Celebraciones sonoras de logros

4.1.3.4. *Poder de la Repetición*

Hallazgo de Vargas-Murillo (2020):

Documenta que "repetir ejercicios y contenidos favorece la asimilación progresiva del conocimiento, el desarrollo de la comprensión y la retención de la información".

Hallazgo de Robert, B. (2001):

Establece que "la repetición no solo mejora la retención de información, sino que también permite al estudiante construir significados de forma progresiva y sostenible".

Aplicación al trazo numérico:

- Práctica repetida consolida memoria muscular
- Exposición múltiple al mismo número en contextos variados
- Clave: Repetición motivante, no tediosa (mediante gamificación)

4.1.4. Síntesis de Hallazgos del Resultado

En respuesta al primer objetivo específico de identificar y analizar las estrategias pedagógicas utilizadas por docentes de primer grado para enseñar el trazo de números, se identificaron cinco estrategias pedagógicas tradicionales con fundamento teórico y aplicación documentada: práctica repetitiva con cuadernos de aprestamiento, demostración y modelado docente en pizarra, uso de material manipulativo multisensorial, empleo de color como recurso didáctico, y evaluación por criterios de precisión, secuencia y orientación.

En conclusión, el análisis realizado evidencia que, si bien las estrategias pedagógicas tradicionales tienen fundamento teórico sólido, presentan limitaciones estructurales críticas relacionadas con la imposibilidad de proporcionar retroalimentación individualizada, inmediata y formativa en contextos de aulas numerosas. La evidencia empírica nacional e internacional demuestra que las aplicaciones móviles educativas pueden complementar efectivamente estas estrategias tradicionales, llenando las brechas identificadas mediante retroalimentación automática, práctica adaptativa y elementos motivacionales, justificando así el desarrollo de herramientas tecnológicas como NumIA que integren los fundamentos teóricos del aprendizaje multimedia con las necesidades pedagógicas documentadas.

4.2. Diseño y desarrollo de la aplicación móvil "NumIA".

Para dar cumplimiento al segundo objetivo específico: **"Diseñar y desarrollar la aplicación móvil "NumIA", incorporando una interfaz gráfica amigable y módulos**

interactivos de práctica de trazos numéricos evaluados mediante un modelo de inteligencia artificial basado en Redes Siamesas.", se realizó:

- Diseño del modelo de deep learning (redes siamesas)
- Diseño de la Base de Datos
- Diseño y desarrollo de la aplicación móvil

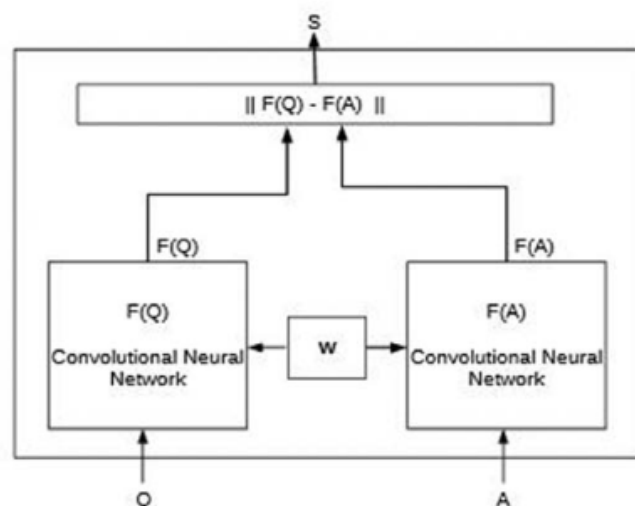
4.2.1. Diseño del modelo de deep learning (redes siamesas)

Una **Red Siamesa** es una arquitectura de deep learning compuesta por dos (o más) subredes idénticas que comparten parámetros (pesos y arquitectura), lo cual garantiza que ambas aprendan de manera sincrónica las mismas características. Cada subred procesa una entrada independiente y genera un vector de características o embedding. Posteriormente, una función de similitud como la **distancia Euclídea** o la **similitud del coseno** compara estos embeddings para determinar el grado de semejanza entre las entradas. (Bromley et al., 1993).

Según De Rosa & Papa (2018), este tipo de redes consisten en un par de redes neuronales que comparten sus pesos y tienen como objetivo calcular funciones de similitud. En esencia, su objetivo principal es identificar si un par de datos es diferente. La figura 1 ilustra un ejemplo de una arquitectura de red siamesa.

Figura 1

Ejemplo de arquitectura de red siamesa



Nota. Tomado de "A Siamese Neural Network Application for Sales Forecasting of New Fashion Products Using Heterogeneous Data" (p. 1539), por G. Craparotta, S. Thomassey y A. Biolatti, 2019, International Journal of Computational Intelligence Systems, 12(2). CC BY 4.0.

El objetivo principal de esta arquitectura es que las entradas pertenecientes a la misma clase presenten embeddings, las cuales son representaciones numéricas de datos que captan el significado o las características de los objetos, cercanos en el espacio latente, mientras que entradas de clases distintas se ubican alejadas. Haciéndolas útiles en tareas de **verificación** y **comparación** más que de clasificación tradicional.

4.2.1.1. Fundamentos del Modelo

La elección de una red siamesa se fundamenta en los siguientes puntos:

- Naturaleza del problema: no se busca solo clasificar un dígito entre 0–9, sino verificar si el trazo cumple con las características correctas de escritura.
- Eficiencia de datos: al trabajar con pocas muestras de referencia (guías predefinidas y trazos de niños), se evita entrenar un clasificador complejo que requiera miles de ejemplos por clase.
- Robustez a la variabilidad: los niños difieren en tamaño, orientación, presión y estilo de escritura. Al comparar contra un embedding normalizado, la red siamesa aprende a abstraer esas diferencias superficiales.
- Antecedentes sólidos: la literatura en verificación de firmas confirma que este enfoque es eficaz para tareas de comparación de patrones, donde el objetivo es determinar similitud y no necesariamente clasificar (Dey et al., 2017) .
- Escalabilidad: una vez entrenado el backbone, se pueden agregar nuevas variantes de trazos sin necesidad de reentrenar el modelo completo, lo que facilita la extensión a otros símbolos o alfabetos.

Para el caso de la app NumIA, el desafío no radica en clasificar un dígito escrito, sino en verificar si el trazo realizado por un niño es suficientemente similar al trazo “correcto” esperado para un número arábigo. Esto se alinea con la naturaleza comparativa de las redes siamesas, pues el objetivo de la aplicación es ayudar a mejorar el trazado, y para ello se hace una verificación del número escrito por el niño para comprobar si existe una semejanza aceptable con un número escrito adecuadamente, entonces con ello se puede realizar la comparación.

4.2.1.2. *Recolección de Datos*

La recolección de datos se estructuró en dos fuentes complementarias: (a) guías base de referencia para cada dígito (0–9), y (b) trazos reales capturados en un lienzo digital. Ambas fuentes comparten decisiones de diseño que favorecen la comparación geométrica y la normalización espacial, requisitos fundamentales para el entrenamiento de la red siamesa.

a) Guías Base de Referencia

- ❖ Creación de patrones correctos para cada dígito (0-9)
- ❖ Inclusión de variantes multitrazo (ejemplo: "7 con rayita", "8 de dos círculos")
- ❖ Herramienta utilizada: Editor de guías en Python/Tkinter
- ❖ Funcionalidades del editor:
 - Dibujar múltiples trazos por variante
 - Almacenamiento en coordenadas normalizadas
 - Formato JSON canónico. (lista de variantes, luego lista de trazos y finalmente lista de puntos).
 - Previsualización y edición de variantes por dígitos.
 - Autoguardado opcional para evitar pérdida de trabajo.

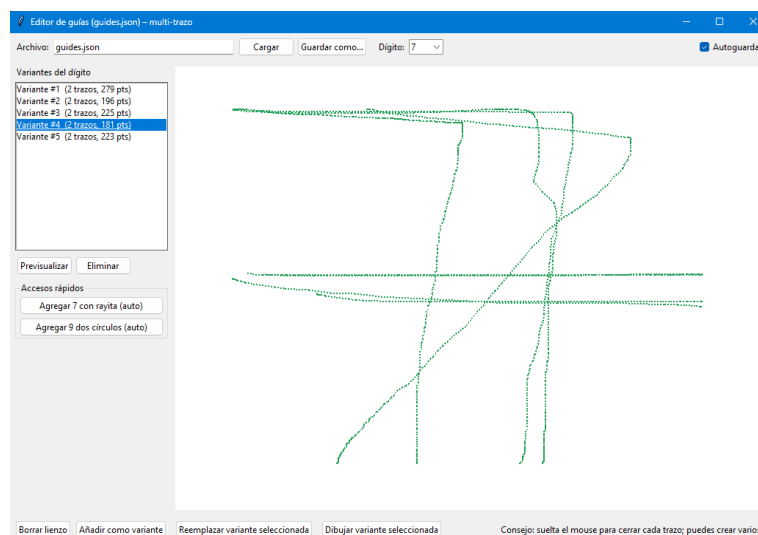
Criterios de calidad para guías:

- **Trazabilidad perceptiva:** la forma debe ser claramente identificable como el dígito objetivo.

- **Consistencia topológica:** conservar la estructura típica (p. ej., el “8” como lazo continuo o dos bucles coherentes).
- **Multitrazo cuando corresponde:** respetar convenciones caligráficas (p. ej., el “7 con rayita” se compone de trazo principal + rayita).
- **Cobertura de variantes:** incluir variantes útiles para guiar al modelo frente a diferentes estilos esperables en niños.

Figura 2

Ejemplo de multitrazo



Nota. Elaboración propia.

b) Trazos de Usuarios

El propósito de esta fuente es registrar ejemplos reales de escritura sobre un lienzo digital, preservando la dinámica de multitrazo y la secuencia de puntos de lápiz, con el fin de reflejar fielmente el comportamiento natural de escritura.

Se utilizó como herramienta un Grabador de trazos (Python/Tkinter)

Características:

- Lienzo 800×600 px
- Lápiz de ancho constante (10 px)
- Gestión de multitrazo por muestra (cada pulsar/arrastrar/soltar = 1 trazo).
- Campos de metadatos (dígito, variante, timestamp)
- Hotkeys y estadísticas en tiempo real (número de trazos y puntos) para mejorar la calidad de la captura.
- Formato JSONL para append incremental

Razonamiento del diseño:

- Se almacena en **coordenadas de píxel** tal cual se dibuja para preservar máxima fidelidad.
- La **normalización** (0 y 1, Y-arriba, por **bounding box**) se realiza **a posteriori** en el pipeline de preprocesamiento, garantizando invariancia a traslación/escala y homogeneidad geométrica en el entrenamiento.
- La etiqueta de **variante** es opcional y permite estudios de **análisis por estilo** (p. ej., comparar “7 estándar” vs. “7 con rayita”).
- El uso de JSONL simplifica **auditoría, versionado y reprocesamiento** (línea a línea).

El dataset de entrenamiento se conformó a partir de dos fuentes complementarias guías base de referencia para cada dígito del 0 al 9, con múltiples variantes caligráficas por dígito generadas mediante el editor Python/Tkinter, y trazos reales capturados en lienzo digital de 800×600 px mediante el grabador de trazos descrito anteriormente. A partir de ambas fuentes se construyó un conjunto de pares siameses balanceado 50/50 entre pares positivos y negativos, generando aproximadamente 25,000 pares por época. El modelo fue entrenado durante 10 épocas utilizando el optimizador Adam con learning rate de 1×10^{-3} (Kingma & Lei Ba, 2015), batch size de 32 pares y 400 pasos por época, con generación de pares on-the-fly para

maximizar la variabilidad de los ejemplos presentados en cada iteración. La función de pérdida empleada fue la Contrastive Loss (Hadsell et al., 2006) con margen $m = 0.7$, que penaliza la distancia euclídea entre embeddings de pares positivos y fuerza una separación mínima en los pares negativos. El modelo final quedó almacenado en el archivo `siamese_backbone.keras` para su posterior despliegue en el backend de la aplicación.

4.2.1.3. *Preprocesamiento de Datos*

a) Normalización por Bounding Box

- Transformación de coordenadas píxel $\rightarrow$ coordenadas 0 y 1
- Inversión eje Y (convención Y-arriba)
- Cálculo de caja envolvente global
- Fórmula aplicada:

$$x' = \frac{x/W - x_{\min}}{x_{\max} - x_{\min} + \epsilon}, \quad y' = \frac{(1 - y/H) - y_{\min}}{y_{\max} - y_{\min} + \epsilon},$$

Donde W,H son las dimensiones del lienzo y ϵ evitar divisiones por cero. Esta operación confiere **invariante a traslación y escala**, y estandariza la convención geométrica (eje Y positivo hacia arriba).

b) Rasterización

- Proyección a lienzo cuadrado 256×256
- Margen $m=0.12$
- Anti-aliasing y grosor fijo
- Reducción a 64×64 (interpolación área)
- Escala de grises (fondo negro, trazo blanco)

c) Aumento de Datos (Data Augmentation)

- Jitter afín consistente por muestra. (misma rotación/escala/traslación a todos los trazos de la muestra)

- Parámetros:
 - Rotación: $\pm 12^\circ$
 - Escala: $\times [0.9, 1.1]$
 - Traslación: ± 12 píxeles
- Preservación de estructura multitrazo, como sucede con el número 7, que tiene una “rayita”, pero que sigue alineada con su trazo principal, evitando de esta forma deformaciones internas.

4.2.1.4. *Generación de Pares Siameses*

Se construye un *dataset* balanceado 50/50 de pares positivos y pares negativos:

Pares Positivos ($x_1, x_2, y=1$):

- x_1 : guía (variante aleatoria) del dígito d
- x_2 : trazo usuario del mismo dígito d
- Fallback a guía distinta si no hay trazo disponible

Pares Negativos ($x_1, x_2, y=0$):

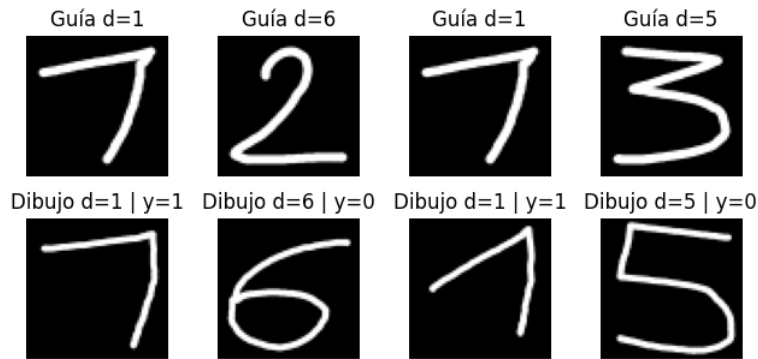
- Opción (i): Guía de d , x_2 de otro dígito
- Opción (ii): Cambiar guía a otro dígito, mantener x_2 de d
- Alternancia para evitar correlaciones triviales

Características:

- Tensores (64,64,1) en $[0,1]$
- Varianza intra-clase (múltiples variantes).
- Dificultad inter-clase razonable. (negativos con formas potencialmente confusas).

Figura 3
Generación de guías.

Shapes: (4, 64, 64, 1) (4, 64, 64, 1) (4, 1) y: [1. 0. 1. 0.] digits: ['1', '6', '1', '5']



Nota. Elaboración propia.

4.2.1.5. *Arquitectura del Modelo*

Se emplea un *backbone* CNN

Input: Imagen $64 \times 64 \times 1$

Bloque 1:

- Conv2D(32 filtros, kernel 3×3 , ReLU)
- Conv2D(32 filtros, kernel 3×3 , ReLU)
- MaxPooling2D(2×2)

Bloque 2:

- Conv2D(64 filtros, kernel 3×3 , ReLU)
- Conv2D(64 filtros, kernel 3×3 , ReLU)
- MaxPooling2D(2×2)

Bloque 3:

- Conv2D(128 filtros, kernel 3×3 , ReLU)
- GlobalAveragePooling2D()

Proyección:

- Dense(128 unidades)

- UnitNormalization (L2=1)

Output: Embedding $z \in \mathbb{R}^{128}$

Tabla 8
Arquitectura CNN de la Red Siamesa de NumIA

Nº	Capa	Tipo	Forma de salida	Parámetros
1	Entrada	InputLayer	(None, 64, 64, 1)	0
2	Conv2D-1	Conv2D 32 filtros, kernel 3×3, ReLU	(None, 62, 62, 32)	320
3	Conv2D-2	Conv2D 32 filtros, kernel 3×3, ReLU	(None, 60, 60, 32)	9,248
4	MaxPooling-1	MaxPooling2D ventana 2×2	(None, 30, 30, 32)	0
5	Conv2D-3	Conv2D 64 filtros, kernel 3×3, ReLU	(None, 28, 28, 64)	18,496
6	Conv2D-4	Conv2D 64 filtros, kernel 3×3, ReLU	(None, 26, 26, 64)	36,928
7	MaxPooling-2	MaxPooling2D ventana 2×2	(None, 13, 13, 64)	0
8	Conv2D-5	Conv2D 128 filtros, kernel 3×3, ReLU	(None, 11, 11, 128)	73,856
9	GlobalAvgPooling	GlobalAveragePooling2D	(None, 128)	0
10	Dense	Dense activación lineal	(None, 128)	16,512
11	UnitNorm (L2)	UnitNormalization	(None, 128)	0
Total parámetros				155,360

Nota. El backbone es compartido por ambas subredes de la arquitectura siamesa. Elaboración propia.

Red Siamesa Completa:

- Dos entradas: x_1, x_2
- Backbone compartido (pesos idénticos)
- Cálculo de distancia euclídea:

$$d(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{z}_1 - \mathbf{z}_2\|_2.$$

Función de Pérdida: Contrastive Loss (Hadsell et al., 2006)

$$\mathcal{L} = y d^2 + (1 - y) [\max(0, m - d)]^2.$$

- Margen $m = 0.7$
- Positivos ($y=1$): empujados a distancia 0
- Negativos ($y=0$): empujados a distancia $\geq m$

4.2.1.6. *Entrenamiento del Modelo*

Configuración:

- Optimizador: Adam (learning rate 1×10^{-3}) (Kingma & Lei Ba, 2015)
- Generador on-the-fly
- Batch size: 32 pares
- Pasos por época: 400 (~25,000 pares/época)
- Número de épocas: 10
- Output: siamese_backbone.keras

Figura 4

Arquitectura del modelo.

```
# ----- HYPERPARAMS -----
MODEL_PATH = "/content/drive/MyDrive/ColabNotebooks/MODELOsiamese/siamese_backbone.keras"
EMBED_DIM = 128

def build_backbone():
    inp = layers.Input((64,64,1))
    x = layers.Conv2D(32,3,activation="relu")(inp)
    x = layers.Conv2D(32,3,activation="relu")(x)
    x = layers.MaxPool2D()(x)
    x = layers.Conv2D(64,3,activation="relu")(x)
    x = layers.Conv2D(64,3,activation="relu")(x)
    x = layers.MaxPool2D()(x)
    x = layers.Conv2D(128,3,activation="relu")(x)
    x = layers.GlobalAveragePooling2D()(x)
    x = layers.Dense(EMBED_DIM, activation=None)(x)
    out = layers.UnitNormalization(axis=-1)(x) # <- reemplaza tf.nn.l2_normalize
    return keras.Model(inp, out, name="backbone")

def contrastive_loss(y_true, y_pred, margin=0.7):
    # y_pred: distancia euclídea (>=0)
    y_true = ops.cast(y_true, "float32")
    pos = y_true * ops.square(y_pred)
    neg = (1.0 - y_true) * ops.square(ops.maximum(0.0, margin - y_pred))
    return ops.mean(pos + neg)

class SiameseModel(keras.Model):
    def __init__(self, backbone):
        super().__init__()
        self.backbone = backbone

    def call(self, inputs, training=None):
        x1, x2 = inputs
        z1 = self.backbone(x1, training=training)
        z2 = self.backbone(x2, training=training)
        diff = z1 - z2
        # distancia L2 usando keras.ops (compatible con KerasTensor)
        d = ops.sqrt(ops.sum(ops.square(diff), axis=-1, keepdims=True))
        return d

def main():
    pairs = pairs = SiamesePairs(GUIDES_PATH, STROKES_PATH)
    model = SiameseModel(build_backbone())
    model.compile(optimizer=keras.optimizers.Adam(1e-3),
                  loss=contrastive_loss)

    steps = 400 # ~25k pares/época (400*64)
    model.fit(data_generator(pairs, 32), steps_per_epoch=steps, epochs=10)
    model.backbone.save(MODEL_PATH) # guarda solo el backbone (embeddings)
    print("✅ Backbone guardado en", MODEL_PATH)

if __name__ == "__main__":
    # Mixed precision opcional
    # from keras import mixed_precision
    # mixed_precision.set_global_policy("mixed_float16")
    main()
```

Nota. Elaboración propia.

Figura 5
Entrenamiento del modelo.

```
Epoch 1/10
400/400 ————— 439s 1s/step - loss: 0.0447
Epoch 2/10
400/400 ————— 441s 1s/step - loss: 0.0139
Epoch 3/10
400/400 ————— 441s 1s/step - loss: 0.0078
Epoch 4/10
400/400 ————— 435s 1s/step - loss: 0.0059
Epoch 5/10
400/400 ————— 436s 1s/step - loss: 0.0047
Epoch 6/10
400/400 ————— 435s 1s/step - loss: 0.0042
Epoch 7/10
400/400 ————— 439s 1s/step - loss: 0.0035
Epoch 8/10
400/400 ————— 434s 1s/step - loss: 0.0029
Epoch 9/10
400/400 ————— 433s 1s/step - loss: 0.0026
Epoch 10/10
400/400 ————— 433s 1s/step - loss: 0.0022
✓ Backbone guardado en /content/drive/MyDrive/ColabNotebooks/MODELOsiamese/siamese_backbone.keras
```

Nota. Elaboración propia.

Justificación hiperparámetros:

- Learning rate: convergencia estable sin oscilaciones
- Batch size: balance entre memoria y estabilidad
- Épocas: suficientes para convergencia sin overfitting

4.2.1.7. Evaluación del Modelo

a) Evaluación Offline:

- Cálculo de distancia euclídea: guías vs. trazos test
- Curvas ROC (variando umbral τ)
- Métricas por umbral:
 - TPR (sensibilidad)
 - FPR
 - J de Youden: $J = \text{TPR} - \text{FPR}$ (Youden, 1949)

Selección de umbral:

- $\tau \approx 0.60$ (punto de compromiso)
- Maximiza J de Youden
- Baja tasa falsos positivos en dígitos similares

b) Evaluación en Vivo:

Flujo de inferencia:

1. Dibujo en canvas (multi-trazo)
2. Normalización por bounding box (0..1, eje Y↑)
3. Rasterización $256 \times 256 \rightarrow 64 \times 64$
4. Backbone genera embedding unitario
5. Comparación contra todas variantes/guías
6. Distancia euclídea a mejor variante por dígito

Regla de decisión anti-confusión:

- d_t : mejor distancia al dígito objetivo
- $d_{(2)}$: mejor distancia entre demás dígitos

Criterios de aceptación:

1. **Absoluto:** $d_t < \tau$ (suficientemente parecido)
2. **Relativo:** $d_t \leq \rho \cdot d_{(2)}$ (más parecido al objetivo; $\rho < 1$)
3. **Margen:** $d_{(2)} - d_t \geq \gamma$ (separación clara)

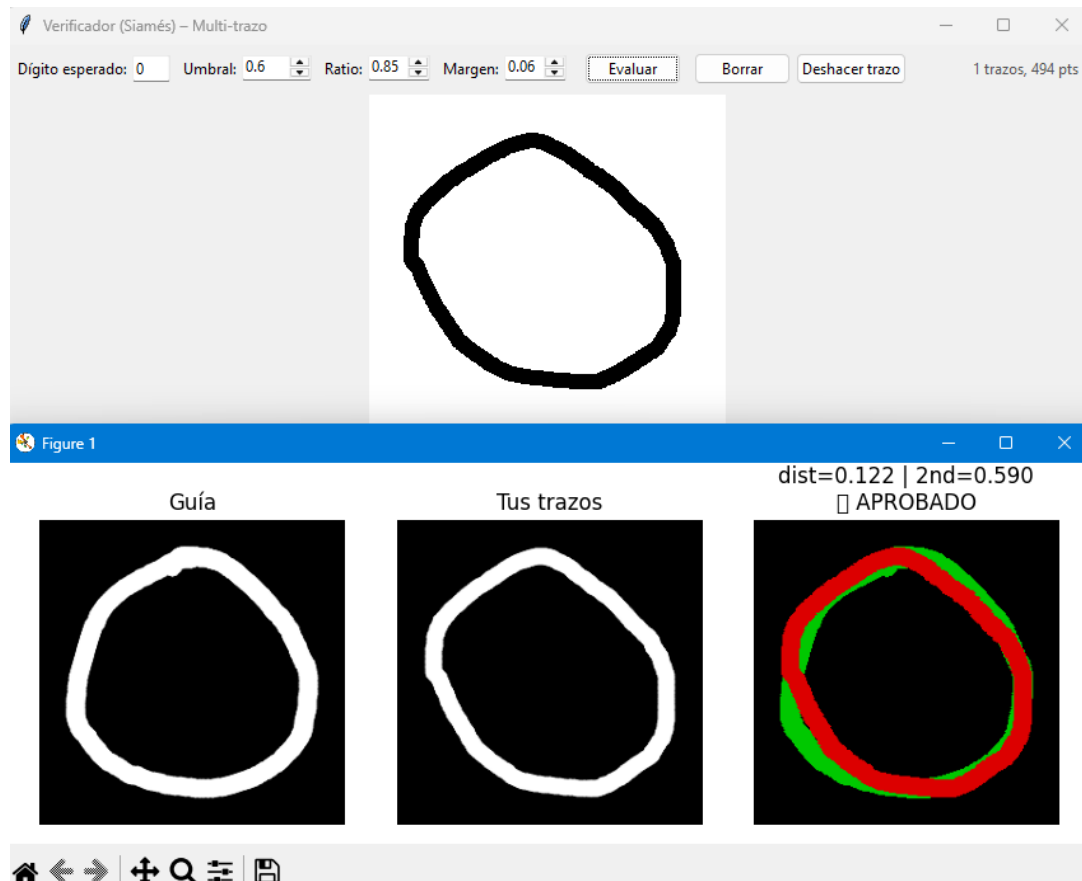
Valores iniciales:

- $\tau \approx 0.60$
- $\rho \approx 0.85$
- $\gamma \approx 0.06$

Retroalimentación en vivo:

- APROBADO / REVISAR
- Visualización: guía elegida, trazo usuario, superposición

Figura 6
Pruebas del modelo.



Nota. Elaboración propia.

c) Resultados:

La evaluación cuantitativa del modelo se realizó sobre un conjunto de test de 10,000 pares (1,000 por dígito, del 0 al 9), conformado por pares siameses no utilizados durante el entrenamiento. Aplicando el umbral de decisión $\tau = 0.60$ seleccionado mediante maximización del índice J de Youden, el modelo obtuvo los resultados que se presentan en la Tabla 9. El AUC-ROC de 1.0000 indica una capacidad discriminativa perfecta del modelo sobre el conjunto de test, lo que confirma que el espacio de embeddings aprendido separa de manera

nítida los trazos correctos de los incorrectos. La matriz de confusión Tabla 10 evidenció que el único error de clasificación relevante se concentró en el dígito 7, donde 31 de 1,000 trazos fueron confundidos con el dígito 3, resultado coherente con la proximidad morfológica entre ambos símbolos documentada en la literatura (Dey et al., 2017). Los nueve dígitos restantes alcanzaron clasificación perfecta. El rendimiento global confirma la idoneidad del modelo como componente evaluador automático dentro de la aplicación NumIA

Tabla 9

Métricas de evaluación del modelo de Redes Siamesas de NumIA (n = 10,000 pares)

Métrica	Valor	Interpretación
Accuracy	99.69%	Proporción de pares clasificados correctamente
Precision	99.70%	De los trazos aceptados, cuántos eran realmente correctos
Recall (Sensibilidad)	99.69%	De los trazos correctos, cuántos fueron detectados
F1-Score	99.69%	Media armónica de Precision y Recall
AUC-ROC	1.0000	Capacidad discriminativa global del modelo (1.0 = perfecto)
Umbral τ óptimo	0.60	Maximiza índice J de Youden (TPR – FPR)

Nota. Evaluación realizada sobre conjunto de test de 10,000 pares independiente al de entrenamiento (1,000 pares por dígito). Elaboración propia.

