



Universidad  
**Norbert Wiener**

Powered by **Arizona State University**

**FACULTAD DE CIENCIAS DE LA SALUD**  
**PROGRAMA ACADÉMICO DE MEDICINA HUMANA**

**Tesis**

Factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024

**Para optar el Título Profesional de**  
Médico Cirujano

**Presentado por:**

**Autor:** Diaz Carrillo, Jimmy Alejandro

**Código ORCID:** <https://orcid.org/0000-0001-9422-5902>

**Asesora:** Dra. Díaz Barrientos, Galina

**Código ORCID:** <https://orcid.org/0000-0002-8502-5842>

**Lima – Perú**

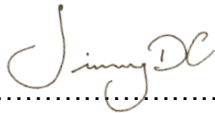
**2025**

 Universidad Norbert Wiener	<b>DECLARACIÓN JURADA DE AUTORIA Y DE ORIGINALIDAD DEL TRABAJO DE INVESTIGACIÓN</b>	
	<b>CÓDIGO: UPNW-GRA-FOR-033</b>	<b>VERSIÓN: 01</b> REVISIÓN: 01

Yo, **Jimmy Alejandro Diaz Carrillo** egresado de la Facultad de **Ciencias de la Salud** y Escuela Académica Profesional de **Medicina Humana** de la Universidad privada Norbert Wiener declaro que el trabajo de investigación **“Factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024”** Asesorado por el docente: **Dra. Galina Díaz Barrientos** DNI 40441748 ORCID 0000-0002-8502-5842 tiene un índice de similitud de 10 (diez) % con código oid: 14912:438889399 verificable en el reporte de originalidad del software Turnitin.

Así mismo:

1. Se ha mencionado todas las fuentes utilizadas, identificando correctamente las citas textuales o paráfrasis provenientes de otras fuentes.
2. No he utilizado ninguna otra fuente distinta de aquella señalada en el trabajo.
3. Se autoriza que el trabajo puede ser revisado en búsqueda de plagios.
4. El porcentaje señalado es el mismo que arrojó al momento de indexar, grabar o hacer el depósito en el turnitin de la universidad y,
5. Asumimos la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión en la información aportada, por lo cual nos sometemos a lo dispuesto en las normas del reglamento vigente de la universidad.



.....  
 Bch. Jimmy Alejandro Diaz Carrillo  
 DNI: 74289105



.....  
 Dra. Galina Díaz Barrientos  
 DNI: 40441748

Lima, 24 de enero de 2025

**Dedicatoria**

Para mi madre Iraida,  
la orquídea más bonita del jardín,  
quien me brinda su fortaleza y ejemplo  
para seguir adelante en la carrera  
como en la vida.

Para mi padre Gregorio,  
luz y agua de las flores,  
quien siempre me enseña a sonreír  
y que cualquier  
problema tiene solución.

Para mi hermana Arely,  
la flor jazmín de la India,  
quien es siempre mi confiable  
primera opinión.

Gracias por hacerme Médico.

El autor

## **Agradecimiento**

A la Dra. Galina Diaz Barrientos, por su tiempo, gentileza y guía en la elaboración de la presente tesis.

A los médicos y médicas de mi panel experto, que desinteresadamente me brindaron su ayuda.

A mi familia extensa, quienes me brindaron siempre el apoyo y ejemplo necesario para ser una buena persona y médico.

A los amigos y amigas que me ayudaron con su apoyo para la conclusión de esta investigación. Así como su soporte para hacer los días de universidad más felices.

A las mascotas que me acompañaron y acompañan, por su infinita lealtad y por regalarme tantas sonrisas.

El autor

## Índice general

\_Toc191303254

<b>Declaración jurada de autoría y originalidad del trabajo.....</b>	<b>II</b>
<b>Dedicatoria.....</b>	<b>IV</b>
<b>Agradecimiento .....</b>	<b>V</b>
<b>Índice de tablas.....</b>	<b>VIII</b>
<b>Índice de figuras .....</b>	<b>IX</b>
<b>Resumen .....</b>	<b>X</b>
<b>Abstract .....</b>	<b>XI</b>
<b>Introducción.....</b>	<b>XII</b>
<b>CAPÍTULO I: EL PROBLEMA .....</b>	<b>1</b>
<b>1.1. Planteamiento del problema .....</b>	<b>1</b>
<b>1.2. Formulación del problema .....</b>	<b>3</b>
<b>1.2.1. Problema general.....</b>	<b>3</b>
<b>1.2.2. Problemas específicos.....</b>	<b>4</b>
<b>1.3. Objetivos de la investigación.....</b>	<b>4</b>
<b>1.3.1. Objetivo general .....</b>	<b>4</b>
<b>1.3.2. Objetivos específicos .....</b>	<b>4</b>
<b>1.4. Justificación de la investigación .....</b>	<b>5</b>
<b>1.4.1. Teórica.....</b>	<b>5</b>
<b>1.4.2 Metodológica.....</b>	<b>6</b>
<b>1.4.3. Práctica .....</b>	<b>7</b>
<b>1.5. Delimitaciones de la investigación.....</b>	<b>8</b>
<b>1.5.1. Temporal.....</b>	<b>8</b>
<b>1.5.2. Espacial.....</b>	<b>8</b>
<b>1.5.3. Recursos.....</b>	<b>9</b>
<b>1.6. Limitaciones de la investigación.....</b>	<b>9</b>
<b>CAPITULO II: MARCO TEORICO.....</b>	<b>11</b>
<b>2.1. Antecedentes.....</b>	<b>11</b>
<b>2.2. Bases teóricas .....</b>	<b>23</b>
<b>2.2.1. Razonamiento clínico .....</b>	<b>23</b>
<b>2.2.2. Guiones de la enfermedad .....</b>	<b>34</b>
<b>2.2.3. Prueba de concordancia de guiones .....</b>	<b>35</b>
<b>2.2.4. Uso de Inteligencia Artificial en Educación Médica .....</b>	<b>41</b>
<b>2.3. Formulación de hipótesis.....</b>	<b>47</b>
<b>2.3.1. Hipótesis General .....</b>	<b>47</b>

2.3.2. Hipótesis específicas .....	48
<b>CAPÍTULO III: METODOLOGÍA .....</b>	<b>50</b>
3.1. Método de la investigación .....	50
3.2. Enfoque de la investigación .....	50
3.3. Tipo de investigación.....	50
3.4. Diseño de la investigación.....	50
3.5 Proceso de Diseño, Revisión y Aplicación del Script Concordance Test (SCT) .....	50
3.6. Población, muestra y muestreo .....	68
3.6. Variables y operacionalización .....	70
3.7. Técnicas e instrumentos de recolección de datos.....	75
3.7.1. Técnica.....	75
3.7.2. Descripción de instrumento .....	75
3.8. Plan de procesamiento y análisis de datos .....	80
3.9. Aspectos éticos.....	82
<b>CAPÍTULO IV: RESULTADOS Y DISCUSIÓN .....</b>	<b>84</b>
4.1. Resultados .....	84
4.1.1. Presentación de resultados descriptivos .....	84
4.1.2. Prueba de hipótesis .....	95
4.1.3. Discusión .....	99
<b>CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES.....</b>	<b>108</b>
5.1. Conclusiones.....	108
5.2. Recomendaciones.....	109
<b>REFERENCIAS .....</b>	<b>111</b>
<b>ANEXOS.....</b>	<b>129</b>
<b>Anexo 1: Matriz de consistencia.....</b>	<b>129</b>
<b>Anexo 2: Instrumento.....</b>	<b>131</b>
<b>Anexo 3: Validez del instrumento .....</b>	<b>157</b>
<b>Anexo 4: Aprobación del comité de ética .....</b>	<b>163</b>
<b>Anexo 5: Formato de consentimiento informado.....</b>	<b>164</b>

## Índice de tablas

<b>Tabla 1.</b> Una guía para generar ítems de SCT en educación médica. Reproducida de: A Prompt for Generating Script Concordance Test Using ChatGPT, Claude, and Llama Large Language Model Chatbots.....	46
<b>Tabla 2.</b> Temario objetivo para la estructura del SCT .....	53
<b>Tabla 3.</b> Temario final v1 para la estructura del SCT.....	57
<b>Tabla 4.</b> Esquema de corrección para el panel de expertos para la valoración del SCT .....	67
<b>Tabla 5.</b> Coeficiente de concordancia de Kendall (W) calculado con SPSS .....	78
<b>Tabla 6.</b> Cálculo del alfa de Cronbach utilizando SPSS .....	78
<b>Tabla 7.</b> Cálculo de la correlación ítem-total para todo el conjunto del test con SPSS .....	79
<b>Tabla 8.</b> Características sociodemográficas de estudiantes internos de una escuela de medicina humana, Universidad Privada Norbert Wiener, Lima 2024.....	84
<b>Tabla 9.</b> Características académicas de estudiantes internos de una escuela de medicina humana, Universidad Privada Norbert Wiener, Lima 2024 .....	85
<b>Tabla 10.</b> Características entorno asistenciales de estudiantes internos de una escuela de medicina humana, Universidad Privada Norbert Wiener, Lima 2024 .....	87
<b>Tabla 11.</b> Cálculo del Alfa de Cronbach para los estudiantes .....	91
<b>Tabla 12.</b> Test de normalidad Kolmogorov-Smirnov para el panel experto y los estudiantes.....	91
<b>Tabla 13.</b> Análisis descriptivo de la distribución de las puntuaciones SCT para el panel experto y los estudiantes .....	92
<b>Tabla 14.</b> Análisis bivariado de los factores sociodemográficos asociados al desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024 .....	95
<b>Tabla 15.</b> Análisis bivariado de los factores académicos asociados al desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.....	96
<b>Tabla 16</b> Análisis bivariado de los factores entorno asistenciales asociados al desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024 .....	99

## Índice de figuras

<b>Figura 1.</b> Ejemplo de Ítem SCT. Traducción de: To what extent is the Script Concordance Test a valid measure of clinical reasoning for Advanced Paediatric Life Support Training? .....	38
<b>Figura 2.</b> Script Concordance Test Generator (ChatGPT).....	47
<b>Figura 3.</b> Proceso de Revisión y Validación del Script Concordance Test .....	63
<b>Figura 4.</b> Cohorte aprobado/no aprobado total .....	88
<b>Figura 5.</b> Distribución de frecuencias de las puntuaciones de los estudiantes .....	89
<b>Figura 6.</b> Dimensión 1: clínico-médica .....	89
<b>Figura 7.</b> Dimensión 2: clínico-quirúrgica.....	90
<b>Figura 8.</b> Distribución de las puntuaciones de los miembros del panel experto y de los estudiantes.....	93
<b>Figura 9.</b> Distribución de frecuencias de las puntuaciones del panel experto vs estudiantes vs ChatGPT-4o.....	94

## Resumen

**Introducción:** La tesis analiza el razonamiento clínico en internos de medicina mediante el "Script Concordance Test" (SCT). Este instrumento evalúa habilidades en contextos de incertidumbre clínica, midiendo cómo las decisiones de los estudiantes se alinean con las de expertos. **Objetivo:** Analizar los factores del razonamiento clínico asociados a su desempeño académico en el SCT en estudiantes de una escuela de Medicina Humana en Lima durante el 2024. **Métodos:** Se adoptó un enfoque cuantitativo con un diseño no experimental, de tipo observacional y transversal. Se incluyeron 87 internos de medicina, evaluando variables sociodemográficas, académicas y de entorno asistencial. El SCT fue diseñado utilizando inteligencia artificial (ChatGPT), lo que permitió generar casos clínicos validados por expertos. Se midieron variables sociodemográficas, académicas y entorno asistenciales mediante un análisis bivariado. **Resultados:** El 74.7% de los estudiantes aprobó el SCT, aunque el 25.3% reprobó y ChatGPT superó a estudiantes como a la mayoría de los expertos. La participación en prácticas clínicas extracurriculares tuvo un impacto significativo en el desempeño académico del SCT ( $p=0.030$ ). **Conclusión:** El SCT demostró ser una herramienta efectiva y con evidencia confiable para evaluar el razonamiento clínico, identificándose una asociación significativa con la participación en prácticas clínicas extracurriculares.

**Palabras clave:** Razonamiento clínico, Script Concordance Test, educación médica, inteligencia artificial

## Abstract

**Introduction:** The thesis analyzes clinical reasoning in medical interns using the Script Concordance Test (SCT). This instrument assesses skills in contexts of clinical uncertainty, measuring how students' decisions align with those of experts. **Objective:** To analyze the factors of clinical reasoning associated with their academic performance in the SCT in students of a school of Human Medicine in Lima during 2024. **Methods:** A quantitative approach with a non-experimental, observational and cross-sectional design was adopted. Eighty-seven medical interns were included, evaluating sociodemographic, academic and care environment variables. The SCT was designed using artificial intelligence (ChatGPT), which allowed the generation of clinical cases validated by experts. Sociodemographic, academic and care environment variables were measured by bivariate analysis. **Results:** 74.7% of the students passed the SCT, although 25.3% failed and ChatGPT outperformed both students and most experts. Participation in extracurricular clinical practice had a significant impact on SCT academic performance ( $p=0.030$ ). **Conclusion:** The SCT proved to be an effective tool with reliable evidence to assess clinical reasoning, identifying a significant association with participation in extracurricular clinical practices.

**Keywords:** clinical reasoning, Script Concordance Test, medical education, artificial intelligence.

## Introducción

El razonamiento clínico es un proceso fundamental en la formación médica, pues permite a los profesionales analizar, diagnosticar y tomar decisiones en escenarios clínicos complejos e inciertos. La presente investigación aborda este aspecto crítico mediante la implementación del *Script Concordance Test* (SCT), una herramienta diseñada para evaluar la capacidad de los internos de medicina de enfrentar casos clínicos con información ambigua o incompleta. Este estudio tiene como propósito identificar los factores académicos, sociodemográficos y del entorno asistencial asociados al desempeño en el SCT en internos de una universidad privada de Lima, Perú. Asimismo, introduce el uso innovador de inteligencia artificial para el diseño de viñetas clínicas, buscando optimizar tanto la evaluación como la formación en razonamiento clínico.

Esta tesis consta de cinco capítulos:

En el Capítulo I, se presentan el problema de investigación, los objetivos y la justificación, resaltando la relevancia de fortalecer las competencias críticas en los futuros médicos.

El Capítulo II desarrolla el marco teórico, analizando los conceptos esenciales del razonamiento clínico, los antecedentes relevantes y la importancia del SCT como herramienta evaluativa.

El Capítulo III detalla la metodología, explicando el diseño observacional, los procedimientos para la validación del SCT y la integración de inteligencia artificial en la creación de casos clínicos.

El Capítulo IV expone los resultados y discusión, destacando cómo los factores estudiados influyen en el desempeño y el potencial de la inteligencia artificial como complemento educativo.

Finalmente, en el Capítulo V, se presentan conclusiones y recomendaciones, enfocadas en mejorar las estrategias de enseñanza y evaluación del razonamiento clínico, sentando las bases para un modelo replicable en contextos similares.

## **CAPÍTULO I: EL PROBLEMA**

### **1.1. Planteamiento del problema**

En la formación médica, el desarrollo y la evaluación del razonamiento clínico (RC) son fundamentales para la toma de decisiones en la práctica clínica. Aunque la comunidad médica reconoce su importancia,<sup>1</sup> muchas instituciones carecen de un currículo estructurado y de docentes capacitados para enseñarlo<sup>2</sup>. La mayoría de los educadores en medicina no poseen formación docente formal, desarrollando sus competencias pedagógicas de manera autodidacta<sup>1,3</sup>. Además, deben gestionar actividades asistenciales, de investigación y administrativas, lo que limita su tiempo para capacitarse en la enseñanza del razonamiento clínico<sup>3</sup>.

Un estudio de Kononowicz et al. (2020), realizado en 76 países, reveló que solo el 28% de los encuestados reportó la existencia de un curso de razonamiento clínico en sus programas, mientras que el 85% consideró necesaria su implementación<sup>2</sup>. Asimismo, el 66% identificó la falta de reconocimiento de su importancia y el 58% señaló la escasez de docentes capacitados como barreras clave para su enseñanza<sup>2</sup>. Tambeh et al. (2023) también encontraron que los docentes carecen de formación especializada en RC, lo que dificulta la enseñanza de estas habilidades<sup>1</sup>. Así la mayoría de las capacitaciones (91.6%) están dirigidas a educadores de estudiantes en etapas clínicas, limitando su integración progresiva<sup>1</sup>.

Un problema adicional es la confusión conceptual sobre el razonamiento clínico, dificultando su enseñanza y evaluación<sup>4</sup>. Este desafío se agrava por la distribución desigual de la literatura sobre RC, con una predominancia del 87.4% de artículos provenientes de

Norteamérica y Europa, mientras que Sudamérica representa solo el 0.8%<sup>5</sup>. Esta disparidad sesga la enseñanza y evaluación del RC, limitando su contextualización en diversas realidades<sup>5</sup>. Diversos métodos se han propuesto para desarrollar y evaluar el RC, como la reflexión estructurada, la autoexplicación, la generación de diagnósticos diferenciales y la enseñanza basada en esquemas<sup>6</sup>. Sin embargo, su efectividad varía debido a diferencias metodológicas y a la falta de estudios comparativos<sup>6</sup>. Entre estos métodos, los guiones de enfermedad destacan como una herramienta ampliamente respaldada, siendo la Prueba de Concordancia de Guiones una de las más utilizadas para evaluar el RC<sup>8</sup>. No obstante, la implementación del SCT enfrenta desafíos, como la creación de preguntas que reflejen la complejidad clínica y la variabilidad en su aplicación según el contexto cultural y educativo<sup>9</sup>.

El uso del SCT varía significativamente según el país. En Suiza, se emplea de manera integral en la evaluación de competencias clínicas, mientras que en Austria y Alemania su implementación es más limitada<sup>10</sup>. En Estados Unidos, un estudio multicéntrico evidenció diferencias significativas en el desempeño en SCT entre estudiantes de distintas universidades,<sup>11</sup> lo que sugiere que su efectividad puede depender del contexto académico específico. De manera similar, Nouh et al. (2012) encontraron variaciones en la progresión del RC entre residentes de cirugía general en distintas universidades de Canadá,<sup>12</sup> reflejando diferencias en la formación y aplicación del SCT dentro de un mismo país. Además, Nomura et al. (2021) exploraron la adaptación del Pediatric Emergency Medicine Script Concordance Test (PEM-SCT) al japonés, destacando la importancia de considerar factores socioculturales en la evaluación del razonamiento clínico<sup>13</sup>. Estos estudios resaltan la necesidad de contextualizar el SCT para maximizar su efectividad en distintos entornos educativos y culturales.

En Perú, la evaluación y desarrollo del RC en internos de medicina presentan deficiencias significativas. No existe un sistema estandarizado de evaluación durante el internado, lo que

dificulta medir la capacidad de los internos para tomar decisiones en escenarios clínicos inciertos. La enseñanza del RC en el internado se basa mayormente en la experiencia práctica sin metodologías estructuradas, generando formación heterogénea. Los métodos de evaluación predominantes, como los exámenes de opción múltiple o casos clínicos orales, priorizan la memorización sobre la toma de decisiones clínicas. A nivel nacional, el Examen Nacional de Medicina (ENAM) es la única evaluación estandarizada para internos, pero presenta múltiples limitaciones en la medición del RC. Ya que se basa en preguntas de opción múltiple que priorizan conocimientos teóricos sobre la toma de decisiones clínicas. Así su aplicación tardía impide realizar intervenciones pedagógicas oportunas, limitando la corrección de deficiencias antes del egreso de los internos.