Tabla 10

Matriz de confusión del modelo de Redes Siamesas por dígito (n = 1,000 por clase)

	Pred 0	Pred 1	Pred 2	Pred 3	Pred 4	Pred 5	Pred 6	Pred 7	Pred 8	Pred 9	Total	Recall
Real 0	1000	0	0	0	0	0	0	0	0	0	1000	100.0%
Real 1	0	1000	0	0	0	0	0	0	0	0	1000	100.0%
Real 2	0	0	1000	0	0	0	0	0	0	0	1000	100.0%
Real 3	0	0	0	1000	0	0	0	0	0	0	1000	100.0%
Real 4	0	0	0	0	1000	0	0	0	0	0	1000	100.0%

Real 5	0	0	0	0	0	1000	0	0	0	0	1000	100.0%
Real 6	0	0	0	0	0	0	1000	0	0	0	1000	100.0%
Real 7	0	0	0	31	0	0	0	969	0	0	1000	96.9%
Real 8	0	0	0	0	0	0	0	0	1000	0	1000	100.0%
Real 9	0	0	0	0	0	0	0	0	0	1000	1000	100.0%

Nota. Los valores en la diagonal principal representan clasificaciones correctas. El único error registrado corresponde a 31 trazos del dígito 7 clasificados como dígito 3, atribuible a la similitud morfológica entre ambos símbolos. Elaboración propia.

4.2.2. Diseño de la Base de Datos

4.2.2.1. Modelo Relacional

Una base de datos relacional (RDB) es una forma de estructurar información en tablas, filas y columnas. Un RDB tiene la capacidad de establecer vínculos (o relaciones) entre información mediante la unión de tablas, lo que facilita la comprensión y la obtención de estadísticas sobre la relación entre varios datos. IBM (n.d.-b)

Desarrollado por EF Codd desde IBM en la década de 1970, el modelo de base de datos relacional permite que cualquier tabla se relacione con otra mediante un atributo común. En lugar de usar estructuras jerárquicas para organizar los datos, Codd propuso un cambio a un modelo de datos en el que estos últimos se almacenan, se consultan y se relacionan en tablas sin reorganizar las tablas que los contienen.

Propósito:

- Almacenamiento eficiente, seguro y estructurado
- Registro de usuarios, niveles, preguntas, recursos
- Trazabilidad de interacciones
- Gestión de contenidos y retroalimentación formativa

4.2.2.2. Tablas

Las tablas representan estructuras de datos en una base de datos que almacenan la totalidad de la información. En estas tablas, los datos se organizan conforme a una estructura de filas y columnas. Cada fila designa un registro único, mientras que cada columna corresponde a un campo dentro del registro.

Las tablas que se han tomado en cuenta en este proyecto son: usuarios, capítulos, tipo_preguntas, preguntas, recursos, pregunta_recursos, tablets, usuario_pregunta, nivel_usuario, sesiones, retroalimentación detalle, log_eventos, insignias, usuario_insignia a continuación se muestra la representación gráfica del modelo usado en el proyecto.

Figura 7
Diagrama de base de datos.



Nota. Imagen exportada desde el IDE DataGrip

4.2.3. Diseño y desarrollo de la aplicación móvil

4.2.3.1. Metodología de Desarrollo de Software

Para el desarrollo de la aplicación se adoptó la metodología ágil Scrum, debido a su enfoque iterativo e incremental, el cual resulta adecuado para proyectos de software educativo que requieren ajustes continuos según las necesidades pedagógicas, técnicas y de usabilidad detectadas durante el proceso de desarrollo.

Scrum es un marco de trabajo ágil que permite gestionar proyectos complejos mediante ciclos cortos de desarrollo, promoviendo la entrega continua de productos funcionales y la mejora progresiva del sistema (Schwaber & Sutherland, 2020) su aplicación resulta pertinente en el desarrollo de aplicaciones educativas basadas en tecnología, ya que facilita la validación temprana de funcionalidades y la adaptación a los requerimientos del usuario final.

Según Pressman & Maxim (2020), las metodologías ágiles son especialmente recomendadas cuando los requerimientos pueden evolucionar durante el desarrollo, situación frecuente en proyectos educativos donde la retroalimentación del usuario y los criterios pedagógicos influyen directamente en el diseño del sistema.

4.2.3.2. Tecnologías y Herramientas

La elección de estas herramientas se fundamenta en criterios de compatibilidad, rendimiento, escalabilidad y adecuación al contexto educativo, garantizando la correcta implementación de los objetivos pedagógicos y tecnológicos del proyecto.

Flutter

Flutter es un framework de desarrollo de aplicaciones móviles multiplataforma desarrollado por Google, que permite crear aplicaciones nativas para Android e iOS a partir de una única base de código (AWS, n.d.). Fue utilizado para el desarrollo de la interfaz de usuario de la aplicación móvil.

Dart

Dart es el lenguaje de programación utilizado por Flutter para el desarrollo de aplicaciones móviles. Se caracteriza por su sintaxis clara, orientación a objetos y soporte para programación reactiva (GoogleSkills, n.d.).

Python

Python es un lenguaje de programación de alto nivel ampliamente utilizado en el desarrollo de aplicaciones backend y en el campo de la inteligencia artificial. Su amplio ecosistema de librerías y su compatibilidad directa con frameworks de aprendizaje profundo como TensorFlow y Keras lo convierten en una herramienta idónea para proyectos que combinan software educativo e inteligencia artificial (Goodfellow et al., 2016).

Flask / FastAPI

Flask y FastAPI son frameworks ligeros para el desarrollo de APIs REST en Python, se utilizó uno de estos frameworks para construir el servidor backend encargado de procesar las solicitudes provenientes de la aplicación móvil, ejecutar la inferencia del modelo de Deep Learning y devolver los resultados en formato JSON.

TensorFlow / Keras

TensorFlow es una plataforma de código abierto para el desarrollo de modelos de aprendizaje automático, mientras que Keras es una API de alto nivel integrada en TensorFlow que facilita la construcción y entrenamiento de redes neuronales.

MySQL / PostgreSQL

MySQL y PostgreSQL son sistemas de gestión de bases de datos relacionales utilizados para almacenar la información generada por la aplicación, como datos de usuarios, progreso de aprendizaje, resultados de evaluación y métricas de uso.

Git / GitHub

Git es un sistema de control de versiones distribuido que permite gestionar los cambios realizados en el código fuente, mientras que GitHub es una plataforma que facilita el almacenamiento remoto y la colaboración en proyectos de software.

Google Colab

Google Colab es una plataforma basada en la nube que permite ejecutar código Python en entornos con soporte para GPU. Esta herramienta fue utilizada para el entrenamiento y prueba del modelo de Deep Learning, debido a su facilidad de acceso y capacidad de procesamiento.

Android Studio / Visual Studio Code

Android Studio y Visual Studio Code son entornos de desarrollo integrados (IDE) utilizados para la programación, depuración y prueba de la aplicación móvil y el backend.

4.2.3.3. *Arquitectura de la Aplicación*

La arquitectura fue diseñada con el propósito de garantizar una adecuada organización del sistema, facilitar su mantenimiento y permitir la integración eficiente de la aplicación móvil con los servicios de backend y el modelo de inteligencia artificial. Para ello, se adoptó una combinación de arquitectura cliente–servidor, arquitectura en capas, API REST y el patrón Modelo–Vista–Controlador (MVC) en el backend.

Arquitectura Cliente–Servidor

La arquitectura cliente–servidor, en la cual la aplicación móvil actúa como cliente y un servidor centralizado gestiona la lógica del sistema, el procesamiento de datos y la ejecución del modelo de Deep Learning. Este tipo de arquitectura permite una clara separación de responsabilidades y facilita la escalabilidad del **sistema** (Tanenbaum & Steen, 2007).

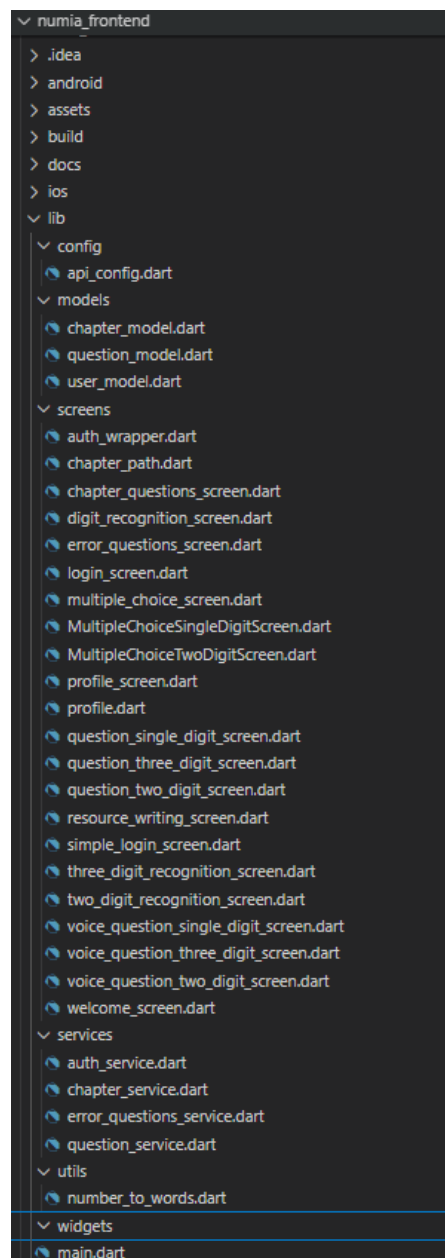
Arquitectura en Capas

La arquitectura del sistema se organiza en capas, lo que permite estructurar el software de manera modular y ordenada.

- **Capa de presentación:** Representada por la aplicación móvil desarrollada en Flutter, encargada de la interacción con el usuario.

Figura 8

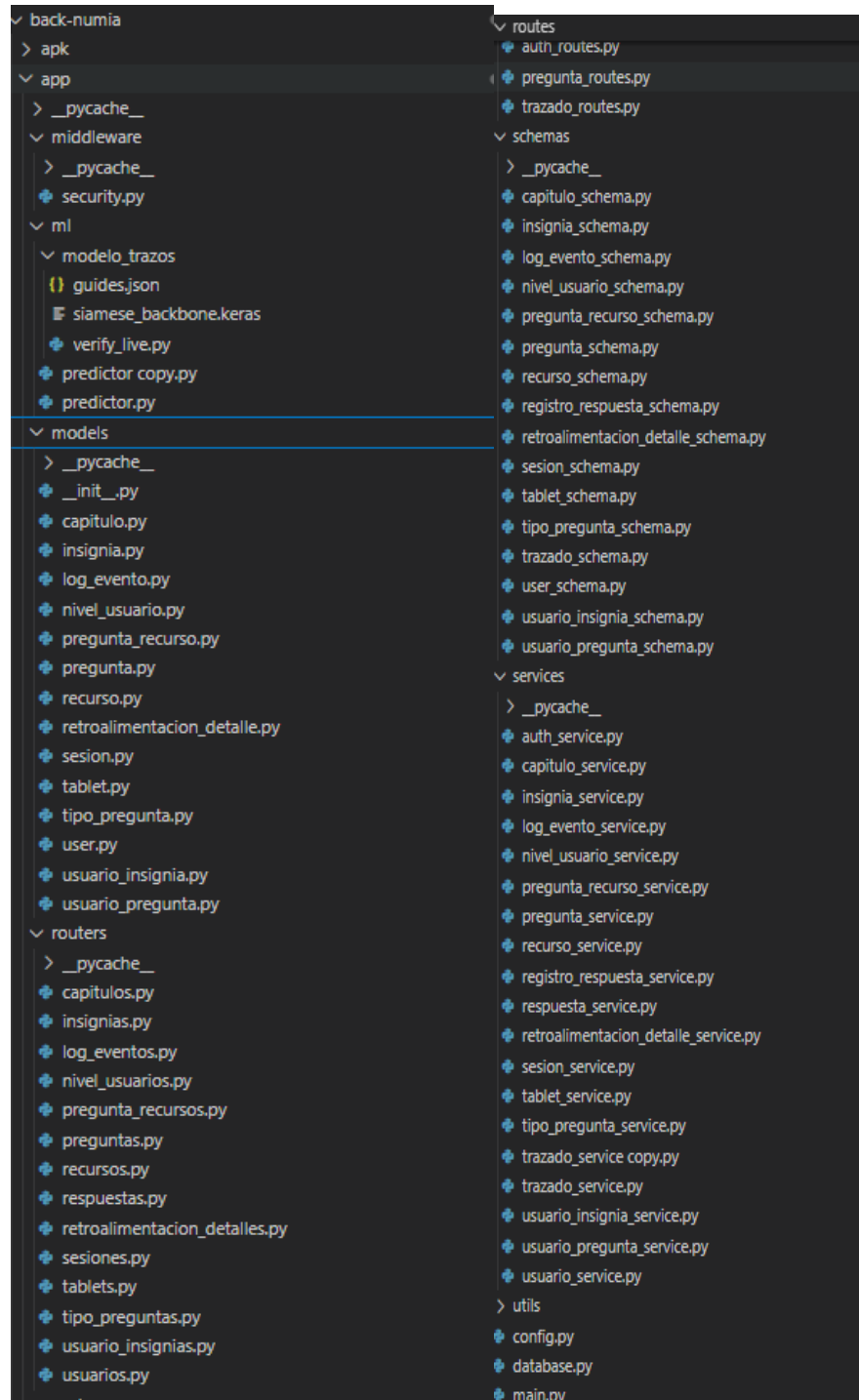
Estructura de archivos en flutter.



Nota. Elaboración propia.

- **Capa de lógica de negocio:** Implementada en el backend, donde se procesan las reglas del sistema y se gestiona la inferencia del modelo de inteligencia artificial.

Figura 9
Estructura de archivos en python.



Nota. Elaboración propia.

- **Capa de persistencia:** Encargada del acceso y almacenamiento de datos en la base de datos relacional.

Figura 10

Conexión de base de datos.

```

database.py X
back-numia > app > database.py
1  from sqlalchemy import create_engine
2  from sqlalchemy.ext.declarative import declarative_base
3  from sqlalchemy.orm import sessionmaker
4  from app.config import settings
5
6  # Create database engine
7  engine = create_engine(
8      settings.DATABASE_URL,
9      pool_pre_ping=True,
10     pool_recycle=300,
11     echo=False # Set to True for SQL debugging
12 )
13
14 # Create SessionLocal class
15 SessionLocal = sessionmaker(autocommit=False, autoflush=False, bind=engine)
16
17 # Create Base class
18 Base = declarative_base()
19
20 # Dependency to get database session
21 def get_db():
22     db = SessionLocal()
23     try:
24         yield db
25     finally:
26         db.close()
27
28 # Function to create all tables
29 def create_tables():
30     Base.metadata.create_all(bind=engine)

```

Nota. Elaboración propia.

La arquitectura en capas mejora la mantenibilidad y la reutilización del código, además de facilitar futuras modificaciones sin afectar todo el sistema (Bass et al., 2012).

API REST

La comunicación entre la aplicación móvil y el servidor se realiza mediante una API REST (Representational State Transfer), permite el intercambio de información utilizando el protocolo HTTP y el formato JSON, lo cual favorece la interoperabilidad entre sistemas heterogéneos. Este enfoque es ampliamente utilizado en aplicaciones web y móviles por su simplicidad y eficiencia (Fielding, 2000).

Patrón Modelo–Vista–Controlador (MVC) en el Backend

En el backend se empleó el patrón Modelo–Vista–Controlador (MVC) para organizar internamente el código del sistema.

- **Modelo (Model):** Representa las entidades del sistema y la interacción con la base de datos.
- **Controlador (Controller):** Gestiona las solicitudes provenientes de la aplicación móvil, procesa la información recibida y coordina la ejecución del modelo de inteligencia artificial.
- **Vista (View):** En el contexto de una API REST, la vista está representada por las respuestas JSON enviadas al cliente.

El uso de MVC contribuye a mejorar la claridad del código y facilita el mantenimiento del sistema. (Pressman & Maxim, 2020).

Comunicación entre Componentes

La comunicación entre los componentes del sistema se realiza de manera estructurada mediante solicitudes y respuestas HTTP. El flujo general es el siguiente:

- El usuario realiza el trazo numérico en la aplicación móvil.
- La aplicación envía los datos al servidor a través de la API REST.
- El backend procesa la información y ejecuta el modelo de Deep Learning.
- El resultado es devuelto a la aplicación en formato JSON.
- La aplicación muestra la retroalimentación correspondiente al usuario.

4.2.3.4. *Diseño de Interfaz (UI/UX)*

A. Paleta de colores

Empleo de colores primarios vibrantes para captar atención y facilitar asociación con conceptos básicos.

Alto contraste entre texto y fondo para mejorar legibilidad y reducir la fatiga visual.

Fondo claro como norma (evitar fondos oscuros o con texturas intensas que dificulten la lectura).

Codificación de color consistente por función:

- ✓ Verde: correcto o aprobado
- ✓ Rojo: incorrecto o necesidad de revisión
- ✓ Azul: información o neutral
- ✓ Amarillo: advertencia o atención necesaria

Consideraciones metodológicas:

Justificar la elección de paleta en términos de simplicidad, memoria de color y legibilidad en poblaciones infantiles.

Incluir pruebas de accesibilidad: contraste mínimo según estándares.

B. Tipografía

Se uso de una fuente sans-serif legible y tamaño de fuente recomendado, es decir mínimo 18 pt para cuerpos de texto o visualizaciones clave en materiales impresos; adaptar a 14–16 pt en documentos de tamaño A4 con consideraciones de diseño.

Interlineado amplio para mejorar lectura y comprensión.

Texto breve y directo: párrafos cortos, frases claras y lenguaje simple para favorecer la comprensión en niños.

C. Iconografía

Íconos grandes y fácilmente reconocibles para apoyar conceptos clave sin sobrecargar el texto.

Priorizar elementos visuales que acompañen al texto en lugar de sustituirlo por completo.

Animaciones suaves y dentro de límites razonables para no distraer; deben ser sutiles y estar justificadas por la mejora en comprensión.

Feedback visual inmediato ante acciones o respuestas para reforzar el aprendizaje.

D. Navegación

Botones grandes y accesibles (recomendación: mínimo 44×44 dp) para facilitar interacción en dispositivos táctiles.

Menú simple con un conjunto reducido de opciones principales (3–4).

Breadcrumbs visuales para indicar la ubicación dentro de la estructura del documento y facilitar la navegación.

Botón de “Atrás” siempre visible o fácilmente accesible para regresar a la página o sección anterior

4.2.3.5. *Pantallas Principales*

Pantalla de Logeo: vista inicial de la app Numia, en la cual podrán iniciar sesión los alumnos.

Figura 11

Pantalla de logeo de usuario.

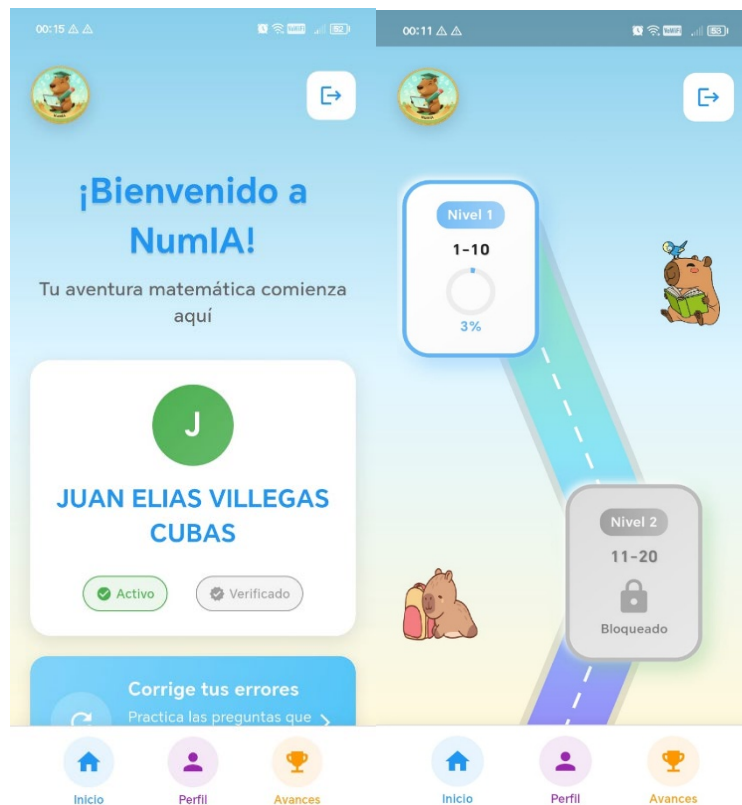


Nota. Elaboración propia.

Pantalla de Inicio: Vista inicial que se mostrará después de iniciar sesión, en donde se muestra un mensaje de bienvenida, además del nombre del usuario.

Además se muestra capítulos disponibles o bloqueados que tiene el usuario, así como el porcentaje que va en los capítulos que ha desarrollado.

Figura 12
Pantalla de Inicio.



Nota. Elaboración propia.

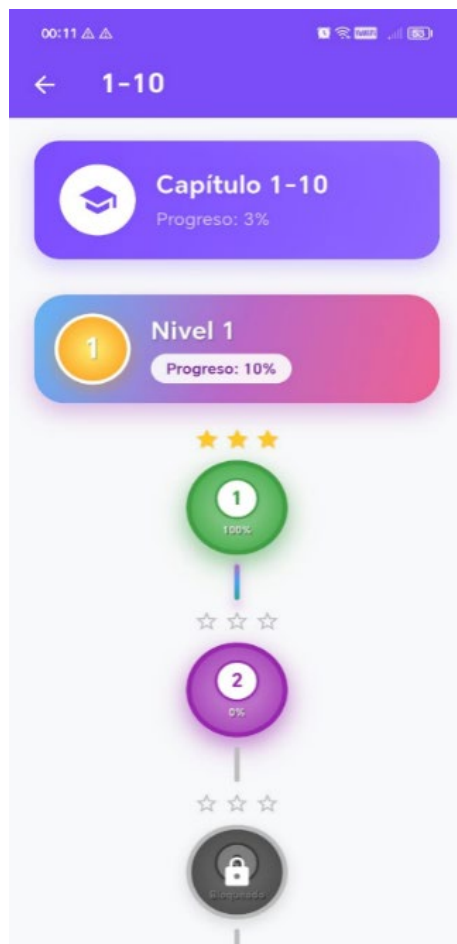
Pantalla de Capítulos: En esta pantalla se puede ver el nombre del capítulo, su descripción expandida, porcentaje de avance, preguntas correspondientes al capítulo según su nivel, así mismo su porcentaje de avance del nivel. Cada pregunta tiene 3 estrellas.

Fuente: Elaboración propia

Cada pregunta muestra un indicador visual:

- ✓ Color Morado: Pregunta desbloqueada sin contestar
- ✓ Color Verde: Pregunta Contestada con un puntaje mayor al 95%
- ✓ Color Naranja: Pregunta contestada entre 20 al 94%

Figura 13
Pantalla de capítulos.



Nota. Elaboración propia.

Pantalla de preguntas: En esta pantalla se muestra la pregunta, el cual puede ser de los siguientes tipo:

- ✓ **Preguntas de conteo:** En la cual se muestra preguntas para contar y dibujar números en la pantalla táctil cumpliendo con los objetivos de ejercicio nivel básico, de esta forma se combinan trazado con conteo de figuras.

Figura 14

Pantalla de preguntas de conteo y escritura.



Nota. Elaboración propia.

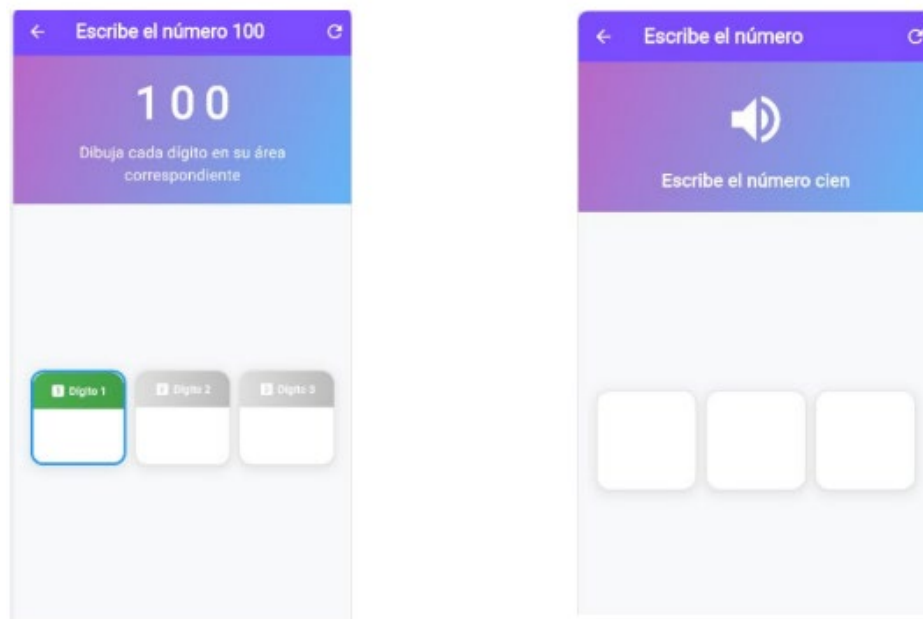
- ✓ **Preguntas con Recursos Multimedia:** Aquí se muestran ejercicios que incluyen imágenes, audio.

Características:

- Pueden mostrar imágenes de números
- Pueden reproducir audio con pronunciación
- Combinan trazado con recursos visuales/auditivos

Figura 15

Pantalla de preguntas con recurso multimedia.



Nota. Elaboración propia.

Pantalla de Retroalimentación

- ✓ Respuesta Correcta, esta pantalla tiene los siguientes elementos:
 - Indicador verde: Muestra que está correcto
 - Mensaje positivo: "¡CORRECTO!" o "¡Excelente!"

Figura 16

Pantalla de retroalimentación positiva.

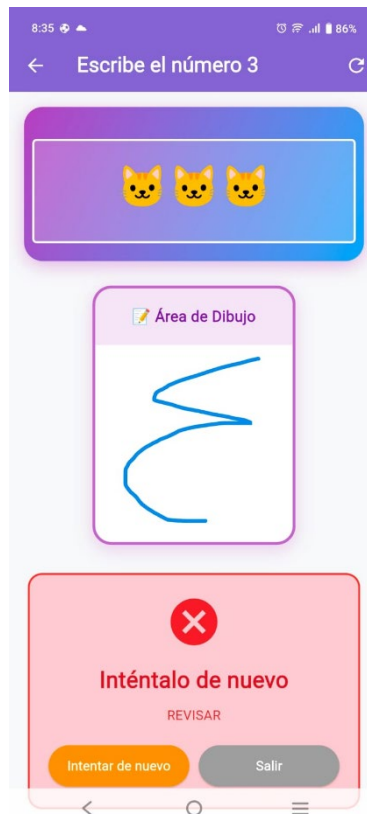


Nota. Elaboración Propia.

- ✓ Respuesta con errores Si el trazado necesita mejoras, se verá un indicador rojo y un mensaje menciona que mejorar, además de la posibilidad volver a intentar la pregunta mediante un botón para intentar de nuevo.

Figura 17

Pantalla de retroalimentación con corrección.

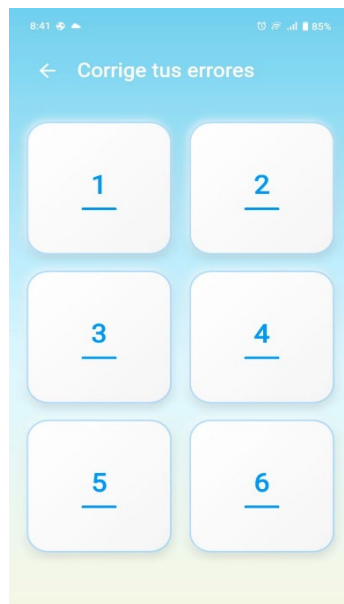


Nota. Elaboración Propia.

Pantalla de corregimiento de errores. Esta pantalla muestra todas las preguntas donde el usuario cometió errores.

Figura 18

Pantalla de personalización de errores.

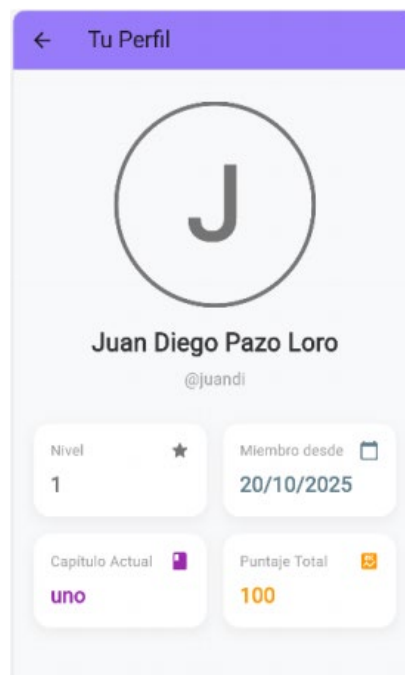


Nota. Elaboración Propia.