Dado que el internado es la última etapa de formación antes del ejercicio profesional, es crucial analizar su desempeño en el razonamiento clínico, ya que esto permitirá identificar áreas de mejora en su enseñanza y evaluar la viabilidad del SCT como herramienta de medición estandarizada en el país. Por ello la presente investigación contribuirá al fortalecimiento de la educación médica, determinando los factores asociados al desempeño académico en el SCT en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en 2024.

## **1.2. Formulación del problema**

### **1.2.1. Problema general**

PG: ¿Cuáles son los factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?

### **1.2.2. Problemas específicos**

Pe<sub>1</sub>: ¿Cuáles son los factores sociodemográficos del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?

Pe<sub>2</sub>: ¿Cuáles son los factores académicos del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?

Pe<sub>3</sub>: ¿Cuáles son los factores entorno asistenciales del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?

### **1.3. Objetivos de la investigación**

#### **1.3.1. Objetivo general**

OG: Determinar los factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

#### **1.3.2. Objetivos específicos**

Oe<sub>1</sub>: Determinar los factores sociodemográficos del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

Oe<sub>2</sub>: Determinar los factores académicos del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

Oe<sub>3</sub>: Determinar los factores entorno asistenciales del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

## **1.4. Justificación de la investigación**

### **1.4.1. Teórica**

Este estudio amplía el conocimiento sobre el Script Concordance Test al evaluar su eficacia en internos de medicina de una universidad privada en Lima, Perú. Se adapta su aplicación al contexto académico y sanitario peruano, considerando las particularidades del sistema de salud y las necesidades formativas de los futuros médicos en escenarios exigentes. Así su aplicación en español es especialmente relevante, dado que existe poca evidencia sobre su uso en países hispanohablantes, lo que limita su impacto en la educación médica<sup>14-16</sup>.

Además, la controversia por los resultados de la ENCAPS 2024 ha evidenciado la necesidad de evaluaciones más precisas en la formación médica. En este contexto, el SCT se presenta como una herramienta objetiva para medir el razonamiento clínico y la toma de decisiones en entornos de alta demanda. Su integración de teoría y práctica permite evaluar con mayor profundidad la capacidad de los estudiantes para resolver problemas complejos. Este estudio busca validar su uso y consolidarlo como un modelo replicable en la educación médica y exámenes nacionales, mejorando la preparación de futuros médicos, especialmente en atención primaria y zonas rurales con acceso limitado a especialistas.

Además, este estudio introduce el uso de ChatGPT en la creación de casos clínicos para el SCT, explorando el potencial de la inteligencia artificial como un recurso innovador en la educación médica. La IA permite generar escenarios clínicos variados, dinámicos y adaptados a diferentes niveles de complejidad, enriqueciendo el proceso de evaluación y brindando a los estudiantes experiencias más realistas<sup>17-18</sup>. Su implementación no solo optimiza la elaboración de casos, sino que también sienta las bases para futuras investigaciones sobre su aplicación en la enseñanza y evaluación médica.

En este contexto, la combinación del SCT con inteligencia artificial no solo representa un avance metodológico, sino también una oportunidad para mejorar la calidad de la educación médica y, en consecuencia, la atención en salud. Al establecer estándares más precisos para medir el razonamiento clínico, este estudio contribuye a la formación de médicos mejor preparados para responder a los desafíos del ejercicio profesional.

#### **1.4.2 Metodológica**

En el aspecto metodológico de este estudio se fundamenta en la implementación del Script Concordance Test como una herramienta validada y confiable para evaluar el razonamiento clínico. El SCT ha demostrado ser eficaz en diversos contextos educativos y clínicos, proporcionando una evaluación precisa del razonamiento clínico en escenarios complejos y con información incompleta, lo cual es particularmente relevante para la atención primaria y en el contexto de la práctica médica en el SERUMS. Este estudio cumple con todos los criterios pertinentes para su realización, asegurando su relevancia y aplicabilidad en la formación de internos de medicina. Además, el SCT será validado mediante consultas con médicos que hayan terminado el SERUMS y con aquellos que estén por terminarlo, y luego revisado por un panel de expertos para garantizar la mayor calidad en su aplicación. El SCT puede ser administrado de manera virtual, lo que facilita su implementación en diferentes

entornos, haciéndolo una herramienta accesible sin comprometer la calidad de la evaluación del razonamiento clínico.

En este contexto, la incorporación de ChatGPT para la formulación de las viñetas clínicas no plantea problemas, ya que su uso será con fines investigativos y no calificadores, siendo además evaluado por expertos, optimizando el proceso de generación de casos sin afectar la validez del SCT.<sup>17</sup> Esta metodología, además, es replicable en diferentes ámbitos educacionales gracias a su accesibilidad y flexibilidad. Por último, el diseño del estudio al ser no experimental con enfoque cuantitativo, es adecuado para analizar el razonamiento clínico en situaciones reales, y utilizará técnicas rigurosas de recolección e interpretación de datos, manteniendo altos estándares éticos en la investigación. Esta metodología no solo garantizará resultados válidos y confiables, sino que también ofrecerá un marco replicable para futuras investigaciones, contribuyendo al mejoramiento continuo de la educación en el área de la salud en Perú.

#### **1.4.3. Práctica**

Este estudio se justifica en el ámbito práctico porque sus resultados pueden ser aplicados para resolver problemas específicos y mejorar la situación actual de la formación médica de los internos de medicina en Perú, quienes luego serán fundamentales en la atención primaria, especialmente en áreas rurales y de difícil acceso. Al implementar el Script Concordance Test para evaluar el razonamiento clínico, la investigación responde a la necesidad de fortalecer la capacidad de los médicos en formación para tomar decisiones clínicas efectivas en contextos complejos e inciertos. Los hallazgos permitirán diseñar y adaptar programas de formación médica que directamente abordan estas carencias, mejorando así las competencias clínicas de los futuros médicos y, en consecuencia, la calidad de la atención en las comunidades más vulnerables del país.

Además, la utilización de ChatGPT para la elaboración de viñetas clínicas ofrece una metodología económica, accesible y fácilmente replicable en otros contextos educativos y clínicos. Este enfoque no solo resuelve problemas prácticos en el contexto peruano, sino que también tiene el potencial de ser adaptado a diferentes sistemas de salud, mejorando la formación médica en diversas regiones. De este modo, el estudio tiene un impacto positivo no solo en la mejora inmediata de la formación médica en Perú, sino que también establece un modelo que puede ser adoptado y adaptado en otros contextos, contribuyendo así a una educación médica más efectiva y a una formación en salud más integral.

## **1.5. Delimitaciones de la investigación**

### **1.5.1. Temporal**

La delimitación temporal del estudio abarcó el período comprendido entre febrero del 2024 a enero del 2025, centrandó la investigación en los internos de la Escuela Académico Profesional de Medicina Humana de la Universidad Norbert Wiener.

De manera general, este periodo de meses aseguró la obtención de datos dentro de un contexto claramente definido, permitiendo una evaluación detallada de los factores asociados al razonamiento clínico en este grupo específico de internos, lo que garantizó que las conclusiones fueran pertinentes y aplicables al contexto actual de la formación médica en Perú.

### **1.5.2. Espacial**

La delimitación espacial del estudio fue presencial para los internos, realizándose en las instalaciones de la universidad privada Norbert Wiener en Lima, Perú. Esta se llevó a cabo en coordinación con la Escuela Académico Profesional de Medicina Humana, aprovechando la ocasión en que los internos participaron en un examen presencial organizado por la misma

institución, aplicándose el Test de Concordancia de Script inmediatamente después de finalizar el examen regular.

El grupo de consultores y panel de expertos participó en modalidad híbrida, lo que permitió la inclusión de expertos tanto de manera presencial como remota, facilitando la participación sin limitaciones geográficas. Esta estrategia garantizó la recolección de datos de manera precisa y detallada en un entorno controlado para los internos, al tiempo que ofreció flexibilidad para la evaluación por parte de los expertos, proporcionando una visión integral y representativa en la evaluación de los factores asociados al razonamiento clínico.

### **1.5.3. Recursos**

La delimitación de recursos para este estudio incluyó la inversión en recursos humanos, específicamente en un panel de expertos que fue financiado por el autor. Además, se dispuso de los fondos necesarios para cubrir los costos relacionados con la obtención de datos y el procesamiento estadístico. La ejecución e incorporación de la muestra para el estudio contó con el apoyo de la Universidad Privada Norbert Wiener, destacando su accesibilidad y espacios. Dado que la prueba se estructuró y aplicó de forma híbrida, los gastos en materiales físicos fueron mínimos, lo que aseguró un manejo eficiente de los recursos a lo largo de la investigación.

### **1.6. Limitaciones de la investigación**

El presente estudio tiene algunas limitaciones que deben considerarse al interpretar sus resultados. El tamaño de la muestra fue menor al esperado, ya que participaron 87 de los 123 internos, lo que puede afectar la representatividad de los hallazgos. Además, al haberse realizado en una única universidad privada en Lima, se limita la posibilidad de generalización a otras instituciones con enfoques distintos. El diseño transversal impidió evaluar la evolución del razonamiento clínico a lo largo del internado. Aunque hay evidencia

de que el desempeño en el SCT mejora entre años académicos, se dispone de menos información sobre cambios dentro del mismo año de formación. Hubo diferencias en la aplicación del SCT, ya que los internos lo realizaron presencialmente y el panel de expertos de manera virtual, lo que pudo generar variaciones en concentración, tiempo de respuesta o interpretación de los casos clínicos. La generación automatizada de ítems del SCT presentó errores como el uso de pruebas diagnósticas inadecuadas, redundancia de conceptos, desbalances en la dificultad de los ítems y escenarios clínicos alejados del contexto de los internos, lo que requirió modificaciones manuales.

También se identificaron posibles sesgos. El sesgo de selección limita la generalización de los resultados al haberse evaluado solo una universidad. Además, el sesgo de fatiga pudo haber afectado el rendimiento de los internos en los ítems finales debido a la extensión del test y el nivel de concentración requerido.

## **CAPITULO II: MARCO TEORICO**

### **2.1. Antecedentes**

#### **Antecedentes internacionales**

Kiyak et al. (2024) llevaron a cabo un estudio descriptivo observacional titulado *"A Prompt for Generating Script Concordance Test Using ChatGPT, Claude, and Llama Large Language Model Chatbots"*, con el objetivo de "proponer un prompt genérico para la generación de ítems del Script Concordance Test utilizando modelos de lenguaje grandes (LLMs) como ChatGPT, Claude, y Llama, que pueda adaptarse a diversas áreas de la medicina y diferentes etapas de la educación médica". El estudio desarrolló un prompt genérico basado en estudios previos y lo aplicó en cuatro modelos de lenguaje grandes para generar ítems de SCT, evaluando su capacidad para crear casos clínicos centrados en el razonamiento clínico bajo condiciones de incertidumbre. Los resultados mostraron que el prompt es efectivo para generar ítems de SCT en cuestión de segundos, aunque se observaron limitaciones, como la necesidad de revisión manual por expertos para asegurar la precisión clínica. Los autores concluyeron que los LLMs pueden generar ítems de SCT de manera rápida y efectiva, aunque es necesaria la validación y revisión por expertos para asegurar la calidad de los ítems generados.<sup>19</sup>

Hudon et al. (2024) realizaron un estudio de métodos mixtos, titulado *"Using ChatGPT in Psychiatry to Design Script Concordance Tests in Undergraduate Medical Education: Mixed Methods Study"*, un artículo original realizado en Quebec, Canadá, con el objetivo de "explorar cómo los Script Concordance Tests generados por ChatGPT se comparan con los

SCTs producidos por expertos clínicos en tres aspectos clave: el escenario, las preguntas clínicas y la opinión experta". La metodología incluyó la evaluación de seis SCTs (tres generados por ChatGPT y tres por expertos) por 102 participantes, utilizando un marco conceptual predefinido. Los resultados mostraron que no hubo diferencias significativas en los escenarios ( $P=.84$ ), preguntas clínicas ( $P=.99$ ), y opinión experta ( $P=.07$ ) entre los SCTs generados por ChatGPT y los creados por expertos. Los autores concluyeron que ChatGPT es una herramienta prometedora para la creación de materiales educativos, aunque es necesaria una mayor validación para garantizar la precisión y eficacia educativa.<sup>18</sup>

Nalini et al. (2024) llevaron a cabo un estudio observacional descriptivo titulado "*Comparison between MCQ, Extended matching questions (EMQ) and Script concordance test for assessment among first-year medical students – A pilot study*", un artículo original realizado en un colegio médico en el sureste de India, con el objetivo de "evaluar la viabilidad de utilizar EMQ y SCT para la evaluación formativa en estudiantes de primer año de medicina y comparar los puntajes de los estudiantes en un multiple-choice question (MCQ) con los de EMQ y SCT en los temas de etiología, manejo y fisiopatología del shock". La metodología incluyó un análisis de 52 estudiantes de primer año que tomaron una evaluación formativa compuesta por 20 MCQs, 6 EMQs, y 2 SCTs, después de una intervención educativa basada en simulación. Los resultados mostraron que los puntajes promedio en MCQ fueron significativamente más altos que en EMQ y SCT ( $P=0.0018$  y  $P=0.00325$ , respectivamente), con un promedio de  $55.6\pm 14.6$  para MCQ,  $41.9\pm 26$  para EMQ, y  $42.6\pm 17.3$  para SCT. Además, bajo el tema de manejo, el puntaje promedio en MCQ fue  $56.4\pm 29.3$  comparado con  $30.1\pm 29.7$  en EMQ ( $P<0.001$ ). La mayoría de los estudiantes consideró que las EMQ evaluaban mejor el contenido impartido, mientras que las SCT evaluaban mejor la aplicación clínica. Los autores sugieren que, aunque los estudiantes obtuvieron mejores resultados en herramientas de evaluación familiares como el MCQ, las

nuevas herramientas como EMQ y SCT son más efectivas para evaluar habilidades de razonamiento clínico, aunque requieren una mayor familiarización y tiempo de ajuste para los estudiantes.<sup>20</sup>

Naylor et al. (2024) realizaron un estudio de cohortes prospectivo titulado *"The Impact of Script Concordance Testing on Clinical Decision-Making in Paramedic Education"*, un artículo original llevado a cabo en la Medical University of Lublin, Polonia. El objetivo del estudio fue "investigar la efectividad del Script Concordance Test en mejorar las habilidades de razonamiento clínico en la educación de paramédicos". La metodología incluyó la aplicación del SCT a dos cohortes de estudiantes de paramédico antes y después de un curso de primeros auxilios especializados, evaluando el impacto del curso sobre sus habilidades de toma de decisiones clínicas. Los resultados indicaron una diferencia media del 14% entre los puntajes del SCT y los MCQ, con las calificaciones en MCQ siendo en promedio un 14% más altas ( $76.2 \pm 12.3$  vs.  $62.2 \pm 14.8$ ,  $P < 0.001$ ). El análisis de fiabilidad mostró un ICC de 0.90 antes del curso y 0.45 después del curso, lo que sugiere una moderada consistencia en los resultados del SCT. El estudio subraya que, aunque el SCT no mostró una correlación significativa con los MCQ, fue bien recibido por los estudiantes y es considerado una herramienta valiosa para mejorar el razonamiento clínico.<sup>21</sup>

Ross et al. (2023) realizaron un estudio de cohortes observacional titulado *"Clinical reasoning in undergraduate paramedicine: utilisation of a script concordance test"*, un artículo original llevado a cabo en Monash University, Australia. El objetivo del estudio fue "determinar si el Test de Concordancia de Guiones es una opción viable y confiable para evaluar el razonamiento clínico en estudiantes de paramedicina de pregrado". La metodología incluyó la administración del SCT a 82 estudiantes de tercer año y la comparación de sus resultados con un panel de referencia compuesto por 14 paramédicos experimentados. Los resultados mostraron que el puntaje medio de los estudiantes fue

significativamente menor que el del panel de referencia ( $65.6 \pm 8.4$  frente a  $80 \pm 5$ ,  $P < 0.001$ ). El SCT demostró una buena consistencia interna con un alfa de Cronbach de 0.79. Además, el 59% de los estudiantes consideró que el SCT reflejaba adecuadamente el razonamiento clínico. Los autores señalaron que, aunque el SCT es una herramienta útil para evaluar el razonamiento clínico, se requiere más investigación para optimizar las técnicas pedagógicas que mejoren esta habilidad en los estudiantes de paramedicina.<sup>22</sup>

Redmond et al. (2022) realizaron un estudio transversal titulado "*Development and validation of a script concordance test to assess biosciences clinical reasoning skills: A cross-sectional study of 1st year undergraduate nursing students*", un artículo original llevado a cabo en tres universidades de habla inglesa: Irlanda, Reino Unido y Australia. El objetivo del estudio fue "desarrollar y validar un Test de Concordancia de Guiones para evaluar las habilidades de razonamiento clínico en ciencias biosanitarias en estudiantes de primer año de enfermería". La metodología incluyó la construcción de un SCT con 36 ítems relacionados con los sistemas cardiovascular, digestivo, renal y respiratorio, que fue administrado a 47 estudiantes de primer año de enfermería. Los resultados mostraron que el SCT tenía una alta consistencia interna, con un alfa de Cronbach de 0.83. El puntaje medio fue significativamente menor en los estudiantes ( $50.78 \pm 8.89$ ) en comparación con los expertos ( $79.8 \pm 12.17$ ), con una diferencia significativa ( $P < 0.0005$ ,  $d = 3$ ). Los autores destacan que el SCT es un método innovador, fiable y válido para evaluar las habilidades de razonamiento clínico en estudiantes de enfermería, aunque identifican áreas de mejora, como la claridad de las instrucciones y la estructura de las preguntas.<sup>23</sup>

Iglesias Gómez et al. (2022) realizaron un estudio observacional descriptivo titulado "*Clinical reasoning evaluation using script concordance test in primary care residents*", un artículo original llevado a cabo en el Centro de Salud Alcantarilla-Sangonera, Servicio Murciano de Salud, Murcia, España. El objetivo del estudio fue "diseñar y validar un Test

de Concordancia de Guiones para la evaluación del razonamiento clínico en residentes de atención primaria pediátrica". La metodología incluyó la elaboración, corrección y validación del SCT con 115 ítems, de los cuales se redujeron a 67 ítems tras la validación por un panel de 5 expertos. Se aplicó el SCT a 102 residentes de pediatría y 43 expertos en atención primaria, con los resultados mostrando que los expertos obtuvieron un puntaje medio de 77.3 (DE 5.4) frente a los residentes que obtuvieron un promedio de 71.7 (DE 7.8), con una diferencia significativa ( $P < 0.00005$ ). El alfa de Cronbach para la consistencia interna fue de 0.74. Los autores concluyen que el SCT diseñado cumple con los criterios de fiabilidad y validez para discriminar entre residentes y expertos, y sugieren su uso para complementar la evaluación del razonamiento clínico en la formación de residentes.<sup>24</sup>

Kün-Darbois et al. (2022) realizaron un estudio prospectivo observacional titulado "*Script concordance test acceptability and utility for assessing medical students' clinical reasoning: a user's survey and an institutional prospective evaluation of students' scores*", un artículo original llevado a cabo en la Facultad de Medicina de la Universidad de Angers, Francia. El objetivo del estudio fue "evaluar la aceptación y utilidad del Test de Concordancia de Guiones en la evaluación del razonamiento clínico de los estudiantes de medicina, y comparar las puntuaciones obtenidas con SCT frente a las modalidades de evaluación estándar". La metodología incluyó una encuesta en línea dirigida a 596 estudiantes de medicina y 41 profesores, con una tasa de respuesta del 33%, y la comparación de las puntuaciones de los estudiantes en tres pruebas sucesivas de SCT y Casos Clínicos Progresivos (PCC) durante tres años. Los resultados mostraron que las puntuaciones de los PCC aumentaron significativamente cada año ( $P < 0.001$ ), mientras que las puntuaciones de SCT solo aumentaron entre la primera y segunda prueba ( $P = 0.004$ ), y se mantuvieron estables después ( $P = 0.770$ ). Además, las puntuaciones de PCC fueron consistentemente más altas que las de SCT en la segunda y tercera prueba ( $P < 0.001$ ). Los autores destacaron que,

a pesar de la resistencia inicial tanto de estudiantes como de profesores, el SCT podría ser una herramienta valiosa para la formación médica avanzada, aunque se requiere una mayor capacitación para mejorar la adhesión.<sup>8</sup>

Moghadami et al. (2021) realizaron un ensayo controlado aleatorizado titulado *"Teaching clinical reasoning to undergraduate medical students by illness script method: a randomized controlled trial"*, un artículo original llevado a cabo en la Facultad de Medicina de la Universidad de Ciencias Médicas de Shiraz, Irán. El objetivo del estudio fue "evaluar la efectividad del método de guiones de enfermedades para enseñar habilidades de razonamiento clínico en estudiantes de medicina de pregrado". La metodología incluyó 100 estudiantes de cuarto año, divididos en un grupo de intervención y un grupo de control. Los estudiantes del grupo de intervención participaron en un taller basado en guiones de enfermedades, mientras que el grupo de control recibió enseñanza tradicional. Los resultados mostraron que el grupo de intervención obtuvo puntajes significativamente más altos en las pruebas posteriores al taller en comparación con el grupo de control, tanto en el conocimiento ( $15.74 \pm 2.47$  frente a  $14.38 \pm 2.59$ ,  $P = 0.009$ ) como en el Test de Concordancia de Guiones ( $6.12 \pm 1.95$  frente a  $4.54 \pm 1.56$ ,  $P = 0.0001$ ). Los autores concluyen que el método de guiones de enfermedades es eficaz para mejorar las habilidades de razonamiento clínico en estudiantes de medicina, y sugieren su implementación en los planes de estudio de educación médica.<sup>25</sup>

Nomura et al. (2021) realizaron un estudio transversal titulado *"Creating Clinical Reasoning Assessment Tools in Different Languages: Adaptation of the Pediatric Emergency Medicine Script Concordance Test to Japanese"*, un artículo original llevado a cabo en Japón, específicamente en Hirosaki University, el National Center for Child Health and Development, y el Tokyo Metropolitan Children's Medical Center. El objetivo del estudio fue "desarrollar y validar una versión japonesa del Test de Concordancia de Guiones en

Medicina de Emergencias Pediátricas (Jpem-SCT) para evaluar las habilidades de razonamiento clínico de los médicos en formación". La metodología incluyó la adaptación del SCT original al contexto japonés mediante un proceso de traducción y validación que involucró a 75 médicos residentes. Los resultados mostraron que el puntaje promedio de los participantes fue de 68.6% (DE 9.8), con un alfa de Cronbach de 0.70, lo que indica una fiabilidad adecuada. Los autores concluyen que la versión japonesa del SCT es un instrumento fiable y válido para evaluar el razonamiento clínico en emergencias pediátricas y puede ser utilizado en la educación clínica en Japón.<sup>13</sup>

Cohen Aubart et al. (2020) realizaron un estudio comparativo multicéntrico titulado "*Are Script Concordance Tests Suitable for the Assessment of Undergraduate Students? A Multicenter Comparative Study*", un artículo original realizado en varias instituciones médicas en Francia, incluyendo el Hospital Pitié-Salpêtrière y el Hospital Bichat. El objetivo del estudio fue "evaluar la idoneidad de los Test de Concordancia de Guiones para la evaluación de estudiantes de medicina de pregrado". La metodología incluyó la implementación de 20 casos de SCT en estudiantes de dos programas de medicina y la comparación de los resultados con los obtenidos en preguntas de opción múltiple (MCQ). Los resultados mostraron que 551 de los 883 estudiantes (62%) completaron los SCTs, obteniendo un puntaje medio agregado de 11.54 (DE 3.29) sobre 20. La correlación entre los resultados de SCT y MCQ fue significativa, y el 86% de los estudiantes consideró que los SCTs eran más difíciles que los MCQs, aunque no hubo diferencias significativas en los puntajes medios entre ambos tipos de prueba. Los autores concluyen que los SCTs son una opción viable para la evaluación de estudiantes de medicina, aunque su implementación puede ser percibida como más desafiante por los estudiantes.<sup>26</sup>

Peyrony et al. (2020) realizaron un estudio observacional multicéntrico titulado "*Impact of Panelists' Experience on Script Concordance Test Scores of Medical Students*", un artículo

original llevado a cabo en varias universidades en Francia, incluyendo Paris Diderot y Paris Descartes. El objetivo del estudio fue "comparar los puntajes del Test de Concordancia de Guiones de estudiantes de medicina según el nivel de experiencia del panel de expertos". La metodología incluyó la administración de un SCT de 30 ítems a 935 estudiantes de nueve universidades francesas, con un panel de referencia compuesto por 75 expertos divididos en tres grupos según su experiencia: residentes, médicos no experimentados, y médicos experimentados. Los resultados mostraron que los puntajes de SCT de los estudiantes fueron más bajos cuando el panel de expertos estaba compuesto por médicos no experimentados en comparación con los residentes (67.1 vs 69.1,  $p < 0.0001$ ) y cuando el panel estaba compuesto por médicos experimentados en comparación con los no experimentados (65.4 vs 67.1,  $p < 0.0001$ ). El alfa de Cronbach varió entre 0.43 y 0.53, lo que indica una consistencia interna aceptable. Los autores sugieren que la experiencia de los expertos tiene un impacto limitado en los puntajes de SCT de los estudiantes, lo que permite la inclusión de expertos con menor experiencia en los paneles de evaluación sin comprometer la validez de la prueba.<sup>27</sup>

Ottolini et al. (2020) realizaron un estudio de métodos mixtos titulado "*Pediatric Hospitalists' Performance and Perceptions of Script Concordance Testing for Self-Assessment*", un artículo original llevado a cabo en varias instituciones en los Estados Unidos, incluyendo el Barbara Bush Children's Hospital en Maine y el Children's National Medical Center en Washington, DC. El objetivo del estudio fue "comparar los niveles de cognición expresados por pediatras al responder preguntas de Script Concordance Test versus preguntas de opción múltiple (MCQ), y analizar sus percepciones sobre la utilidad del SCT para la autoevaluación formativa". La metodología incluyó la participación de 12 pediatras que respondieron a preguntas de SCT y MCQ, utilizando un ejercicio de "pensar en voz alta" seguido de entrevistas cualitativas. Los resultados mostraron que un 74% de los

comentarios de los pediatras fueron codificados como procesos cognitivos de alto nivel al responder SCT, en comparación con solo un 19% al responder MCQ (chi-cuadrado  $P < 0.00001$ ). Los autores concluyeron que las preguntas de SCT son más efectivas para estimular el razonamiento clínico y la toma de decisiones bajo condiciones de incertidumbre, lo que las hace útiles para la práctica deliberada en la medicina hospitalaria pediátrica.<sup>28</sup>

Sadeghi et al. (2019) realizaron un estudio observacional multicéntrico titulado "*Combination of Different Clinical Reasoning Tests in a National Exam*", un artículo original llevado a cabo en el marco de la novena Olimpiada Médica Nacional en Irán, organizada en la Universidad de Ciencias Médicas Shahid Beheshti en Teherán, bajo la supervisión directa del Ministerio de Salud y Educación Médica. El objetivo del estudio fue "evaluar la confiabilidad de la combinación de diferentes pruebas de razonamiento clínico, incluyendo el Test de Concordancia de Guiones, el Key Features (KF), el Clinical Reasoning Problem (CRP), y el Comprehensive Integrative Puzzle (CIP) en un examen nacional". La metodología incluyó la participación de 339 estudiantes de medicina de alto rendimiento de 60 escuelas médicas, quienes fueron evaluados con una combinación de las cuatro pruebas mencionadas. Los resultados mostraron que la confiabilidad combinada de las pruebas, medida mediante el alfa de Cronbach, fue de 0.815, con los valores de confiabilidad individuales de 0.76 para el SCT, 0.81 para el KF, 0.80 para el CRP y 0.92 para el CIP. El puntaje total promedio fue de  $169.92 \pm 41.54$  de un máximo de 240 puntos. Los autores concluyen que la combinación de diferentes pruebas de razonamiento clínico es un método confiable para evaluar esta habilidad en exámenes de alta exigencia, y recomiendan su uso en escuelas médicas para exámenes nacionales.<sup>29</sup>

Lineberry et al. (2019) realizaron un estudio prospectivo titulado "*Experts' responses in script concordance tests: a response process validity investigation*", un artículo original llevado a cabo en varias instituciones, incluyendo la Universidad de Kansas y la Universidad

de Illinois en Chicago, Estados Unidos. El objetivo del estudio fue "explorar los procesos de respuesta de los expertos en un panel de calificación de un Test de Concordancia de Guiones para comprender mejor sus divergentes creencias sobre cómo nuevos datos clínicos alteran la idoneidad de las acciones propuestas dentro de casos simulados de pacientes". La metodología incluyó la participación de 10 gastroenterólogos argentinos que sirvieron como panel de expertos en un SCT existente, quienes recontestaron 15 casos nueve meses después de su participación inicial y luego respondieron preguntas que exploraban su razonamiento. Los resultados mostraron que la consistencia de las respuestas de los expertos a lo largo del tiempo fue limitada, con un kappa ponderado lineal de 0.33 (IC del 95%: 0.21-0.46), lo que indica una fiabilidad moderada. Los autores concluyen que los procesos de respuesta de los expertos no fueron consistentes con la interpretación clásica de los puntajes de SCT, lo que sugiere la necesidad de reconsiderar cómo se utilizan estos puntajes para guiar evaluaciones en el aprendizaje.<sup>30</sup>

Karila et al. (2017) realizaron un artículo de revisión titulado "*Le test de concordance de script: un outil pédagogique multimodal*", llevado a cabo en varias instituciones de la Université Paris Sud y el Hôpital Universitaire Paul-Brousse en Francia. El objetivo del estudio fue "hacer una revisión específica de la literatura sobre las bases teóricas, la metodología y la utilización del Test de Concordancia de Guiones en el campo de la salud". La metodología incluyó una revisión de 62 artículos seleccionados mediante bases de datos como PubMed, EMBASE, y PsycINFO, utilizando palabras clave en inglés y francés relacionadas con el SCT. Los resultados destacaron que el SCT es una herramienta pedagógica fiable y válida que permite discriminar los niveles de práctica entre estudiantes, internos y médicos especialistas, con un coeficiente alfa de Cronbach superior a 0.70 en varias aplicaciones. Los autores concluyeron que el SCT es un instrumento valioso para evaluar el razonamiento clínico en situaciones de incertidumbre y recomiendan su

integración en la formación médica en Francia a partir de 2017, especialmente para las Épreuves Classantes Nationales informatisées (ECNi) que comenzarán en 2020.<sup>9</sup>

Subra et al. (2017) realizaron un estudio longitudinal observacional titulado "*Reliability and validity of the script concordance test for postgraduate students of general practice*", un artículo original llevado a cabo en la Facultad de Medicina de Toulouse, Francia. El objetivo del estudio fue "demostrar la fiabilidad y validez del Test de Concordancia de Guiones en el contexto de la medicina general y evaluar si esta herramienta pudiera ser utilizada para monitorear el progreso en el razonamiento clínico de los estudiantes de posgrado". La metodología incluyó la administración de un SCT de 135 ítems a 90 estudiantes al inicio y cada seis meses durante tres años de su formación, comparando sus resultados con los de un panel de 20 expertos en medicina general. Los resultados mostraron que el puntaje promedio de los expertos fue significativamente más alto que el de los estudiantes (76.7 frente a 68.9-73.1,  $p < 0.001$ ) en todas las evaluaciones, y el alfa de Cronbach varió entre 0.65 y 0.83, indicando una fiabilidad aceptable. Los autores concluyen que el SCT es una herramienta fiable y válida para evaluar el razonamiento clínico en medicina general, sugiriendo su implementación en los programas de formación para monitorear el desarrollo de competencias clínicas.<sup>31</sup>

### **Antecedentes regionales**

Echeverry-Díaz y Restrepo-Escobar (2023) realizaron un estudio observacional titulado "*Prueba de razonamiento clínico script: Una mirada desde la práctica educativa en medicina interna*", con el objetivo de "explorar el uso y efectividad del Test de Concordancia de Guiones en la evaluación del razonamiento clínico en residentes de medicina interna". La metodología incluyó la aplicación del SCT a una muestra de 20 residentes de medicina interna en el Hospital Universitario de La Samaritana en Bogotá, Colombia, comparando sus

respuestas con un panel de 5 expertos en la misma especialidad. Se diseñaron 10 casos clínicos específicos, y los resultados mostraron una concordancia promedio de 65% entre los residentes y el panel de expertos, con un rango de variabilidad significativo entre los especialistas (coeficiente de variación del 18%). Además, se observó que solo el 40% de los residentes alcanzó un nivel de concordancia considerado aceptable (70% o más), mientras que la desviación estándar de los puntajes de los residentes fue de 12 puntos. Los autores concluyeron que, aunque el SCT es un método prometedor para evaluar el razonamiento clínico, su aplicación presenta retos significativos, especialmente en la construcción y validación de casos clínicos, sugiriendo la necesidad de estudios adicionales para optimizar su uso en la educación médica.<sup>33</sup>

Hamui et al. (2018) realizaron un estudio observacional descriptivo titulado "*Prueba de concordancia de script: primera experiencia en pediatría de alcance nacional*", un artículo original llevado a cabo en varias instituciones en Argentina, incluyendo la Sociedad Argentina de Pediatría y el Hospital de Pediatría SAMIC "Prof. Dr. Juan Pedro Garrahan" en Buenos Aires. El objetivo del estudio fue "describir la primera experiencia de aplicación de un Test de Concordancia de Guiones a nivel nacional en residentes de Pediatría, utilizando internet en tiempo real". La metodología incluyó la participación de 268 residentes de tercer año, de 56 sedes, quienes completaron el examen en un promedio de  $46.1 \pm 27.1$  minutos, obteniendo un puntaje promedio de  $65.3 \pm 7.47$  sobre 100 puntos. La confiabilidad del instrumento, medida mediante el coeficiente alfa de Cronbach, fue de 0.49, indicando una fiabilidad limitada. Los autores concluyen que, a pesar de las limitaciones en la confiabilidad, la experiencia mostró la factibilidad y aplicabilidad del SCT en la formación de posgrado en Pediatría a nivel nacional en Argentina.<sup>14</sup>

Roberti et al. (2016) realizaron un estudio cuantitativo, analítico y descriptivo titulado "*Script concordance test in medical schools in Brazil: possibilities and limitations*", un

artículo original llevado a cabo en la Escuela de Medicina de la Universidad Federal de Goiás (UFG), en Goiânia, Brasil. El objetivo del estudio fue "analizar la aplicación del Test de Concordancia de Guiones entre los estudiantes de medicina en diferentes etapas de su formación". La metodología incluyó la participación de 159 estudiantes, quienes completaron un SCT compuesto por 10 casos clínicos en medicina interna, con cinco ítems cada uno, utilizando una escala Likert de cinco puntos. Los resultados mostraron que los estudiantes en las fases preclínica y clínica obtuvieron puntajes promedio del 51.6% y 63.4% del puntaje máximo posible, respectivamente. El alfa de Cronbach del instrumento fue de 0.64, indicando una fiabilidad limitada. Los autores concluyen que la construcción y aplicación del SCT es difícil y puede no ser viable en unidades con recursos limitados, sugiriendo que estas dificultades podrían limitar su uso como método de evaluación.<sup>15</sup>

## **2.2. Bases teóricas**

### **2.2.1. Razonamiento clínico**

#### **Concepto y terminologías**

El razonamiento clínico es esencial para la práctica médica, actuando como su columna vertebral.<sup>34,35</sup> Según el enfoque teórico adoptado, el concepto de 'razonamiento clínico' puede recibir distintas definiciones.<sup>36</sup> Su importancia radica en su función orientadora, que abarca desde el diagnóstico y análisis de casos clínicos hasta la elección de intervenciones terapéuticas y la estimación del pronóstico del paciente.<sup>37</sup>

En la literatura, el término 'razonamiento clínico' ha sido interpretado y definido de múltiples formas, reflejando la complejidad inherente al proceso de toma de decisiones en contextos clínicos. Young et al.,<sup>5</sup> en su investigación, identificaron 110 términos diferentes utilizados para describir el razonamiento clínico, lo que subraya la diversidad conceptual en torno a

este aspecto crucial de la práctica profesional. Para facilitar la comprensión de esta diversidad, dichos términos se han clasificado en seis categorías generales:

1. **Habilidades de razonamiento:** Se refiere a las capacidades necesarias para realizar un razonamiento clínico efectivo, abarcando términos como habilidades clínicas y habilidades cognitivas.
2. **Rendimiento del razonamiento:** Relacionado con los objetivos que se buscan alcanzar mediante el razonamiento clínico, utilizando términos como competencia, perspicacia o pericia.
3. **Proceso de razonamiento:** Centrado en el "cómo" se lleva a cabo el razonamiento clínico, describiendo los procesos y métodos que lo componen, como el razonamiento analítico, la intuición o las heurísticas.
4. **Resultado del razonamiento:** Enfocado en el "qué" se obtiene de un proceso de razonamiento, incluyendo los resultados como un diagnóstico o un plan de tratamiento, así como la calidad de estos resultados, por ejemplo, precisión y calidad, y también los posibles errores o fallos, como sesgos o errores.
5. **Contexto del razonamiento:** Se refiere al lugar o entorno en el que se lleva a cabo el razonamiento, más allá de la mente del clínico. Incluye los factores que pueden afectar ese proceso, tales como los métodos participativos, la toma de decisiones colaborativa y la conciencia situacional, que contempla las influencias del entorno o las circunstancias en la cognición.
6. **Finalidad/objetivo del razonamiento:** Se enfoca en el "por qué" del razonamiento clínico, es decir, las razones para tratar al paciente, determinar un tratamiento o proponer un diagnóstico.