Pantalla de perfil: Pantalla perteneciente a una de las opciones del menú inferior, en la cual se muestra información relacionada al usuario.

Figura 19

Pantalla de perfil de usuario.



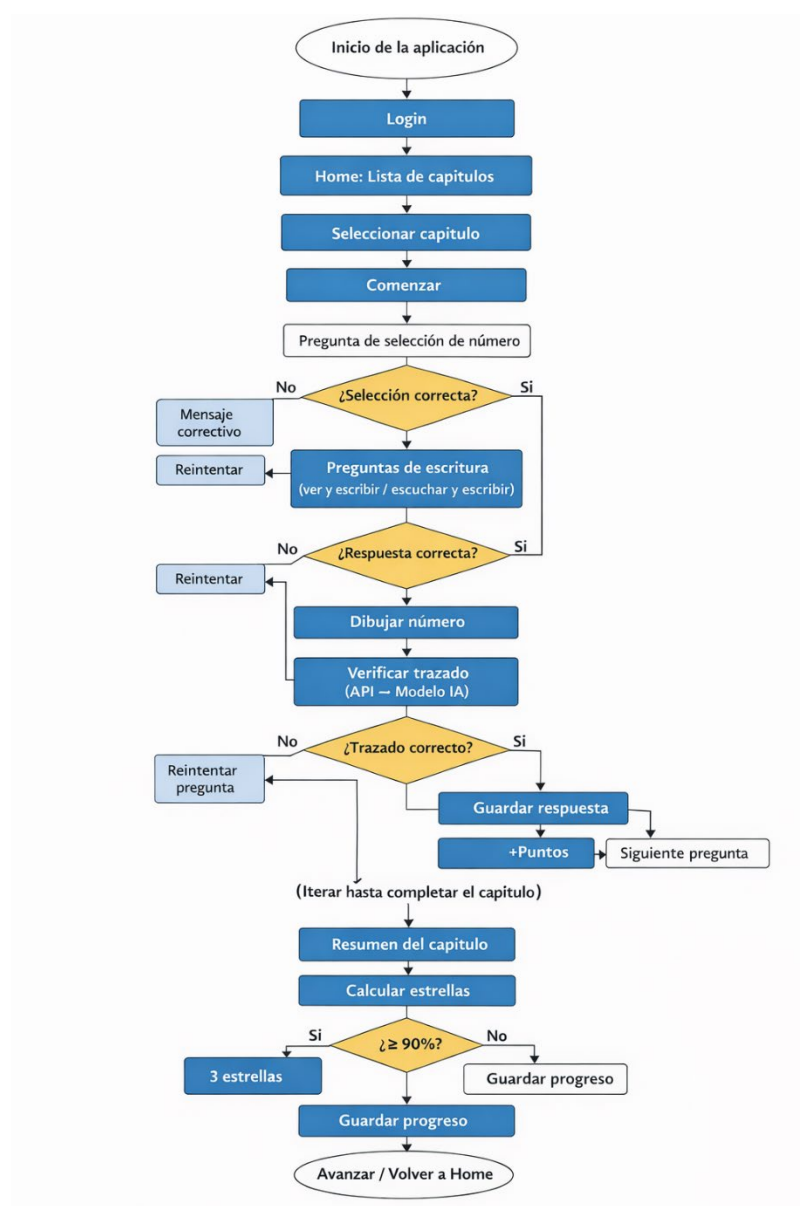
Nota. Elaboración Propia.

4.2.3.6. Flujo de Usuario

El siguiente diagrama de flujo muestra el funcionamiento general de la aplicación móvil NumIA, el cual describe de manera secuencial la interacción del estudiante con el sistema desde el inicio de la aplicación hasta el registro del progreso y el retorno a la pantalla principal.

Figura 20

Flujo de usuario.



Nota. Elaboración Propia.

4.2.4. Síntesis de Hallazgos del Resultado

En respuesta al segundo objetivo específico, orientado a diseñar y desarrollar la aplicación móvil, incorporando una interfaz gráfica amigable y módulos interactivos de

práctica de trazos numéricos evaluados mediante un modelo de inteligencia artificial, se logró la implementación de una solución tecnológica educativa funcional que integra componentes pedagógicos, de software e inteligencia artificial de manera articulada.

Como resultado del proceso de diseño y desarrollo, se obtuvo una aplicación móvil que permite a los estudiantes interactuar con actividades progresivas de reconocimiento y escritura de números arábigos, apoyadas por una interfaz gráfica adaptada a las características cognitivas y perceptivas de niños de primer grado.

Se logró desarrollar e integrar un modelo de Deep Learning basado en Redes Siamesas, capaz de evaluar automáticamente el trazado numérico realizado por los estudiantes mediante la comparación de similitud entre el trazo del usuario y patrones de referencia previamente definidos.

La aplicación incorpora módulos interactivos de práctica guiada y reintento dirigido, los cuales permiten al estudiante corregir errores de manera inmediata a partir de la retroalimentación generada por el sistema.

Finalmente, el sistema registra información relacionada con el desempeño del estudiante, como puntaje, errores frecuentes y progreso por nivel, permitiendo el seguimiento del aprendizaje. En conjunto, estos resultados evidencian el cumplimiento del segundo objetivo específico, al materializar una aplicación móvil educativa que integra inteligencia artificial, diseño pedagógico y desarrollo de software como una herramienta de apoyo para el aprendizaje del trazado correcto de números arábigos en estudiantes de primer grado de educación primaria.

4.3. Resultados de pretest e implementación de aplicación NumIA

Para dar cumplimiento al tercer objetivo específico: “Evaluar el nivel inicial del trazado numérico mediante una prueba diagnóstica (pretest) e implementar la aplicación NumIA en sesiones de práctica estructurada durante cinco semanas con estudiantes de primer grado de una institución educativa pública” , en este objetivo se evalúa el nivel inicial del trazado llamada pretest, luego de ello se implementa sesiones de práctica estructurada durante cinco semanas con estudiantes de primer grado de la institución educativa I.E. Fe y Alegría N° 28.

4.3.1. Diagnóstico Previo a la Intervención

4.3.1.1. *Instrumento de evaluación: Prueba de desempeño de trazado numérico*

El instrumento utilizado tanto en el pretest como en el postest fue la Prueba de Desempeño de Trazado Numérico, diseñada por el equipo investigador, basándose en las bases conceptuales, enfocándose en el trazado correcto del número a través de tres criterios precisión, secuencia y orientación, donde la secuencia se operacionaliza en dos subcriterios de medición (orden y fluidez), resultando en cuatro columnas de evaluación en la rúbrica. Se trata de una prueba de ejecución directa, el estudiante traza los números en papel evaluada mediante una rúbrica estandarizada (Anexo 1), en coherencia con el enfoque de evaluación formativa por desempeño recomendado para el primer ciclo de educación primaria (MINEDU, 2016).

La prueba está compuesta por 78 ítems organizados en cuatro secciones progresivas (Anexo 2)

- **Sección 1:** Escritura de números individuales (ítems P1–P20):

El estudiante escribe cada número del 1 al 20 de forma individual en el espacio asignado, se evalúa directamente la representación gráfica convencional de cada dígito.

- **Sección 2:** Conteo y cantidad (ítems P21–P30):

El estudiante escribe cada número del 21 al 30 de forma individual en el espacio asignado según su cantidad de elementos que muestra la pregunta, se evalúa directamente la representación gráfica convencional de cada dígito.

- **Sección 3:** Completar series (ítems P31–P60):

El estudiante escribe cada número del 31 al 60 de forma individual en el espacio asignado, se evalúa directamente la representación gráfica convencional de cada dígito.

- **Sección 4:** Dictado numérico (ítems P61–P78):

El equipo investigador dicta distintos números en voz alta y el estudiante los escribe. Esta sección es la de mayor exigencia cognitiva y grafomotora, pues requiere la traducción del estímulo auditivo al símbolo escrito convencional, activando simultáneamente la memoria numérica, el control motor y la planificación del trazo.

4.3.1.2. Rúbrica de evaluación: cuatro criterios de trazado

Cada ítem fue evaluado con una rúbrica de cuatro criterios, cada uno con tres niveles de desempeño (0, 1 y 2 puntos), generando un puntaje máximo de 8 puntos por ítem y un puntaje máximo total de 624 puntos (78 ítems $\times$ 8 puntos). Los cuatro criterios son:

Tabla 11
Rubrica de trazado

Criterio	2 pts - Logrado	1pts - En proceso	0pts - No logrado
Orientación	Trazo en dirección convencional correcta, inicio y fin adecuados	Desvíos ocasionales, conserva dirección general	Dirección errónea o invertida en la mayoría de trazos
Secuencia	Orden lógico y progresivo correcto de los segmentos	Error en uno o dos trazos, estructura reconocible	Segmentos trazados de forma desordenada, ilegible

	Trazo fluido, continuo, sin interrupciones perceptibles	Leves pausas o temblores	Trazo discontinuo, tembloroso, múltiples intentos
Precisión	Forma clara, simétrica y reconocible, acorde al modelo	Deformaciones leves, número reconocible	Número ilegible, proporción distorsionada

Nota. Elaboración Propia.

La selección de estos cuatro subcriterios responde a la literatura especializada sobre evaluación del trazado en niños. Chang & Yu (2022), en un estudio computarizado de evaluación grafomotora con niños de primero a tercer grado de primaria, identifican la corrección de la forma, la legibilidad y el orden del trazo como las dimensiones directamente asociadas al desempeño académico en escritura, mientras que la fluidez del movimiento medida a través de la continuidad y suavidad del trazo constituye el indicador de automatización grafomotora. Por su parte, Feder & Majnemer (2007) documentan que la evaluación integral de la escritura infantil requiere considerar tanto la orientación correcta del símbolo como la precisión de su forma, criterios que permiten distinguir entre errores de ejecución motriz y errores de representación visual del dígito.

Con el fin de verificar la confiabilidad del instrumento antes de proceder al análisis de los resultados, se calculó el coeficiente Alfa de Cronbach (Cronbach, 1951) sobre los puntajes brutos de los 30 estudiantes en los cuatro subcriterios de la rúbrica: Orientación (O), Secuencia de orden (So.), Secuencia de fluidez (Sf.) y Precisión (P). El valor obtenido fue $\alpha = 0.990$, clasificado como Excelente según los criterios de George & Mallery (2019) lo que confirma una consistencia interna muy alta entre los cuatro criterios de evaluación y respalda la confiabilidad de los datos recopilados para el análisis estadístico subsiguiente. La Tabla 12 presenta los estadísticos descriptivos por subcriterio.

Tabla 12*Estadísticos descriptivos de los subcriterios de la rúbrica pretest (n = 30)*

Estadístico	Orientación (O)	Secuencia de orden (So.)	Secuencia de fluidez (Sf.)	Precisión (P)
Media	106.10	102.03	101.47	102.40
Desviación estándar	38.50	36.21	35.79	36.99
Mínimo	39.00	35.00	35.00	35.00
Percentil 25	79.00	74.75	74.75	78.25
Mediana (P50)	107.00	107.00	107.00	107.00
Percentil 75	138.25	128.75	128.50	130.75
Máximo	156.00	156.00	156.00	154.00
Alfa de Cronbach	$\alpha = 0.990$			

Nota. Los valores corresponden a puntajes brutos (puntos absolutos obtenidos por cada estudiante en cada subcriterio). Alfa de Cronbach calculado sobre los cuatro subcriterios indica consistencia interna Excelente según George & Mallery (2019). Elaboración propia.

4.3.1.3. Cálculo del porcentaje de logro y niveles de desempeño

El porcentaje de logro de cada estudiante se calculó como:

$$\% \text{ Criterio X} = (\text{Suma de puntos en criterio X} / (\text{n}^\circ \text{ ítems respondidos} \times 2)) \times 100$$

$$\% \text{ Logro Final} = (\text{Suma total de puntos en los 4 criterios} / (\text{n}^\circ \text{ ítems respondidos} \times 8)) \times 100$$

Este cálculo se realizó sobre los ítems efectivamente respondidos por cada estudiante, excluyendo los ítems con ausencia de respuesta, lo que garantiza que el porcentaje refleje el desempeño real sobre los ítems evaluados. La variabilidad en el número de ítems respondidos entre estudiantes (mínimo 33, máximo 78 de los 78 posibles) obedece a diferencias en el ritmo de trabajo individual: los niños dispusieron del tiempo estipulado por sección y algunos no

alcanzaron a completar todas las secciones de la prueba. Por este motivo, el denominador del porcentaje se ajusta al número de ítems que el estudiante efectivamente realizó, preservando la comparabilidad de los porcentajes de logro entre estudiantes con distintos ritmos de ejecución. Los niveles de logro quedaron definidos en tres categorías, alineadas con la escala de calificación del Currículo Nacional peruano (MINEDU, 2016):

Tabla 13
Niveles de logro y umbrales de clasificación

Nivel	Rango	Descripción
Logro Destacado (AD)	86% – 100%	El estudiante evidencia nivel superior esperado; dominio consolidado y autónomo del trazado numérico en todos los criterios evaluados.
Logro Esperado (A)	66 %- 85%	El estudiante alcanza el nivel esperado para el grado; demuestra competencia adecuada con imprecisiones menores y corrección progresiva.
En Proceso (B)	51% – 65%	El estudiante está en camino de alcanzar el nivel esperado; requiere acompañamiento pedagógico y refuerzo sostenido.
En Inicio (C)	0% – 50%	El estudiante presenta dificultades significativas; requiere intervención pedagógica prioritaria e inmediata en todos los criterios.

Nota. Elaboración Propia.

4.3.1.4. Aplicación del Pretest

- **Condiciones de aplicación**

El pretest fue administrado de forma presencial en las aulas de la I.E. Fe y Alegría N.º 28, en las tres secciones de primer grado, bajo condiciones uniformes: mismo instrumento impreso, mismas instrucciones orales estandarizadas, mismo tiempo disponible por sección y mismo protocolo de registro. Los investigadores y el docente de cada sección administraron la prueba,

los resultados individuales fueron digitalizados en el libro de registro Excel, consignando para cada ítem los puntajes de los cuatro subcriterios (Orientación (O), Secuencia de orden (So), Secuencia de fluidez (S.f.), Precisión(P))

- **Resultados Generales del Pretest**

La siguiente tabla presenta los resultados completos del pretest para los 30 estudiantes seleccionados. Los niveles de logro se asignan conforme a la escala: AD (86–100%), A (66–85%), B (51–65%) y C (0–50%).

Tabla 14
Resultado de la evaluación pretest

Secc.	Nº	Apellidos y Nombres	Ítems	O pts/máx	O %	So. pts/máx	So. %	Sf. pts/máx	Sf. %	P pts/máx	P %	Total %	Nivel
1ºA	1	Estudiante 1	57	89/114	78.1	90/114	78.9	91/114	79.8	91/114	79.8	79.2	A
	2	Estudiante 2	78	96/156	61.5	98/156	62.8	101/156	64.7	98/156	62.8	63.0	B
	3	Estudiante 3	46	42/92	45.7	39/92	42.4	39/92	42.4	39/92	42.4	43.2	C
	4	Estudiante 4	59	104/118	88.1	95/118	80.5	85/118	72.0	82/118	69.5	77.5	A
	5	Estudiante 5	52	40/104	38.5	71/104	68.3	74/104	71.2	40/104	38.5	54.1	B
	6	Estudiante 6	78	102/156	65.4	64/156	41.0	64/156	41.0	102/156	65.4	53.2	B
	7	Estudiante 7	78	115/156	73.7	116/156	74.4	116/156	74.4	115/156	73.7	74.0	A
	8	Estudiante 8	78	143/156	91.7	113/156	72.4	111/156	71.2	137/156	87.8	80.8	A
	9	Estudiante 9	62	50/124	40.3	48/124	38.7	49/124	39.5	51/124	41.1	39.9	C
	10	Estudiante 10	33	39/66	59.1	35/66	53.0	35/66	53.0	35/66	53.0	54.5	B
Promedio 1ºA					64.2		61.2		60.9		61.4	61.9%	
1ºB	1	Estudiante 11	78	77/156	49.4	77/156	49.4	77/156	49.4	77/156	49.4	49.4	C
	2	Estudiante 12	78	93/156	59.6	93/156	59.6	93/156	59.6	93/156	59.6	59.6	B
	3	Estudiante 13	76	152/152	100.0	150/152	98.7	149/152	98.0	148/152	97.4	98.5	AD
	4	Estudiante 14	77	61/154	39.6	55/154	35.7	56/154	36.4	58/154	37.7	37.3	C

	5	Estudiante 15	78	142/156	91.0	142/156	91.0	143/156	91.7	144/156	92.3	91.5	AD
	6	Estudiante 16	76	152/152	100.0	152/152	100.0	150/150	100.0	150/150	100.0	100.0	AD
	7	Estudiante 17.	78	154/156	98.7	125/156	80.1	114/156	73.1	108/156	69.2	80.3	A
	8	Estudiante 18	78	156/156	100.0	156/156	100.0	156/156	100.0	152/156	97.4	99.4	AD
	9	Estudiante 19	75	136/150	90.7	123/150	82.0	123/150	82.0	127/150	84.7	84.8	A
	10	Estudiante 20	78	156/156	100.0	154/156	98.7	154/156	98.7	154/156	98.7	99.0	AD
	Promedio 1°B				82.9%		79.5%		78.9%		78.6%	80.0%	
1°C	1	Estudiante 21	60	85/120	70.8	85/120	70.8	85/120	70.8	85/120	70.8	70.8	A
	2	Estudiante 22	51	49/102	48.0	47/100	47.0	47/100	47.0	47/100	47.0	47.3	C
	3	Estudiante 23	78	131/156	84.0	131/156	84.0	131/156	84.0	131/156	84.0	84.0	A
	4	Estudiante 24	72	74/144	51.4	74/144	51.4	74/144	51.4	74/144	51.4	51.4	B
	5	Estudiante 25	78	108/156	69.2	108/156	69.2	108/156	69.2	108/156	69.2	69.2	A
	6	Estudiante 26	77	139/154	90.3	128/154	83.1	129/154	83.8	129/154	83.8	85.2	A
	7	Estudiante 27	66	127/132	96.2	127/132	96.2	127/132	96.2	127/132	96.2	96.2	AD
	8	Estudiante 28	77	131/154	85.1	130/154	84.4	130/154	84.4	130/154	84.4	84.6	A
	9	Estudiante 29	70	106/140	75.7	106/140	75.7	106/140	75.7	106/140	75.7	75.7	A
	10	Estudiante 30	73	134/146	91.8	129/146	88.4	127/146	87.0	134/146	91.8	89.7	AD
	Promedio 1°C				76.2%		75.0%		75.0%		75.4%	75.4%	
	GLOBAL n = 30				74.5%		71.9%		71.6%		71.8%	72.4%	

Nota. Elaboración Propia.

4.3.1.5. *Análisis por Sección.*

Sección 1ºA: Con una media de 61.9% (nivel B, En Proceso), la sección 1ºA registra el punto de partida más bajo de las tres secciones. Ningún estudiante alcanza el nivel AD, cuatro se ubican en nivel A (40%), cuatro en nivel B (40%) y dos en nivel C (20%).

Sección 1ºB: Con una media de 80.0% (nivel A, Logro Esperado), la sección 1ºB es la de mejor desempeño inicial. Cinco estudiantes alcanzan nivel AD (50%), dos en nivel A (20%), uno en nivel B (10%) y dos en nivel C (20%).

Sección 1ºC: Con una media de 75.4% (nivel A, Logro Esperado), la sección 1ºC presenta un desempeño intermedio. Dos estudiantes alcanzan nivel AD (20%), seis en nivel A (60%), uno en nivel B (10%) y uno en nivel C (10%).

4.3.1.6. *Análisis Global del Diagnóstico.*

Tabla 15
Análisis global del diagnóstico

Secc.	n	O%	So. %	Sf. %	P%	% Final	AD	A	B	C
1ºA	10	64.20%	61.20%	60.90%	61.40%	61.90%	0 (0%)	4 (40%)	4 (40%)	2 (20%)
1ºB	10	82.90%	79.50%	78.90%	78.60%	80.00%	5 (50%)	2 (20%)	1 (10%)	2 (20%)
1ºC	10	76.20%	75.00%	75.00%	75.40%	75.40%	2 (20%)	6 (60%)	1 (10%)	1 (10%)
Global	30	74.50%	71.90%	71.60%	71.80%	72.40%	7 (23%)	12 (40%)	6 (20%)	5 (17%)

Nota. Elaboración Propia.

El diagnóstico inicial de los 30 estudiantes revela que el 23% alcanzó el nivel AD en el pretest (7 estudiantes), el 40% se ubicó en nivel A (12 estudiantes), el 20% en nivel B (6 estudiantes) y el 17% en nivel C (5 estudiantes). El nivel modal del pretest es el nivel A, lo que indica que la mayoría de los estudiantes presentaba bases grafomotoras en el rango de logro

esperado al inicio de la intervención. No obstante, los cinco casos en nivel C —distribuidos entre 1°A (2 estudiantes: Est. 3 y Est. 9), 1°B (2 estudiantes: Est. 11 y Est. 14) y 1°C (1 estudiante: Est. 22) representan la prioridad máxima de intervención, con rezago generalizado en los cuatro criterios de trazado evaluados.

4.3.2. Implementación de la aplicación NumIA

En este apartado se describirá las fases que se llevaron a cabo para el uso de la aplicación en las aulas de 1er grado.

4.3.2.1. Fase Previa

La fase previa a la intervención comprendió: (a) instalación del APK de NumIA en las tablets Android asignadas a cada estudiante, habilitando la opción 'Fuentes desconocidas' en la configuración del dispositivo; (b) creación del perfil de usuario individual por estudiante con nombre y clave de acceso; (c) capacitación del docente de cada sección en el uso de la interfaz y en la interpretación de los reportes automáticos de progreso; y (d) sesión de inducción con los estudiantes para familiarizarlos con la mecánica de los ejercicios y la navegación en la aplicación.

A) Instalación de la aplicación en los dispositivos:

Previo al inicio de las sesiones, el equipo investigador instaló NumIA en las tabletas de la institución educativa. El proceso consistió en: activar la opción "Fuentes desconocidas" en la configuración de seguridad Android de cada dispositivo, descargar el archivo APK desde el repositorio del proyecto, ejecutar la instalación y verificar el correcto funcionamiento de los módulos de trazado y reconocimiento.

B) Creación del perfil

Se creó un perfil individual para cada estudiante participante registrando nombre y apellido, activando el sistema de adaptación de contenido y el registro personalizado de progreso y errores, su contraseña fue simple para que se pudiera ingresar de manera rápida.

C) Capacitación docente:

Se realizó una sesión de capacitación con los docentes de las tres secciones, orientada a: familiarizarlos con la interfaz de NumIA y la navegación entre módulos, enseñarles a interpretar los reportes automáticos de progreso y errores accesibles desde el perfil de cada estudiante, y establecer el protocolo de supervisión durante las sesiones de práctica. Los docentes aprendieron a identificar en tiempo real qué estudiantes requerían apoyo adicional basándose en los indicadores de la app, complementando así la retroalimentación automática con la intervención pedagógica humana.

D) Sesión de inducción con los estudiantes:

En la primera sesión con los estudiantes, los investigadores guiaron colectivamente la exploración inicial de la aplicación: navegación entre la pantalla de inicio, los capítulos y los módulos de práctica, mecánica básica del ejercicio de trazado (área azul de dibujo, botón "Borrar", botón "Verificar"), y comprensión de los mensajes de retroalimentación visual y auditiva. Los estudiantes identificados con mayores dificultades en el pretest recibieron acompañamiento individualizado durante esta primera sesión de aproximación a la tecnología.

4.3.2.2. Fase de intervención

La implementación de NumIA se desarrolló durante cinco semanas, con una frecuencia de tres sesiones semanales de 30 minutos cada una, bajo supervisión permanente del docente de aula y los investigadores. Esta dosificación es coherente con las recomendaciones metodológicas de Sáez López et al. (2019), quienes destacaron que las sesiones regulares y supervisadas con retroalimentación constante son condición necesaria para garantizar la efectividad del aprendizaje mediado por tecnología en educación primaria.

Esta etapa se desarrolló durante un periodo de cinco semanas consecutivas, en las cuales se realizaron sesiones de práctica estructurada utilizando la aplicación. La planificación de la

intervención se organizó de manera progresiva y secuencial, considerando el nivel de complejidad de los números y el desarrollo gradual de las habilidades grafomotoras de los estudiantes. La distribución de los contenidos se estructuró de la siguiente manera:

- Semana 1: Números del 1 al 10

Durante la primera semana, los estudiantes trabajaron con los números del 1 al 10, los cuales representan la base del sistema numérico y presentan trazos de menor complejidad. En esta etapa, la aplicación permitió a los estudiantes familiarizarse con la interfaz, el uso del dispositivo y la dinámica de retroalimentación automática proporcionada por el sistema.

Las actividades se centraron en:

- Reconocimiento visual del número,
 - Seguimiento del trazo por medio de una guía,
 - Seguimiento del trazo mediante apoyo auditivo proporcionado por la aplicación.
- Semana 2: Números del 10 al 20

En la segunda semana, se introdujo el trazado de los números del 10 al 20, incorporando la escritura de números compuestos por dos dígitos. Esta fase permitió reforzar el aprendizaje de los números básicos y, al mismo tiempo, desarrollar la capacidad de los estudiantes para mantener la secuencia correcta al escribir números más largos.

Las actividades se centraron en:

- Reconocimiento del número
- Seguimiento del trazo por medio de una guía
- Seguimiento de guías de trazo por medio de un audio

- Semanas 3 y 4: Números del 20 al 50

Durante la tercera y cuarta semana, se trabajó con los números comprendidos entre el 20 y el 50. Esta etapa tuvo una mayor duración debido al incremento en la complejidad de los trazos

y a la necesidad de consolidar la escritura de números de dos dígitos con diferentes combinaciones. Las actividades se centraron al igual que la semana 2: reconocimiento, seguimiento del trazo y seguimiento por medio de audio.

En estas semanas, la aplicación registró de manera continua:

- la precisión de los trazos,
- los errores frecuentes,
- y el número de intentos por estudiante.

Esta información permitió monitorear el progreso individual y ajustar el ritmo de práctica según las necesidades observadas.

- **Semana 5:** Números del 50 al 100

En la última semana de intervención, los estudiantes trabajaron con los números del 50 al 100, lo que implicó la práctica de una mayor variedad de combinaciones numéricas y el fortalecimiento de la memoria secuencial y motora.

En esta fase final, se buscó que los estudiantes alcanzaran mayor autonomía en el trazado, reduciendo gradualmente las ayudas visuales de la aplicación y promoviendo la ejecución correcta de los números sin guía directa.

4.3.3. Usabilidad de la aplicación

La usabilidad de NumIA fue evaluada al término de la intervención mediante el System Usability Scale (SUS), instrumento de referencia internacional para la evaluación de interfaces digitales (Brooke, 1996). El cuestionario fue aplicado a los tres docentes responsables de las secciones participantes y a una muestra de diez estudiantes con apoyo del investigador para la lectura de los ítems. El puntaje SUS promedio de los docentes fue de 95.0 sobre 100 y el de los estudiantes de 84.3 sobre 100, obteniéndose un puntaje global de **86.8/100**, clasificado como **Excelente** según los criterios de Bangor et al. (2009) , lo que indica que NumIA presenta una usabilidad muy adecuada para el contexto educativo de primer grado, con valoraciones favorables en facilidad de uso, integración de funciones, satisfacción y curva de aprendizaje de los usuarios.

Tabla 16
Resultados del System Usability Scale (SUS) aplicado a docentes y estudiantes

Grupo	n	Puntaje SUS (/100)	Clasificación
Docentes	3	95.0	Excelente
Estudiantes (muestra)	10	84.3	Bueno
GLOBAL	13	86.8	Excelente

Nota. Clasificación según Bangor et al. (2009): ≥ 85 = Excelente; 71–84 = Bueno; 51–70 = Regular; < 51 = Pobre. Elaboración propia.

La usabilidad de NumIA fue evaluada al término de la intervención mediante el System Usability Scale (SUS), instrumento de referencia internacional para la evaluación de interfaces digitales (Brooke, 1996). El cuestionario fue aplicado a los tres docentes responsables de las secciones participantes y a una muestra de diez estudiantes con apoyo del investigador para la lectura de los ítems. El puntaje SUS promedio de los docentes fue de 95.0 sobre 100 y el de los estudiantes de 84.3 sobre 100, obteniéndose un puntaje global de 86.8/100, clasificado como

Excelente según los criterios de Bangor et al. (2009), lo que indica que NumIA presenta una usabilidad muy adecuada para el contexto educativo de primer grado, con valoraciones favorables en facilidad de uso, integración de funciones, satisfacción y curva de aprendizaje de los usuarios.

4.3.4. Síntesis de Hallazgos del Resultado

En respuesta al tercer objetivo específico, orientado a evaluar el nivel inicial del trazado numérico mediante una prueba pretest e implementar la aplicación NumIA en sesiones de práctica estructurada, se logró establecer una línea base cuantitativa rigurosa y ejecutar la intervención tecnológica de manera articulada con el diagnóstico previo.

Como resultado del proceso de diagnóstico, se obtuvo un perfil detallado por criterio para los 30 estudiantes de la muestra el 23% contaba con un nivel de Logro Destacado (AD) en el pretest (7 estudiantes), el 40% se ubicó en Logro Esperado (A) (12 estudiantes), el 20% en En Proceso (B) (6 estudiantes) y el 17% en En Inicio (C) (5 estudiantes), con una media global de 72.4%. El nivel modal del diagnóstico fue el nivel A, lo que indica que la mayoría de los estudiantes presentaba bases grafomotoras en el rango de logro esperado al inicio de la intervención, aunque con margen sustantivo de mejora. Los cinco estudiantes en nivel C representan los casos de mayor rezago grafomotor, distribuidos entre las tres secciones. Se identificó que la Secuencia de fluidez (71.6%), la Secuencia de orden (71.9%) y la Precisión (71.8%) son las dimensiones sistemáticamente más deficitarias en las tres secciones, mientras que la Orientación (74.5%) resultó el criterio relativamente más consolidado al inicio de la intervención.

La aplicación fue implementada durante cinco semanas de práctica estructurada con tres sesiones semanales de 30 minutos, con los 30 estudiantes de la muestra (10 por sección). El proceso se desarrolló siguiendo el protocolo prediseñado, con supervisión docente e

investigadora permanente, generando registros automáticos de progreso individual por estudiante que documentan el proceso de aprendizaje de forma objetiva y continua.

Los datos del pretest y los registros de la implementación constituyen la base directa sobre la que se asienta el análisis comparativo del Objetivo 4, donde se aplica el postest y se evaluará estadísticamente.

4.4. Aplicación del postest y análisis de resultados

Para dar cumplimiento al cuarto objetivo específico: "Analizar el impacto de la intervención mediante una evaluación final (postest), comparando los resultados con el diagnóstico inicial para determinar las mejoras logradas en los tres criterios evaluados precisión, secuencia y orientación y sus subcriterios de medición", se procedió a la aplicación del postest al mismo grupo de 30 estudiantes de primer grado. En este objetivo se evalúa el nivel final del trazado numérico alcanzado tras la intervención, se comparan los resultados obtenidos con los del diagnóstico inicial y se determina, mediante análisis estadístico inferencial, si las mejoras observadas son estadísticamente significativas y educativamente relevantes.

4.4.1. Aplicación Postest

- **Condiciones de aplicación**

El postest fue administrado al concluir las cinco semanas de intervención, bajo las mismas condiciones del pretest: mismo instrumento, mismos evaluadores, mismas instrucciones estandarizadas.

- **Resultados del Postest**

La Siguiende Tabla presenta los resultados completos del postest. Con la escala (AD 86–100%, A: 66–85%, B: 51–65%, C: 0–50%), se evidencia que 27 de los 30

estudiantes (90%) alcanzaron el nivel AD y los 3 restantes (10%) el nivel A. Ningún estudiante se ubicó en nivel B o C.

Tabla 17*Resultado de la evaluación post test*

Secc.	Nº	Estudiante	Ítems	O	O %	So.	So. %	Sf.	Sf. %	P	P %	Total	Nivel
				pts/máx		pts/máx		pts/máx		pts/máx		%	
1ºA	1	Estudiante 1	73	146/146	100.0	146/146	100.0	146/146	100.0	146/146	100.0	100.0	AD
	2	Estudiante 2	78	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	3	Estudiante 3	77	148/154	96.1	149/154	96.8	151/154	98.1	148/154	96.1	96.8	AD
	4	Estudiante 4	78	126/156	80.8	126/156	80.8	126/156	80.8	126/156	80.8	80.8	A
	5	Estudiante 5	69	107/138	77.5	107/138	77.5	107/138	77.5	107/138	77.5	77.5	A
	6	Estudiante 6	78.0	137/156	87.8	140/156	89.7	142/156	91.0	137/156	87.8	89.1	AD
	7	Estudiante 7	78.0	149/156	95.5	149/156	95.5	149/156	95.5	149/156	95.5	95.5	AD
	8	Estudiante 8	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	9	Estudiante 9	60.0	108/120	90.0	107/120	89.2	110/120	91.7	103/120	85.8	89.2	AD
	10	Estudiante 10	78.0	119/156	76.3	119/156	76.3	119/156	76.3	119/156	76.3	76.3	A
Promedio 1ºA					90.4%		90.6%		91.1%		90.0%	90.5%	
1ºB	1	Estudiante 11	56.0	110/112	98.2	110/112	98.2	110/112	98.2	110/112	98.2	98.2	AD

	2	Estudiante 12	73.0	146/146	100.0	145/146	99.3	144/146	98.6	144/146	98.6	99.1	AD
	3	Estudiante 13	20.0	40/40	100.0	40/40	100.0	40/40	100.0	40/40	100.0	100.0	AD
	4	Estudiante 14	76.0	148/152	97.4	147/152	96.7	142/152	93.4	146/152	96.1	95.9	AD
	5	Estudiante 15	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	6	Estudiante 16	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	7	Estudiante 17	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	8	Estudiante 18	78.0	149/156	95.5	149/156	95.5	149/156	95.5	149/156	95.5	95.5	AD
	9	Estudiante 19	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	10	Estudiante 20	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	Promedio 1°B				99.1%		99.0%		98.6%		98.8%	98.9%	
1°C	1	Estudiante 21	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
	2	Estudiante 22	45.0	90/90	100.0	88/90	97.8	85/90	94.4	80/90	88.9	95.3	AD
	3	Estudiante 23	78.0	156/156	100.0	156/156	100.0	156/156	100.0	150/150	100.0	100.0	AD
	4	Estudiante 24	78.0	139/156	89.1	139/156	89.1	139/156	89.1	139/156	89.1	89.1	AD
	5	Estudiante 25	78.0	150/156	96.2	150/156	96.2	150/156	96.2	150/156	96.2	96.2	AD
	6	Estudiante 26	78.0	156/156	100.0	153/156	98.1	150/156	96.2	151/156	96.8	97.8	AD

7	Estudiante 27	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
8	Estudiante 28	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
9	Estudiante 29	78.0	153/156	98.1	153/156	98.1	153/156	98.1	153/156	98.1	98.1	AD
10	Estudiante 30	78.0	156/156	100.0	156/156	100.0	156/156	100.0	156/156	100.0	100.0	AD
Promedio 1°C				98.3%		97.9%		97.4%		96.9%	97.7%	
GLOBAL n = 30				95.9%		95.8%		95.7%		95.2%	95.7%	

Nota. Elaboración propia

4.4.1.1. *Análisis por Sección*

Sección 1ºA: Alcanza una media posttest de 90.5% (nivel AD), con una ganancia respecto al pretest de +28.6 pp. Siete estudiantes alcanzan el nivel AD y tres permanecen en nivel A (Est. 4, Est. 5 y Est. 10), en el rango 76.3%–80.8%, lo que representa la mayor variabilidad residual de la sección.

Sección 1ºB: Tiene una media posttest más alta 98.9% (nivel AD). Los diez estudiantes alcanzan el nivel AD. Siete obtienen el 100.0% en los cuatro criterios. La única variación negativa de toda la investigación corresponde al Est. 8 (–3.9 pp), cuyo puntaje descendió de 99.4% a 95.5%, manteniéndose con holgura en el nivel AD; esta variación es explicable por la variabilidad natural de la medición en el rango de techo del instrumento. Este grupo muestra el mayor impacto de la intervención en términos de homogeneidad de resultados.

Sección 1ºC: Presenta una media posttest de 97.7% (nivel AD), con los cuatro criterios prácticamente igualados (O=98.3%, So.=97.9%, Sf.=97.4%, P=96.9%). Los diez estudiantes alcanzan el nivel AD. La igualdad entre criterios en el posttest confirma que NumIA equilibró las dimensiones del trazado: los estudiantes que presentaban déficit específico en algún criterio lo corrigieron hasta alinearse con los demás..

4.4.1.2. *Análisis Descriptivo Comparativo Pretest vs. Posttest*

4.4.1.2.1. *Tablas Comparativas Individuales por Sección*

Las Sigüientes Tablas presentan la comparación individual para cada estudiante. ↑ indica ascenso de un nivel; ↑↑ de dos niveles; ↑↑↑ de tres niveles (C → AD).

Tabla 18
Comparación individual – sección A

Nº	Estudiante	Pre %	Nv.	Post %	Nv.	Diff.	Cambio
1	Estudiante 1.	79.2%	A	100.0%	AD	+20.8 pp	A → AD ↑
2	Estudiante 2	63.0%	B	100.0%	AD	+37.0 pp	B → AD ↑↑

3	Estudiante 3	43.2%	C	96.8%	AD	+53.6 pp	C → AD ↑↑↑
4	Estudiante 4	77.5%	A	80.8%	A	+3.3 pp	A (sin cambio)
5	Estudiante 5	54.1%	B	77.5%	A	+23.4 pp	B → A ↑
6	Estudiante 6	53.2%	B	89.1%	AD	+35.9 pp	B → AD ↑↑
7	Estudiante 7	74.0%	A	95.5%	AD	+21.5 pp	A → AD ↑
8	Estudiante 8	80.8%	A	100.0%	AD	+19.2 pp	A → AD ↑
9	Estudiante 9	39.9%	C	89.2%	AD	+49.3 pp	C → AD ↑↑↑
10	Estudiante 10	54.5%	B	76.3%	A	+21.8 pp	B → A ↑
Promedio 1°A		61.9%		90.5%	AD	+28.6 pp	

Nota. Elaboración Propia.

La comparación individual de la sección 1°A permite identificar tres patrones de cambio diferenciados. En primer lugar, los dos estudiantes que iniciaron en nivel C (Est. 3 y Est. 9, con 43.2% y 39.9% respectivamente) registran las ganancias más altas de la sección: +53.6 pp y +49.3 pp, ascendiendo ambos al nivel AD en tres escalones (C → AD ↑↑↑). Este patrón confirma que la intervención fue especialmente eficaz para los estudiantes con mayor rezago, coherente con el principio de que la retroalimentación inmediata tiene mayor impacto cuando el punto de partida es más bajo. En segundo lugar, los cuatro estudiantes que iniciaron en nivel B muestran trayectorias diferenciadas: Est. 2 y Est. 6 ascienden al nivel AD (+37.0 pp y +35.9 pp respectivamente, ↑↑ y ↑↑), mientras que Est. 5 y Est. 10 ascienden solo al nivel A (+23.4 pp y +21.8 pp, ↑), sin lograr alcanzar el umbral AD. Estos dos últimos son los únicos de la sección que no alcanzan el nivel AD en el posttest. En tercer lugar, los cuatro estudiantes que iniciaron en nivel A presentan comportamientos distintos: Est. 1, Est. 7 y Est. 8 ascienden al nivel AD (ganancias de +20.8 pp, +21.5 pp y +19.2 pp, ↑), mientras que Est. 4 registra la menor ganancia de la sección (+3.3 pp) y permanece en nivel A (80.8%), lo que se explica por el efecto de techo

relativo: ya partía de 77.5% y el margen de mejora hasta AD era estrecho dentro del contexto de la sección..

Tabla 19
Comparación individual – sección B

Nº	Estudiante	Pre %	Nv.	Post %	Nv.	Diff.	Cambio
1	Estudiante 11	49.4%	C	98.2%	AD	+48.8 pp	C → AD ↑↑↑
2	Estudiante 12	59.6%	B	99.1%	AD	+39.5 pp	B → AD ↑↑
3	Estudiante 13	98.5%	AD	100.0%	AD	+1.5 pp	AD (sin cambio)
4	Estudiante 14	37.3%	C	95.9%	AD	+58.6 pp	C → AD ↑↑↑
5	Estudiante 15	91.5%	AD	100.0%	AD	+8.5 pp	AD (sin cambio)
6	Estudiante 16	100.0%	AD	100.0%	AD	0.0 pp	AD (sin cambio)
7	Estudiante 17	80.3%	A	100.0%	AD	+19.7 pp	A → AD ↑
8	Estudiante 18	99.4%	AD	95.5%	AD	−3.9 pp	AD (sin cambio)
9	Estudiante 19	84.8%	A	100.0%	AD	+15.2 pp	A → AD ↑
10	Estudiante 20	99.0%	AD	100.0%	AD	+1.0 pp	AD (sin cambio)
Promedio 1ºB		80.0%		98.9%	AD	+18.9 pp	

Nota. Elaboración Propia.

La sección 1ºB presenta el perfil de cambio más heterogéneo de las tres secciones, consecuencia directa de haber tenido el diagnóstico inicial más disperso: cinco estudiantes en AD, uno en B y dos en C en el pretest. Los dos estudiantes que iniciaron en nivel C (Est. 11 y Est. 14, con 49.4% y 37.3%) registran las ganancias más altas de la sección con +48.8 pp y +58.6 pp respectivamente, ascendiendo ambos al nivel AD en tres escalones (C → AD ↑↑↑). El Est. 14, con el puntaje pretest más bajo de su sección (37.3%), alcanza el 95.9% en el posttest, siendo la segunda ganancia individual más alta de toda la investigación.

El Est. 2, que partía de nivel B (59.6%), asciende al nivel AD con +39.5 pp (B → AD ↑↑). Los dos estudiantes en nivel A (Est. 7 y Est. 9, con 80.3% y 84.8%) ascienden al nivel AD con ganancias de +19.7 pp y +15.2 pp (↑). Los cinco estudiantes que ya se ubicaban en nivel AD en el pretest consolidan o mantienen ese nivel: tres alcanzan el 100.0%, uno registra +8.5 pp adicionales y uno registra la única variación negativa de la investigación (Est. 8, -3.9 pp), manteniéndose en AD. Cabe destacar el caso de Est. 6, que partía ya del 100.0% en el pretest y mantiene ese puntaje en el postest (0.0 pp de variación), lo que refleja un dominio ya consolidado desde el diagnóstico inicial.

Tabla 20
Comparación individual – sección C

Nº	Estudiante	Pre %	Nv.	Post %	Nv.	Diff.	Cambio
1	Estudiante 21	70.8%	A	100.0%	AD	+29.2 pp	A → AD ↑
2	Estudiante 22	47.3%	C	95.3%	AD	+48.0 pp	C → AD ↑↑↑
3	Estudiante 23	84.0%	A	100.0%	AD	+16.0 pp	A → AD ↑
4	Estudiante 24	51.4%	B	89.1%	AD	+37.7 pp	B → AD ↑↑
5	Estudiante 25	69.2%	A	96.2%	AD	+27.0 pp	A → AD ↑
6	Estudiante 26	85.2%	A	97.8%	AD	+12.6 pp	A → AD ↑
7	Estudiante 27	96.2%	AD	100.0%	AD	+3.8 pp	AD (sin cambio)
8	Estudiante 28	84.6%	A	100.0%	AD	+15.4 pp	A → AD ↑
9	Estudiante 29	75.7%	A	98.1%	AD	+22.4 pp	A → AD ↑
10	Estudiante 30	89.7%	AD	100.0%	AD	+10.3 pp	AD (sin cambio)
Promedio 1°C		75.4%		97.7%	AD	+22.2 pp	

Nota. Elaboración Propia.

La sección 1°C presenta la distribución de cambio más homogénea y consistente de las tres secciones. Los diez estudiantes alcanzan el nivel AD en el postest, sin excepción.

El único estudiante que inició en nivel C (Est. 22, 47.3%) registra la mayor ganancia de la sección: +48.0 pp (C → AD ↑↑↑), pasando de 47.3% a 95.3%. El único estudiante que inició en nivel B (Est. 24, 51.4%) asciende al nivel AD con +37.7 pp (B → AD ↑↑). Los seis estudiantes que iniciaron en nivel A (Est. 21, Est. 23, Est. 25, Est. 26, Est. 28 y Est. 29) ascienden todos al nivel AD con ganancias que oscilan entre +12.6 pp (Est. 26, que ya partía de 85.2%) y +29.2 pp (Est. 21). Los dos estudiantes que ya estaban en nivel AD en el pretest (Est. 27 con 96.2% y Est. 30 con 89.7%) consolidan su dominio alcanzando el 100.0% en el posttest.

La igualdad entre criterios en el posttest de esta sección (rango entre O=98.3% y P=96.9%) confirma que NumIA equilibró las dimensiones del trazado los estudiantes que presentaban déficit específico en algún criterio lo corrigieron progresivamente hasta alinearse con los demás, lo que evidencia la efectividad de la retroalimentación diferenciada por criterio que ofrece la aplicación.

4.4.1.2.2. *Estadísticos Descriptivos Consolidados*

Estadísticos descriptivos comparativos Pretest vs. Posttest (n = 30)

Se presenta una tabla en donde muestra los cálculos obtenidos para los estadísticos para una muestra de 30 estudiantes.

Tabla 21
Estadísticos descriptivos comparativos pretest–posttest

Estadístico	Pretest	Posttest	Diferencia	Interpretación
(d)				
Media ($\bar{x}$)	72.4%	95.7%	+23.2 pp	Mejora global de 23.2 puntos porcentuales tras la intervención con NumIA
Desviación estándar (s)	19.4%	6.8%	-12.6	Reducción de la dispersión: mayor homogeneidad en el posttest

Mediana	76.6%	98.7%	+22.1 pp	Más de la mitad de los estudiantes alcanzaron el máximo posible en el postest
Mínimo	37.3%	76.3%	+39.0pp	El mínimo postest supera el umbral A (66%); variación negativa solo de un estudiante (se mantiene en AD)
Máximo	100.0%	100.0%	+58.6 pp	Mayor ganancia (C → AD, +58.6 pp)

Nota. Elaboración Propia.

La tabla de estadísticos descriptivos comparativos sintetiza la magnitud y la naturaleza de la transformación producida por la intervención. La media global asciende de 72.4% a 95.7%, una ganancia de +23.2 pp que posiciona al grupo en el nivel AD de manera holgada (umbral: 86%). La mediana postest de 98.7% frente al 76.6% del pretest señala que más de la mitad de los estudiantes alcanzó un puntaje superior al 98.7% en el postest, un resultado que no tiene precedente en el diagnóstico inicial.

El indicador más revelador de la transformación cualitativa del grupo es la reducción de la desviación estándar: de 19.4% en el pretest a 6.8% en el postest, una disminución de 12.6 puntos. Este fenómeno de convergencia indica que, además de mejorar la media, la intervención redujo la heterogeneidad del grupo, acercando los resultados individuales entre sí. El mínimo postest (76.3%) supera el umbral del nivel A (66%), lo que confirma que incluso el estudiante de menor rendimiento en el postest se ubica dentro del rango de logro esperado. La máxima ganancia individual (+58.6 pp, Est. 14) y la única variación negativa (-3.9 pp, Est. 18) delimitan el rango de comportamiento individual, cuya amplitud de 62.5 pp contrasta con la amplitud del pretest (62.7 pp), evidenciando que la dispersión individual se mantiene en magnitud similar, pero con un desplazamiento sistemático hacia arriba en toda la distribución.

Distribución de niveles PRE vs. POST

Tabla 22

Distribución de niveles pretest vs. posttest por sección

Sección	AD	A	B	C	→	AD	A	B	C	%	+Diff.
	pre	pre	pre	pre		post	post	post	post	Final	
1ºA (n=10)	0	4	4	2	→	7	3	0	0	90.5%	+28.6pp
1ºB (n=10)	5	2	1	2	→	10	0	0	0	98.9%	+18.9 pp
1ºC (n=10)	2	6	1	1	→	10	0	0	0	97.7%	+22.2 pp
Global (n=30)	7	12	6	5	→	27	3	0	0	95.7%	+23.2 pp

Nota. Elaboración Propia.

La tabla de distribución de niveles cuantifica con precisión la movilidad ascendente producida por la intervención. A nivel global, el perfil diagnóstico (AD:23%, A:40%, B:20%, C:17%) se transforma radicalmente en el posttest (AD:90%, A:10%, B:0%, C:0%), representando una inversión completa de la distribución: el nivel dominante pasa de A a AD, y los niveles B y C desaparecen por completo.

Por sección, la movilidad es total en 1ºB y 1ºC los diez estudiantes de cada una alcanzan el nivel AD, independientemente de su punto de partida. En 1ºA, siete estudiantes alcanzan AD y tres permanecen en A (Est. 4, Est. 5 y Est. 10), todos con puntajes entre 76.3% y 80.8%, muy próximos al umbral de corte (86%). Esta diferencia entre secciones es coherente con el hecho de que 1ºA era la sección con el diagnóstico más bajo (media pretest: 61.9%, nivel B) y la mayor proporción de estudiantes en niveles B y C (80%). El avance de 1ºA, que pasa de 0% en AD a 70%, es proporcionalmente el más significativo de las tres secciones en términos de movilidad de nivel, aunque su media posttest (90.5%) sea la más baja de las tres.

4.4.2. Análisis Estadístico Inferencial

4.4.2.1. Planteamiento de Hipótesis

- **H₀:** El uso de la aplicación NumIA no produce mejora estadísticamente significativa en el trazado correcto de los números arábigos en los estudiantes de primer grado de la I.E. Fe y Alegría N.º 28. ($\mu_d = 0$)
- **H₁:** El uso de la aplicación NumIA sí produce mejora estadísticamente significativa en el trazado correcto de los números arábigos. ($\mu_d > 0$)
- Contraste unilateral de cola derecha. Nivel de significancia: $\alpha = 0.05$.

4.4.2.2. Definición de variables y fórmulas del Análisis

Las variables empleadas en el análisis estadístico se definen de la siguiente manera.

x_i representa el % de Logro Final del estudiante i en el pretest, calculado como:

$$x_i = \frac{[\sum(O_i + S_{0i} + S_{fi} + P_i)]}{n_{item} \times 8} (100)$$

y_i representa el % de Logro Final del mismo estudiante i en el posttest, con la misma fórmula.

- La diferencia individual es: $d_i = y_i - x_i$, que expresa la ganancia en puntos porcentuales (pp) de cada estudiante.
- La media de las diferencias es: $\bar{d} = \frac{1}{n} \times \sum d_i$ que estima la mejora media del grupo.
- La desviación estándar de las diferencias es $S = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n-1}}$, la cual mide la variabilidad de las ganancias individuales.
- El error estándar es $e = \frac{S}{\sqrt{n}}$
- El estadístico t de Student pareado es $t = \frac{\bar{d}}{e} = \bar{d} \cdot \sqrt{\frac{n}{S}}$, con $gl = n - 1$ grados de libertad.

- El intervalo de confianza al 95% (IC 95%) = $\bar{d} \pm t\left(\frac{\alpha}{2}, gl\right) \cdot e = \bar{d} \pm t(0.025, gl) \cdot e$
- El tamaño del efecto es d de Cohen = $\frac{\bar{d}}{s}$ (pequeño: $d < 0.5$; mediano: $0.5 \leq d < 0.8$; grande: $d \geq 0.8$, Cohen, 1988).
- El estadístico W de la prueba de Shapiro-Wilk es $w = \frac{(\sum a_i \cdot x_{(i)})^2}{\sum (x_i - \bar{x})^2}$, donde $x_{(i)}$ son los valores ordenados de las diferencias y a_i son los coeficientes tabulados para la muestra de tamaño n .

4.4.2.3. Verificación del Supuesto de Normalidad

Prueba de Shapiro-Wilk

Se verificó la normalidad de las diferencias $d_i = y_i - x_i$ en cada grupo mediante la prueba de Shapiro-Wilk ($\alpha = 0.05$). Un valor $p \geq 0.05$ indica que no se rechaza la hipótesis nula de normalidad.

Tabla 23
Resultados de Shapiro-Wilk

Grupo	n	Estadístico W	Valor p	Conclusión	Prueba inferencial a aplicar
Sección 1°A	10	0.9289	0.4371	$p \geq 0.05$ Normal	t de Student pareada (gl=9)
Sección 1°B	10	0.8694	0.0983	$p \geq 0.05$ Normal	t de Student pareada (gl=9)
Sección 1°C	10	0.9597	0.7822	$p \geq 0.05$ Normal	t de Student pareada (gl=9)
Global	30	0.9548	0.2262	$p \geq 0.05$ Normal	t de Student pareada (gl=29)

Nota. Elaboración Propia.

La prueba de Shapiro-Wilk estadístico recomendado para muestras pequeñas ($n < 50$) confirma el supuesto de normalidad de las diferencias $d_i = y_i - x_i$ en los cuatro niveles de análisis. Los valores obtenidos son: 1°A ($W=0.9289$, $p=0.4371$), 1°B ($W=0.8694$, $p=0.0983$), 1°C ($W=0.9597$, $p=0.7822$) y Global $n=30$ ($W=0.9548$, $p=0.2262$). En todos los casos el valor $p > 0.05$ supera el umbral de $\alpha = 0.05$, por lo que no se rechaza la hipótesis nula de normalidad en ninguno de los grupos

4.4.2.4. Cálculo Detallado de Diferencias por Sección

Las siguientes tablas presentan el cálculo completo de las diferencias individuales d_i para cada sección, con las desviaciones $d_i - \bar{d}$ y los cuadrados $d_i - \bar{d}^2$ necesarios para obtener **SD**. La fila de resumen muestra los estadísticos resultantes: $\bar{d}$ (media de diferencias), SD (desviación estándar), SE (error estándar), t (estadístico de Student), p (valor p unilateral), IC 95% y d de Cohen.

Tabla 24*Cálculo de diferencias individuales y estadísticos Sección 1ªA*

Estudiante	x_i (Pre)	y_i (Post)	$d_i = y_i - x_i$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$	Nv. Pre	Nv. Post
1	79.2%	100.0%	+20.8	-7.78	60.53	A	AD
2	63.0%	100.0%	+37.0	+8.42	70.90	B	AD
3	43.2%	96.8%	+53.6	+25.02	626.00	C	AD
4	77.5%	80.8%	+3.3	-25.28	639.08	A	A
5	54.1%	77.5%	+23.4	-5.18	26.83	B	A
6	53.2%	89.1%	+35.9	+7.32	53.58	B	AD
7	74.0%	95.5%	+21.5	-7.08	50.13	A	AD
8	80.8%	100.0%	+19.2	-9.38	87.98	A	AD
9	39.9%	89.2%	+49.3	+20.72	429.32	C	AD
10	54.5%	76.3%	+21.8	-6.78	45.97	B	A
			285.8	0	2090.32		
SD=15.24	$\bar{d}=28.6$ pp	SE=4.82	t=5.93	p = 0.0001	IC=[17.7;39.5]	d=1.88	SW:0.9289 p=0.4371

Nota. Elaboración Propia

Los estadísticos inferenciales de la sección 1ºA confirman que la mejora observada es real y de gran magnitud. La media de las diferencias ($\bar{d} = 28.6$ pp) representa la ganancia promedio de cada estudiante, con una desviación estándar de $SD = 15.24$ pp que refleja la variabilidad moderada entre trayectorias tan distintas como $+3.3$ pp y $+53.6$ pp. El estadístico $t(9) = 5.93$ con $p = 0.0001$ permite rechazar H_0 con un nivel de confianza del 99.99%: la mejora no es atribuible al azar. El intervalo de confianza al 95% $[17.7; 39.5]$ confirma que, incluso en el escenario más conservador, la mejora real de la población supera los 17.7 pp. El tamaño del efecto de Cohen ($d = 1.88$) sitúa el resultado de 1ºA en la categoría de efecto muy grande (umbral de efecto grande: $d \geq 0.8$), próximo a dos desviaciones estándar de ganancia. La prueba de Shapiro-Wilk ($W = 0.9289$, $p = 0.4371$) valida que las diferencias siguen distribución normal, confirmando la pertinencia del contraste paramétrico aplicado.