### **Teoría Cognitiva: procesos duales**

En medicina, la toma de decisiones clínicas precisa y efectiva es fundamental para el desempeño profesional. Este proceso se basa en modelos de razonamiento cognitivo, como la teoría del razonamiento dual de Daniel Kahneman y Amos Tversky, que distingue entre dos sistemas de pensamiento: el Sistema 1, rápido e intuitivo, basado en el reconocimiento de patrones, y el Sistema 2, más analítico y deliberado.<sup>38</sup>

El Sistema 1 opera en áreas cerebrales primitivas, como la corteza insular, vinculadas a las emociones.<sup>39</sup> Especialidades como la radiología y la dermatología dependen de este mecanismo, como se evidenció en estudios de fMRI, donde los radiólogos activaron regiones cerebrales profundas al interpretar imágenes rápidamente.<sup>40</sup>

Por otro lado, el Sistema 2 emplea un enfoque hipotético-deductivo, analizando signos, síntomas y hallazgos clínicos para alcanzar un diagnóstico.<sup>41</sup> Es más frecuente en médicos inexpertos y estudiantes, mientras que los expertos alternan entre ambos sistemas según la complejidad del caso.<sup>1</sup>

### **Modelos de razonamiento clínico**

Los modelos de razonamiento clínico se moldean según las particularidades de los profesionales que los utilizan. Con frecuencia, estos modelos están impregnados de estrategias y significados que se alinean con los valores, intereses, y expectativas de los clínicos, adaptándose a la especialidad y el entorno en el que ejercen su práctica.<sup>36</sup>

- **Modelo de razonamiento diagnóstico**

El modelo biomédico postula que las teorías científicas pueden aplicarse de manera directa en la práctica médica. Este enfoque se centra en identificar la patología del paciente y vincularla con los sistemas corporales que no están funcionando correctamente.<sup>42</sup>

- **Modelo de razonamiento clínico propuesto por de Charlin**

Este modelo busca descomponer y simplificar los complejos procesos cognitivos que tienen lugar en la mente del terapeuta.<sup>43</sup> Se fundamenta principalmente en la teoría de los Illness scripts y en la investigación sobre calificadores semánticos, ambas previamente discutidas.<sup>44,45</sup>

- **Modelo de razonamiento clínico propuesto por Durning**

Este modelo, de desarrollo más reciente, busca integrar las teorías y avances previos en la investigación sobre el razonamiento clínico.<sup>46</sup> Se distingue por su flexibilidad al representar tanto interacciones lineales como no lineales en contextos clínicos reales. Su objetivo es capturar la complejidad inherente al razonamiento clínico y avanzar hacia la posibilidad de evaluarlo en entornos bien definidos.<sup>47</sup>

### **Componentes**

En la práctica médica, el razonamiento clínico se compone de elementos esenciales que permiten a los profesionales abordar y resolver eficazmente desafíos diagnósticos y terapéuticos. Este proceso integra desde la recopilación de datos y formulación de hipótesis hasta la representación precisa del problema clínico y la elaboración de un diagnóstico diferencial. Cada una de estas etapas, junto con la justificación diagnóstica y planificación del tratamiento, guía la toma de decisiones médicas, optimizando la calidad de la atención al paciente.<sup>48</sup> A continuación, se analizan los componentes clave de este proceso fundamental en la medicina.

- **Recopilación de información**<sup>49,50</sup>

Es el proceso mediante el cual se recopila información para formular o modificar hipótesis. De naturaleza dinámica, abarca la anamnesis, examen físico, pruebas de laboratorio e imágenes, así como la revisión de la historia clínica. También puede realizarse de forma

implícita mediante la observación. La selección de datos se basa en representaciones del conocimiento sobre la patología, como esquemas o guiones clínicos.

- **Generación de hipótesis**<sup>51,52</sup>

Proceso preliminar, analítico o no analítico, en el que el médico identifica posibles patologías que expliquen los hallazgos clínicos. La generación de hipótesis activa representaciones del conocimiento sobre la enfermedad en un ciclo continuo donde la recolección de datos y la formulación de hipótesis se retroalimentan. Por ejemplo, las hipótesis iniciales guían la obtención de nueva información, permitiendo refinarlas o plantear nuevas.

- **Representación del problema**<sup>51,53</sup>

Es una representación mental activa que abarca todos los aspectos importantes del caso, incluidos los hallazgos clínicos del paciente y las dimensiones biopsicosociales. Esta representación puede ser comunicada a través de un resumen que incorpora calificativos semánticos y los hallazgos clave del caso.

- **Diagnóstico Diferencial**<sup>54,55</sup>

Una lista de hipótesis diagnósticas que refleja las mejores clasificaciones de la representación del problema. A medida que se ajustan la confianza y la evidencia en torno a estas hipótesis, emerge un diagnóstico principal.

- **Dirigir o trabajar el diagnóstico**<sup>56</sup>

Un diagnóstico en el que la probabilidad de que un médico considere que un paciente tiene una enfermedad específica ha superado el umbral necesario para justificar la realización de pruebas adicionales o iniciar un tratamiento, aun cuando el diagnóstico no sea concluyente.

- **Justificación Diagnóstica**<sup>54,57</sup>

Es el proceso de emplear la evidencia recogida, como los hallazgos clínicos más relevantes, para identificar uno o varios diagnósticos probables y justificar su selección mediante la comparación con otras posibles alternativas diagnósticas. Esta justificación se comunica, de manera oral o escrita, cuando es necesario en un contexto social, aunque no siempre forma parte del razonamiento clínico en su fase inicial.

- **Manejo y tratamiento**<sup>56,58</sup>

Las acciones derivadas del razonamiento clínico incluyen la elaboración de un pronóstico, el manejo y tratamiento del paciente, la implementación de medidas preventivas, y el alivio de los síntomas (incluyendo la mejora de la calidad de vida), así como la justificación de estas intervenciones.

### **Razonamiento clínico en la medicina**

Existen diversas teorías sobre el razonamiento clínico en medicina, las cuales dependen del nivel de experiencia del médico<sup>59</sup>. El razonamiento hipotético-deductivo (HDR)<sup>60-61</sup> consiste en generar y evaluar hipótesis clínicas de manera iterativa<sup>62</sup>. Es común en médicos principiantes o expertos ante casos inusuales y se apoya en el "chunking", un proceso que integra fragmentos de conocimiento para optimizar el razonamiento<sup>63</sup>. Sin embargo, al inicio, es un método lento y detallado que requiere esfuerzo<sup>62</sup>.

El reconocimiento de patrones, utilizado por médicos experimentados en situaciones comunes<sup>64</sup>, permite recuperar información clínica rápidamente a partir de experiencias previas<sup>62</sup>. Cuanta mayor sea la similitud con casos anteriores, más eficiente será el proceso<sup>65</sup>.

Por último, el escaneo sistemático es una revisión meticulosa de los sistemas corporales, considerada la fase inicial del razonamiento clínico. Se basa en un protocolo estructurado para recopilar información de manera completa<sup>62</sup>.

### **Factores asociados al razonamiento clínico**

El razonamiento clínico puede estar influenciado por factores sociodemográficos, entre los cuales destaca el sexo como un predictor relevante. Estudios indican que las estudiantes de medicina tienen enfoque más detallado y exhaustivo en la identificación de características clínicas relevantes, plantean un mayor número de diagnósticos diferenciales como opciones de tratamiento, y obtienen además mejores resultados en pruebas de razonamiento clínico en comparación con sus pares masculinos. Estas diferencias sugieren que el sexo puede influir en la manera en que se aborda y organiza la información clínica, destacando un enfoque más exhaustivo y detallado en mujeres.<sup>66,67</sup>

En cuanto a los factores académicos, el antecedente de estudios previos desempeña un rol crucial en el desarrollo de las capacidades diagnósticas en el campo del razonamiento clínico.<sup>67</sup> Aquellos con formación en otras áreas suelen mostrar un rendimiento superior en las primeras etapas de la carrera, aunque esta diferencia se equilibra con el tiempo.<sup>67</sup> La experiencia en investigación, por su parte, contribuye al razonamiento clínico al fortalecer el pensamiento crítico y el uso de la medicina basada en evidencia, herramientas esenciales para la toma de decisiones informada.<sup>68</sup> Además, la recepción de mentoría académica fomenta la confianza y precisión en los juicios diagnósticos, y las prácticas clínicas extracurriculares ofrecen una exposición temprana a situaciones reales.<sup>69,70</sup> Así estos dos factores permiten a los estudiantes practicar y refinar su capacidad de razonamiento clínico en entornos controlados.<sup>69,70</sup>

Finalmente, los factores del entorno asistencial, como la supervisión médica en el internado y la retroalimentación del personal médico, son esenciales para un aprendizaje efectivo en la toma de decisiones clínicas usando el razonamiento clínico.<sup>71</sup> La supervisión cercana y la retroalimentación continua brindan la oportunidad de afinar su razonamiento clínico

mediante el ajuste constante de las habilidades.<sup>71</sup> Además, la ubicación geográfica del establecimiento de salud y el nivel de atención, especialmente en áreas rurales o con menos recursos, presentan escenarios clínicos complejos que requieren un razonamiento clínico avanzado para adaptarse a las limitaciones y responder adecuadamente a los desafíos diagnósticos en estos contextos.<sup>72</sup>

### **2.2.1.1 Modelos de Evaluación del razonamiento clínico**

Actualmente, existen diversas herramientas para evaluar el razonamiento clínico, pero la literatura al respecto está dispersa entre diferentes especialidades médicas, lo que dificulta su selección e implementación por parte de los educadores.<sup>48</sup> Además, estas evaluaciones suelen diseñarse para contextos específicos.<sup>73</sup>

Daniel et al. (2019)<sup>48</sup> revisaron 538 estudios clave, seleccionados de 14,709 registros, sobre métodos de evaluación del razonamiento clínico en estudiantes de medicina, residentes y médicos. La mayoría de estos estudios provienen de Estados Unidos, Europa y Canadá. Identificaron tres principales categorías de evaluación: en el aula, en entornos clínicos simulados y en entornos clínicos reales.

#### **Evaluaciones en el aula**

1. **Las preguntas de opción múltiple (MCQ)** presentan un escenario clínico seguido de hasta cinco respuestas posibles. Estas pueden estructurarse para seleccionar la mejor respuesta, combinar alternativas, determinar si cada opción es verdadera o falsa, o realizar emparejamientos.<sup>48</sup>

2. **Las preguntas de emparejamiento ampliadas (EMQ)** son similares a las MCQ, en cuanto también utilizan un escenario clínico, pero requieren elegir una única respuesta de una lista más extensa de opciones, aplicable a varias preguntas.<sup>74,75</sup>

3. **Preguntas de respuesta corta o larga (ensayo):** Este enfoque presenta un caso clínico seguido de una o más preguntas que solicitan respuestas en formato de texto libre, las cuales pueden variar desde unas pocas palabras hasta respuestas más elaboradas en varias oraciones.<sup>76,77</sup>

4. **Preguntas de redacción modificadas (MEQ):** Este enfoque proporciona información sobre un caso de manera cronológica.<sup>78,79</sup> Tras cada sesión, los estudiantes deben documentar una decisión en un formato de texto libre antes de continuar con la siguiente parte del caso.<sup>48</sup>

5. **Problemas de gestión de pacientes (PMP):** Estos involucran escenarios clínicos detallados donde se presentan múltiples opciones de diagnóstico y tratamiento.<sup>80,81</sup> Los estudiantes deben elegir entre diferentes alternativas de acción, y las consecuencias de esas acciones (como los resultados de un electrocardiograma) se revelan a medida que se avanza en el caso.<sup>48</sup>

6. **Exámenes de características clave (KFE):** Evalúan la toma de decisiones clínicas mediante escenarios con preguntas centradas en aspectos fundamentales del caso, como un dolor de cabeza súbito en la hemorragia subaracnoidea.<sup>82-83 48</sup>

7. **Pruebas de concordancia de guiones:** Presentan escenarios clínicos con incertidumbre para simular el razonamiento clínico.<sup>84-85</sup> Los participantes evalúan cómo ciertos factores afectan la probabilidad de una respuesta, comparando sus elecciones con las de un panel de expertos.

8. **Puzzles clínicos o integrales:** Estructurados en cuadrícula, presentan hallazgos clínicos en columnas y diagnósticos en filas.<sup>86-87</sup> Los estudiantes deben comparar elementos para identificar coincidencias y formar guiones de enfermedad para cada diagnóstico.

9. **Los mapas conceptuales** son una herramienta esquemática mediante la cual los estudiantes representan sus conocimientos en un dominio específico y la organización de esos conocimientos a través de una ilustración gráfica.<sup>88,89</sup>

10. **Los exámenes orales** son realizados por uno o varios profesores, con o sin guión, y tienen como objetivo evaluar el razonamiento clínico, la toma de decisiones y los valores profesionales.<sup>90,91</sup>

### **Evaluaciones en entornos clínicos simulados**

1. **Exámenes clínicos (OSCE):** Evalúan las habilidades clínicas mediante estaciones donde los estudiantes realizan tareas con pacientes estandarizados, observadores y notas escritas, aunque no se centran exclusivamente en el razonamiento clínico.<sup>92-93</sup>

2. **Simulación mejorada con tecnología:** Utiliza herramientas que replican escenarios clínicos, desde maniqués de alta fidelidad hasta avatares en realidad virtual, permitiendo una interacción dinámica en la evaluación.<sup>94-95</sup>

### **Evaluaciones en entornos clínicos**

1. **La observación directa,** consiste en la evaluación del desempeño clínico de los estudiantes por un observador, generalmente un docente, en entornos reales.<sup>96</sup> Se emplean herramientas como el mini-CEX, aunque no todas están enfocadas en el razonamiento clínico.<sup>96-97</sup>

2. **Las evaluaciones globales** son formularios de evaluación completados por el profesorado,<sup>98,99</sup> que representan juicios integrales del preceptor sobre el desempeño en el razonamiento clínico, basados en observaciones tanto directas como indirectas.

3. **Los PCO (Presentaciones Orales de casos)** son informes estructurados en los que los estudiantes presentan casos clínicos de manera verbal.<sup>45,100</sup> La evidencia diagnóstica y

terapéutica es evaluada en función de cómo el estudiante selecciona y organiza los datos, así como en la elaboración de la evaluación y el plan de manejo del caso.

4. **Las notas escritas** permiten estructurar la información clínica en un informe, evaluando el razonamiento clínico en secciones como el resumen del caso, diagnóstico diferencial, justificación y plan de manejo, similar a los PCO.<sup>101</sup>

5. **El CSR (Recuerdo Estimulado por Gráficos)** es un método híbrido que combina la revisión de la documentación clínica con un examen oral, donde un evaluador analiza el razonamiento del estudiante. Usa gráficos y elementos visuales para estimular el recuerdo y mejorar la toma de decisiones diagnósticas.<sup>102-103</sup>

6. **El pensamiento en voz alta** es una técnica en la que se asigna a los estudiantes una tarea específica y se les solicita que verbalicen sus pensamientos de manera no filtrada mientras realizan la tarea.<sup>104,105</sup> Esta técnica puede aplicarse simultáneamente durante la realización de la tarea o de forma diferida, justo después de completarla.

7. **El SRL-M (Microanálisis del Aprendizaje Autorregulado)** es un protocolo de entrevista estructurada que recopila información en tiempo real sobre los pensamientos, acciones y emociones de los estudiantes al enfrentar una tarea clínica.<sup>106-107</sup> Combinado con el pensamiento en voz alta, evalúa la metacognición y autorregulación del aprendizaje.

#### **2.2.1.2 Razonamiento probabilístico**

El razonamiento probabilístico, o bayesiano, es un tipo de razonamiento inductivo que aplica la ley de probabilidades condicionales de Bayes para cuantificar la incertidumbre.<sup>108</sup> En la práctica clínica, consiste en ajustar la confianza en una hipótesis inicial (a priori) a medida que se incorpora nueva información relevante, como signos clínicos, síntomas o resultados de pruebas diagnósticas. Esto transforma la hipótesis en una nueva estimación (a posteriori). Así, cuando un Test de Concordancia de Guiones (TCS) evalúa el impacto de nueva

información en una hipótesis, está aplicando un enfoque basado en el razonamiento bayesiano.<sup>109</sup>

### **2.2.2. Guiones de la enfermedad**

El concepto de "guión" fue introducido en la literatura médica por Clancey et al. (1983)<sup>110</sup> solo unos años después de que el concepto general de guión surgiera en la literatura de psicología cognitiva. La teoría de los guiones ofrece una explicación de cómo la información se almacena y se recupera en la mente humana, influenciando la interpretación que los individuos hacen de los objetos y eventos en su entorno.<sup>111</sup> Aplicada a las ciencias de la salud, esta teoría sugiere que el conocimiento médico se organiza en estructuras especializadas conocidas como "guiones", las cuales conectan información clínica y fisiopatológica relevante sobre diversas categorías diagnósticas, enfermedades específicas, o incluso sobre pacientes individuales.<sup>44</sup>

El "guión de la enfermedad" descrito por Feltovich y Barrows en 1984,<sup>112</sup> se presenta como una herramienta que conecta dos enfoques: el diagnóstico como un proceso de razonamiento y el diagnóstico como un proceso de categorización rápida. Estos guiones comienzan a formarse desde el primer encuentro clínico y, con el tiempo, se actualizan, reestructuran, adaptan, podan y refinan con la experiencia acumulada.<sup>65</sup>

El marco conceptual de la teoría de los guiones sostiene que el cerebro humano interpreta el mundo al comparar las características de los modelos mentales con las estructuras de escenas reales, verificando consistencias e inconsistencias, patrones e irregularidades.<sup>25</sup> Basándose en la teoría de los guiones, los médicos expertos realizan diagnósticos al considerar diagnósticos diferenciales mediante la comparación y contraste de características clave.<sup>25</sup> Estos expertos activan redes de conocimiento preexistentes, denominadas "scripts de enfermedades".<sup>113</sup> El método de guionización de enfermedades emplea un marco teórico que

organiza el conocimiento diagnóstico en distintas categorías, como la epidemiología, la fisiopatología de las enfermedades, los síntomas, los signos clínicos, las intervenciones, las características definatorias y las características discriminatorias, todo lo cual conduce a un diagnóstico preciso.<sup>112-114</sup>

### **2.2.3. Prueba de concordancia de guiones**

Entre los avances recientes en la evaluación del razonamiento clínico se encuentra la Prueba de Concordancia de Guiones, desarrollada a finales de la década de 1990 por Charlin et al., la SCT se basa en la teoría de los guiones y está diseñada para medir la interpretación de datos clínicos en casos mal definidos.<sup>115</sup> Tanto su formato como su sistema de puntuación son innovadores.