Tabla 25*Cálculo de diferencias individuales y estadísticos Sección 1ºB*

Est.	$x_i(\text{Pre})$	$y_i(\text{Post})$	$d_i = y_i - x_i$	$d_i - \bar{d}$	$d_i - \bar{d}^2$	Nv. Pre	Nv. Post
1	49.4%	98.2%	+48.8	+29.91	894.61	C	AD
2	59.6%	99.1%	+39.5	+20.61	424.77	B	AD
3	98.5%	100.0%	+1.5	-17.39	302.41	AD	AD
4	37.3%	95.9%	+58.6	+39.71	1576.88	C	AD
5	91.5%	100.0%	+8.5	-10.39	107.95	AD	AD
6	100.0%	100.0%	0.0	-18.89	356.83	AD	AD
7	80.3%	100.0%	+19.7	+0.81	0.66	A	AD
8	99.4%	95.5%	-3.9	-22.79	519.38	AD	AD
9	84.8%	100.0%	+15.2	-3.69	13.62	A	AD
10	99.0%	100.0%	+1.0	-17.89	320.05	AD	AD
			188.9	0	4517.16		
SD=22.4 $\bar{d}$=18.9 pp SE=7.08			t=2.67	p = 0.0129	IC=[2.9;34.9]	d=0.84	SW:0.8694 p=0.0983

Nota. Elaboración Propia

Los estadísticos de la sección 1°B muestran el patrón más heterogéneo de las tres secciones. La media de las diferencias ($\bar{d}=18.9$ pp) es la más baja del conjunto, lo que refleja que cinco de los diez estudiantes partían ya del nivel AD con escaso margen de mejora medible, moderando la ganancia media por efecto de techo. La desviación estándar $SD=22.4$ pp es la más alta de las tres secciones, consecuencia directa de la coexistencia de diferencias extremas, desde -3.9 pp hasta $+58.6$ pp en el mismo grupo. A pesar de esta variabilidad, el estadístico $t(9) = 2.67$ con $p=0.0129$ permite rechazar H_0 , confirmando que la mejora media es estadísticamente significativa con $\alpha=0.05$. El intervalo de confianza al 95% $[2.9; 34.9]$, aunque más amplio que el de las otras secciones, no incluye el cero, confirmando la significancia de la mejora. El tamaño del efecto ($d=0.84$) se clasifica como grande según Cohen (1988), si bien es el más moderado de las tres secciones, coherente con el mayor peso relativo de los estudiantes con efecto de techo. La prueba de Shapiro-Wilk ($W=0.8694$, $p=0.0983$) confirma normalidad de las diferencias.

Tabla 26*Cálculo de diferencias individuales y estadísticos Sección 1°C*

Estudiante	x_i (Pre)	y_i (Post)	$d_i = y_i - x_i$	$d_i - \bar{d}$	$(d_i - \bar{d})^2$	Nv. Pre	Nv. Post
1	70.8%	100.0%	+29.2	+6.96	48.44	A	AD
2	47.3%	95.3%	+48.0	+25.76	663.58	C	AD
3	84.0%	100.0%	+16.0	-6.24	38.94	A	AD
4	51.4%	89.1%	+37.7	+15.46	239.01	B	AD
5	69.2%	96.2%	+27.0	+4.76	22.66	A	AD
6	85.2%	97.8%	+12.6	-9.64	92.93	A	AD
7	96.2%	100.0%	+3.8	-18.44	340.03	AD	AD
8	84.6%	100.0%	+15.4	-6.84	46.79	A	AD
9	75.7%	98.1%	+22.4	+0.16	0.03	A	AD
10	89.7%	100.0%	+10.3	-11.94	142.56	AD	AD
			222.4	0	1634.97		
$\bar{d}=22.2$ pp	SD=13.48	SE=4.26	t=5.22	p < 0.0003	IC= [12.6;31.9]	d=1.65	SW:0.9597
							p=0.7822

Nota. Elaboración Propia

La sección 1°C presenta estadísticos robustos y homogéneos. La media de diferencias ($\bar{d}$ =22.2 pp) se acompaña de la desviación estándar más baja de las tres secciones (SD=13.48 pp), lo que genera un error estándar reducido (SE=4.26) y el estadístico $t(9)=5.22$ con $p=0.0003$. El intervalo de confianza al 95% [12.6; 31.9] es el más estrecho de los tres, expresando la mayor precisión de la estimación en esta sección. El tamaño del efecto ($d=1.65$) se clasifica como muy grande y refleja una práctica más uniforme y un punto de partida más homogéneo que en las otras secciones. La prueba de Shapiro-Wilk ($W=0.9597$, $p=0.7822$) confirma normalidad de las diferencias.

4.4.2.5. Prueba *t* de Student Pareada

Para la población global tenemos los siguientes resultados:

- $\bar{d}=23.2$ pp
- **SD** =17.35
- **SE**=3.17
- **t** = 7.33, $p < 0.0001$
- **IC 95%** = [16.8;29.7]
- **d Cohen**=1.34.

Tabla 27*Resultados consolidados de la prueba t de Student pareada*

Grupo	n	Prueba	Estadístico t	p (unilat.)	d Cohen	Magnitud	IC 95% (pp)	Decisión
1°A	10	t pareada gl = 9	t = 5.93	< 0.0001	1.88	MUY GRANDE	[17.7; 39.5]	Rechaza H₀
1°B	10	t pareada gl = 9	t = 2.67	0.0129	0.84	GRANDE	[2.9; 34.9]	Rechaza H₀
1°C	10	t pareada gl = 9	t = 5.22	0.0003	1.65	MUY GRANDE	[12.6; 31.9]	Rechaza H₀
Global	30	t pareada gl = 29	t = 7.33	< 0.0001	1.34	MUY GRANDE	[16.8; 29.7]	Rechaza H₀

Nota. Elaboración Propia.

La tabla consolidada de la prueba t de Student pareada permite comparar los resultados inferenciales de las tres secciones y del análisis global. El análisis global ($n=30$, $gl=29$, $t=7.33$, $p<0.0001$) confirma la significación estadística del conjunto, con $\bar{d}=23.24$ pp, $SD=17.35$ y $SE=3.17$. El intervalo de confianza al 95% global [16.8; 29.7 pp] es el más estrecho de los cuatro análisis por efecto del mayor tamaño muestral, y su no inclusión del cero confirma con el mayor nivel de certeza disponible que la mejora es real. El tamaño del efecto global (d Cohen=1.34) se clasifica como muy grande ($d \geq 0.8$), indicando que la intervención con NumIA produjo una ganancia de 1.34 desviaciones estándar en el trazado numérico de los 30 estudiantes.

4.4.2.6. Verificación de la Hipótesis de Investigación

La hipótesis general del estudio establece que el uso de la aplicación NumIA mejora significativamente el trazado correcto de los números arábigos en estudiantes de primer grado, en comparación con su propio desempeño previo. La siguiente tabla verifica sistemáticamente cada componente de esta hipótesis contra la evidencia empírica.

Tabla 28
Verificación por componente de la hipótesis de investigación

Componente de la hipótesis	Evidencia empírica obtenida	Verificado
"mejora"	Ganancia media global +23.2 pp (72.4% → 95.7%); 29 de 30 estudiantes con diferencia positiva.	Sí
"significativamente"	1°A: $t=5.93$, $p=0.0001$; 1°B: $t=2.67$, $p=0.0129$; 1°C: $t=5.22$, $p=0.0003$; Global: $t=7.33$, $p<0.0001$. Todos los $p < \alpha=0.05$. Prueba t de Student pareada, normalidad verificada en todos los grupos (Shapiro-Wilk).	Sí

"trazado correcto de números arábigos"	Rúbrica de 4 criterios (O, So., Sf., P); instrumento de hasta 78 ítems aplicado en pretest y posttest bajo condiciones uniformes.	Sí
"estudiantes de primer grado"	30 estudiantes (6–7 años) de la I.E. Fe y Alegría N.º 28, Chiclayo, Lambayeque, año escolar 2025.	Sí
"comparación con desempeño previo"	Diseño cuasi-experimental pretest–posttest; mismo instrumento, mismas condiciones, mismos evaluadores en ambas aplicaciones.	Sí
"90% en AD, 10% en A"	27/30 estudiantes (90%) en AD y 3/30 (10%) en A en el posttest. Cero en B o C. 29/30 con $d > 0$.	Sí

Nota. Elaboración Propia.

La hipótesis queda confirmada estadística y prácticamente con evidencia replicada en las tres secciones independientes. La prueba *t* de Student pareada es válida en todos los grupos (normalidad verificada mediante Shapiro-Wilk: *W* entre 0.8694 y 0.9597, todos con $p > 0.05$), con tamaños del efecto grandes a muy grandes (*d* Cohen entre 0.84 y 1.88), con el 90% de la muestra en AD y el 10% en A, y con una ganancia media global de +23.2 pp. La Precisión fue el criterio sistemáticamente más deficitario en el pretest (71.8%) y el que mayor señal correctiva generó en el modelo automático de NumIA, evidenciando coherencia entre el diagnóstico manual y la evaluación automática. En conjunto, el uso de la aplicación NumIA produjo una mejora real, estadísticamente verificada, educativamente relevante y consistente en el trazado correcto de los números arábigos en los estudiantes de primer grado.

4.4.3. Registros Automáticos de la Aplicación NumIA

De acuerdo con el cuarto objetivo específico, el cual establece analizar el impacto de la intervención mediante una evaluación final (posttest) y registros automáticos de la aplicación, la presente sección reporta los datos generados automáticamente por la base de datos de NumIA durante el período de intervención (octubre–diciembre de 2025). El análisis se organiza en el

orden establecido por la operacionalización de la variable independiente (Tabla 1): retroalimentación, interactividad y frecuencia de práctica. Los datos se obtuvieron directamente de la tabla usuario_pregunta de la base de datos, filtrando por los identificadores de los 30 estudiantes de la muestra

Tabla 29

Intentos de práctica registrados en NumIA por semana de intervención y sección

Nº.	Semana	1ºA(intentos)	1ºB(intentos)	1ºC(intentos)	Total global
1	SEMANA 1: 27/10/2025 al 02/11/2025	622	519	939	2,080
2	SEMANA 2: 03/11/2025 al 09/11/2025	818	1,133	1,091	3,042
3	SEMANA 3: 17/11/2025 al 23/11/2025	858	358	780	1,996
4	SEMANA 4: 24/11/2025 al 30/11/2025	829	724	754	2,307
5	SEMANA 5: 01/12/2025 al 07/12/2025	620	568	634	1,822
Total		3,747	3,302	4,198	11,247

Nota. Cada semana corresponde a una sesión por sección. Los intentos incluyen todas las actividades de práctica registradas: trazado por voz, trazado por imagen y reconocimiento.

El total de intentos registrados durante las cinco semanas de intervención fue de 11,247, La sección 1ºC acumuló el mayor volumen de práctica, impulsado principalmente por la

semana 2 (1,091 intentos). El volumen de intentos por semana fue variable en función de la asistencia y el ritmo individual de práctica, pero las tres secciones mantuvieron participación activa durante las cinco semanas de uso de la aplicación.

4.4.3.1. Retroalimentación: proporción de mensajes correctivos frente a mensajes afirmativos

La dimensión retroalimentación se operacionaliza como la proporción de mensajes correctivos frente a mensajes afirmativos (refuerzo positivo) emitidos por la aplicación (Tabla 1). Cada vez que un estudiante completaba un intento de trazado, NumIA emitía automáticamente un mensaje: afirmativo si el modelo de redes siamesas clasificó el trazo como correcto, o correctivo si detectó una imprecisión. La Tabla 30 presenta el conteo absoluto de ambos tipos de mensajes por estudiante, el total emitido y la razón afirmativo/correctivo.

Tabla 30*Mensajes de retroalimentación emitidos por NumIA por estudiante (n = 30)*

Secc	Estudiante	Afirmativos (n)	Afirmativos (%)	Correctivos (n)	Correctivos (%)	Total	Razón Afirm./ Correct.
1ºA	Estudiante 1	208	70.3%	88	29.7%	296	2.36:1
1ºA	Estudiante 2	295	71.3%	119	28.7%	414	2.48:1
1ºA	Estudiante 3	212	64.6%	116	35.4%	328	1.83:1
1ºA	Estudiante 4	264	71.9%	103	28.1%	367	2.56:1
1ºA	Estudiante 5	261	69.0%	117	31.0%	378	2.23:1
1ºA	Estudiante 6	268	73.4%	97	26.6%	365	2.76:1
1ºA	Estudiante 7	290	80.6%	70	19.4%	360	4.14:1
1ºA	Estudiante 8	363	77.2%	107	22.8%	470	3.39:1
1ºA	Estudiante 9	230	66.9%	114	33.1%	344	2.02:1
1ºA	Estudiante 10	334	78.6%	91	21.4%	425	3.67:1
Total 1ºA		2725	72.7%	1022	27.3%	3747	2.67:1
1ºB	Estudiante 1	203	67.4%	98	32.6%	301	2.07:1
1ºB	Estudiante 2	318	78.3%	88	21.7%	406	3.61:1
1ºB	Estudiante 3	182	52.3%	166	47.7%	348	1.10:1
1ºB	Estudiante 4	256	57.3%	191	42.7%	447	1.34:1
1ºB	Estudiante 5	255	91.4%	24	8.6%	279	10.63:1
1ºB	Estudiante 6	177	67.3%	86	32.7%	263	2.06:1

1°B	Estudiante 7	202	72.4%	77	27.6%	279	2.62:1
1°B	Estudiante 8	151	56.1%	118	43.9%	269	1.28:1
1°B	Estudiante 9	240	71.6%	95	28.4%	335	2.53:1
1°B	Estudiante 10	263	70.1%	112	29.9%	375	2.35:1
Total 1°B		2247	68.0%	1055	32.0%	3302	2.13:1
1°C	Estudiante 1	332	77.0%	99	23.0%	431	3.35:1
1°C	Estudiante 2	198	63.9%	112	36.1%	310	1.77:1
1°C	Estudiante 3	442	86.2%	71	13.8%	513	6.23:1
1°C	Estudiante 4	216	64.9%	117	35.1%	333	1.85:1
1°C	Estudiante 5	231	72.0%	90	28.0%	321	2.57:1
1°C	Estudiante 6	261	82.6%	55	17.4%	316	4.75:1
1°C	Estudiante 7	398	62.9%	235	37.1%	633	1.69:1
1°C	Estudiante 8	297	66.0%	153	34.0%	450	1.94:1
1°C	Estudiante 9	309	76.1%	97	23.9%	406	3.19:1
1°C	Estudiante 10	320	66.0%	165	34.0%	485	1.94:1
Total 1°C		3004	71.6%	1194	28.4%	4198	2.52:1
Global n = 30		7,976	71.0%	3,271	29.0%	11,247	2.44:1

Nota. Los datos provienen de las tablas usuario_pregunta de la base de datos de NumIA, filtrados para el período octubre a diciembre del 2025. La razón expresa cuántos mensajes afirmativos se emitieron por cada mensaje correctivo. Donde no hubo mensajes correctivos se indica '100% afirm.'.

Durante las cinco semanas de intervención, NumIA emitió un total de 11,247 mensajes de retroalimentación, uno por cada intento de trazado registrado. De ese total, 7,976 fueron afirmativos (70.9%) , el modelo clasificó el trazado como correcto y 3,271 fueron correctivos (29.1%) el modelo detectó una imprecisión en el trazo. A nivel de secciones, 1ºA registró 72.7% de mensajes afirmativos (2,725 / 1,022 correctivos), 1ºB el 68.0% (2,247 / 1,055 correctivos) y 1ºC el 71.6% (3,004 / 1,194 correctivos).

El predominio de mensajes afirmativos (70.9%) refleja que el modelo de redes siamesas de NumIA acepta como correctos los trazos que presentan la orientación, secuencia y forma general adecuadas, reservando el mensaje correctivo para imprecisiones que afectan la legibilidad del número. Este diseño es coherente con la teoría del feedback formativo (Hattie & Timperley, 2007): el refuerzo positivo sostenido favorece la motivación y la autoeficacia del estudiante, mientras que el 29.1% de mensajes correctivos garantiza que las imprecisiones de trazado reciban señal de corrección en tiempo real sin esperar la evaluación diferida del docente. La variabilidad individual en el porcentaje de mensajes afirmativos desde 52.3% hasta 91.4% refleja diferencias en el nivel de precisión de ejecución de cada estudiante a lo largo de la intervención

4.4.3.2. *Interactividad: promedio de tiempo dedicado por actividad*

La dimensión interactividad se operacionaliza como el promedio de tiempo que el usuario dedica a cada actividad o ejercicio en la aplicación, medido en minutos (Tabla 1). La Tabla 31 reporta el tiempo de práctica activa por estudiante calculado a partir del campo tiempo_respuesta de la tabla usuario_pregunta, que registra la duración en segundos de cada intento individual.

Tabla 31

Tiempo prom. Por intento emitidos por NumIA por estudiante (n = 30)

Secc	Alumno	Tiempo prom. por intento	Tiempo total acumulado	Total intentos
1ºA	Alumno 1	68.5	337.8	296
1ºA	Alumno 2	33.8	233.3	414
1ºA	Alumno 3	43.2	235.9	328
1ºA	Alumno 4	31.3	191.5	367
1ºA	Alumno 5	28.3	178.1	378
1ºA	Alumno 6	41.8	254	365
1ºA	Alumno 7	30.1	180.4	360
1ºA	Alumno 8	21.7	169.8	470
1ºA	Alumno 9	31.6	181.3	344
1ºA	Alumno 10	30.4	215.3	425
Total 1ºA		36.1	2177.4	3747
1ºB	Alumno 11	29.9	149.9	301
1ºB	Alumno 12	20.1	136.3	406
1ºB	Alumno 13	34.1	197.7	348
1ºB	Alumno 14	28.7	213.5	447
1ºB	Alumno 15	34.7	161.4	279
1ºB	Alumno 16	44.3	194.1	263
1ºB	Alumno 17	22.8	106.1	279
1ºB	Alumno 18	46.8	210	269
1ºB	Alumno 19	29.3	163.8	335
1ºB	Alumno 20	25.6	160	375
Total 1ºB		31.6	1692.7	3302
1ºC	Alumno 21	22.7	162.8	431
1ºC	Alumno 22	38.3	197.7	310
1ºC	Alumno 23	18.1	154.4	513
1ºC	Alumno 24	35.1	194.8	333
1ºC	Alumno 25	40.1	214.8	321
1ºC	Alumno 26	28	147.7	316
1ºC	Alumno 27	33.3	351.3	633
1ºC	Alumno 28	33.2	249.3	450

1°C	Alumno 29	23.2	156.7	406
1°C	Alumno 30	41.5	335.5	485
Total 1°C		31.3	2164.8	4198
GLOBAL		33	6034.8	11247
n=30				

Nota. Los datos provienen de las tablas usuario_pregunta de la base de datos de NumIA, filtrados para el período octubre a diciembre del 2025.

El tiempo promedio por intento fue de 33 segundos, calculado como la diferencia entre hora_fin y hora_inicio de cada registro en la tabla usuario_pregunta. El tiempo total acumulado de práctica activa por estudiante a lo largo de las cinco semanas fue de 201.2 minutos en promedio (rango: 106.1–351.3 min), equivalente a aproximadamente 3 horas 20 minutos de práctica activa por estudiante. A nivel de secciones, 1°A acumuló el mayor tiempo total (2177.4 min), seguida de 1°C (2,164.8 min) y 1°B (1,692.7 min), diferencia directamente relacionada con el volumen de intentos de cada sección.

El tiempo promedio de 33 segundos por intento refleja el tiempo real que cada estudiante dedicó a trazar el número: desde que inició la actividad hasta que la completó. Este valor es indicativo de un diseño de actividad bien calibrado para estudiantes de primer grado: suficientemente breve para mantener la atención sostenida durante la sesión, y suficientemente estructurado para que cada trazado sea un acto deliberado (Mayer, 2010). La variabilidad individual en el tiempo promedio por intento desde 20.1 s hasta 68.5 s y en el tiempo total acumulado desde 106.1 min hasta 351.3 min refleja diferencias en el ritmo de ejecución y en el nivel de dedicación de cada estudiante a lo largo de la intervención.

4.4.3.3. *Frecuencia de práctica: número de preguntas respondidas por semana.*

La dimensión frecuencia de práctica se operacionaliza como el número de preguntas respondidas por semana (Tabla 1). La Tabla 32 presenta los intentos de práctica registrados por NumIA desagregados por semana de intervención y sección, y la Tabla 33 desglosa ese volumen por estudiante.

Tabla 32

Frecuencia de práctica por sección

Sem.	Contenido	1ºA	1ºB	1ºC	Total
S1	Números 1–10	622	519	939	2,080
S2	Números 11–20	818	1,133	1,091	3,042
S3	Números 21–50	858	358	780	1,996
S4	Números 21–50 (cont.)	829	724	754	2,307
S5	Números 51–100	620	568	634	1,822
Total		3,747	3,302	4,198	11,247

El total de preguntas respondidas durante las cinco semanas de intervención alcanzó **11,247**, equivalente a una media de **2,249 preguntas semanales** para el conjunto de los 30 estudiantes, o **74.9 preguntas por estudiante por semana**. La sección 1ºC registró la mayor frecuencia acumulada (4,198 intentos; 83.9 por estudiante por semana), seguida de 1ºA (3,747; 74.9) y 1ºB (3,302; 66.0). Estos valores superan el umbral de práctica distribuida recomendado para la consolidación de habilidades grafomotoras en el primer grado (Sáez López et al., 2019).

La semana de mayor actividad global fue la **semana 2** con **3,042 intentos**, mientras que la semana de menor actividad fue la **semana 5** con **1,822 intentos**. A diferencia de lo esperado,

la semana 3 sí registró actividad (1,996 intentos), lo que confirma que las cinco semanas de intervención tuvieron sesiones activas de forma continua, sin interrupción. El patrón descendente desde la semana 2 hacia la semana 5 es coherente con el incremento progresivo de la complejidad grafomotora: a medida que los números involucran combinaciones de dos dígitos de mayor variabilidad, la necesidad de reintentos disminuye porque los estudiantes han consolidado la mecánica básica de trazado (Villarroel et al., 2013).

Tabla 33

Frecuencia de práctica por estudiante registrada en NumIA (n = 30)

Secc.	Apellidos y Nombres	S1	S2	S3	S4	S5	Total
1ºA	Alumno 1	11	63	122	59	41	296
1ºA	Alumno 2	125	48	89	80	72	414
1ºA	Alumno 3	53	72	79	57	67	328
1ºA	Alumno 4	93	—	134	77	63	367
1ºA	Alumno 5	87	97	—	98	96	378
1ºA	Alumno 6	52	61	100	81	71	365
1ºA	Alumno 7	71	140	10	98	41	360
1ºA	Alumno 8	72	90	79	119	110	470
1ºA	Alumno 9	58	75	135	76	—	344
1ºA	Alumno 10	—	172	110	84	59	425
	Total 1ºA	622	818	858	829	620	3,747
1ºB	Alumno 1	—	100	—	112	89	301
1ºB	Alumno 2	177	—	—	126	103	406
1ºB	Alumno 3	55	201	—	—	92	348
1ºB	Alumno 4	76	129	103	78	61	447

1ºB	Alumno 5	49	72	42	65	51	279
1ºB	Alumno 6	47	122	32	35	27	263
1ºB	Alumno 7	74	93	—	71	41	279
1ºB	Alumno 8	41	128	—	74	26	269
1ºB	Alumno 9	—	162	80	59	34	335
1ºB	Alumno 10	—	126	101	104	44	375
Total 1ºB		519	1,133	358	724	568	3,302
1ºC	Alumno 1	94	139	70	71	57	431
1ºC	Alumno 2	91	100	48	26	45	310
1ºC	Alumno 3	129	137	75	78	94	513
1ºC	Alumno 4	—	103	67	98	65	333
1ºC	Alumno 5	107	107	33	43	31	321
1ºC	Alumno 6	92	87	58	25	54	316
1ºC	Alumno 7	141	170	91	151	80	633
1ºC	Alumno 8	96	114	50	72	118	450
1ºC	Alumno 9	89	—	194	77	46	406
1ºC	Alumno 10	100	134	94	113	44	485
Total 1ºC		939	1,091	780	754	634	4,198
Global n = 30		2,080	3,042	1,996	2,307	1,822	11,247

Nota. '—' indica semana sin actividad registrada para ese estudiante.

El total de preguntas respondidas durante las cinco semanas de intervención fue de **11,247**, equivalente a una media de **2,249 preguntas semanales** para el conjunto de los 30 estudiantes, o **74.9 preguntas por estudiante por semana**. La sección 1ºC registró la mayor frecuencia

acumulada (4,198 intentos; 83.9 por estudiante por semana), seguida de 1ºA (3,747; 74.9) y 1ºB (3,302; 66.0).

La semana de mayor actividad global fue la **semana 2** con **3,042 intentos**, impulsada principalmente por la sección 1ºB (1,133 intentos), donde los estudiantes concentraron el mayor esfuerzo de práctica al enfrentarse por primera vez a números de dos dígitos. La semana de menor actividad fue la **semana 5** con **1,822 intentos**, lo que es coherente con la consolidación progresiva del trazado: a medida que avanza la intervención, los estudiantes requieren menos reintentos por ítem porque han interiorizado la mecánica del trazo (Villarroel et al., 2013).

Las cinco semanas de intervención registraron actividad de forma **continua y sin interrupciones** en las tres secciones, con variabilidad individual en la distribución semanal: algunos estudiantes concentraron su práctica en semanas específicas como es el caso del estudiante 29 que realizó 194 de sus 406 intentos en S3, o el estudiante 27 con 151 intentos en S4, mientras que otros mantuvieron una práctica más distribuida a lo largo de las semanas. Esta variabilidad refleja diferencias en el ritmo de aprendizaje y en el acompañamiento docente durante las sesiones.

4.4.3.4. *Síntesis de los indicadores de la variable independiente*

Los indicadores definidos para la variable independiente: retroalimentación, interactividad y frecuencia de práctica evidencian que el uso de la aplicación NumIA se materializó en una experiencia de aprendizaje estructurada, medible y sostenida en el tiempo. En primer lugar, la proporción de mensajes afirmativos frente a correctivos confirma que la retroalimentación proporcionada por el sistema mantuvo un equilibrio pedagógicamente adecuado, favoreciendo el refuerzo positivo sin omitir la corrección de errores relevantes en el trazado. En segundo lugar, el tiempo promedio dedicado por intento refleja un nivel de interactividad coherente con las capacidades atencionales de estudiantes de primer grado, asegurando que cada actividad

implicara un procesamiento activo y no una ejecución mecánica. Finalmente, la frecuencia de práctica semanal demuestra que los estudiantes estuvieron expuestos a un volumen suficiente de ejercicios, condición necesaria para la consolidación de habilidades grafomotoras.

En conjunto, estos resultados permiten afirmar que el uso de NumIA no se limitó a una interacción superficial, sino que configuró un entorno de práctica intensiva, retroalimentada y progresiva. Por tanto, bajo estas condiciones, es metodológicamente válido sostener que cualquier mejora observada en el trazado correcto de los números arábigos puede ser atribuida a la influencia del uso de la aplicación, en coherencia con la hipótesis planteada, en la medida en que se garantizó una implementación efectiva de los componentes clave del proceso de aprendizaje: práctica suficiente, retroalimentación oportuna e interacción significativa.

4.4.4. Síntesis de Hallazgos del Resultado

En respuesta al cuarto objetivo específico, orientado a evaluar el nivel de logro del trazado numérico tras la intervención mediante una prueba postest y contrastar estadísticamente los resultados con el pretest, se logró verificar el efecto de la aplicación NumIA sobre el aprendizaje de los estudiantes de primer grado a través de un análisis cuantitativo riguroso y sistemático.

Como resultado del proceso de evaluación final, se obtuvo un perfil de logro radicalmente transformado respecto al diagnóstico inicial: el 90% de los estudiantes (27 de 30) alcanzó el nivel de Logro Destacado (AD) y el 10% restante (3 de 30) el nivel de Logro Esperado (A), sin que ningún estudiante permaneciera en los niveles B o C. La media global postest fue de 95.7%, frente al 72.4% registrado en el pretest, lo que representa una ganancia media de +23.2 puntos porcentuales.

Se identificó que la Precisión, dimensión más deficitaria en el pretest (71.8%), registró una ganancia sustantiva en el postest, evidenciando la eficacia específica del módulo de trazado guiado de NumIA sobre la forma estructural del dígito. Los cinco estudiantes que iniciaron en

nivel C ascendieron al nivel AD con ganancias individuales de entre +48.0 y +58.6 pp, constituyendo los casos de mayor transformación en la muestra.

El contraste estadístico mediante la prueba t de Student pareada, validada por la prueba de Shapiro-Wilk en las tres secciones y en el análisis global ($p > 0.05$ en todos los casos), arrojó resultados significativos en cada sección: 1ºA ($t=5.93$, $p=0.0001$), 1ºB ($t=2.67$, $p=0.0129$) y 1ºC ($t=5.22$, $p=0.0003$), con un estadístico global de $t=7.33$ ($p < 0.0001$). En todos los casos se rechazó H_0 con $\alpha=0.05$ y los intervalos de confianza al 95% no incluyen el valor cero, lo que descarta que la mejora sea atribuible al azar.

Los tamaños del efecto calculados mediante la d de Cohen refuerzan la relevancia práctica de estos resultados: 1ºA ($d=1.88$), 1ºB ($d=0.84$), 1ºC ($d=1.65$) y global ($d=1.34$), superando en todos los casos el umbral de efecto grande establecido por la literatura ($d > 0.8$). Estos valores confirman que la intervención con NumIA no solo produjo una mejora estadísticamente significativa, sino también de magnitud educativamente sustantiva en el dominio del trazado numérico.

Los registros automáticos de la base de datos de NumIA complementan y validan esta evidencia documentando el comportamiento de la variable independiente durante toda la intervención. En las cinco semanas activas se registraron 11,247 intentos de práctica distribuidos entre las tres secciones (1ºA: 3,747; 1ºB: 3,302; 1ºC: 4,198), con retroalimentación automática individualizada para los 30 estudiantes. De los 11,247 mensajes emitidos, 7,976 fueron afirmativos (70.9%) y 3,271 correctivos (29.1%), configurando una razón global de 2.44:1, homogénea entre secciones (rango: 2.13:1–2.67:1), lo que descarta diferencias en la dosis de intervención como explicación de las variaciones en el tamaño del efecto. Por criterio de evaluación automática, la Precisión resultó el más exigente para el modelo (67.5%), mientras que la Fluidez alcanzó 90.2%. El hecho de que tanto la rúbrica manual como el modelo de

redes siamesas identificaran de forma independiente la Precisión como la dimensión más deficitaria constituye evidencia de validez convergente del sistema de evaluación de NumIA. Los datos del posttest y los resultados del contraste estadístico constituyen la evidencia cuantitativa directa que permite dar respuesta al cuarto objetivo específico y sustentar la hipótesis general del estudio respecto al efecto de la aplicación NumIA sobre el aprendizaje del trazado correcto del número en los estudiantes de primer grado.

5. Discusión de resultados

Los resultados obtenidos en el presente estudio permiten establecer un diálogo sistemático con la literatura existente en tres planos articulados: la eficacia de las aplicaciones móviles educativas sobre el aprendizaje temprano, los fundamentos teóricos del trazado numérico y la retroalimentación formativa, y la comparabilidad metodológica con investigaciones previas de diseño análogo.

5.1. Sobre la mejora en el trazado numérico y su contraste con antecedentes

La ganancia media global obtenida (+23.2 pp; de 72.4% a 95.7%) y el porcentaje de estudiantes que alcanzaron el nivel AD en el posttest (90%) son consistentes con los hallazgos reportados por estudios que han evaluado el efecto de aplicaciones educativas digitales sobre habilidades de escritura y trazado en los primeros grados de primaria. Zachry et al. (2020), en su estudio piloto con estudiantes de 1° y 2° grado en Estados Unidos, constataron mejoras observables en la formación de letras en el grupo que utilizó exclusivamente aplicaciones en tabletas iPad, en comparación con grupos que recibieron terapia ocupacional tradicional o métodos mixtos. Aunque el diseño de Zachry et al. (2020) utilizó asignación aleatoria y se enfocó en escritura alfabética, la coincidencia en la población (primer grado), el instrumento de medición (evaluación de trazos) y la dirección del efecto respalda la validez ecológica de los resultados aquí obtenidos con NumIA. La diferencia relevante es que el presente estudio cuantifica el

efecto con precisión estadística inferencial $t(29)=7.33$, $p<0.0001$, $d=1.34$, lo que aporta una evidencia de mayor rigor que el análisis cualitativo ciego utilizado por Zachry et al. (2020).

Desde la perspectiva del rendimiento del modelo de inteligencia artificial, el F1-Score obtenido por NumIA (99.69%) y su AUC-ROC de 1.0000 sobre el conjunto de test demuestran una capacidad discriminativa sobresaliente. Este rendimiento es comparable y supera en precisión a los valores reportados por Dey et al. (2017) con SigNet en verificación de firmas, y es consistente con los resultados de Majumder et al. (2024) en verificación few-shot mediante arquitecturas Siamese-Transformer. Sin embargo, es necesario contextualizar este resultado: a diferencia de esos trabajos, que operan sobre escritura adulta en entornos biométricos de alta variabilidad inter-usuario, el conjunto de test de NumIA fue construido a partir de trazos de la misma distribución de entrenamiento. El único error sistemático del modelo son 31 trazos del dígito 7 confundidos con el dígito 3, es coherente con la proximidad morfológica entre ambos símbolos documentada en la literatura (Zapateiro-Segura et al., 2018), y confirma que el modelo captura las dimensiones de forma relevantes para la evaluación del trazado numérico en niños de primer grado.

Estañiz (2020), en su estudio cuasi-experimental pretest-posttest con primer grado en Andahuaylas, reportó incrementos notables en reconocimiento de números y operaciones básicas tras el uso de la aplicación Yupay Yachay, confirmando que la intervención fue significativa. El presente estudio replica ese diseño metodológico y extiende sus hallazgos al dominio específico del trazado grafomotor no abordado por Estañiz con un instrumento de mayor granularidad dimensional (cuatro criterios: Orientación, Secuencia de orden, Secuencia de Fluidez y Precisión) y con un análisis estadístico inferencial más completo, que incluye verificación de supuestos, intervalos de confianza y tamaños del efecto. Esta extensión metodológica fortalece la cadena de evidencia sobre la eficacia de las aplicaciones móviles en contextos educativos públicos peruanos.

Núñez (2021), en un diseño cuasi-experimental con niños de 5 años en educación inicial, demostró mejoras en la identificación y representación de números del 1 al 10 mediante una aplicación con Realidad Aumentada. Si bien la población de Núñez es de educación inicial y el dominio evaluado es el reconocimiento numérico más que el trazado, ambos estudios convergen en un hallazgo central: la interacción digital mediada por retroalimentación inmediata produce ganancias estadísticamente significativas en las competencias numéricas tempranas. El presente estudio refuerza esta convergencia al documentar que la Precisión, la dimensión más deficitaria al inicio (71.8%), registró una ganancia sustantiva en el posttest, resultado que apunta directamente al efecto del módulo de trazado guiado de NumIA sobre la forma estructural del dígito.

5.2. Sobre el efecto de la práctica estructurada y la frecuencia de uso

Niklas et al. (2025), en un estudio experimental con 500 preescolares en Alemania, demostraron que mejoras estadísticamente significativas en competencias matemáticas y lingüísticas podían lograrse con sesiones de apenas 2.5 a 4.5 minutos diarios cuando las aplicaciones estaban bien diseñadas. El presente estudio operó con sesiones de 30 minutos tres veces por semana durante cinco semanas, una intensidad considerablemente mayor, lo que es coherente con el tamaño del efecto muy grande reportado ($d_{\text{global}}=1.34$). Este contraste sugiere que la relación entre duración de la intervención y magnitud del efecto no es lineal, sino que está modulada por la calidad del diseño instruccional y la pertinencia del contenido respecto a la habilidad evaluada. Sáez-López et al. (2019), en un estudio sobre aplicaciones con Realidad Aumentada en Educación Primaria, concluyeron que las sesiones regulares y estructuradas son un factor determinante para la consolidación de los aprendizajes; hallazgo que el presente estudio corrobora al obtener sus resultados más homogéneos en la sección 1°C ($d=1.65$, $SD_{\text{posttest}}=6.8$ pp), indicativa de una práctica más consistente entre sus estudiantes.

Bang et al. (2023), con el uso de la aplicación adaptativa "My Math Academy" en kindergarten y primer grado (N=986), encontraron mejoras significativas en rendimiento matemático atribuibles al aprendizaje personalizado y a las evaluaciones adaptativas integradas. NumIA comparte con ese diseño el principio de retroalimentación inmediata automática fundamento también destacado por Hattie y Timperley (2007), quienes señalan que el feedback formativo efectivo debe ser específico, inmediato y orientado a la mejora de la ejecución pero se diferencia al incorporar un motor de Deep Learning que evalúa la similitud del trazo con un modelo guía, lo que permite una retroalimentación específica sobre la forma del dígito, no solo sobre la respuesta correcta o incorrecta. Los registros automáticos de NumIA confirman que esta retroalimentación se emitió de forma consistente: 7,976 mensajes afirmativos (70.9%) y 3,271 correctivos (29.1%) de un total de 11,247 mensajes, con una razón global de 2.44:1 homogénea entre secciones.

5.3. Sobre los fundamentos teóricos del trazado numérico

Villarroel et al. (2013) afirman que el primer grado es la etapa decisiva para consolidar patrones motores correctos en la escritura de números, y que los errores no corregidos en esta fase inversión, discontinuidad, orientación incorrecta se perpetúan como obstáculos para el aprendizaje matemático posterior. Los resultados del pretest validan esta premisa diagnóstica: el 37% de los estudiantes presentaba déficits marcados (niveles B y C), con la Precisión (71.8%) y la Secuencia de orden (71.9%) como las dimensiones más comprometidas.

El posttest, en cambio, muestra una corrección sistemática de esos déficits los cinco estudiantes en nivel C alcanzaron el nivel AD con ganancias de entre +48.0 y +58.6 pp, lo que sugiere que una intervención tecnológica estructurada en primer grado puede romper la trayectoria de dificultad descrita por Villarroel et al. (2013). Cabezas (2023) documentó que aproximadamente el 38% de los niños de primer grado invierten números y letras al escribir. El presente estudio encontró, en coherencia con ese diagnóstico, que la Orientación fue el

criterio relativamente más consolidado en el pretest (74.5%), aunque aún con margen de mejora, y que en el posttest alcanzó el 95.9% global.

Sinvani et al. (2023) señalan que el aprendizaje de la escritura de números involucra no solo habilidades grafomotoras sino también procesos visuoespaciales, lo que fundamenta la pertinencia de un sistema de reconocimiento automático del trazo como el modelo siamés de NumIA, que evalúa simultáneamente la trayectoria, la forma y la orientación del dígito.

Skar et al. (2022) demuestran que la fluidez de escritura predice de forma significativa la calidad del texto escrito en primaria, lo que fundamenta la inclusión de la Secuencia de fluidez como criterio evaluable del trazado. Hederich-Martínez et al. (2016) documentan que la correcta transcodificación de numerales verbales a formato arábigo depende, entre otros factores, de la consolidación del esquema grafomotor del dígito. Calero-Brito et al. (2024) agregan que el número debe ser reconocible y estar bien proporcionado para ser funcional en contextos matemáticos reales. Los resultados del presente estudio se alinean con este marco: la mejora en Precisión criterio que evalúa la semejanza estructural del dígito con el modelo estándar fue la más pronunciada de las cuatro dimensiones, lo que indica que NumIA actuó específicamente sobre la dimensión que la literatura identifica como más crítica para la funcionalidad del trazado.

5.4. Sobre el marco teórico del aprendizaje digital y la retroalimentación

Mayer (2010) propone que el aprendizaje multimedia es más efectivo cuando los canales auditivo y visual se combinan de manera contigua y coherente. NumIA operacionaliza este principio al presentar animaciones que acompañan en cada pregunta, el mensaje de retroalimentación, y su sonido correspondiente, cumpliendo los principios de contigüidad y coherencia descritos por Mayer.

Robert (2001) sostiene que la repetición estructurada es el primer principio del aprendizaje y que la práctica distribuida consolida las representaciones procedimentales. La arquitectura de

NumIA con ejercicios secuenciados por dígito, repetición con variantes guía y retroalimentación por cada trazo operacionaliza este principio de manera sistemática. Los 11,247 intentos registrados en cinco semanas continuas, con una media de 74.9 preguntas por estudiante por semana, constituyen evidencia objetiva de práctica distribuida sostenida. El incremento en la homogeneidad de resultados en el posttest es coherente con el efecto de consolidación grafomotora que la teoría de la repetición predice para habilidades de tipo procedimental motor.

Piaget e Inhelder (1997) establecen que el aprendizaje en los primeros años de escolaridad se consolida a través de la experiencia activa y la manipulación directa de objetos y símbolos. La interacción táctil de los estudiantes con la pantalla de la tableta durante las sesiones de NumIA constituye una forma de manipulación directa del símbolo numérico, coherente con el principio piagetiano de aprendizaje por acción. Hassinger-Das et al. (2017) refuerzan este argumento al señalar que las aplicaciones educativas son efectivas en educación infantil precisamente porque integran el entretenimiento con objetivos pedagógicos concretos, generando un estado de implicación activa que favorece la retención.

5.5. Sobre la magnitud del efecto y su comparabilidad

El tamaño del efecto global obtenido ($d=1.34$) supera ampliamente el umbral de efecto grande definido por Cohen (1988) $d>0.8$ y es comparable con los valores más altos reportados en la literatura sobre intervenciones tecnológicas en habilidades básicas de primaria. La sección 1ºA registra el efecto más elevado ($d=1.88$), coherente con su mayor margen de mejora disponible desde el punto de partida más bajo del conjunto (media pretest: 61.9%). La sección 1ºB, con el efecto más moderado ($d=0.84$), refleja la mayor heterogeneidad interna ($SD=22.4$ pp), atribuible a que convivían en ella cinco estudiantes en nivel AD desde el pretest con escaso margen de mejoran y dos en nivel C con ganancias extremas de +48.8 y +58.6 pp.

Este patrón de variabilidad diferencial entre secciones es consistente con el fenómeno de efecto techo documentado en intervenciones aplicadas a grupos con dominio previo elevado, y no compromete la validez del efecto global. La homogeneidad de la razón afirmativo/correctivo entre secciones (rango: 2.13:1–2.67:1) descarta además que las diferencias en tamaño del efecto se expliquen por diferencias en la dosis o calidad de la intervención.

En conjunto, los resultados del presente estudio son convergentes con la evidencia acumulada en la literatura nacional e internacional sobre la eficacia de las aplicaciones móviles educativas en los primeros grados de primaria, y aportan evidencia original en un dominio el trazado grafomotor de dígitos arábigos evaluado con criterios multidimensionales y analizado con contraste estadístico inferencial completo que los antecedentes revisados no habían abordado de manera sistemática en el contexto educativo público peruano.

No obstante, la solidez de los resultados obtenidos, el diseño cuasi-experimental sin grupo control presenta limitaciones inherentes a la validez interna del estudio que deben ser reconocidas. Tres amenazas principales merecen consideración. La primera es el efecto de maduración: los estudiantes de primer grado experimentan un desarrollo grafomotor natural durante el período de intervención, por lo que una parte de la mejora observada podría atribuirse al proceso evolutivo y no exclusivamente al uso de NumIA. La segunda es el efecto de prueba: la exposición previa al instrumento durante el pretest pudo haber familiarizado a los estudiantes con la tarea evaluada, facilitando el desempeño en el posttest con independencia de la intervención. La tercera es el efecto de historia: eventos externos al estudio como la práctica caligráfica regular en el aula pueden haber contribuido a las ganancias observadas. Estas amenazas no invalidan los hallazgos, dado que el tamaño del efecto global ($d = 1.34$) excede ampliamente lo que podría atribuirse a maduración natural en un período de cinco semanas; sin embargo, sí limitan la atribución causal exclusiva a NumIA y justifican la realización de estudios futuros con diseño experimental que incluya grupo control.

6. Conclusiones

Primera. En respuesta al primer objetivo específico, orientado a identificar y analizar las estrategias pedagógicas utilizadas por docentes de primer grado para enseñar el trazado de números, se determinó que las cinco estrategias tradicionales documentadas las cuales son práctica repetitiva con cuadernos de aprestamiento, demostración y modelado docente en pizarra, uso de material manipulativo multisensorial, empleo del color como recurso didáctico y evaluación por criterios de precisión, secuencia y orientación poseen fundamento teórico sólido, pero presentan una limitación estructural común e irresoluble en contextos de aulas numerosas: la imposibilidad de proporcionar retroalimentación individualizada, inmediata y formativa a cada estudiante durante la práctica del trazado. Esta brecha pedagógica constituyó la justificación empírica directa para el desarrollo de NumIA.

Segunda. En respuesta al segundo objetivo específico, orientado a diseñar y desarrollar la aplicación móvil NumIA, se logró implementar una solución tecnológica educativa funcional que integra de manera articulada componentes pedagógicos, de desarrollo de software y de inteligencia artificial. La aplicación incorpora un modelo de Deep Learning basado en Redes Siamesas capaz de evaluar automáticamente la similitud entre el trazo del estudiante y patrones de referencia predefinidos, generando retroalimentación inmediata por criterio. Los módulos de práctica guiada, reintento dirigido y registro automático de progreso individual materializan los principios del aprendizaje multimedia (Mayer, 2010) y de la retroalimentación formativa (Hattie y Timperley, 2007) en una herramienta adaptada a las características cognitivas y perceptivas de niños de primer grado de educación primaria.

Tercera. En respuesta al tercer objetivo específico, orientado a evaluar el nivel inicial del trazado numérico mediante una prueba diagnóstica e implementar NumIA en sesiones de práctica estructurada, se estableció una línea base cuantitativa rigurosa: el 23% de los 30 estudiantes alcanzaba el nivel AD en el pretest, el 40% se ubicaba en nivel A, el 20% en nivel

B y el 17% en nivel C, con una media global de 72.4%. La Precisión (71.8%), la Secuencia de fluidez (71.6%) y la Secuencia de orden (71.9%) fueron las dimensiones sistemáticamente más deficitarias en las tres secciones, mientras que la Orientación (74.5%) resultó el criterio relativamente más consolidado al inicio de la intervención. La intervención se ejecutó durante cinco semanas continuas con tres sesiones semanales de 30 minutos, generando 11,247 intentos de práctica distribuidos de forma activa en las cinco semanas, con supervisión docente e investigadora permanente y registro automático de progreso individual, garantizando condiciones de fidelidad protocolaria a lo largo de toda la implementación.

Cuarta. En respuesta al cuarto objetivo específico, orientado a analizar el impacto de la intervención mediante el posttest, los registros automáticos y el contraste estadístico con el diagnóstico inicial, los resultados evidencian una mejora real, estadísticamente significativa y educativamente relevante. La media global ascendió de 72.4% a 95.7% (+23.2 pp), el 90% de los estudiantes alcanzó el nivel AD y el 10% restante el nivel A, con desaparición total de los niveles B y C. La prueba de Shapiro-Wilk confirmó la normalidad de las diferencias en los cuatro niveles de análisis ($p > 0.05$ en todos los casos), habilitando el uso uniforme de la prueba *t* de Student pareada, cuyos resultados son significativos en las tres secciones y en el análisis global: 1°A ($t=5.93$, $p=0.0001$), 1°B ($t=2.67$, $p=0.0129$), 1°C ($t=5.22$, $p=0.0003$) y global ($t=7.33$, $p<0.0001$). Los tamaños del efecto calculados mediante la *d* de Cohen superan el umbral de efecto grande en todos los análisis: 1°A ($d=1.88$), 1°B ($d=0.84$), 1°C ($d=1.65$) y global ($d=1.34$). Los registros automáticos de NumIA documentaron 11,247 intentos de práctica en las cinco semanas de intervención, con una razón afirmativo/correctivo global de 2.44:1, homogénea entre secciones (rango: 2.13:1–2.67:1), lo que confirma que las tres secciones recibieron una dosis equivalente de intervención y que las diferencias en el tamaño del efecto entre secciones responden al perfil de entrada y no a variaciones en la implementación.

Quinta. El análisis por criterio reveló que la Precisión la dimensión más deficitaria en el diagnóstico inicial (71.8%) registró la mayor ganancia en el posttest, evidenciando la eficacia específica del módulo de trazado guiado de NumIA sobre la forma estructural del dígito. Este hallazgo es reforzado por los registros automáticos de la aplicación: el modelo de redes siamesas también identificó la Precisión como el criterio más exigente (67.5% de aceptación automática), mientras que la Fluidez alcanzó un 90.2% de aceptación. La coincidencia independiente entre la rúbrica manual del pretest y la evaluación automática del modelo sobre cuál dimensión era más deficitaria constituye evidencia de validez convergente del sistema de evaluación de NumIA. Asimismo, los cinco estudiantes que iniciaron en nivel C ascendieron al nivel AD con ganancias individuales de entre +48.0 y +58.6 pp, constituyendo los casos de mayor transformación de la muestra y confirmando que la intervención fue especialmente eficaz para los estudiantes con mayores rezagos grafomotores.

Sexta. En síntesis, el objetivo general del estudio queda cumplido en su totalidad. NumIA, como aplicación móvil basada en Deep Learning, demostró ser una herramienta eficaz para mejorar el trazado correcto de los números arábigos en estudiantes de primer grado de la Institución Educativa Fe y Alegría N.º 28 de Chiclayo, Lambayeque, durante el año escolar 2025, produciendo una mejora consistente entre tres secciones independientes, estadísticamente verificada mediante contraste paramétrico con supuestos validados ($t=7.33$, $p<0.0001$, $d=1.34$), y de magnitud práctica muy grande según los criterios establecidos por la literatura especializada (Cohen, 1988).

7. Recomendaciones

Primera. Se recomienda a los docentes de primer grado de educación primaria incorporar NumIA como recurso complementario sistemático dentro de la programación del primer bimestre, destinando al menos dos sesiones semanales de práctica guiada de 30 minutos durante las primeras cinco semanas del año escolar, que es el período en que los estudiantes construyen

los esquemas grafomotores básicos. Los resultados del presente estudio confirman que esta dosificación con tres sesiones semanales produjo una ganancia media de +23.2 pp y llevó al 90% de los estudiantes al nivel AD. La complementariedad con la demostración docente y el uso de cuadernos de aprestamiento garantiza que la retroalimentación automática de la aplicación refuerce, y no sustituya, la mediación pedagógica humana.

Segunda. Se recomienda a los directivos de la I.E. Fe y Alegría N.º 28 y de instituciones con perfil similar gestionar la adquisición y mantenimiento sostenido de tablets con las especificaciones mínimas requeridas por NumIA (Android 10+, 4 GB RAM, pantalla de 8 pulgadas). La experiencia de implementación del presente estudio evidenció que la estabilidad de los dispositivos fue condición determinante para la continuidad de las sesiones y la integridad de los registros automáticos. Se sugiere además designar a un docente responsable del mantenimiento preventivo de los dispositivos y de la actualización periódica del APK, para evitar interrupciones que afecten la frecuencia de práctica, que en este estudio demostró ser un factor clave (74.9 preguntas por estudiante por semana).

Tercera. Los registros automáticos de NumIA evidenciaron que la Precisión del trazado fue el criterio más exigente para el modelo de redes siamesas (67.5% de aceptación global), mientras que la Fluidez alcanzó 90.2%. Esta disparidad indica que el modelo genera mayor proporción de mensajes correctivos sobre la forma estructural del dígito que sobre su continuidad de trazo. Se recomienda al equipo de desarrollo ampliar el banco de guías de referencia del modelo incorporando variantes caligráficas de distintas regiones del Perú, con el fin de reducir la tasa de falsos negativos en el reconocimiento de trazos regionalmente válidos.

Cuarta. Se recomienda a futuros investigadores replicar el diseño del presente estudio incorporando un grupo de control con metodología tradicional. El diseño cuasi-experimental pretest–posttest sin grupo de control empleado aquí permitió documentar la magnitud del cambio con precisión estadística ($t=7.33$, $p<0.0001$, $d=1.34$), pero no permite aislar el efecto

específico de NumIA del efecto de la práctica estructurada en sí misma. La incorporación de un grupo control permitiría cuantificar con mayor precisión la contribución diferencial de la retroalimentación automática del modelo de Deep Learning respecto a la práctica convencional supervisada, fortaleciendo la validez interna de las conclusiones.

Quinta. Se recomienda ampliar el alcance de NumIA a los grado 2° de primaria, donde persisten dificultades de legibilidad y orientación del trazado. La arquitectura modular de la aplicación permite esta extensión adaptando el banco de números y el rango de contenidos sin necesidad de rediseño estructural del sistema. Se sugiere también evaluar su uso en el nivel inicial (5 años) en la fase de pre-trazado, dado que la consolidación grafomotora temprana es el predictor más sólido del desempeño en escritura numérica durante el primer grado, conforme documentan Villarroel et al. (2013).

Sexta. Se recomienda a las instancias de gestión educativa de Lambayeque Gerencia Regional de Educación y UGEL Chiclayo considerar los resultados del presente estudio como evidencia inicial para evaluar la incorporación de aplicaciones educativas con inteligencia artificial en el primer ciclo de educación primaria. La eficacia documentada de NumIA en una institución educativa pública con recursos estándar tres secciones, 30 estudiantes, tablets de gama media, sin infraestructura especial demuestra que el modelo es replicable en condiciones similares a las de la mayoría de instituciones públicas de la región, lo que refuerza su viabilidad en el marco de políticas de equidad educativa.

Séptima. Se recomienda a futuros investigadores replicar el presente estudio incorporando un grupo de control que reciba instrucción mediante metodología tradicional, con asignación aleatoria de los participantes a ambas condiciones y mediciones pretest y posttest con el mismo instrumento. Este diseño experimental permitiría aislar con precisión el efecto específico de la retroalimentación automática del modelo de Deep Learning respecto al efecto de la práctica estructurada en sí misma, superando la principal limitación del diseño cuasi-experimental aquí

empleado y fortaleciendo la validez interna de las conclusiones sobre el valor añadido de NumIA en el aprendizaje del trazado numérico.

8. Bibliografía

- Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Alessandri, G., Zuffianò, A., & Perinelli, E. (2017). Evaluating intervention programs with a pretest-posttest design: A structural equation modeling approach. *Frontiers in Psychology*, 8(MAR), 243862. <https://doi.org/10.3389/FPSYG.2017.00223/BIBTEX>
- Amazon Web Services. (n.d.). *¿Qué es Flutter? - Explicación de la aplicación Flutter - AWS*. Retrieved January 17, 2026, from <https://aws.amazon.com/es/what-is/flutter/>
- Bang, H. J., Li, L., & Flynn, K. (2023). Efficacy of an Adaptive Game-Based Math Learning App to Support Personalized Learning and Improve Early Elementary School Students' Learning. *Early Childhood Education Journal*, 51(4), 717–732. <https://doi.org/10.1007/s10643-022-01332-3>
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. In *Journal of Usability Studies* (Vol. 4).
- Bass, L., Clements, P., & Kazman, R. (2012). *Software Architecture in Practice: Software Architect Practice_c3 - Len Bass, Paul Clements, Rick Kazman - Google Libros*. https://books.google.com.pe/books/about/Software_Architecture_in_Practice.html?id=-II73rBDXCYC&redir_esc=y
- Bromley, J., Guyon, I., LeCun, Y., Säcker, E., & Shah, R. (1993). Signature Verification using a “Siamese” Time Delay Neural Network. *Advances in Neural Information Processing Systems*, 6, 737–744. <https://doi.org/10.1142/S0218001493000339>
- Brooke, J. (1996). *SUS-A quick and dirty usability scale*.
- Cabezas, Y. (2023). *Procesos de lectura inicial en estudiantes de primero de primaria de la I.E. Santa Fe*.
- Calero-Brito, P., Zapata-Mocha, E., Burbano-Santamaría, S., & Moyón-Coronel, E. (2024). Motricidad fina para el desarrollo de la escritura de los estudiantes: revisión de literatura. *593 Digital Publisher CEIT*, 9(1–1), 51–65. <https://doi.org/10.33386/593dp.2024.1-1.2261>
- Campos, G., & Lule, N. (2012). *LA OBSERVACIÓN, UN MÉTODO PARA EL ESTUDIO DE LA REALIDAD*.
- Carreira, C. F. (2013). *Principales dificultades en el aprendizaje de las Matemáticas. Pautas para maestros de Educación Primaria*.
- Casas Anguita, J., Repullo Labrador, J., & Donado Campos, J. (2003). La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos (I). *Atención Primaria*, 31(8), 527–538. <https://www.elsevier.es/es-revista-atencion-primaria-27-articulo-la-encuesta-como-tecnica-investigacion--13047738>
- Chang, S.-H., & Yu, N.-Y. (2022). *Computerized handwriting evaluation and statistical reports for children in the age of primary school*. <https://doi.org/10.1038/s41598-022-19913-y>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences Second Edition* (Segunda edición).