La SCT se emplea para evaluar la competencia en "razonamiento clínico" de los examinados en contextos de incertidumbre clínica.<sup>116</sup> Se describen breves escenarios clínicos y se proporciona información adicional paso a paso. A medida que se recibe esta nueva información, el examinado debe tomar decisiones diagnósticas, de seguimiento o terapéuticas.<sup>117</sup> Utilizando una escala Likert de 5 puntos, que va de -2 a +2, el evaluado debe indicar en qué medida la información adicional refuerza o debilita la hipótesis de enfermedad planteada en el escenario.<sup>118</sup> Los resultados de los examinados se comparan posteriormente con las evaluaciones realizadas por un panel de expertos; la respuesta considerada como "estándar de oro" es aquella que obtiene la mayor cantidad de puntos y en la que la mayoría de los expertos coincide.<sup>116</sup> A la respuesta modal se le asigna un punto completo, mientras que a las demás respuestas proporcionadas por los miembros del panel se les otorga un crédito parcial.<sup>119</sup>

El SCT ha sido diseñado con el objetivo de evaluar la capacidad de razonamiento clínico de los candidatos cuando se enfrentan a decisiones complejas, características de las tres fases

clave de la toma de decisiones clínicas: diagnóstico, investigación y tratamiento.<sup>116</sup> El enfoque del SCT es único porque combina dos características clave y valiosas<sup>120</sup>: se basa en escenarios clínicos auténticos, lo que aumenta el realismo de la evaluación mediante el uso de herramientas de alta fidelidad como videos, audios e imágenes clínicas.<sup>121</sup> A diferencia de otros métodos de evaluación que requieren un gran esfuerzo y se centran en el pensamiento de orden superior, el SCT es eficiente y viable.<sup>122</sup> Además, su aplicación en entornos virtuales está plenamente justificada en el contexto de la educación médica.<sup>9</sup>

## **Componentes**

### **1. Casos y preguntas**

Los casos, presentados como breves escenarios, plantean situaciones clínicas desafiantes resumidas en unas pocas oraciones. En algunos casos, el escenario concluye con una oración que describe explícitamente el problema a resolver o un aspecto a investigar, aunque, en la mayoría de las ocasiones, el problema está implícito. Aunque el SCT se utiliza frecuentemente para abordar cuestiones de diagnóstico, tratamiento o manejo clínico, también puede adaptarse para evaluar aspectos más específicos, como dilemas éticos o cuestiones de profesionalismo.<sup>123</sup>

El número ideal de ítems o casos necesarios para alcanzar un nivel de confiabilidad de 0.8 fue evaluado en tres estudios. Un estudio empleó la fórmula de Spearman-Brown para determinar cuántos casos eran necesarios para alcanzar una confiabilidad de 0.8, considerando que cada caso tenía uno o dos elementos; el número óptimo encontrado fue de 36 casos.<sup>121</sup> Otros dos estudios investigaron cuántos ítems eran necesarios para lograr una confiabilidad de 0.8, obteniendo resultados de 130 ítems<sup>124</sup> y entre 48 y 102 ítems,<sup>125</sup> respectivamente. Se ha demostrado que agregar más ítems en lugar de aumentar el número de casos es una estrategia más efectiva para mejorar la confiabilidad de las pruebas. Esta

opción también resulta más viable, ya que un mayor número de ítems en menos casos reduce la carga de trabajo tanto para los diseñadores de las pruebas como el tiempo de lectura requerido para los examinados.<sup>125</sup> No obstante, hay un efecto de techo cuando se superan tres o cuatro ítems por caso.<sup>119</sup>

## **2. Escala Likert en SCT**

La escala del test es uniforme, con valores negativos a la izquierda, 0 como posición neutral y valores positivos a la derecha, minimizando la ambigüedad y errores de medición. En investigación y tratamiento, términos como "contraindicado" e "indicado" resaltan la incertidumbre e introducen factores legales y de riesgo-beneficio, reflejando escenarios reales<sup>123</sup>.

Inicialmente, el SCT usaba una escala Likert de 7 puntos, pero actualmente emplea una de 5. Su puntuación tradicional pondera las respuestas en relación con la respuesta modal, con variantes que consideran el número total de respuestas del panel o la distancia respecto a la respuesta modal o media<sup>119</sup>.

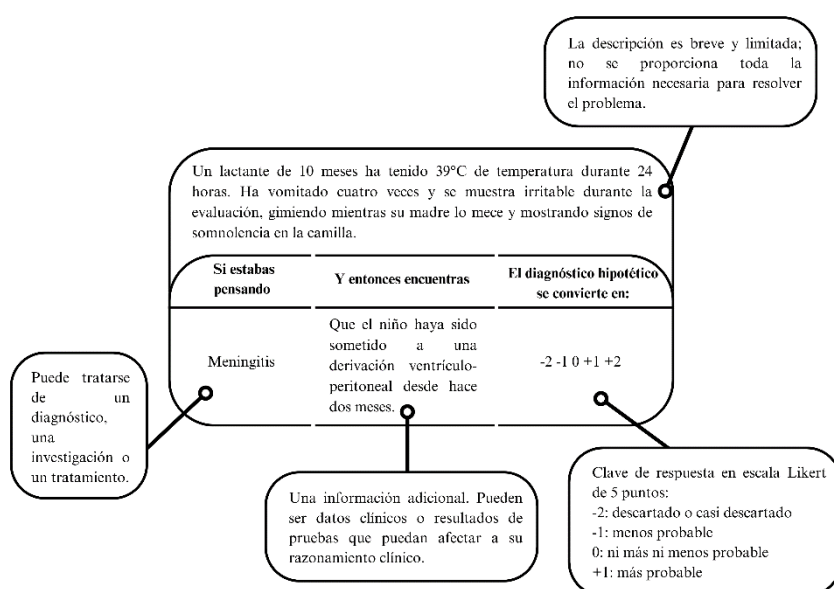
## **3. Puntuación**

La puntuación en el SCT (figura 1) se basa en la comparación de las respuestas de los examinados con las de un panel de referencia compuesto por médicos con experiencia en el área evaluada.<sup>123</sup> Los miembros de este panel completan la prueba de forma individual, y sus respuestas se utilizan para construir la clave de puntuación.<sup>117</sup> Los créditos asignados a cada pregunta se derivan de las respuestas proporcionadas por el panel de referencia.<sup>123</sup>

Para cada respuesta, el crédito se calcula dividiendo el número de miembros del panel que seleccionaron esta opción por el valor modal de la pregunta. Por ejemplo, si en una pregunta específica, quince miembros del panel eligieron "-2," dos optaron por "-1" y uno seleccionó "0," el crédito para la opción "-2" sería 1 (15/15), el crédito para "-1" sería 0.13 (2/15), y

para "0" sería 0.06 (1/15). Las opciones no seleccionadas, como "+1" y "+2," recibirán un crédito de 0. Este método asegura que todas las preguntas tengan el mismo valor máximo (1) y mínimo (0). Las puntuaciones obtenidas en cada pregunta se suman para obtener una puntuación total de la prueba, que luego se divide por el número de preguntas y se multiplica por 100 para obtener un puntaje porcentual.<sup>123</sup>

**Figura 1.** Ejemplo de Ítem SCT. Traducción de: To what extent is the Script Concordance Test a valid measure of clinical reasoning for Advanced Paediatric Life Support Training?<sup>126</sup>



#### **4. Tamaño del panel, composición y reclutamiento**

Según Gagnon et al. (2005),<sup>127</sup> para exámenes de alta relevancia, se requieren 15 miembros en el panel para obtener estimaciones de fiabilidad aceptables, como un coeficiente alfa de Cronbach adecuado. Aunque el valor de estas estimaciones mejora con paneles más grandes, superar los 20 miembros sólo ofrece una mejora marginal. Para exámenes de menor importancia, como evaluaciones formativas dentro de una rotación clínica, se pueden utilizar paneles más pequeños. No obstante, paneles con menos de 10 personas están asociados con un mayor margen de error en las estimaciones de fiabilidad.<sup>123</sup>

La composición del panel es otro factor crucial. El principio fundamental del SCT es comparar el desempeño de los estudiantes o residentes con un grupo que represente legítimamente la profesión o especialidad a la que aspiran. Por lo tanto, los paneles deben estar formados por médicos con experiencia clínica general en el campo, en lugar de especialistas en áreas muy específicas. La composición del panel debe adaptarse al objetivo de la evaluación, y es importante mantener el anonimato de sus miembros, ya que el SCT no está diseñado para proporcionar puntuaciones individuales sobre el rendimiento de los expertos.<sup>123</sup>

La composición del panel de referencia, ya sea especializado o incluyendo médicos de atención primaria, influye en los resultados del SCT y en la clasificación de los estudiantes.<sup>128</sup>

### **SCT y otras pruebas de evaluación**

El Test de Concordancia de Guiones fomenta un 74% de procesos cognitivos de alto nivel, como aplicar, analizar, evaluar y crear, mientras que las preguntas de opción múltiple (MCQ) solo alcanzan un 19%. Un análisis estadístico ( $\chi^2=205.51$ ,  $p<0.00001$ ) demuestra una diferencia significativa entre ambas metodologías. Aunque las MCQ son populares por su facilidad de administración y sus sólidas propiedades psicométricas, su capacidad para representar escenarios clínicos complejos es limitada, ya que suelen evaluar conocimiento teórico estándar sin incorporar la incertidumbre propia del razonamiento clínico.<sup>28</sup>

Las preguntas de respuesta corta (SAQs) evalúan el razonamiento clínico de manera directa, pero el SCT ofrece una evaluación más contextualizada y estandarizada. A diferencia de los exámenes orales estructurados, donde los examinadores pueden influir en la evaluación, el SCT minimiza la subjetividad al compararse con un estándar establecido por expertos. Además, es más estructurado y confiable que los exámenes orales no estructurados, los

cuales pueden estar sujetos a sesgos. Mientras que el Examen Clínico Objetivo Estructurado (OSCE) evalúa habilidades prácticas y de comunicación en un entorno simulado, el SCT se enfoca en la toma de decisiones clínicas bajo incertidumbre. Las MCQs y SAQs miden principalmente conocimiento teórico, mientras que el SCT evalúa la aplicación del razonamiento clínico en escenarios complejos y ambiguos.<sup>10 116 117 129 130</sup>

En un estudio de Sadeghi et al. (2019)<sup>29</sup>, se compararon cuatro pruebas de evaluación del razonamiento clínico: Key Feature (KF), Script Concordance Test (SCT), Clinical Reasoning Problem (CRP) y Comprehensive Integrative Puzzle (CIP). Estas fueron aplicadas a 339 estudiantes de medicina de alto rendimiento en 60 facultades de Irán. La fiabilidad de cada prueba y su combinación fue medida mediante el alfa de Cronbach, obteniendo un valor global de 0.815, lo que indica alta consistencia interna. A nivel individual, las fiabilidades fueron: 0.81 (KF), 0.76 (SCT), 0.80 (CRP) y 0.92 (CIP). El puntaje total promedio en el examen fue de  $169.92 \pm 41.54$  sobre 240 puntos. Por tipo de prueba, los promedios fueron:  $37.99 \pm 7.73$  (KF),  $19.70 \pm 5.46$  (SCT),  $52.94 \pm 17.36$  (CRP) y  $59.28 \pm 16.86$  (CIP). Las correlaciones entre cada prueba y el puntaje total mostraron valores de 0.803 (KF), 0.631 (SCT), 0.885 (CRP) y 0.887 (CIP), siendo esta última la más alta. La correlación más baja se observó entre SCT y CRP (0.473), lo que sugiere que estas pruebas evalúan distintos aspectos del razonamiento clínico.

La Técnica de Presentación de Casos (SNAPPS) facilita la expresión del razonamiento clínico<sup>131</sup>, mientras que el SCT evalúa la gestión de la incertidumbre y la aplicación del conocimiento en situaciones complejas. La autoexplicación fomenta el procesamiento activo y la construcción del conocimiento<sup>132</sup>, complementando al SCT al fortalecer la aplicación del razonamiento. Por su parte, la reflexión estructurada se enfoca en el desarrollo y perfeccionamiento del razonamiento clínico durante el aprendizaje<sup>133</sup>.

#### **2.2.4. Uso de Inteligencia Artificial en Educación Médica**

En los últimos años, la inteligencia artificial (IA), especialmente modelos de lenguaje generativo como ChatGPT, ha ganado relevancia en el ámbito médico. Aunque sigue en desarrollo, ha generado preocupaciones sobre su impacto en la integridad académica, particularmente en la redacción de trabajos y asistencia a estudiantes<sup>134</sup>.

En medicina, se prevé que ChatGPT mejore la precisión diagnóstica, amplíe diagnósticos diferenciales y reduzca errores médicos<sup>135</sup>. En exámenes estandarizados, su precisión en el USMLE Step 1 era del 65%<sup>136</sup>, aumentando al 90% con ChatGPT-4<sup>137</sup>. En el ENAM (Perú), GPT-4 alcanzó un 86,1%, superando el 77,2% de GPT-3.5<sup>138</sup>.

Uno de los desafíos más significativos en la educación médica actual es abordar los problemas cognitivos que los médicos enfrentan hoy y en el futuro. Aunque la búsqueda de información se ha vuelto más accesible gracias a internet, resolver casos clínicos complejos sigue siendo una tarea desafiante.<sup>39</sup>

#### **Aplicaciones clínicas y educativas**

La inteligencia artificial (IA) se define como el desarrollo de sistemas capaces de realizar tareas que normalmente requieren inteligencia humana<sup>39</sup>. ChatGPT, el modelo de IA más relevante en medicina, es un transformador generativo preentrenado (GPT) diseñado para reconocer y generar lenguaje escrito<sup>39</sup>. Analiza preguntas y genera respuestas basadas en patrones lingüísticos encontrados en la web<sup>39</sup>.

Actualmente, se emplea en la educación de pregrado, posgrado, formación clínica y educación médica continua<sup>139</sup>. Además, la IA se integra con tecnologías como realidad virtual, 5G, computación en la nube, big data, dispositivos portátiles e Internet de las cosas para mejorar la educación médica<sup>140-142</sup>.

Siguiendo esta misma línea, ya existe evidencia que respalda el uso de la inteligencia artificial para desarrollar métodos de evaluación, como las preguntas de opción múltiple (MCQ),<sup>143-146</sup> demostrando su validez y comparabilidad con las elaboradas por seres humanos en términos de calidad, dificultad y capacidad de discriminación.<sup>147</sup> Además, la IA ha demostrado ser efectiva en la creación de pruebas de concordancia de guiones, produciendo resultados equivalentes a los generados por un panel de expertos compuesto por psiquiatras.<sup>18</sup> Para estandarizar y facilitar la creación de estas evaluaciones, se han desarrollado chatbots automatizados a través de ChatGPT: Case-based MCQ Generator<sup>148</sup> y Script Concordance Test generation.<sup>19</sup>

La inteligencia artificial (IA) se aplica en medicina en áreas como radiología, diagnóstico, odontología, oftalmología, cardiología y psiquiatría<sup>139</sup>. Ante el avance de la medicina inteligente, universidades han incorporado cursos sobre informática, codificación y algoritmos para atender futuras demandas clínicas<sup>149-153</sup>. Su uso destaca en la formación de expertos y educación continua, especialmente en diagnóstico y radiología, donde el big data y la computación en la nube optimizan la toma de decisiones<sup>154-156</sup>. Se enfatiza que los médicos deben adoptar la IA para mejorar su precisión y eficiencia mediante formación profesional y autoaprendizaje<sup>142-149</sup>.

Aunque un número limitado de estudios en la educación de pregrado y posgrado describe el uso de la IA en métodos como la clase invertida y la realidad virtual para mejorar las habilidades prácticas de los estudiantes,<sup>152,153,157</sup> en China la literatura se centra en el futuro de la IA en la educación médica, influenciada por políticas estatales como el Plan de Acción de Educación Superior de 2018 para la Innovación en Inteligencia Artificial (HEAI).<sup>140-</sup>

142,154,158

### **Ramas de la IA en la educación medica**

En particular, Nagi et al. (2023)<sup>159</sup> describen tres tecnologías clave que están transformando la educación médica: el aprendizaje automático y profundo (Machine Learning/Deep Learning), el entrenamiento en habilidades robóticas (Robotic Skills Training), y la realidad virtual (Virtual Reality), cada una de ellas con un impacto significativo en diferentes áreas de la formación médica:

### **1. Aprendizaje Automático/Profundo (Machine Learning/Deep Learning):**

El aprendizaje automático y profundo optimiza el análisis de datos médicos para diagnósticos y predicciones clínicas. En educación médica, mejora la precisión diagnóstica y la toma de decisiones, como en el diagnóstico de enfermedades oculares mediante imágenes médicas. También se aplica en la planificación preoperatoria, permitiendo a los estudiantes adquirir experiencia antes de situaciones reales.

### **2. Entrenamiento en Habilidades Robóticas (Robotic Skills Training):**

El entrenamiento en habilidades robóticas permite a los estudiantes practicar procedimientos quirúrgicos complejos en un entorno simulado. Es especialmente útil en cirugía y ortopedia, donde ayuda a mejorar la precisión, perfeccionar técnicas y reducir errores en intervenciones reales.

### **3. Realidad Virtual (Virtual Reality):**

La realidad virtual crea entornos inmersivos que replican situaciones médicas, permitiendo a los estudiantes practicar de manera segura. Es especialmente útil en cirugía, donde facilita el perfeccionamiento de técnicas, y en salud conductual, donde ayuda a desarrollar empatía y mejorar la comunicación mediante simulaciones realistas.

## **Beneficios y limitaciones de la IA en la educación medica**

Preiksaitis y Rose (2023)<sup>160</sup> sintetizaron en su estudio las oportunidades y limitaciones de la inteligencia artificial (IA) generativa en la educación médica, resaltando la necesidad de un enfoque equilibrado para optimizar su impacto educativo y minimizar sus riesgos. En este sentido, la IA ofrece diversos beneficios, como la interacción sinérgica entre sus ventajas y limitaciones, la personalización del aprendizaje a partir de datos de pruebas estandarizadas, la síntesis de investigaciones para generar contenido actualizado, el fortalecimiento de habilidades de redacción y organización de ideas, y la creación de materiales innovadores y diversos. No obstante, su implementación también plantea desafíos, como la posibilidad de errores o "alucinaciones", la necesidad de construir modelos conceptuales más sólidos, la verificación constante de la información para fomentar un aprendizaje más crítico, la importancia de seguir investigando su impacto en la educación médica y la adopción de un enfoque proactivo que permita gestionar sus riesgos mientras se maximiza su potencial educativo.

#### **2.2.4.1 Generación del test de concordancia de guiones mediante ChatGPT**

La generación de un Test de Concordancia de Guiones mediante ChatGPT se centra en diseñar situaciones clínicas complejas que reflejan la realidad médica, desarrollando preguntas que capturen la incertidumbre inherente a la toma de decisiones en contextos de salud. Este proceso incluye la creación de escenarios clínicos realistas, la formulación de preguntas que representen la incertidumbre clínica y el establecimiento de una escala de respuestas que permita evaluar cómo se alinean las decisiones de los participantes con las de un panel de expertos.<sup>19</sup> Comprendiendo este marco teórico, Alexandre et al. (2024)<sup>18</sup> compararon los SCT generados por ChatGPT con aquellos producidos por psiquiatras clínicos, evaluando aspectos como el escenario, las preguntas clínicas y la opinión de los expertos. Los resultados indicaron que no hubo diferencias significativas entre los SCT generados por ChatGPT y los desarrollados por expertos en los componentes evaluados.