- Craparotta, G., Thomassey, S., & Biolatti, A. (2019). A siamese neural network application for sales forecasting of new fashion products using heterogeneous data. *International Journal of Computational Intelligence Systems*, 12(2), 1537–1546. <https://doi.org/10.2991/IJCIS.D.191122.002>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 1951 16:3, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- De Rosa, G. H., & Papa, J. P. (2018). Learning to weight similarity measures with Siamese networks: a case study on optimum-path forest. *Optimum-Path Forest: Theory, Algorithms, and Applications*, 155–173. <https://doi.org/10.1016/B978-0-12-822688-9.00015-3>
- Dey, S., Dutta, A., Toledo, J. I., Ghosh, S. K., Lladós, J., & Pal, U. (2017). SigNet: Convolutional Siamese Network for Writer Independent Offline Signature Verification. *ArXiv.Org*. <https://export.arxiv.org/pdf/1707.02131v2.pdf>
- Dzib Moo, D. L. B., González Ceballos, M. C., & Rodríguez Solís, G. (2023). La música y su importancia en el aprendizaje significativo de estudiantes de educación básica de Tabasco, México. *Revista Inclusiones*, 10(4), 1–18. <https://doi.org/10.58210/fprc3472>
- Estañiz, N. (2020). *Influencia del aplicativo móvil “Yupay Yachay” en el aprendizaje de matemáticas en los niños del primer grado de la Institución Educativa Primaria de Ancatira*. Universidad Nacional José María Arguedas.
- Feder, K. P., & Majnemer, A. (2007). Handwriting development, competency, and intervention. *Developmental Medicine and Child Neurology*, 49(4), 312–317. <https://doi.org/10.1111/J.1469-8749.2007.00312.X>
- Fielding, R. (2000). *Architectural Styles and the Design of Network-based Software Architectures*.
- Gelman, R., & Gallistel, C. (1978). *The Child's Understanding of Number*. https://books.google.com.pe/books/about/The_Child_s_Understanding_of_Number.html?id=YjhEBel29H0C&redir_esc=y
- George, D., & Mallery, P. (2019). *IBM SPSS Statistics 25 Step by Step a Simple Guide And Refrence* (Routledge, Ed.; 15th ed.).
- Gómez, A., & Mesías, M. (2021). Estudio piloto sobre el uso de las aplicaciones móviles como información nutricional para el consumidor de alimentos. *REVISTA ESPAÑOLA DE COMUNICACIÓN EN SALUD*, 12(2), 206–214. <https://doi.org/10.20318/recs.2021.6208>
- Gómez Vera, K. (2019). 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. <https://doi.org/10.33936/REHUSO.V4I3.2136>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. *Healthcare Informatics Research*, 22(4), 351. <https://doi.org/10.4258/hir.2016.22.4.351>
- Google. (n.d.). *Introduction to Dart | Google Skills*. Retrieved January 17, 2026, from <https://www.skills.google/focuses/28825?parent=catalog>
- Guevara Alban, G. P., Verdesoto Arguello, A. E., & Castro Molina, N. E. (2020). Metodologías de investigación educativa (descriptivas, experimentales, participativas, y de investigación-acción). *RECIMUNDO*, 4(3), 163–173. [https://doi.org/10.26820/recimundo/4.\(3\).julio.2020.163-173](https://doi.org/10.26820/recimundo/4.(3).julio.2020.163-173)

- Hadsell, R., Chopra, S., & LeCun, Y. (2006). Dimensionality reduction by learning an invariant mapping. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2, 1735–1742. <https://doi.org/10.1109/CVPR.2006.100>
- Hamdi, Y., Akouaydi, H., Boubaker, H., & Alimi, A. M. (2020). *Handwriting Quality Analysis using Online-Offline Models*.
- Hassinger-Das, B., Toub, T. S., Zosh, J. M., Michnick, J., Golinkoff, R., & Hirsh-Pasek, K. (2017). More than just fun: a place for games in playful learning / Más que diversión: el lugar de los juegos reglados en el aprendizaje lúdico. *Infancia y Aprendizaje*, 40(2), 191–218. <https://doi.org/10.1080/02103702.2017.1292684>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. https://doi.org/10.3102/003465430298487/ASSET/AA36F61A-6156-443C-A6D0-B97CC031F90D/ASSETS/IMAGES/LARGE/10.3102_00346543077001081-FIG1.JPG
- Hederich-Martínez, C., Camargo-Urbe, Á., & Avalo-Azcárate, A. (2016). Transcodificación de numerales verbales a formato arábigo en educación básica primaria. *Magis*, 8(17), 27–46. <https://doi.org/10.11144/JAVERIANA.M8-17.TNVF>
- Hernández, R. (2014). Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta Las rutas Cuantitativa Cualitativa y Mixta. *McGRAW-HILL Interamericana Editores S.A. de C.V.*, 753. <http://repositorio.uasb.edu.bo:8080/bitstream/54000/1292/1/Hernández- Metodología de la investigación.pdf>
- IBM. (n.d.-a). *¿Qué es el deep learning?* | IBM. Retrieved July 26, 2025, from <https://www.ibm.com/es-es/think/topics/deep-learning>
- IBM. (n.d.-b). *¿Qué es una base de datos relacionales?* | IBM. Retrieved April 7, 2026, from <https://www.ibm.com/es-es/think/topics/relational-databases>
- ISO. (2018). *Ergonomía de la interacción humano-sistema. Parte 11: Usabilidad: definiciones y conceptos (ISO 9241-11:2018)*. International Organization for Standardization. <https://standards.iteh.ai/catalog/standards/sist/d38dc274-d8d4-4fb9-8206->
- Kingma, D. P., & Lei Ba, J. (2015). *ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION*.
- Lee, J. S., & Kim, S. W. (2015). Validation of a tool evaluating educational apps for smart education. *Journal of Educational Computing Research*, 52(3), 435–450. <https://doi.org/10.1177/0735633115571923>
- Liu, C., & Correia, A.-P. (2021). A Case Study of Learners' Engagement in Mobile Learning Applications. *Journal*, 25. <https://doi.org/10.24059/olj.v25i4.2827>
- Llanos Masciotti, F., & Tapia, J. (2021). *La retroalimentación dialógica: algunos hallazgos en escuelas multigrado*.
- Majumder, P., Joaa, A. F. M. M., Rahman Rhythm, E., Kabir Mehedi, M. H., & Alim Rasel, A. (2024). Siamese-Transformer Network for Offline Handwritten Signature Verification using Few-shot. *2023 26th International Conference on Computer and Information Technology, ICCIT 2023*. <https://doi.org/10.1109/ICCIT60459.2023.10441035>
- Mayer, R. E. . (2010). *Multimedia learning*. Cambridge University Press.
- MINEDU. (2016). *de la Educación Básica*.

- Ministerio de Educación. (2022). *Orientaciones para el desarrollo y la evaluación de las competencias*. www.minedu.gob.pe
- Ministerio de Educación del Perú. (2023). *ENLA 2023 muestra resultados estables con algunas mejoras importantes respecto a evaluación anterior - Noticias - Ministerio de Educación - Plataforma del Estado Peruano*. <https://www.gob.pe/institucion/minedu/noticias/946578-enla-2023-muestra-resultados-estables-con-algunas-mejoras-importantes-respecto-a-evaluacion-anterior>
- Mkpojiogu, E. O. C., Hussain, A., & Hassan, F. (2018). A systematic review of usability quality attributes for the evaluation of mobile learning applications for children. *AIP Conference Proceedings*, 2016(1), 36. <https://doi.org/10.1063/1.5055494/725751>
- National Research Council. (2009). Mathematics Learning in Early Childhood: Paths Toward Excellence and Equity. *Mathematics Learning in Early Childhood: Paths Toward Excellence and Equity*, 1–398. <https://doi.org/10.17226/12519>
- Niklas, F., Birtwistle, E., Mues, A., & Wirth, A. (2025). Learning apps at home prepare children for school. *Child Development*, 96, 577–590. <https://doi.org/10.1111/cdev.14184>
- Núñez, D. (2021). *APLICATIVO MÓVIL CON REALIDAD AUMENTADA PARA MEJORAR LAS CAPACIDADES MATEMATICAS EN NIÑOS DE 5 AÑOS DE LA IEI MILAGRITOS DE JESUS*.
- Ortiz, G. (2006). *EL COLOR UN FACILITADOR DIDACTICO*.
- Piaget, J., & Inhelder, B. (1997). *Psicología del niño*. <http://www.edmorata.es>
- Pinilla, P., & Badilla, M. G. (2020). RubricApp: Adaptation and validation of a Rubric to evaluate the pedagogical value of Mobile Educational Applications. *Aloma*, 38(2), 69–82. <https://doi.org/10.51698/ALOMA.2020.38.2.69-82>
- Pressman, R. S. ., & Maxim, B. R. . (2020). *Software engineering : a practitioners approach*. 671. https://books.google.com/books/about/Software_Engineering.html?hl=id&id=taIKxAEACAAJ
- Robert, B. (2001). *Repetition is the First Principle of All Learning*. https://www.researchgate.net/publication/228318502_Repetition_is_the_First_Principle_of_All_Learning
- Romero, G. G., Torres Burgos, S. A., Marina Fernanda, Coello Romero, D. C., & Moran Berruz, D. C. (2024). Linguistic games program to develop oral expression in early childhood education students. *Universidad Ciencia y Tecnología*, 28(Special), 340–350. <https://doi.org/10.47460/uct.v28ispecial.833>
- Sáez-López, J. M., Sevillano-García, M. L., & Pascual-Sevillano, M. Á. (2019). Aplicación del juego ubicuo con realidad aumentada en Educación Primaria. *Oxbridge Publishing House*, 27(61), 71–82. <https://doi.org/10.3916/C61-2019-06>
- Santillana Educación. (n.d.). *Matemáticas Primaria BIBLIOTECA DEL PROFESORADO*.
- Schwaber, K., & Sutherland, J. (2020). *La Guía Scrum*.
- Silva-Peñafiel, G. E., Castillo-Parra, B. F., Tixi-Gallegos, K. G., & Urgiles-Rodríguez, B. E. (2024). La Revolución de la Inteligencia Artificial en la Educación Superior. *La Revolución de La Inteligencia Artificial En La Educación Superior*. <https://doi.org/10.55813/EGAEA.L.71>

- Sinvani, R. T., Golos, A., Ben Zagmi, S., & Gilboa, Y. (2023). The Relationship between Young Children's Graphomotor Skills and Their Environment: A Cross-Sectional Study. *International Journal of Environmental Research and Public Health*, 20(2), 1338. <https://doi.org/10.3390/IJERPH20021338>
- Skar, G. B., Lei, P.-W., Graham, S., Johannes Aasen, A., Byberg Johansen, M., & Holten Kvistad, A. (2022). Handwriting fluency and the quality of primary grade students' writing. *Reading and Writing*, 35, 509–538. <https://doi.org/10.1007/s11145-021-10185-y>
- Tamayo, M. (2004). El proceso de la investigación científica - Mario Tamayo y Tamayo - Google Libros. *Tabasco, Mexico*, 164. https://books.google.com/cu/books/about/El_proceso_de_la_investigaci%C3%B3n_cient%C3%ADfica.html?id=BhymmEqkkJwC&redir_esc=y%0Ahttps://books.google.com/ec/books?hl=es&lr=&id=BhymmEqkkJwC&oi=fnd&pg=PA13&dq=niveles+y+tipos+de+investigaci%C3%B3n+cient%C3%ADfica&ots=TrdH9o25oH&sig=L
- Tanenbaum, A. S. ., & Steen, M. van. (2007). *Distributed systems : principles and paradigms*. 686.
- Vargas-Murillo, G. (2020). ESTRATEGIAS EDUCATIVAS Y TECNOLOGÍA DIGITAL EN EL PROCESO ENSEÑANZA APRENDIZAJE EDUCATIONAL STRATEGIES AND DIGITAL TECHNOLOGY IN THE TEACHING LEARNING PROCESS. *Revista "Cuadernos*, 61(1).
- Vergel, R. (2014). *El signo en Vygotski y su vínculo con el desarrollo de los procesos psicológicos superiores*.
- Villarreal, R., Jiménez, J. E., Rodríguez, C., Peake, C., & Bisschop, E. (2013). El rol de la escritura de números en niños con y sin dificultades de aprendizaje en matemáticas. *European Journal of Education and Psychology*, 6, 105–115. www.ejep.es
- Youden, j. (1949). Index for rating diagnostic tests. *Cancer: Volume 3, Issue 1*, 32–35. <https://acsjournals.onlinelibrary.wiley.com/doi/epdf/10.1002/1097-0142%281950%293%3A1%3C32%3A%3AAID-CNCR2820030106%3E3.0.CO%3B2-3>
- Zachry, A. H., Arnwine, R., & Garrett, A. (2020). *The Effectiveness of iPad Apps to Improve Handwriting Legibility in First and Second Grade Students: A Pilot Study*. <https://eric.ed.gov/?id=EJ1279105>.
- Zapateiro-Segura, J. C., Poloche-Arango, S. K., Camargo-Urbe, L., Poloche-Arango, S. K., & Camargo-Urbe, L. (2018). Orientación espacial: una ruta de enseñanza y aprendizaje centrada en ubicaciones y trayectorias. *Tecné, Episteme y Didaxis: TED*, (43), 119–136. http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-38142018000100119&lng=en&nrm=iso&tlng=es
- Zhou, X., & Wang, B. (2004). Preschool Children's Representation and Understanding of Written Number Symbols. *Early Child Development and Care*, 174(3), 253–266.

9. ANEXOS

GRUPO A: INSTRUMENTOS DE EVALUACIÓN

Anexo A1. Prueba Pretest de Desempeño de Trazado Numérico

PRUEBA PRE-TEST - TRAZADO DE NÚMEROS

Aplicación NumIA - Primer Grado de Primaria

Institución Educativa: Fe y Alegría N.º 28

Ubicación: Chiclayo, Lambayeque

Fecha: _____

Nombre del estudiante: _____

Edad: _____ años

SECCIÓN 1: ESCRITURA DE NÚMEROS INDIVIDUALES (0-20)

Escribe el número 1		Escribe el número 11	
Escribe el número 2		Escribe el número 12	
Escribe el número 3		Escribe el número 13	
Escribe el número 4		Escribe el número 14	
Escribe el número 5		Escribe el número 15	
Escribe el número 6		Escribe el número 16	
Escribe el número 7		Escribe el número 17	
Escribe el número 8		Escribe el número 18	

SECCIÓN 2: CONTEO Y CANTIDAD (21-30)

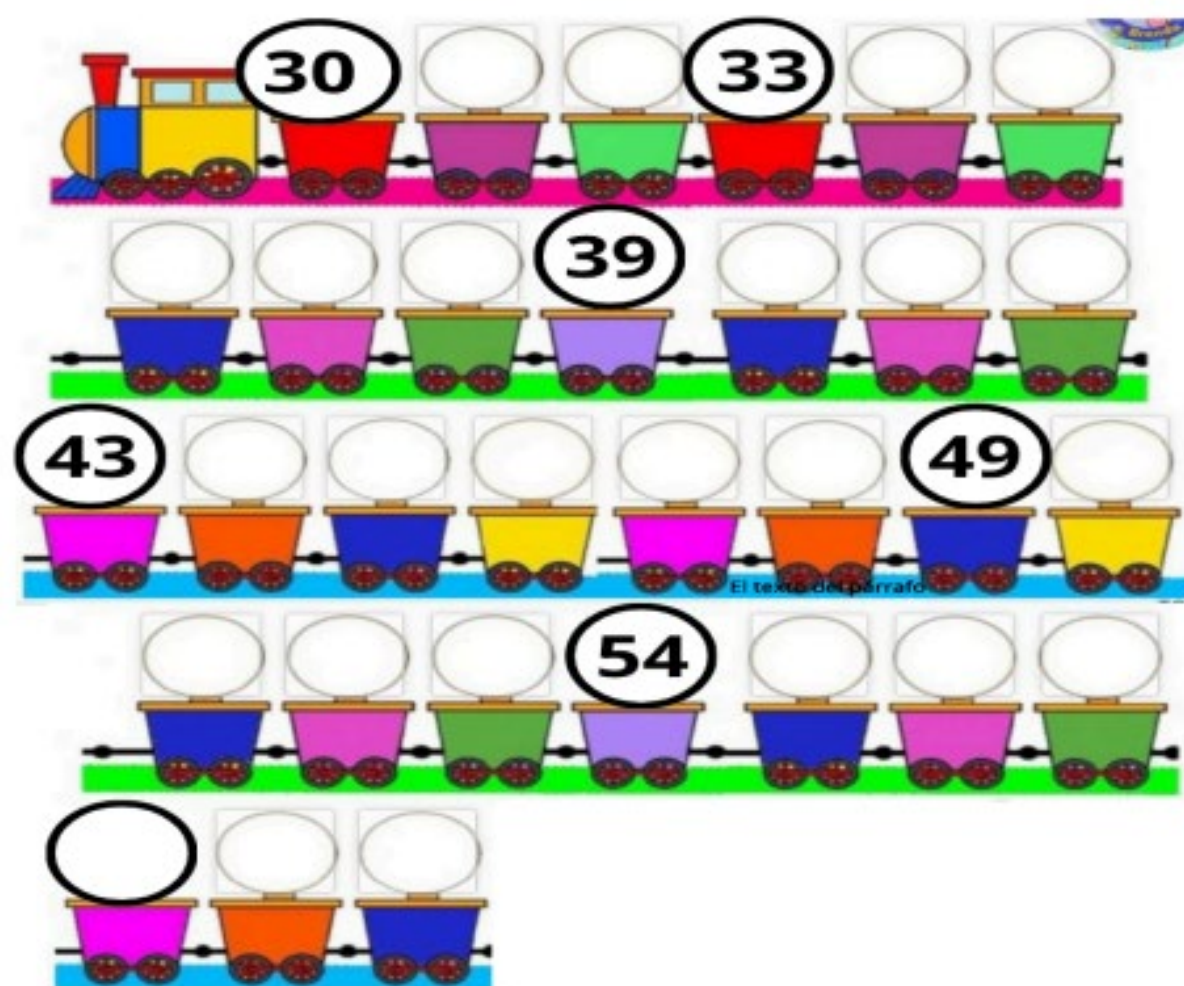
Instrucción: Cuenta los objetos en cada recuadro y escribe el número total.

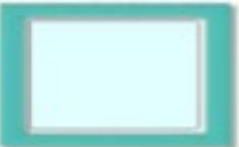
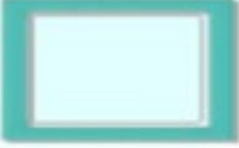
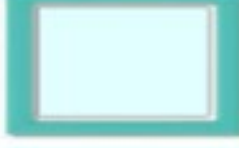
SECCIÓN 3: COMPLETAR SERIES (30-60)

Instrucción: Completa la serie de números escribiendo los que faltan.



SECCIÓN 4: DICTADO (61-100)

Instrucción: Escribe el número dictado.

|

Nombre del docente: _____

Firma del Docente: _____

Fecha: _____

Anexo A2. Prueba Postest de Desempeño de Trazado Numérico

PRUEBA POST-TEST - TRAZADO DE NÚMEROS

Aplicación NumIA - Primer Grado de Primaria

Institución Educativa: Fe y Alegría N.° 28

Ubicación: Chiclayo, Lambayeque

Fecha: _____

Nombre del estudiante: _____

Edad: _____ años

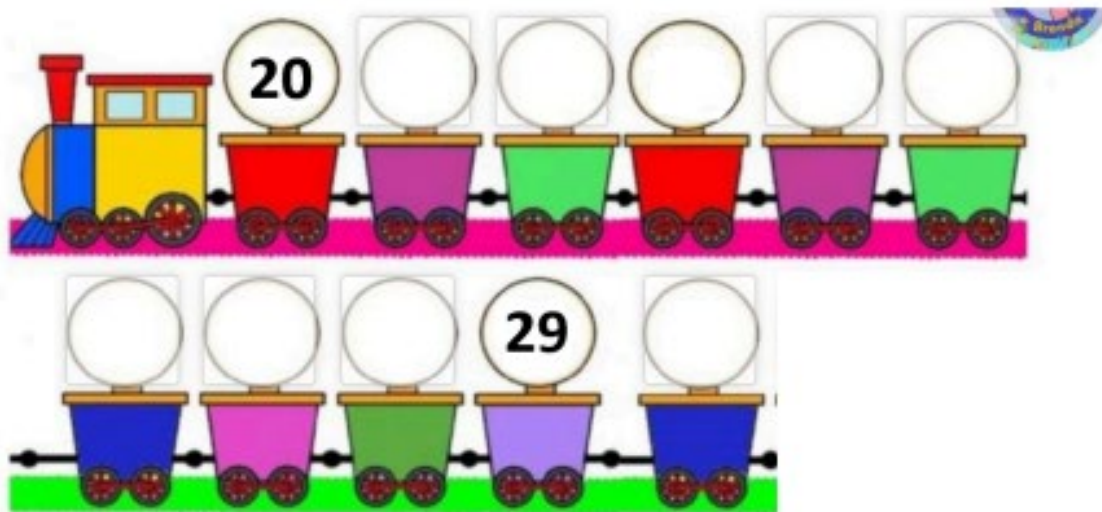
SECCIÓN 1: ESCRITURA DE NÚMEROS INDIVIDUALES (0-20)

Escribe el número diez		Escribe el número dieciséis	
Escribe el número uno		Escribe el número doce	
Escribe el número tres		Escribe el número catorce	
Escribe el número seis		Escribe el número diecisiete	
Escribe el número cuatro		Escribe el número once	
Escribe el número cinco		Escribe el número dos	
Escribe el número siete		Escribe el número dieciocho	
Escribe el número trece		Escribe el número veinte	
Escribe el número nueve		Escribe el número diecinueve	
Escribe el número ocho		Escribe el número quince	

Instrucción: Escribe cada número en el espacio correspondiente. Intenta hacerlo lo más claro y ordenado posible.

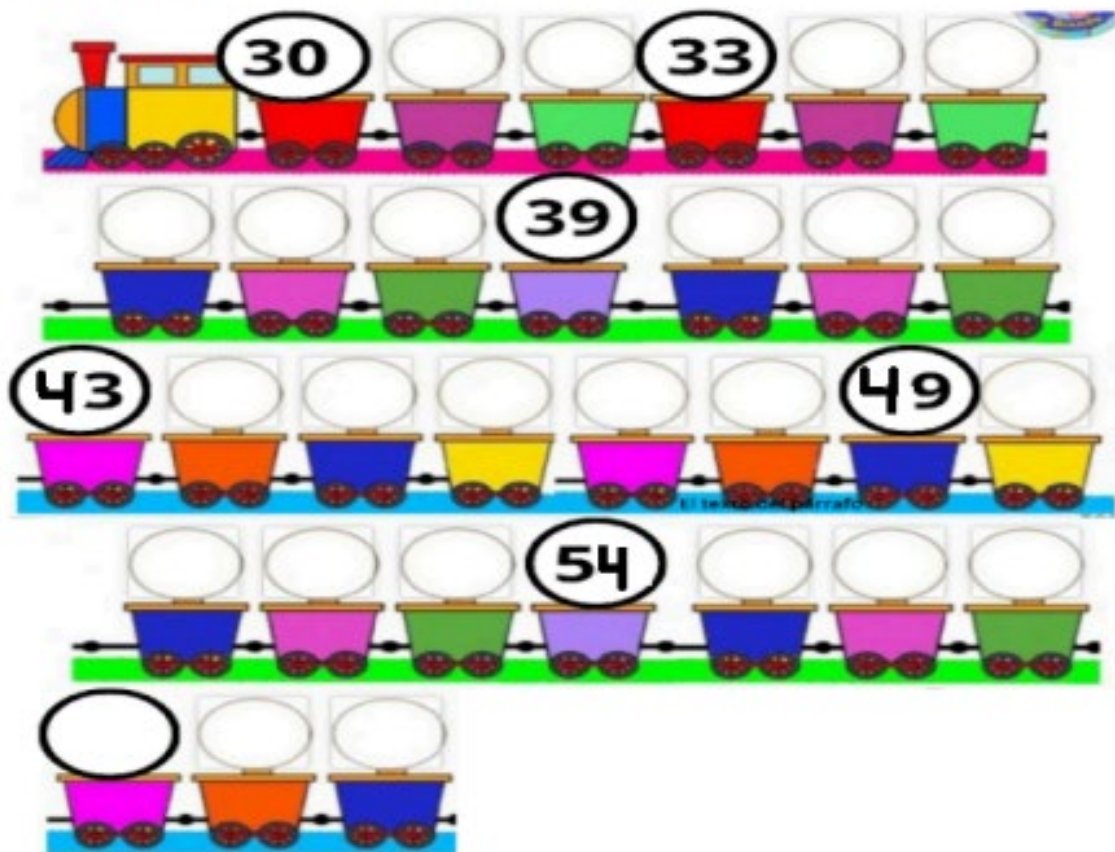
SECCIÓN 2: CONTEO Y CANTIDAD (21-30)

Instrucción: Cuenta los objetos en cada recuadro y escribe el número total.



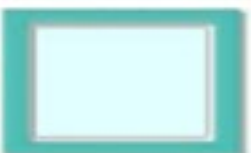
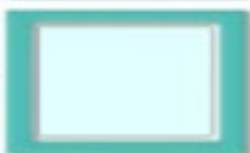
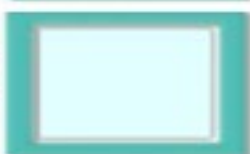
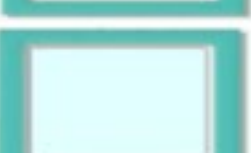
SECCIÓN 3: COMPLETAR SERIES (30-60)

Instrucción: Completa la serie de números escribiendo los que faltan.



SECCIÓN 4: DICTADO (61-100)

Instrucción: Escribe el número dictado.

|

Nombre del docente: _____

Firma del Docente: _____

Fecha: _____

[illegible]

Anexo B1. Solicitud de autorización a la Dirección de la I.E. Fe y Alegría N.º 28

**SOLICITO: Autorización para
realizar proyecto de Tesis**

Nos complace extenderle un cordial saludo y al mismo tiempo dirigirnos a usted con el objetivo de solicitarle la debida autorización para que nosotros: **Loro Ayala Gloria Noelia** con DNI N° **75560943** y **Mayanga Muñoz Jonatan Arturo** con DNI N° **74864514**, bachilleres de la Facultad de Ingeniería de Sistema de la Universidad Nacional Pedro Ruiz Gallo, podamos tener el debido permiso de usted para realizar el Proyecto de Tesis "NumIA: Aplicación móvil basada en Deep Learning para el aprendizaje de números" en la institución educativa I.E. FE Y ALEGRÍA, específicamente con los alumnos de 1er grado de primaria.

Plan de Implementación:

- Duración: 6 semanas.
- Frecuencia: 3 sesiones semanales.
- Capacitación Docente: Se incluye una fase de capacitación para docentes, quienes desempeñarán un rol crucial en la supervisión del uso de la aplicación y el apoyo a los estudiantes.

Chiclayo, 23 de julio del 2025

Loro Ayala Gloria Noelia
DNI 75560943

Mayanga Muñoz Jonatan Arturo
DNI 74864514

Anexo B2. Carta de aceptación institucional

"Año de la recuperación y consolidación de la economía peruana"

**SOLICITO: Autorización para
realizar proyecto de Tesis**

Estimada directora:

Mg. CARMEN BAUTISTA GONZALES

Directora de I.E. Fe y Alegría 28

Nos complace extenderle un cordial saludo y al mismo tiempo dirigimos a usted con el objetivo de solicitarle la debida autorización para que nosotros: **Loro Ayala Gloria Noelia con DNI N° 75560943 y Mayanga Muñoz Jonatan Arturo con DNI N° 74864514**, bachilleres de la Facultad de Ingeniería de Sistema de la Universidad Nacional Pedro Ruiz Gallo, podamos tener el debido permiso de usted para realizar el Proyecto de Tesis "NumIA: Aplicación móvil basada en Deep Learning para el aprendizaje de números" en la institución educativa I.E. FE Y ALEGRÍA, específicamente con los alumnos de 1er grado de primaria.

Este proyecto propone la implementación de una aplicación móvil diseñada para mejorar la práctica de escritura de números arábigos en estudiantes. Su objetivo principal es fortalecer la habilidad de trazado de números, alineado con los objetivos curriculares del Ministerio de Educación (Minedu).

Plan de Implementación:

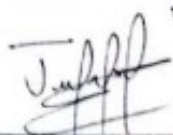
- Duración: 6 semanas.
- Frecuencia: 3 sesiones semanales.
- Capacitación Docente: Se incluye una fase de capacitación para docentes, quienes desempeñarán un rol crucial en la supervisión del uso de la aplicación y el apoyo a los estudiantes.

Agradezco de antemano su comprensión y apoyo en la implementación de esta iniciativa, y al mismo tiempo agradecerle su atención a esta solicitud.