### **Prompt y pasos para la generación de ítems del SCT**

La Tabla 1 presenta el cuestionario utilizado para producir los *SCT*, el cual se basa en fundamentos teóricos y prácticos bien establecidos para su constructo.<sup>18,116,123,161</sup> Al ser una plantilla general adaptable a las necesidades específicas de los usuarios, se solicita que estos completen las secciones relevantes, marcadas en negrita y entre corchetes, con la siguiente información<sup>19</sup>:

- **Grupo Destinatario:** ¿Estos ítems de SCT están destinados a estudiantes de medicina de pregrado, residentes de posgrado, participantes en educación médica continua, u otro grupo específico? La primera sección destacada en negrita debe completarse con esta información.
- **Enfoque de las Preguntas:** La literatura sugiere que cada SCT debe centrarse en un único aspecto, ya sea diagnóstico, investigación o tratamiento.<sup>123</sup> El usuario debe especificar el enfoque requerido en esta sección. ¿Qué aspecto necesita abordar en relación con un problema clínico: diagnóstico, secuencia de investigaciones o tratamiento? Las siguientes tres partes de la plantilla, indicadas en negrita como "[ESCRIBA EL ASPECTO EN EL QUE DEBEN CENTRARSE LAS PREGUNTAS: EJ., diagnóstico, secuencia de investigaciones, tratamiento]", deben completarse de manera coherente con la opción seleccionada. Por ejemplo, si desea crear un SCT sobre tratamiento, debe ingresar "tratamiento" en cada una de estas secciones. Si planea generar otro ítem de SCT enfocado en investigaciones, debe ingresar "secuencia de investigaciones" en cada una de las tres partes.
- **Tema o Problema Clínico:** ¿Cuál es el tema clínico o el problema específico sobre el que debe centrarse el SCT? Esta sección debe completarse con el tema específico o problema clínico que se abordará.

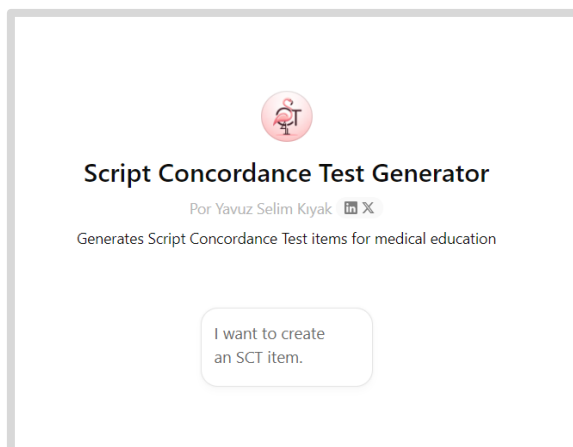
- **Guía Clínica:** Si el usuario prefiere que las preguntas se basen en una guía clínica específica, debe mencionarla en esta sección. Alternativamente, puede completar la sección con "las guías confiables y/o consenso de expertos sobre este problema clínico," lo cual es igualmente válido para la generación de ítems del SCT.

Y para automatizar aún más el proceso de generación del SCT, Kiyak y Emekli desarrollaron el "Script Concordance Test Generator" (figura 2) utilizando ChatGPT. Esta herramienta gratuita elimina la necesidad de copiar y pegar indicaciones detalladas; en su lugar, los usuarios solo deben completar las secciones requeridas.<sup>19</sup>

**Tabla 1.** Una guía para generar ítems de SCT en educación médica. Reproducida de: A Prompt for Generating Script Concordance Test Using ChatGPT, Claude, and Llama Large Language Model Chatbots.<sup>19</sup>

<p>You are a script concordance test (SCT) developer for medical exams for [GENERAL DESCRIPTION OF THE TARGET GROUP, SUCH AS, "undergraduate medical students", "postgraduate medical trainees", "continuing medical education participants", OR A MORE DETAILED INFORMATION, SUCH AS "last year radiology residents in Turkey"].</p> <p>SCT is a method used to assess clinical reasoning and decision-making skills in healthcare students and professionals. It is designed to evaluate how they interpret clinical information and make decisions under conditions of uncertainty. Each scenario is a clinical vignette describing a medical situation. For each scenario, there are multiple possible options or actions that a physician could take or consider in that situation. Each SCT consists of four main components.</p> <ol style="list-style-type: none"> <li>1. A very brief clinical vignette on a clinical problem regarding [TYPE WHAT THE SET OF QUESTIONS SHOULD FOCUS ON: E.G. diagnosis, order of investigation, treatment]. In a single table with four rows, the first row includes the labels:</li> <li>2. First column: ONLY three different key plausible [TYPE WHAT THE SET OF QUESTIONS SHOULD FOCUS ON: E.G. diagnosis, order of investigation, treatment] for the clinical vignette. The label of the column should be "If you were thinking of ...".</li> <li>3. Second column: A new piece of clinical information for each of the initial option to make the situation significantly more complex based on the clinical vignette, such as, new symptoms, previously undisclosed aspects of the patient's history, physical examination findings, and results of tests or previous treatments. The label of the column should be "Then you learn that ...".</li> <li>4. Third column: Asking the examinee to rate on a 5-point scale (from -2 to +2) how this new information impacts or affects the initial option - whether it makes that option more or less appropriate, and to what degree. The label of the column should be "It becomes ... (-2 much less appropriate, -1 less appropriate, 0 no change, +1 more appropriate, +2 much more appropriate)".</li> </ol> <p>So the examinee has to reason through how the new clinical data changes or updates the appropriateness of the initial [TYPE WHAT THE SET OF QUESTIONS SHOULD FOCUS ON: E.G. diagnosis, order of investigation, treatment] option they were given. The 5-point scale allows them to indicate if the new information makes the original option more likely, less likely, or doesn't really change it, and to what extent. The idea is that this tests the student's clinical reasoning abilities by presenting evolving, uncertain medical cases and having them re-evaluate options as new information comes in, which is what happens in real medical practice.</p> <p>Write a SCT for the following clinical problem: [TYPE A CLINICAL PROBLEM OR TOPIC OR LEARNING OBJECTIVE RELATED TO DIAGNOSIS, INVESTIGATION, OR TREATMENT] according to [TYPE THE NAME OF A RELEVANT GUIDELINE OR JUST TYPE "the reliable guidelines and/or expert consensus on this clinical problem"]. Indicate the level of appropriateness in the third column cells, together with a very brief rationale.</p>
---

**Figura 2.** Script Concordance Test Generator (ChatGPT)



## 2.3. Formulación de hipótesis

### 2.3.1. Hipótesis General

#### **Hipótesis de trabajo (Ht)**

Existen factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

#### **Hipótesis nula (Ho)**

No existen factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

## 2.3.2. Hipótesis específicas

### 2.3.2.1. Hipótesis específica 1

- **H<sub>1</sub>:** Los factores sociodemográficos del razonamiento clínico están asociados con el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.
- **H<sub>0</sub>:** No existe una asociación entre los factores sociodemográficos del razonamiento clínico y el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.

### 2.3.2.2. Hipótesis específica 2

- **H<sub>1</sub>:** Los factores académicos del razonamiento clínico están asociados con el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.
- **H<sub>0</sub>:** No existe una asociación entre los factores académicos del razonamiento clínico y el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.

### 2.3.2.3. Hipótesis específica 3

- **H<sub>1</sub>:** Los factores del entorno asistencial del razonamiento clínico están asociados con el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.

- **H<sub>0</sub>:** No existe una asociación entre los factores del entorno asistencial del razonamiento clínico y el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.

## **CAPÍTULO III: METODOLOGÍA**

### **3.1. Método de la investigación**

Se adoptó el método hipotético-deductivo, dado que la investigación parte de una hipótesis que será evaluada mediante el análisis de datos estadísticos. A través de un proceso lógico y estructurado, se interpretarán los resultados obtenidos, permitiendo extraer conclusiones fundamentadas en la evidencia recolectada<sup>162</sup>.

### **3.2. Enfoque de la investigación**

La investigación adoptó un enfoque cuantitativo, dado que se usó variables cuantificables y las cuales no se manipularon<sup>163</sup>.

### **3.3. Tipo de investigación**

Esta investigación se define como investigación básica, ya que busca generar conocimiento sobre los factores que afectan el desempeño académico en una evaluación clínica. Su propósito principal es comprender estos factores sin intervenir directamente en la formación o evaluación de los estudiantes<sup>164</sup>.

### **3.4. Diseño de la investigación**

Este estudio tiene un diseño no experimental, de corte transversal, de alcance correlacional y con un enfoque prospectivo dentro de su clasificación temporal<sup>165,166</sup>.

### **3.5 Proceso de Diseño, Revisión y Aplicación del Script Concordance Test (SCT)**

#### **1) Elección de grupo de patologías del Script Concordance Test**

Lo primero que se realizó fue un temario de patologías como base para delimitar posteriormente el número de ítems y casos. El temario fue desarrollado a partir de estudios peruanos que identifican las patologías más prevalentes en territorios vulnerables a nivel nacional, con un enfoque particular en el contexto del SERUMS.<sup>167,168</sup> Esta elección buscó aprovechar el contexto del SERUMS para desarrollar y validar el Script Concordance Test enfocado en el primer nivel de atención, que permita a los internos de medicina humana practicar situaciones clínicas comunes a este nivel asistencial. Dado que la mayoría de estos internos enfrentarán el SERUMS en su primer año de práctica como médicos generales, la simulación de este entorno no solo resulta pertinente, sino también estratégica para prepararlos en los desafíos específicos del primer nivel de atención. Los casos clínicos y las patologías seleccionadas en el SCT abarcaron situaciones que los internos ya han abordado durante su carrera y su internado, garantizando que todos, independientemente de su rotación en el primer nivel de atención, puedan responder de manera competente. Así, este ejercicio no solo se alinea con las competencias esperadas en su formación, sino que también refuerza su preparación para responder a las necesidades sanitarias del sistema peruano de salud.

Asimismo, el temario estuvo alineado con evaluaciones nacionales en salud,<sup>169,170</sup> principalmente en la Tabla de Especificaciones del ENAM, una evaluación ampliamente reconocida en la formación médica en el Perú.<sup>171</sup> Esto garantizó que los contenidos cubran tanto las competencias teóricas como las habilidades clínicas necesarias para el ejercicio médico, permitiendo que los futuros profesionales puedan enfrentar eficazmente las enfermedades más comunes en el país, con un enfoque en el diagnóstico y tratamiento bajo incertidumbre.

Además de su relevancia en la formación médica, el ENAM se ha consolidado como una herramienta de evaluación de amplio reconocimiento y credibilidad, respaldada por una extensa evidencia.<sup>172,173</sup> Es por ello que su inclusión en el desarrollo del temario aseguró que

los médicos estarán preparados no solo para responder a las demandas del sistema de salud peruano, sino también para desempeñarse en un contexto clínico que exige tanto competencia técnica como la capacidad de tomar decisiones clínicas bajo incertidumbre, un aspecto clave del razonamiento clínico que, como se ha descrito anteriormente, sigue siendo un desafío en la educación médica.

### **Definición del Temario de Ítems**

Por medio de este marco teórico, se procedió a estructurar el temario (tabla 2) en dos áreas fundamentales: clínico-médicas y clínico-quirúrgicas, asignando a cada uno 41 y 27 ítems respectivamente. Inicialmente, se consideró incluir un área de temas transversales; sin embargo, se decidió omitirla debido a que estos temas tienden a enfocarse más en un razonamiento metodológico que clínico y, además, presentan una limitada evidencia fuera de este último enfoque.<sup>174</sup> Por lo tanto, su inclusión habría desviado parte de nuestro objetivo principal en esta tesis: evaluar el razonamiento clínico.

Por lo expuesto, la estructura del temario planteada respondió a la necesidad de cubrir las competencias médicas esenciales en estos ámbitos, asegurando una distribución equilibrada que abarque tanto las patologías de mayor prevalencia como aquellas de manejo crítico en la práctica clínica. En las áreas clínicas, se llevó a cabo una selección rigurosa que abarca el diagnóstico o el manejo específico para cada patología, además de los procedimientos pertinentes, fundamentándose en su relevancia y frecuencia en el contexto del SERUMS.

**Tabla 2.** Temario objetivo para la estructura del SCT**Temario (80 patologías)**

<b>ÁREAS</b>
<b>CLÍNICO MÉDICAS - 41</b>
<b>Medicina</b>
Tuberculosis
Infecciones urinarias
Dengue
Rabia
Hepatitis
Hipertensión arterial
Cardiopatía isquémica
Diarrea aguda
Hipertiroidismo
Hipotiroidismo
Obesidad
Nutrición
Diabetes mellitus
Cefalea
Envenamiento
Intoxicaciones agudas
EPOC
Neumonía adquirida en la comunidad
Lumbalgia
Artrosis
Artritis
<b>Pediatría</b>
Niño febril
Varicela
Rubeola
Sarampión
Escabiosis
Urticaria
Pediculosis
Resfrío común
Enfermedad diarreica aguda
Inmunizaciones
Anemia ferropénica
CRUP
Parasitosis intestinal
Faringitis estreptocócica
<b>Emergencias y críticos adultos</b>

Paro cardio respiratorio
OVACE
<b>Procedimientos</b>
Inyectables
Lavado de oídos
Colocación de sondas
Limpieza de heridas
<b>CLÍNICO QUIRÚRGICAS - 27</b>
<b>Cirugía general</b>
Apendicitis aguda
Obstrucción intestinal baja
Obstrucción intestinal alta
Politraumatismo
Luxaciones
Litiasis urinaria
Quemaduras
Heridas de piel y faneras
Fracturas
Conjuntivitis
Colecistitis aguda
Enfermedad litiásica biliar
Traumatismo abdominal abierto
Traumatismo abdominal cerrado
<b>Gineco obstetricia</b>
Hipertensión inducida por el embarazo
S. HELLP
Trabajo de parto
ITU en embarazo
Hemorragia de la primera mitad del embarazo
Hemorragia de la segunda mitad del embarazo
Diabetes y gestación
Control prenatal
Abuso sexual
Infecciones de transmisión sexual
Planificación familiar
Dismenorrea primaria
Hemorragia Uterina Anormal

### **Número de ítems y casos**

Diversos estudios han demostrado que, para alcanzar una confiabilidad óptima en el SCT, se recomienda incluir entre 60 y 70 ítems. Este rango ha demostrado ser suficiente para

lograr un coeficiente de confiabilidad cercano a 0.80, lo que asegura una evaluación precisa del razonamiento clínico.<sup>32,85,117,175-178</sup> Otros estudios han explorado el número óptimo de ítems o casos necesarios para alcanzar un nivel de confiabilidad adecuado, generalmente establecido en 0.8. Uno de los estudios más citados utilizó la fórmula Spearman-Brown para calcular que 36 casos, cada uno con uno o dos ítems, eran suficientes para lograr esta confiabilidad. Sin embargo, otros estudios han mostrado resultados variados, sugiriendo entre 13 y 102 ítems como el rango óptimo para obtener la misma fiabilidad.<sup>121,124,125</sup>

Otro estudio comparativo analizó el impacto de agregar más casos frente a aumentar la cantidad de ítems por caso en pruebas de análisis de generalización en el contexto de las pruebas SCT. Este estudio concluyó que añadir más ítems dentro de los casos era más eficaz para mejorar la confiabilidad, y que agregar casos adicionales no incrementaba la precisión de manera tan significativa.<sup>125</sup> La evidencia muestra que 25 casos con tres ítems cada uno pueden generar coeficientes de confiabilidad entre 0.75 y 0.86, según estimaciones del alfa de Cronbach.<sup>118</sup> Este enfoque no solo es adecuado en términos de confiabilidad, sino que también se ajusta a las limitaciones cognitivas de los médicos, cuya memoria de trabajo mayormente no puede manejar un número elevado de hipótesis simultáneas.<sup>119,179</sup> Concluyendo por ello que un número reducido de ítems por caso es el más eficiente para evaluar el razonamiento clínico, dado que reduce la carga cognitiva de los examinados sin comprometer la calidad de la evaluación.<sup>119,179</sup> Además, se ha observado que, a partir de tres o cuatro ítems por caso, se produce un efecto de saturación, lo que sugiere que aumentar el número de ítems más allá de este punto no mejora de manera significativa la confiabilidad de la prueba.<sup>125</sup> Por ello se debe recopilar una cantidad razonable de información por caso (3 preguntas/  $\pm$  1 pregunta) para obtener una evaluación confiable de las puntuaciones del SCT.<sup>125</sup>

Por lo que se optó en utilizar entre 2 y 3 ítems, ya que la literatura revisada sugiere que este rango de ítems no solo mejora la confiabilidad de las pruebas, sino que también se adapta a las limitaciones cognitivas de los participantes.

Además, este número de ítems por caso permite mantener un equilibrio adecuado entre la cantidad de información recopilada y la capacidad de los examinados para procesar dicha información. Asimismo, como ya se describió, se ha observado que incluir más de 3 o 4 ítems por caso genera un efecto de saturación que, lejos de aportar, no contribuye significativamente a mejorar la confiabilidad de la prueba. Por lo tanto, este enfoque se alinea con las mejores prácticas en el diseño de pruebas, garantizando que se mantenga la calidad y validez de los resultados obtenidos en el Script Concordance Test.

Con la decisión sobre el número de ítems por caso establecida, se optó por agrupar patologías relacionadas dentro de un mismo caso, basándose en el temario objetivo. Esta estrategia optimiza las ventajas del 'Script Concordance Test', al permitir la inclusión de múltiples diagnósticos en una sola viñeta y combinar elementos de manejo, favoreciendo así una evaluación más integral sin incrementar el número total de ítems. Algunas patologías, que tuvieron un amplio campo para el desarrollo en el ámbito médico de primer nivel y que se caracterizan por su frecuencia y sencillez en la resolución, no se agruparon con otras. En su lugar, se seleccionaron para ocupar dos ítems en un caso, asegurando una evaluación más específica y detallada. Así también se eliminaron los procedimientos de colocación de sondas e inyectables, ya que al hacer pruebas con el "Script Concordance Test Generator", no se pudo obtener una verdadera relación con alguna enfermedad ni adaptarse a una verdadera evaluación del razonamiento clínico. Esto apoyado con el hecho de que en el contexto del SERUMS, se adapta más al mantenimiento de sondas que al de la colocación de estas. Además, teniendo en cuenta que el médico puede y debería tener conocimiento de estos procedimientos, esta función como la de los inyectables está dentro de las

competencias de enfermería. Así también se concluyó en eliminar “conjuntivitis” y reemplazarlo por “otitis” para poder integrarlo con el procedimiento de “lavado de oídos” con el fin de beneficiarnos de esta relación sin aumentar más ítems.

La tabla 3 presenta el resultado del desarrollo y delimitación del temario, donde se obtuvieron 67 ítems distribuidos en 27 casos, de los cuales 13 casos contienen 3 ítems y 14 casos incluyen 2 ítems.

**Tabla 3.** Temario final v1 para la estructura del SCT

**Temario (67 ítems) (27 Casos)**

Áreas	Ítems	Casos
<b>CLÍNICO MÉDICAS</b>	<b>41</b>	<b>17</b>
<b>Medicina</b>	<b>24</b>	<b>10</b>
Tuberculosis - neumonía adquirida en la comunidad - EPOC (3)	Ítem -1 -2 - 3	1
Infecciones urinarias (2)	Ítem - 4 -5	2
Rabia - limpieza de heridas (2)	Ítem - 6 - 7	3
Hepatitis - dengue (2)	Ítem - 8 - 9	4
Hipertensión arterial - cefalea - cardiopatía isquémica (3)	Ítem -10 -11 - 12	5
Diarrea aguda (2)	Ítem - 13 - 14	6
Hipertiroidismo - hipotiroidismo (2)	Ítem - 15 - 16	7
Diabetes mellitus - obesidad - nutrición (3)	Ítem -17 -18 - 19	8
Intoxicaciones agudas - envenenamiento (2)	Ítem - 20 - 21	9
Lumbalgia - artrosis - artritis (3)	Ítem -22 -23 - 24	10
<b>Pediatría</b>	<b>15</b>	<b>6</b>
Niño febril - inmunizaciones (2)	Ítem - 25 - 26	11
Enf. Exantemáticas (varicela - rubéola - sarampión) (3)	Ítem - 27 -28 - 29	12
Enfermedad diarreica aguda - parasitosis intestinal (2)	Ítem - 30 - 31	13
Escabiosis - urticária - pediculosis (3)	Ítem - 32 - 33 - 34	14
Anemia ferropénica (2)	Ítem - 35 - 36	15
CRUP - resfrío común - faringitis estreptocócica (3)	Ítem - 37 - 38 - 39	16
<b>Emergencias y críticos adultos</b>	<b>2</b>	<b>1</b>
Paro cardio respiratorio - OVACE (2)	Ítem - 40 - 41	17
<b>CLÍNICO QUIRÚRGICAS</b>	<b>26</b>	<b>10</b>

<b>Cirugía general</b>	<b>13</b>	<b>5</b>
Apendicitis aguda - colecistitis aguda - enfermedad litiasica biliar (3)	Ítem - 42 - 43 - 44	18
Lavado de oídos - otitis (2)	Ítem - 45 - 46	19
Obstrucción intestinal alta y baja - litiasis urinaria (3)	Ítem - 47 - 48 - 49	20
Politraumatismo - traumatismo abdominal abierto y cerrado (3)	Ítem - 50 - 51 - 52	21
Fracturas - luxaciones (2)	Ítem - 53 - 54	22
<b>Gineco obstetricia</b>	<b>13</b>	<b>5</b>
Hipertensión inducida por el embarazo – S. HELLP (2)	Ítem - 55 - 56	23
Hemorragia de la primera mitad del embarazo - hemorragia de la segunda mitad del embarazo - trabajo de parto (3)	Ítem - 57 - 58 - 59	24
Control prenatal - diabetes y gestación - ITU en embarazo (3)	Ítem - 60 - 61 - 62	25
Infecciones de transmisión sexual - abuso sexual (2)	Ítem - 63 - 64	26
Dismenorrea primaria - hemorragia uterina anormal - planificación familiar (3)	Ítem - 65 - 66 - 67	27

### **Escala Likert**

Para promover una selección más equilibrada de opciones en la escala Likert, se empleó anclajes en los extremos que no sean demasiado absolutos. En lugar de usar términos categóricos como "absolutamente seguro" o "seguro", se utilizará expresiones más flexibles, como "más apropiado o mucho más apropiado", lo que fomentará el uso de toda la escala y evitará sesgos en las respuestas.<sup>116</sup>

Adicionalmente, se usó como método de puntuación una escala de 5 puntos ya que proporciona una mayor fiabilidad y coeficientes de correlación más sólidos en comparación con los métodos que emplean una escala de 3 puntos. Este enfoque en la escala de 5 puntos permite una evaluación más precisa del razonamiento clínico, maximizando la consistencia de las respuestas y mejorando la validez psicométrica de la prueba.<sup>180</sup>

Antes de iniciar con el SCT se presentaron ejemplos prácticos en video al panel de expertos con el formato de las preguntas y la escala Likert utilizada. Permitiendo que puedan entender

y usar toda la escala de manera efectiva, lo que mejorará la precisión y el rendimiento en la prueba.<sup>123</sup>

### **Creación del SCT mediante el “Script Concordance Test Generator”**

Con nuestro temario objetivo y adicional ya completado, se procedió a crear los casos mediante el “Script Concordance Test Generator”, siguiendo los pasos que el mismo GPT recomienda para su uso:

1. **Descripción del grupo objetivo:** Indicar a qué nivel de estudiantes o profesionales de la salud está dirigido el SCT. Por ejemplo:
  - Estudiantes de medicina de pregrado.
  - Residentes de radiología en su último año.
  - Participantes en educación médica continua.
2. **Problema clínico:** Definir el caso o problema médico que se desea evaluar. Puede estar relacionado con el diagnóstico, investigación o tratamiento de un tema específico. Ejemplo:
  - Dolor en el cuadrante superior derecho.
  - Evaluación de la fiebre en un niño.
3. **Foco del SCT:** Definir si el conjunto de preguntas se va a centrar en el diagnóstico, orden de investigación o tratamiento del problema clínico. Esto guiará el tipo de preguntas que se formularán.
4. **Guía clínica o consenso (opcional):** Además en el caso que sea necesario se proporcionará el nombre de una guía clínica peruana específica que se prefiera usar

como referencia para las preguntas. Y en caso no se especifique una se usará: "Guías de práctica clínica y/o consenso de expertos sobre este problema clínico".

Con esta información como base, lo que se hizo inicialmente fue delimitar el grupo objetivo con el propósito de que todo el caso se pueda adaptar a este. Aprovechando la retroalimentación del "Script Concordance Test Generator", se definió junta a ella las condicionantes para el ajuste específico del grupo objetivo, es decir, lo que se describiría literalmente en el GPT:

### **Grupo objetivo**

- **Nivel educativo:** Estudiantes de pregrado en medicina.
- **Campo de enfoque:** Medicina general.
- **Objetivo de evaluación:** Evaluar el razonamiento clínico en el primer nivel de atención, simulando un contexto de SERUMS (Servicio Rural y Urbano Marginal en Salud).
- **Etapas de formación:** Internos de medicina en el último año de carrera, en el último mes antes de finalizar su formación.
- **Contexto de sistema de salud:** Sistema de salud peruano.

Al tener listo el primer y segundo punto (temario), el siguiente paso fue decidir para cada patología, si enfocarnos en el diagnóstico, manejo o describirlos de manera conjunta. Esta elección se realizó en función de la relevancia de cada enfermedad dentro de su contexto clínico. En ciertos casos, se descartó el diagnóstico debido a que el "estándar de oro" para algunas patologías dependía exclusivamente de pruebas complementarias (laboratorio, ECG, radiografías, entre otros). Asimismo, se evitó seleccionar el manejo en aquellas patologías que requerían intervenciones en niveles de atención más especializados. Además, en los

casos donde no fue posible agrupar las patologías y teniendo dos ítems disponibles, se optó por incluir tanto el diagnóstico como el tratamiento.

En el punto cuatro, se decidió utilizar el adjetivo predeterminado en la mayoría de los casos. Este solo se modificó en situaciones donde las guías peruanas presentaban un manejo o diagnóstico considerablemente diferente al de las guías internacionales, o cuando la epidemiología nacional jugaba un papel decisivo. Finalmente, al generar todos los casos e ítems para nuestro SCT, se seleccionaron o ajustaron nuevamente para que se adaptaran lo mejor posible a nuestro estudio. Estos se organizaron en un documento de Word para facilitar el orden y su posterior transferencia a la plataforma “Tally Form” que funcionará como medio para plantear el SCT y formulario de datos.

### **Incorporación de los ítems en Tally Form**

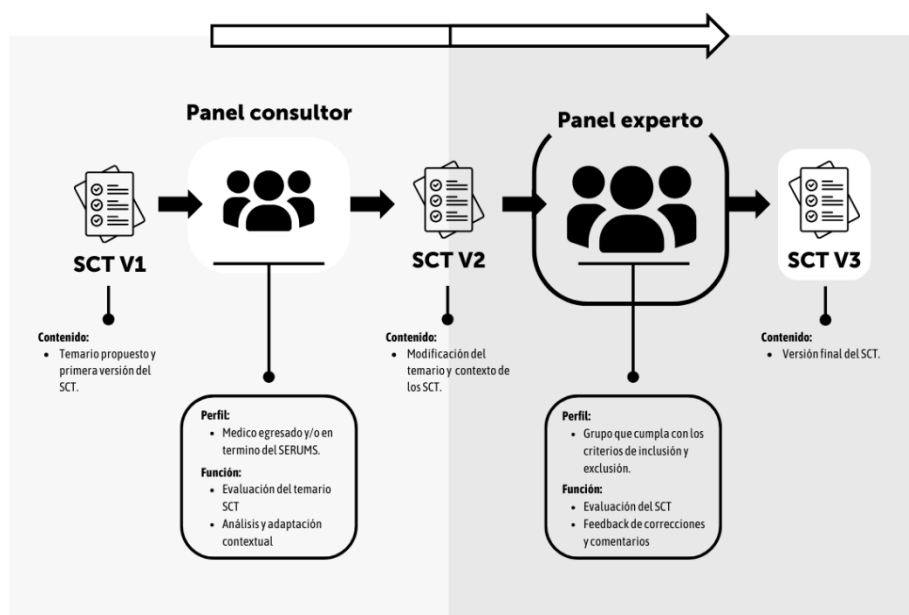
Tras la selección de los casos en nuestro documento en Word, cada caso fue organizado y aleatorizados en páginas individuales en la plataforma Tally Form, que se utilizó en esta fase para la revisión y retroalimentación por parte de consultores y el panel de expertos. Durante esta revisión virtual, los consultores y expertos tuvieron la posibilidad de navegar hacia adelante o hacia atrás entre las páginas, lo que les permitió revisar y corregir sus respuestas para optimizar la precisión en el SCT. La decisión de realizar esta revisión y validación de manera virtual se tomó con el fin de brindar una mayor accesibilidad a los expertos, quienes, en muchos casos, no se encuentran disponibles para reuniones presenciales o resulta complejo coordinar su asistencia en formato físico. Esta modalidad en línea no comprometió la calidad del proyecto, ya que existen evidencias de que el SCT puede aplicarse de manera efectiva y confiable en plataformas digitales, manteniendo así los estándares metodológicos necesarios para la validación<sup>12,16,36,126,176,181-186</sup>.

Para la muestra final, compuesta por los internos de medicina humana, se generó la prueba en formato físico a partir de la plataforma, aprovechando la presentación dinámica de viñetas clínicas que facilita Tally Form.

## **2) Revisión experta y validación de los ítems del SCT**

Como se muestra en la figura 3, la revisión experta y validación de los ítems del Script Concordance Test comenzó con una evaluación preliminar por un panel de consultores, seguida de una revisión por un panel de expertos. Los consultores, médicos generales con experiencia reciente en atención primaria en el contexto del SERUMS, ajustaron inicialmente los ítems para asegurar que reflejen situaciones clínicas comunes y realistas. Posteriormente, el panel de expertos en Medicina Familiar y Comunitaria validó tanto la relevancia clínica como la estructura de las preguntas, garantizando una evaluación precisa del razonamiento clínico en los internos. Esta fase incluyó ajustes en la escala de respuesta y la eliminación de ítems con alta variabilidad, optimizando la confiabilidad y aplicabilidad del SCT en el entorno formativo de los futuros médicos.

**Figura 3.** Proceso de Revisión y Validación del Script Concordance Test. Elaboración propia.



### Panel consultor

El SCT fue revisado inicialmente por un grupo de consultores conformado por médicos generales con experiencia clínica reciente, quienes han completado o están próximos a completar el SERUMS. La función de estos consultores fue evaluar y ajustar el contexto clínico de las preguntas, con el objetivo de que estas reflejen situaciones comunes y realistas del entorno clínico característico del SERUMS. La revisión abordó tanto el contexto de los casos clínicos como la frecuencia de las patologías propias de este periodo de atención en salud primaria.

Se seleccionaron cuatro consultores de distintos niveles de atención y ubicaciones geográficas diversas, con el propósito de integrar una visión amplia y contextualizada en la validación del SCT. La experiencia y diversidad geográfica de estos médicos permitió una

mayor adecuación de los casos clínicos a las realidades epidemiológicas y prácticas a las que los internos de medicina humana se enfrentarán en el contexto del SERUMS.

Luego de la revisión y retroalimentación por parte del panel consultor, se agregaron y modificaron distintos datos en el SCT. Además, se concluyó eliminar 4 ítems (hipertiroidismo, hipotiroidismo, litiasis urinaria y politraumatismo), por distintos motivos como poca frecuencia epidemiológica en el contexto SERUMS, poca correlación entre ítems en el mismo caso e ítems muy similares.

### **Panel de expertos**

Se ha demostrado que paneles con menos de 10 miembros comprometen la confiabilidad del test, mientras que incluir más de 20 expertos no proporciona beneficios adicionales significativos en la fiabilidad.<sup>127</sup> Por esta razón, se envió la invitación para integrar el panel a diversos contactos, estableciendo un mínimo de 10 miembros. Las redes sociales fueron la principal herramienta de comunicación, además de las recomendaciones de los especialistas que iban confirmando su participación. De esta manera, 12 especialistas, cumpliendo con los criterios de inclusión y exclusión, completaron el SCT en un promedio de dos semanas y media.

Durante esta fase, los expertos también evaluaron la escala Likert y los anclajes utilizados para asegurar que promuevan una distribución equilibrada de respuestas a lo largo de la escala. Se limitó el número de ítems con puntajes modales de "0" lo menos posible, para evitar que los participantes recurran a la respuesta neutral. Con esta estrategia se buscará mitigar el impacto de la tendencia conocida como "respuesta de punto medio", donde los participantes tienden a seleccionar la opción neutral en la escala Likert. Además, se procuró asegurar que las respuestas modales se distribuyan de manera equilibrada en toda la escala.<sup>187,188</sup> Se excluyeron los ítems considerados erróneos o cuestionables que presentaron

una alta variabilidad entre expertos, ya que la evidencia demuestra que la eliminación de este tipo de ítems mejora significativamente la confiabilidad.<sup>189</sup>

La tabla 4 presenta la versión del formulario utilizada por los consultores y el panel experto, la cual incluyó un cuestionario anexo para valorar la viñeta clínica y los ítems, adaptando los criterios propuestos por Caire et al. (2004).<sup>23,190,191</sup>

### **Criterios de inclusión al panel experto**

- Médicos especialistas en Medicina Familiar y Comunitaria.
- Recertificación vigente por el Colegio Médico del Perú.

### **Criterios de exclusión al panel experto**

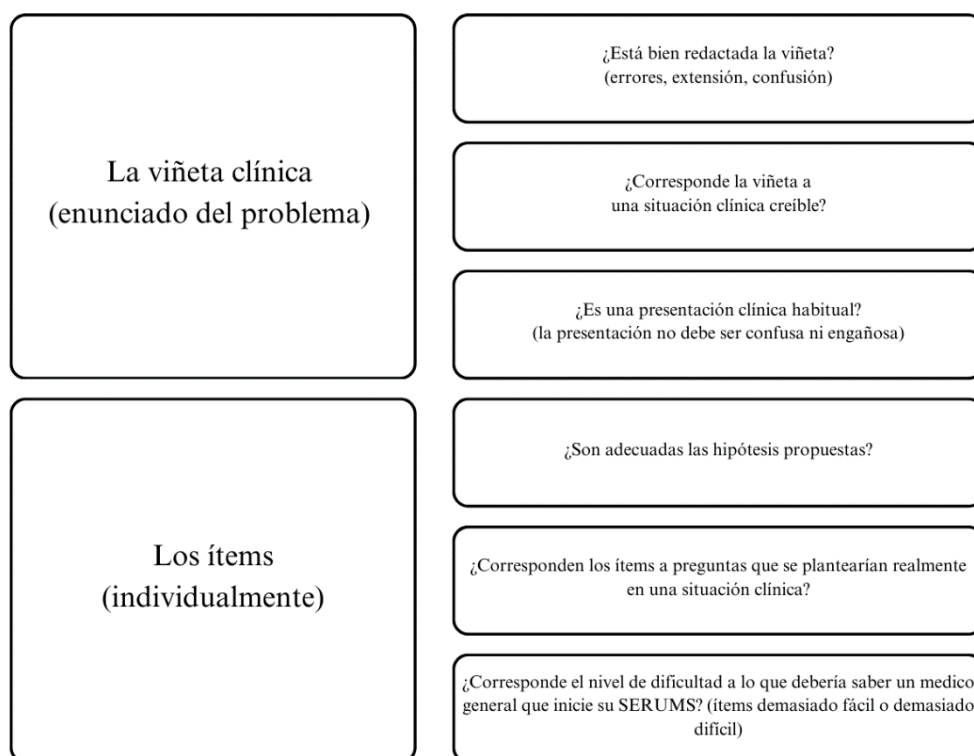
- Menos de 1 año de experiencia en el primer nivel de atención.
- No haber participado en programas de formación continua, cursos de actualización médica o conferencias en los últimos 5 años.

### **Justificación de los criterios de inclusión y exclusión del panel de expertos**

La selección de médicos especialistas en Medicina Familiar y Comunitaria como panel de expertos para este estudio responde a la necesidad de contar con profesionales altamente capacitados en la atención primaria, que posean un conocimiento integral y experiencia en el abordaje de problemas de salud prevalentes en el primer nivel de atención. Esta especialidad garantiza una formación sólida en la evaluación y manejo de casos clínicos comunes y complejos en contextos de atención básica, como el SERUMS, que es el escenario simulado para las viñetas clínicas del Script Concordance Test. Además, los médicos de esta especialidad están habituados a un enfoque centrado en el paciente y poseen habilidades de razonamiento clínico esenciales para evaluar adecuadamente el desempeño académico de los internos en escenarios reales.

Para asegurar la relevancia y calidad de las evaluaciones, se excluyeron del panel aquellos profesionales con menos de un año de experiencia en el primer nivel de atención, ya que una experiencia insuficiente podría limitar su perspectiva en situaciones clínicas habituales en entornos de atención primaria. Asimismo, se estableció como criterio de exclusión la falta de participación en programas de formación continua o eventos académicos en los últimos cinco años, con el fin de contar con expertos que mantengan conocimientos actualizados y estén familiarizados con las últimas guías y protocolos en medicina familiar y comunitaria. Estos criterios buscan asegurar que el panel de expertos esté compuesto por profesionales plenamente cualificados y actualizados, capaces de ofrecer una evaluación precisa y contextualizada del razonamiento clínico de los internos de medicina humana.

Luego de la recepción total de las respuestas de los 12 miembros del panel, estas fueron incorporadas a Microsoft Excel v. 2020 para luego ser analizadas en SPSS v.28. Posteriormente, se inició la revisión manual de cada ítem y respuesta del cuestionario anexo. Durante este proceso, se realizaron modificaciones en varios ítems y viñetas en diferentes secciones de la estructura, como datos de anamnesis, precisión del diagnóstico, examen físico, datos ambiguos y falta parcial de información en el caso. Se acordó con el panel la eliminación de 10 ítems (48, 12, 11, 51, 27, 42, 39, 57, 58, 59), principalmente debido a la alta variabilidad en las respuestas y la ausencia de una mayoría clara que respaldara una decisión común. Entre los ítems eliminados es de importancia mencionar el caso clínico 2, que incluía tres ítems (57, 58, 59), el cual fue eliminado totalmente debido a los errores de contenido identificados tanto en la encuesta anexa como en los comentarios del panel.