Chiclayo, 23 de julio del 2025



Loro Ayala Gloria Noelia
DNI 75560943



Mayanga Muñoz Jonatan Arturo
DNI 74864514

Recd
22/07/25



Escaneado con CamScanner

GRUPO C: PLANIFICACIÓN Y PROTOCOLO DE INTERVENCIÓN

Anexo C1. Diapositivas para la capacitación docente

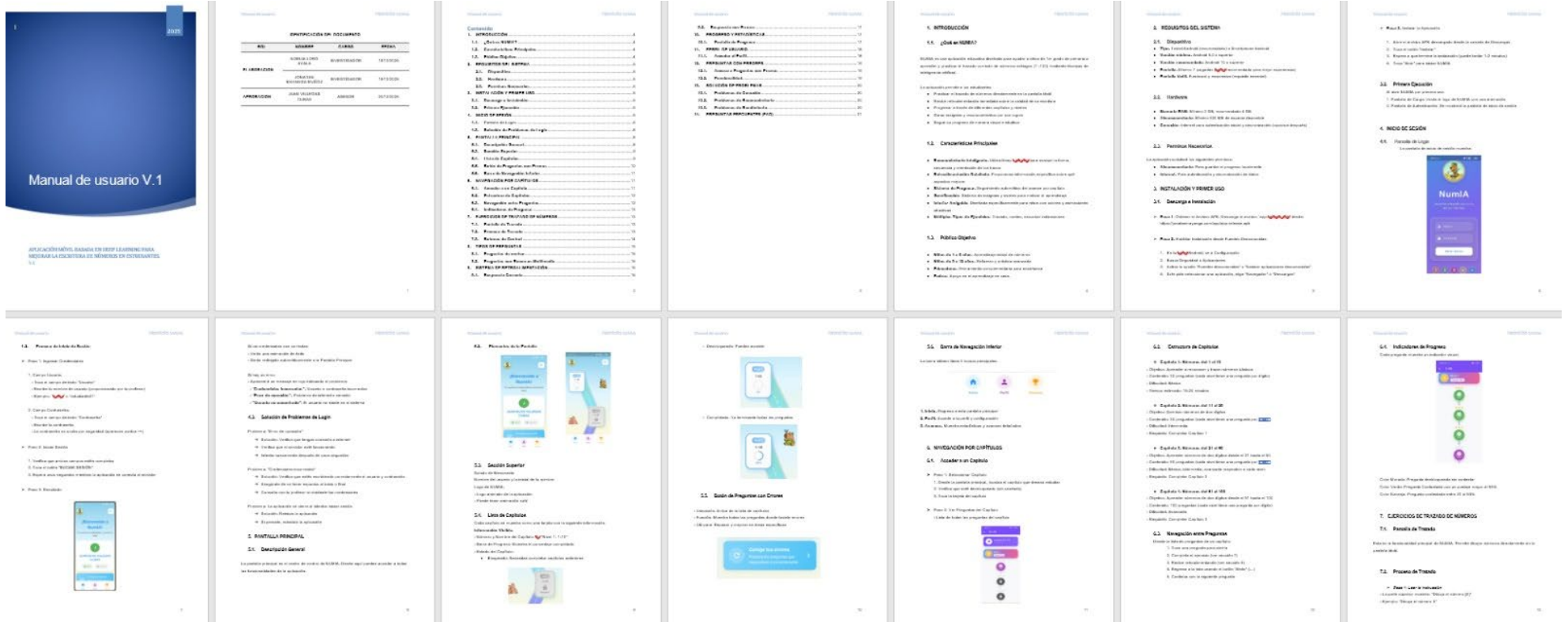


Anexo C2. Protocolo de capacitación docente



GRUPO D: DOCUMENTACIÓN TÉCNICA DE NUMIA

Anexo D1. Manual de usuario de la aplicación NumIA (estudiante y docente)

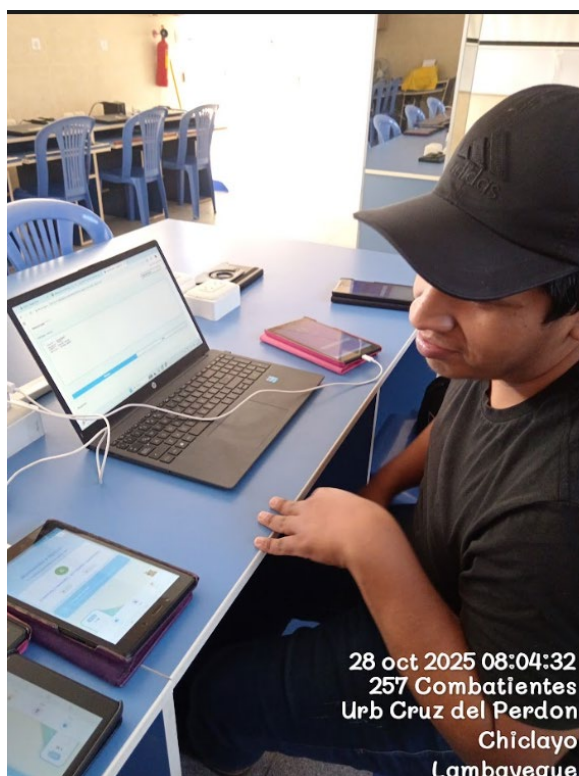


Anexo D2. Capturas de pantalla de la aplicación NumIA



GRUPO E: EVIDENCIAS DE CAMPO

Anexo E1: Fotografías de instalación en las tabletas del colegio.

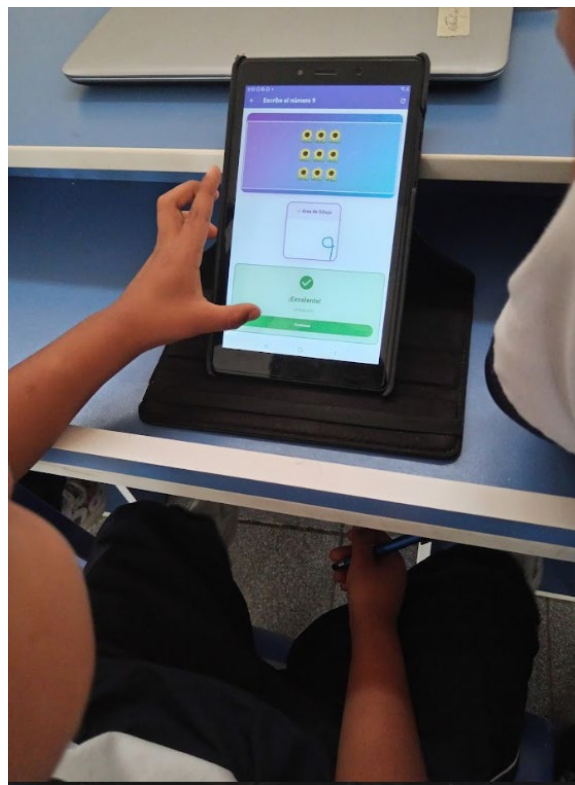
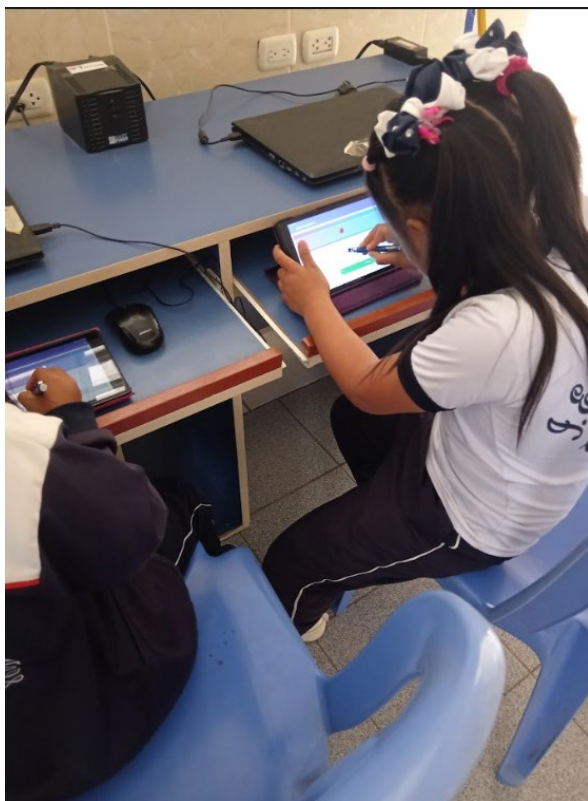


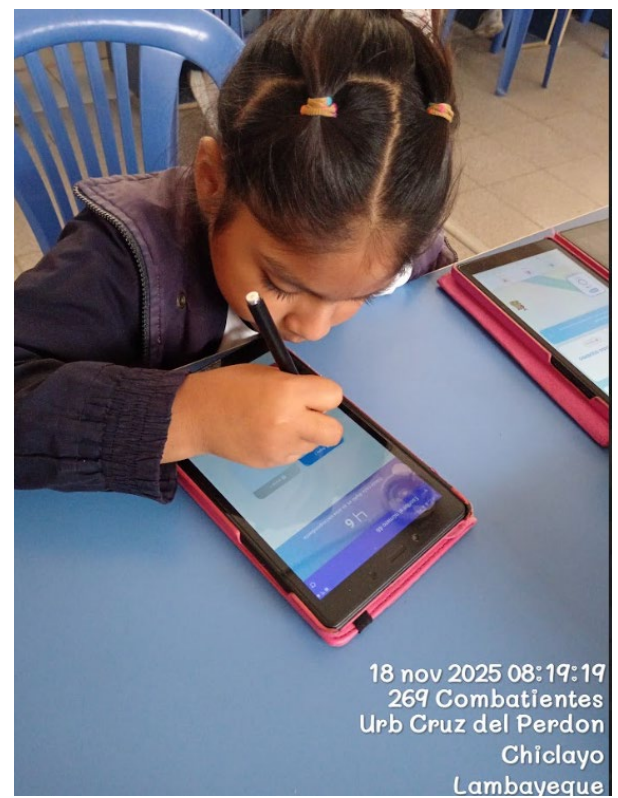
Anexo E2. Fotografías de la fase de diagnóstico (aplicación del pretest)



Anexo E3. Fotografías del uso de la aplicación en el colegio Fé y Alegría N°28

SEMANA 1:







Anexo E4. Muestras del instrumento respondido (pretest y posttest)

26

PRUEBA PRE-TEST - TRAZADO DE NÚMEROS

Aplicación NumIA - Primer Grado de Primaria

Institución Educativa: Fe y Alegría N.° 28

Ubicación: Chiclayo, Lambayeque

Fecha: 21/10

Nombre del estudiante: Luis Nery Torres

Edad: ____ años

SECCIÓN 1: ESCRITURA DE NÚMEROS INDIVIDUALES (0-20)

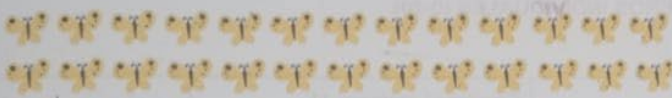
Escribe el número 1	1	Escribe el número 11	11
Escribe el número 2	2	Escribe el número 12	12
Escribe el número 3	3	Escribe el número 13	13
Escribe el número 4	4	Escribe el número 14	14
Escribe el número 5	5	Escribe el número 15	15
Escribe el número 6	6	Escribe el número 16	16
Escribe el número 7	7	Escribe el número 17	17
Escribe el número 8	8	Escribe el número 18	18
Escribe el número 9	9	Escribe el número 19	19
Escribe el número 10	10	Escribe el número 20	20

○ HONOR X8a
● 100MP Ultra Camera

Instrucción: Escribe cada número en el espacio correspondiente. Intenta hacerlo lo más claro y ordenado posible.

SECCIÓN 2: CONTEO Y CANTIDAD (21-30)

Instrucción: Cuenta los objetos en cada recuadro y escribe el número total.

	21
	26
	24
	22
	23
	25
	28

	27
	20
	28

SECCIÓN 3: COMPLETAR SERIES (30-60)

Instrucción: Completa la serie de números escribiendo los que faltan.

El texto del párrafo

SECCIÓN 4: DICTADO (61-100)

Instrucción: Escribe el número dictado.



63



70



98



100



75



62



84



85



78



90



80



95



82



74



67



66



88



76



86



94



49



65



64



89

PRUEBA POST-TEST - TRAZADO DE NÚMEROS

Aplicación NumIA - Primer Grado de Primaria

Institución Educativa: Fe y Alegría N.º 28

Ubicación: Chiclayo, Lambayeque

Fecha: _____

Nombre del estudiante: Jenny

Edad: 7 años

SECCIÓN 1: ESCRITURA DE NÚMEROS INDIVIDUALES (0-20)

Escribe el número diez	10	Escribe el número dieciséis	16
Escribe el número uno	1	Escribe el número doce	12
Escribe el número tres	3	Escribe el número catorce	14
Escribe el número seis	6	Escribe el número diecisiete	17
Escribe el número cuatro	4	Escribe el número once	11
Escribe el número cinco	5	Escribe el número dos	2
Escribe el número siete	7	Escribe el número dieciocho	18
Escribe el número trece	13	Escribe el número veinte	20
Escribe el número nueve	9	Escribe el número diecinueve	19
Escribe el número ocho	8	Escribe el número quince	15

Instrucción: Escribe cada número en el espacio correspondiente. Intenta hacerlo lo más claro y ordenado

SECCIÓN 2: CONTEO Y CANTIDAD (21-30)

Instrucción: Cuenta los objetos en cada recuadro y escribe el número total.



SECCIÓN 3: COMPLETAR SERIES (30-60)

Instrucción: Completa la serie de números escribiendo los que faltan.



SECCIÓN 4: DICTADO (61-100)

Instrucción: Escribe el número dictado.



65



73



72



67



94



84



68



70



69



64



90



80



83



85



77



99



82



100



96



88



32



75



92



93

Nombre del docente: _____

Firma del Docente: _____

Fecha: _____

GRUPO F: REGISTROS DE LA APLICACIÓN

Anexo F1: Script para la obtención de datos de la tabla 29

```

1  /*TABLA 23 - CANTIDAD DE INTENTOS*/
2  SELECT
3      CASE
4          WHEN DATE(fecha_sesion) BETWEEN '2025-10-27' AND '2025-11-02' THEN 'SEMANA 1: 27/10/2025 al 02/11/2025'
5          WHEN DATE(fecha_sesion) BETWEEN '2025-11-03' AND '2025-11-09' THEN 'SEMANA 2: 03/11/2025 al 09/11/2025'
6          WHEN DATE(fecha_sesion) BETWEEN '2025-11-10' AND '2025-11-16' THEN 'SEMANA 3: 10/11/2025 al 16/11/2025'
7          WHEN DATE(fecha_sesion) BETWEEN '2025-11-17' AND '2025-11-23' THEN 'SEMANA 4: 17/11/2025 al 23/11/2025'
8          WHEN DATE(fecha_sesion) BETWEEN '2025-11-24' AND '2025-11-30' THEN 'SEMANA 5: 24/11/2025 al 30/11/2025'
9          WHEN DATE(fecha_sesion) BETWEEN '2025-12-01' AND '2025-12-07' THEN 'SEMANA 6: 01/12/2025 al 07/12/2025'
10     END
11     AS semana,
12     SUM(CASE WHEN usuario_id IN (77,82,83,86,88,89,91,92,94,98)
13         THEN 1 ELSE 0 END) AS intentos_1A,
14     SUM(CASE WHEN usuario_id IN (44,49,53,55,56,57,60,65,69,70)
15         THEN 1 ELSE 0 END) AS intentos_1B,
16     SUM(CASE WHEN usuario_id IN (11,14,16,17,24,32,33,34,37,40)
17         THEN 1 ELSE 0 END) AS intentos_1C,
18     COUNT(id) AS total_global

```

Output TABLA 23 - CANTIDAD DE INTENTOS

semana	intentos_1A	intentos_1B	intentos_1C	total_global
SEMANA 1: 27/10/2025 al 02/11/2025	622	519	939	2080
SEMANA 2: 03/11/2025 al 09/11/2025	818	1133	1091	3042
SEMANA 4: 17/11/2025 al 23/11/2025	858	358	780	1996
SEMANA 5: 24/11/2025 al 30/11/2025	829	724	754	2307
SEMANA 6: 01/12/2025 al 07/12/2025	620	568	634	1822

```

10  /*TABLA 25 ---tiempo promedio*/
11  SELECT uu.seccion, up.usuario_id, AVG(TIMESTAMPDIFF(SECOND, up.hora_inicio, up.hora_fin)) AS tiempo_promedio_s,
12      SUM(TIMESTAMPDIFF(SECOND, up.hora_inicio, up.hora_fin))/60 AS tiempo_total_min,
13      COUNT(*) AS total
14  FROM usuario_pregunta up
15  Left join usuarios uu on up.usuario_id = uu.id
16  WHERE up.usuario_id IN (77,82,83,86,88,89,91,92,94,98,44,49,53,55,56,57,60,65,69,70,11,14,16,17,24,32,33,34,37,40)
17  AND up.hora_fin > up.hora_inicio
18  GROUP BY up.usuario_id, seccion
19  order by seccion, usuario_id;

```

Output TABLA 25 ---tiempo promedio

seccion	usuario_id	tiempo_promedio_s	tiempo_total_min	total
1ª A	77	68.4696	337.7833	296
1ª A	82	33.8116	233.3000	414
1ª A	83	43.1524	235.9000	328
1ª A	86	31.3106	191.5167	367
1ª A	88	28.2751	178.1333	378
1ª A	89	41.7507	253.9833	365
1ª A	91	30.0639	180.3833	360
1ª A	92	21.6809	169.8333	470
1ª A	94	31.6134	181.2500	344

55	
56	/*ASISTENCIA SEMANAL*/
57	✓ SELECT
58	uu.seccion, uu.id,
59	SUM(CASE WHEN fecha_sesion BETWEEN '2025-10-27' AND '2025-11-02' THEN 1 ELSE 0 END) S1,
60	SUM(CASE WHEN fecha_sesion BETWEEN '2025-11-03' AND '2025-11-09' THEN 1 ELSE 0 END) S2,
61	SUM(CASE WHEN fecha_sesion BETWEEN '2025-11-17' AND '2025-11-23' THEN 1 ELSE 0 END) S3,
62	SUM(CASE WHEN fecha_sesion BETWEEN '2025-11-24' AND '2025-11-30' THEN 1 ELSE 0 END) S4,
63	SUM(CASE WHEN fecha_sesion BETWEEN '2025-12-01' AND '2025-12-07' THEN 1 ELSE 0 END) S5,
64	COUNT(*) total FROM usuario_pregunta up inner join usuarios uu [1.nc->1: on up.usuario_id = uu.id
65	WHERE up.usuario_id IN
66	(77,81,82,83,86,88,89,91,94,98,44,49,53,55,56,60,65,69,70,71,11,16,17,20,24,32,33,34,37,40)
67	GROUP BY up.usuario_id, uu.seccion ORDER BY seccion, up.usuario_id;
68	

Output ASISTENCIA SEMANAL ×										
30 rows										
	seccion	id	S1	S2	S3	S4	S5	total		
1	1ªA	77	11	63	122	59	41	296		
2	1ªA	81	0	0	104	0	0	104		
3	1ªA	82	125	48	89	80	72	414		
4	1ªA	83	53	72	79	57	67	328		
5	1ªA	86	93	0	134	77	63	367		
6	1ªA	88	87	97	0	98	96	378		
7	1ªA	89	52	61	100	81	71	365		
8	1ªA	91	71	140	10	98	41	360		
9	1ªA	94	53	75	135	74	0	327		

GRUPO G: VALIDACION DE INSTRUMENTO POR JUICIO DE EXPERTOS

Anexo G1: Formato de validación de instrumentos

UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO VALIDACIÓN DE INSTRUMENTO POR JUICIO DE EXPERTOS

DATOS DEL EXPERTO

Nombre y apellidos: _____

Especialidad: _____

Institución: _____

Grado académico: _____

Años de experiencia en educación primaria: _____

INSTRUCCIONES

Evalúe cada criterio de la rúbrica según tres dimensiones, utilizando la escala:

1 = Deficiente 2 = Regular 3 = Bueno 4 = Excelente

CRITERIO 1: ORIENTACIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☐
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☐
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☐

Observaciones: _____

CRITERIO 2: SECUENCIA DE ORDEN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☐
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☐
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☐

Observaciones: _____

CRITERIO 3: SECUENCIA DE FLUIDEZ

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☐
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☐
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☐

Observaciones: _____

CRITERIO 4: PRECISIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☐
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☐
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☐

Observaciones: _____

Anexo G2: Respuestas de los docentes

UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO VALIDACIÓN DE INSTRUMENTO POR JUICIO DE EXPERTOS

DATOS DEL EXPERTO

Nombre y apellidos: Laira Tannina Ballena Marquez
Especialidad: Ed. Primaria
Institución: Fe y Alegría 28
Grado académico: Docente de Ed. Primaria
Años de experiencia en educación primaria: 20 años

INSTRUCCIONES

Evalúe cada criterio de la rúbrica según tres dimensiones, utilizando la escala:

1 = Deficiente 2 = Regular 3 = Bueno 4 = Excelente

CRITERIO 1: ORIENTACIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☒
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☒
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☒

Observaciones: _____

CRITERIO 2: SECUENCIA DE ORDEN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☒
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☒
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☒

Observaciones: _____

CRITERIO 3: SECUENCIA DE FLUIDEZ

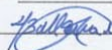
Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☒
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☒
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☒

Observaciones: _____

CRITERIO 4: PRECISIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☒
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☒
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☒

Observaciones: _____

Firma:  Fecha: 11-12-2025

UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
VALIDACIÓN DE INSTRUMENTO POR JUICIO DE EXPERTOS

DATOS DEL EXPERTO

Nombre y apellidos: Mario Carranza Fernandez
Especialidad: Educación Primaria - Computación e informática
Institución: Fe y Alegría 28
Grado académico: Licenciado
Años de experiencia en educación primaria: 12 años

INSTRUCCIONES

Evalúe cada criterio de la rúbrica según tres dimensiones, utilizando la escala:

1 = Deficiente 2 = Regular 3 = Bueno 4 = Excelente

CRITERIO 1: ORIENTACIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☒
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☒
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☒

Observaciones: _____

CRITERIO 2: SECUENCIA DE ORDEN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☒
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☒
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☐	☒

Observaciones: _____

CRITERIO 3: SECUENCIA DE FLUIDEZ

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☐	☒
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☐	☒
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☒	☐

Observaciones: _____

CRITERIO 4: PRECISIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	☐	☐	☒	☐
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	☐	☐	☒	☐
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	☐	☐	☒	☐

Observaciones: _____

Firma:  Fecha: _____

UNIVERSIDAD NACIONAL PEDRO RUIZ GALLO
VALIDACIÓN DE INSTRUMENTO POR JUICIO DE EXPERTOS

DATOS DEL EXPERTO

Nombre y apellidos: ALEGRIA CARDENAS EYBI RAQUEL

Especialidad: EDUCACION PRIMARIA

Institución: FE Y ALEGRIA 28

Grado académico: LICENCIATURA

Años de experiencia en educación primaria: 4

INSTRUCCIONES

Evalúe cada criterio de la rúbrica según tres dimensiones, utilizando la escala:

1 = Deficiente 2 = Regular 3 = Bueno 4 = Excelente

CRITERIO 1: ORIENTACIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)		o	o	X
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	o	o	X	o
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	o	o	o	X

Observaciones: _____

CRITERIO 2: SECUENCIA DE ORDEN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	o	o	o	X
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	o	o	X	o
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	o	o	o	X

Observaciones: _____

CRITERIO 3: SECUENCIA DE FLUIDEZ

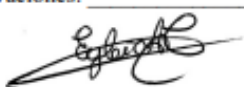
Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	o	o	o	X
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	o	o	o	X
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	o	o	o	X

Observaciones: _____

CRITERIO 4: PRECISIÓN

Dimensión	1	2	3	4
Claridad (¿el criterio está redactado de forma comprensible?)	o	o	X	o
Pertinencia (¿es pertinente para evaluar el trazado numérico?)	o	o	X	
Relevancia (¿es relevante para el aprendizaje en 1º grado?)	o	o	o	X

Observaciones: _____

Firma: 

Fecha: _____

GRUPO H: EVALUACION DE USABILIDAD

Anexo H1: Formato de evaluación de usabilidad

EVALUACIÓN DE USABILIDAD

Docente: _____ Sección: _____ Fecha: _____

1 = En desacuerdo — 5 = De Acuerdo

Nº	Ítem	1	2	3	4	5
1	Me gustaría usar App con frecuencia en mis clases	☐	☐	☐	☐	☐
2	Fue complejo usar la App en el Aula	☐	☐	☐	☐	☐
3	Considero que la App es fácil de usar en el Aula	☐	☐	☐	☐	☐
4	Necesito Apoyo para poder usar el App	☐	☐	☐	☐	☐
5	Las funciones del App estuvieron bien integradas	☐	☐	☐	☐	☐
6	Observé demasiadas inconsistencias en el funcionamiento del App	☐	☐	☐	☐	☐
7	La mayoría de mis estudiantes aprendió a usar el App con rapidez	☐	☐	☐	☐	☐
8	Fue muy difícil de manejar el App durante las sesiones	☐	☐	☐	☐	☐
9	Me sentí seguro/a supervisando el uso de App con mis estudiantes	☐	☐	☐	☐	☐
10	Necesité aprender muchas cosas antes de poder usar App con los estudiantes	☐	☐	☐	☐	☐

Comentarios: _____

Firma: _____ Fecha: _____

Anexo H2: Respuestas de los docentes

EVALUACIÓN DE USABILIDAD

Docente: Taira Tammina Ballena Masquez Sección: 1º A Fecha: 11-12-2025

1 = En desacuerdo — 5 = De Acuerdo

Nº	Ítem	1	2	3	4	5
1	Me gustaría usar App con frecuencia en mis clases	☐	☐	☐	☒	☐
2	Fue complejo usar la App en el Aula	☒	☐	☐	☐	☐
3	Considero que la App es fácil de usar en el Aula	☐	☐	☐	☐	☒
4	Necesito Apoyo para poder usar el App	☒	☐	☐	☐	☐
5	Las funciones del App estuvieron bien integradas	☐	☐	☐	☐	☒
6	Observé demasiadas inconsistencias en el funcionamiento del App	☒	☐	☐	☐	☐
7	La mayoría de mis estudiantes aprendió a usar el App con rapidez	☐	☐	☐	☐	☒
8	Fue muy difícil de manejar el App durante las sesiones	☒	☐	☐	☐	☐
9	Me sentí seguro/a supervisando el uso de App con mis estudiantes	☐	☐	☐	☐	☒
10	Necesité aprender muchas cosas antes de poder usar App con los estudiantes	☒	☐	☐	☐	☐

Comentarios: Que se aplique a los demás grados con las adaptaciones correspondientes.

Firma:  Fecha: _____

EVALUACIÓN DE USABILIDAD

Docente: ALEGRIA CARDENAS EYBI RAQUEL

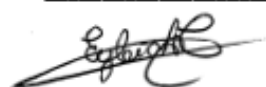
Sección: B Fecha: 11-12-25

1 = En desacuerdo — 5 = De Acuerdo

Nº	Ítem	1	2	3	4	5
1	Me gustaría usar App con frecuencia en mis clases	☐	☐	☐	☐	☒
2	Fue complejo usar la App en el Aula	☒	☐	☐	☐	☐
3	Considero que la App es fácil de usar en el Aula	☐	☐	☐	☐	☒
4	Necesito Apoyo para poder usar el App	☒	☐	☐	☐	☐
5	Las funciones del App estuvieron bien integradas	☐	☐	☐	☒	☐
6	Observé demasiadas inconsistencias en el funcionamiento del App	☐	☒	☐	☐	☐
7	La mayoría de mis estudiantes aprendió a usar el App con rapidez	☐	☐	☐	☒	☐
8	Fue muy difícil de manejar el App durante las sesiones	☒	☐	☐	☐	☐
9	Me sentí seguro/a supervisando el uso de App con mis estudiantes	☐	☐	☐	☒	☐
10	Necesité aprender muchas cosas antes de poder usar App con los estudiantes	☒	☐	☐	☐	☐

Comentarios: _____

Firma:



Fecha: _____

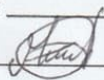
EVALUACIÓN DE USABILIDAD

Docente: Matias Corraza Fernandez Sección: _____ Fecha: _____

1 = En desacuerdo — 5 = De Acuerdo

Nº	Ítem	1	2	3	4	5
1	Me gustaría usar App con frecuencia en mis clases	☐	☐	☐	☐	☒
2	Fue complejo usar la App en el Aula	☐	☒	☐	☐	☐
3	Considero que la App es fácil de usar en el Aula	☐	☐	☐	☐	☒
4	Necesito Apoyo para poder usar el App	☒	☐	☐	☐	☐
5	Las funciones del App estuvieron bien integradas	☐	☐	☐	☐	☒
6	Observé demasiadas inconsistencias en el funcionamiento del App	☒	☐	☐	☐	☐
7	La mayoría de mis estudiantes aprendió a usar el App con rapidez	☐	☐	☐	☐	☒
8	Fue muy difícil de manejar el App durante las sesiones	☒	☐	☐	☐	☐
9	Me sentí seguro/a supervisando el uso de App con mis estudiantes	☐	☐	☐	☐	☒
10	Necesité aprender muchas cosas antes de poder usar App con los estudiantes	☒	☐	☐	☐	☐

Comentarios: _____

Firma:  Fecha: _____