**Tabla 4.** Esquema de corrección para el panel de expertos para la valoración del SCT.Adaptado y traducido de Claire et al. (2004)<sup>190</sup>

### 3) Aplicación final del SCT en la población objetivo

Se proyectó inicialmente, luego del proceso de revisión y refinamiento de las viñetas clínicas en base a la literatura, utilizar aproximadamente 60 ítems. Esta cantidad buscó equilibrar la profundidad y diversidad de los escenarios clínicos sin extender innecesariamente el tiempo de evaluación, evitando así la fatiga y asegurando una concentración óptima de los participantes durante la prueba. Sin embargo, tras el proceso de revisión y eliminación de ítems, se llegó a un total de 53, los cuales fueron considerados los más representativos y adecuados para el propósito de la evaluación. Se esperó que el test dure entre 60 y 90 minutos, lo que ha demostrado ser adecuado para obtener puntuaciones fiables.<sup>125</sup> Aunque se ha visto que el tiempo es muy variado dependiendo de la población objetivo, dificultad de las viñetas clínicas y número de ítems.<sup>14,15,189,192-194</sup> Finalmente el día de la ejecución el

tiempo promedio que usaron los estudiantes fue de 30 min., similar a otros estudios con número de ítems similares a esta investigación, que describieron en tiempos entre 30 y 60 min.<sup>195,196</sup>

### **3.6. Población, muestra y muestreo**

La población del estudio estuvo compuesta por los internos de medicina de la Universidad Norbert Wiener, ubicada en Lima, Perú, con un total de 123 estudiantes disponibles durante el período de estudio, según datos del Portal Transparencia Wiener. Se empleó un muestreo censal, considerando a todos los internos como participantes en el estudio, dado que se buscó analizar a la totalidad de la población objetivo.

No obstante, debido a factores como la disponibilidad y el consentimiento informado, la muestra final consistió en 87 participantes que completaron el Test de Concordancia de Script. Para garantizar la representatividad estadística de esta muestra, se aplicó la fórmula de Fisher para poblaciones finitas de manera inversa. Al verificar los parámetros con un nivel de confianza del 90% ( $Z=1.645$ ) y un margen de error del 5% ( $e=0.05$ ), se determinó que el tamaño mínimo necesario era de 85 participantes.

#### **Criterios de Inclusión:**

- Estudiantes de la Escuela Académico Profesional de Medicina de la Universidad Privada Norbert Wiener que estén cursando el séptimo año de estudios (internado médico) durante el año 2024.

#### **Criterios de Exclusión:**

- Estudiantes que no otorguen su consentimiento informado de manera voluntaria.
- Estudiantes que hayan completado recientemente exámenes de razonamiento clínico en condiciones similares al SCT.

### **Justificación de los criterios de inclusión y exclusión de los participantes**

La selección de los criterios de inclusión se basa en la necesidad de contar con una población homogénea que refleje fielmente el nivel de formación y competencia clínica esperados en el internado médico, etapa avanzada del currículo de medicina. Al incluir solo a los estudiantes de séptimo año de la Universidad Privada Norbert Wiener durante el periodo 2024, se asegura que los participantes tengan una formación común y estén en el mismo estadio de desarrollo profesional. Esto permite que el Test de Concordancia de Script evalúe su razonamiento clínico en un contexto relevante y estandarizado, alineado con los objetivos de formación de la institución y ajustado a las demandas clínicas propias del internado.

Los criterios de exclusión, por otro lado, aparta aquellos estudiantes que no otorguen su consentimiento informado asegurando la participación voluntaria y ética, condición fundamental en toda investigación. Además, la exclusión de estudiantes que hayan completado recientemente pruebas de razonamiento clínico similares al SCT es clave para evitar el efecto de aprendizaje o familiaridad previa con el formato y la estructura de la evaluación. Esto contribuye a mantener la integridad de la prueba, garantizando que los resultados reflejan el razonamiento clínico de los internos de manera auténtica y sin interferencias derivadas de exposiciones previas.

### 3.6. Variables y operacionalización

Variable	Definición conceptual	Definición operacional	Dimensión	Indicador	Escala de medición	Escala Valorativa
Factores asociados al razonamiento clínico	Sexo registrado en el documento nacional de identidad reportado por el interno de medicina humana <sup>197</sup> .	Registro del sexo (masculino o femenino) obtenida a través de un cuestionario.	Factores sociodemográficos	Sexo	Cualitativa nominal	1. Masculino 2. Femenino
	Cantidad de años cumplidos por el interno de medicina humana desde su nacimiento hasta el momento de la evaluación <sup>198</sup> .	Registro de la edad en años cumplidos mediante un cuestionario administrado al interno de medicina humana.				Edad

Existencia de una carrera profesional o estudios superiores completados antes del ingreso a la Escuela Académico Profesional de Medicina Humana en la Universidad Privada Norbert Wiener <sup>199</sup> .	Registro de la carrera profesional o estudios superiores completados antes del ingreso a la Escuela Académico Profesional de Medicina humana en la Universidad Privada Norbert Wiener mediante un cuestionario	Factores académicos	Estudios o carrera previa	Cualitativa nominal	<ol style="list-style-type: none"> <li>1. Ninguna</li> <li>2. Enfermería</li> <li>3. Obstetricia</li> <li>4. Nutrición</li> <li>5. Odontología</li> <li>6. Terapia Física y Rehabilitación</li> <li>7. Psicología</li> <li>8. Otros</li> </ol>
Experiencia en investigación del interno de medicina humana <sup>200</sup> .	Nivel de participación previa en actividades de investigación, identificada a través de un cuestionario.		Experiencia en investigación	Cualitativa ordinal	<ol style="list-style-type: none"> <li>1. No tengo</li> <li>2. Ponencia en congreso</li> <li>3. Participación administrativa</li> <li>4. Artículo en revista</li> <li>5. Varios artículos</li> </ol>
Recepción de mentoría académica recibida de mentores durante la formación en los internos de medicina humana <sup>201</sup> .	Apoyo y orientación académica recibida de mentores durante la formación, evaluada		Participación en mentoría académica	Cualitativa nominal	<ol style="list-style-type: none"> <li>1. Estudiante</li> <li>2. Mentor</li> <li>3. Ninguno</li> </ol>



						<ol style="list-style-type: none"> <li>1. Baja</li> <li>2. Moderada</li> <li>3. Alta</li> </ol>
						<p style="text-align: center;">Pediatría</p> <ol style="list-style-type: none"> <li>1. Baja</li> <li>2. Moderada</li> <li>3. Alta</li> </ol>
						<p style="text-align: center;">Medicina interna</p> <ol style="list-style-type: none"> <li>1. Baja</li> <li>2. Moderada</li> <li>3. Alta</li> </ol>
	Complejidad del establecimiento de salud donde el interno realiza su práctica, clasificado según el nivel de atención <sup>203</sup> .	Clasificación del establecimiento de salud según su nivel de atención, obtenida mediante un cuestionario.		Nivel de atención del establecimiento de salud en el internado		<ol style="list-style-type: none"> <li>1. Primer nivel de atención</li> <li>2. Segundo nivel de atención</li> <li>3. Tercer nivel de atención</li> </ol>

	Localización geográfica del centro de salud donde el interno realiza sus prácticas <sup>204</sup> .	Localización geográfica del centro de salud donde el interno realiza sus prácticas, evaluada mediante un cuestionario.		Ubicación del establecimiento de salud en el internado		<ol style="list-style-type: none"> <li>1. Lima sur</li> <li>2. Lima norte</li> <li>3. Lima centro</li> <li>4. Lima oeste</li> <li>5. Lima este</li> <li>6. Otro departamento</li> </ol>
Desempeño académico en el Script Concordance Test	Capacidad de los internos para demostrar competencia en razonamiento clínico, evaluado mediante el Test de Concordancia de Guiones <sup>189</sup> .	Clasificación binaria de "aprobado" o "no aprobado" según el criterio de 4 desviaciones estándar (DE) por debajo de la media del panel de expertos.	Clínico-médica	Script Concordance Test	Cualitativo nominal	<ol style="list-style-type: none"> <li>1. Aprobado</li> <li>2. Desaprobado</li> </ol>
			Clínico-quirúrgica			

### **3.7. Técnicas e instrumentos de recolección de datos**

#### **3.7.1. Técnica**

La técnica de recolección de datos fue cuantitativa, de tipo encuesta, utilizando el Script Concordance Test, cuyos ítems fueron generados a través del “Script Concordance Test Generator” basado en ChatGPT-4o. Esta herramienta permitió la creación automática de escenarios clínicos diseñados específicamente para evaluar el razonamiento clínico en condiciones de incertidumbre. Esta automatización optimizó el proceso de desarrollo de ítems, manteniendo una alta calidad en la generación de estos.<sup>143</sup>

#### **3.7.2. Descripción de instrumento**

El Script Concordance Test Generator utiliza modelos de lenguaje natural, como ChatGPT, Claude, y Llama, para crear ítems de SCT, siendo seleccionada para este estudio ChatGPT-4o.<sup>143</sup> Este cuestionario estructurado se compone por ítems que se estructuran a partir de escenarios clínicos complejos en los que se presenta una hipótesis inicial, seguida de nueva información que debe ser evaluada. Los participantes califican esta nueva información en una escala Likert de -2 (mucho menos probable) a +2 (mucho más probable), según cómo afecta la hipótesis inicial.<sup>123</sup> Luego, las respuestas obtenidas se compararán con las respuestas proporcionadas por un panel de expertos para evaluar el nivel de concordancia y, así, medir la precisión y el razonamiento clínico del participante en situaciones de incertidumbre.<sup>123</sup> Esta comparación permite garantizar que las respuestas de los participantes sean evaluadas en función de los estándares clínicos establecidos por expertos.<sup>123</sup> El uso de este proceso asegura que la evaluación refleje la toma de decisiones en un contexto clínico real.<sup>123</sup>

Además, el SCT ofrece una medición detallada del desempeño académico al evaluar no solo los conocimientos teóricos de los participantes, sino también su capacidad para aplicar esos

conocimientos en situaciones de duda o ambigüedad clínica. Este enfoque permite valorar la adaptabilidad del estudiante en la toma de decisiones, así como su habilidad para integrar y ajustar hipótesis clínicas ante nueva información, lo cual es fundamental en la formación médica. A través de la comparación con un panel de expertos, el SCT garantiza que los participantes sean evaluados según los estándares clínicos más altos, reflejando su capacidad para manejar y tomar decisiones fundamentadas en evidencia en escenarios clínicos complejos.

Aunque no existen dimensiones explícitas definidas para el Script Concordance Test, este estudio ha seleccionado dos dimensiones clave basadas en su relevancia y adecuación al temario. Estas dimensiones permiten un enfoque integral para evaluar el desempeño académico de los participantes en áreas esenciales de la medicina. Las dimensiones seleccionadas son las siguientes:

**Clínico-médica:** Esta dimensión abarca la capacidad del participante para integrar y aplicar conocimientos médicos en escenarios complejos, como en medicina interna, pediatría y emergencias. En este contexto, se evaluará tanto el manejo clínico como el diagnóstico de diversas condiciones médicas, así como la habilidad para tomar decisiones informadas en situaciones de desconcierto clínico.

**Clínico-quirúrgica:** En esta dimensión, se evaluó las competencias prácticas y decisionales del participante en procedimientos quirúrgicos, con un enfoque en cirugía general y gineco-obstetricia. Al igual que en la dimensión clínico-médica, se valorará el manejo y diagnóstico de las condiciones quirúrgicas, así como la habilidad para tomar decisiones rápidas y seguras durante intervenciones quirúrgicas.

Estas dimensiones han sido elegidas por su alineación con el temario y la naturaleza del SCT, permitiendo una evaluación más precisa del desempeño académico de los participantes

en contextos médicos y quirúrgicos. Este enfoque adoptado garantiza que la evaluación refleje competencias clave para la práctica médica, brindando una medición detallada y práctica del razonamiento clínico en escenarios reales.

### **Punto de cohorte aprobado/no aprobado**

En el contexto de la evaluación mediante el SCT, no existe una clasificación estándar o un punto de cohorte universalmente aceptado para definir si un puntaje es "aprobado" o "no aprobado". Sin embargo, un criterio razonable y validado en estudios previos con estudiantes de pregrado,<sup>192,205</sup> que presentan características similares a nuestro estudio, es establecer el umbral de aprobación/desaprobación en 4 desviaciones estándar por debajo de la media del panel de expertos. Para eso mediante SPSS V.28, obtuvimos la media (77.98) y desviación estándar (6.41) de los puntajes del panel experto. Finalmente, con estos datos se procedió a calcular la cohorte para los puntajes.

#### **Fórmula:**

$$\text{Umbral} = \mu - (4 \cdot \sigma)$$

**Media del panel de expertos ( $\mu$ )**

**Desviación estándar ( $\sigma$ )**

$$\text{Umbral} = 77.98 - (4 \cdot 6.41)$$

$$\text{Umbral} = 52.34$$

El cálculo resultó en un umbral de 52.34, lo que implica que todos los estudiantes con una puntuación mayor o igual a este valor serían considerados aprobados, mientras que aquellos con una puntuación inferior serían clasificados como desaprobados.

### **3.7.3. Validación**

El panel de expertos revisó los ítems generados por el Script Concordance Test Generator para asegurar que los escenarios y las preguntas reflejen adecuadamente las incertidumbres clínicas. Como se muestra en la tabla 5, la validez del panel se evaluó mediante el coeficiente de concordancia W de Kendall, obteniendo un valor de 0.413 con un  $p < 0.001$ . Aunque la concordancia observada fue moderada, el resultado es estadísticamente significativo, lo que indica que los evaluadores lograron una alineación suficiente como para que los resultados sean considerados relevantes y confiables.

**Tabla 5.** Coeficiente de concordancia de Kendall (W) calculado con SPSS.

<b>Estadísticos de prueba</b>	
<b>N</b>	12
<b>W de Kendall<sup>a</sup></b>	.413
<b>Chi-cuadrado</b>	257.842
<b>gl</b>	52
<b>Sig.asin.</b>	<.001

a. Coeficiente de concordancia de Kendall

### 3.7.4 Confiabilidad

Como se muestra en la tabla 6, la confiabilidad del instrumento fue evaluada mediante el coeficiente Alfa de Cronbach en el panel experto, obteniendo un valor de  $\alpha = 0.839$  con los 53 ítems totales, lo que indica una buena consistencia interna entre los ítems del instrumento de medición.

**Tabla 6.** Cálculo del alfa de Cronbach utilizando SPSS.

<b>Estadísticas de fiabilidad</b>	
Alfa de Cronbach	N de elementos
0.839	53

Se llevó a cabo el cálculo de la correlación ítem-total para todo el conjunto del test (tabla 7), con el fin de evaluar la consistencia interna de los ítems en relación con la puntuación total. A pesar de que algunos ítems presentaron correlaciones negativas, estas no alcanzaron un nivel suficientemente alto como para sugerir que su eliminación podría mejorar significativamente el valor del alfa de Cronbach. En otras palabras, aunque los ítems negativos fueron identificados, su eliminación no tuvo un impacto relevante en el aumento de la fiabilidad global del test, lo que sugiere que el alfa de Cronbach se mantiene estable incluso al mantener esos ítems.

**Tabla 7.** Cálculo de la correlación ítem-total para todo el conjunto del test con SPSS.

<b>Estadísticas de total de elemento</b>					
	Media de escala si el elemento se ha suprimido	Varianza de escala si el elemento se ha suprimido	Correlación total de elementos corregida	Correlación múltiple al cuadrado	Alfa de Cronbach si el elemento se ha suprimido
VAR00001	181.5833	317.902	.452	.	.832
VAR00002	183.1667	340.515	-.090	.	.846
VAR00006	180.8333	318.333	.582	.	.831
VAR00007	183.2500	324.386	.256	.	.837
VAR00008	181.5833	316.447	.425	.	.833
VAR00009	181.0833	328.083	.355	.	.836
VAR00010	181.6667	319.697	.438	.	.833
VAR00011	182.3333	325.697	.220	.	.838
VAR00012	181.4167	331.356	.164	.	.838
VAR00013	181.1667	328.697	.156	.	.840
VAR00014	181.7500	317.477	.375	.	.834
VAR00015	181.1667	326.697	.443	.	.835
VAR00016	182.6667	328.970	.150	.	.840
VAR00017	181.4167	325.356	.208	.	.839
VAR00018	182.6667	325.333	.290	.	.836
VAR00019	182.6667	319.333	.358	.	.834
VAR00020	180.7500	333.295	.224	.	.838
VAR00021	181.2500	318.023	.399	.	.833
VAR00023	180.7500	327.114	.553	.	.834
VAR00024	183.1667	332.152	.102	.	.840
VAR00025	180.9167	332.083	.154	.	.839

VAR00028	182.0833	329.538	.251	.	.837
VAR00029	181.6667	303.879	.633	.	.826
VAR00030	181.2500	314.023	.564	.	.830
VAR00031	182.4167	324.992	.229	.	.838
VAR00032	181.0000	327.818	.212	.	.838
VAR00034	182.5000	340.091	-.094	.	.843
VAR00035	182.5000	332.091	.107	.	.840
VAR00036	182.7500	321.114	.673	.	.831
VAR00037	181.5000	337.364	-.003	.	.841
VAR00038	181.4167	322.265	.437	.	.833
VAR00039	180.6667	331.333	.333	.	.837
VAR00040	182.9167	318.083	.690	.	.830
VAR00041	181.6667	312.061	.594	.	.829
VAR00042	181.5833	330.083	.253	.	.837
VAR00043	181.5833	321.174	.321	.	.835
VAR00044	180.9167	324.992	.277	.	.836
VAR00046	182.1667	325.061	.328	.	.835
VAR00047	181.5833	337.356	-.023	.	.844
VAR00048	183.1667	325.788	.278	.	.836
VAR00049	180.6667	326.970	.570	.	.834
VAR00050	182.7500	313.841	.478	.	.831
VAR00051	180.8333	330.515	.377	.	.836
VAR00052	181.0000	332.000	.192	.	.838
VAR00054	181.0000	335.455	.132	.	.839
VAR00055	181.9167	324.992	.255	.	.837
VAR00056	181.4167	340.811	-.106	.	.844
VAR00057	180.7500	320.023	.525	.	.832
VAR00058	181.0833	337.902	-.015	.	.840
VAR00059	182.8333	325.606	.283	.	.836
VAR00061	182.4167	329.538	.159	.	.839
VAR00062	182.2500	318.932	.317	.	.836
VAR00063	181.5000	337.000	-.007	.	.842

### 3.8. Plan de procesamiento y análisis de datos

El procesamiento y análisis de datos en esta investigación se llevó a cabo mediante un enfoque cuantitativo, basado en la recolección de datos a través del Script Concordance Test aplicado a internos de medicina humana.

Los datos recolectados físicamente fueron transcritos y almacenados inicialmente en Microsoft Excel v. 2020, donde se organizaron y realizaron las primeras limpiezas de datos, como la detección de valores atípicos o datos faltantes. Posteriormente, a partir de las respuestas en la escala de Likert proporcionadas por cada estudiante, se procedió a calcular su puntaje individual utilizando el método descrito previamente en el marco teórico de este estudio. Dado que no existe de manera pública una herramienta automatizada para obtener estos puntajes, se desarrolló un procedimiento más eficiente empleando Excel v. 2020. Este método incluyó la aplicación de fórmulas, razones y macros, lo que permitió agilizar el cálculo de los puntajes de manera más rápida y precisa.

Los datos codificados fueron exportados a SPSS v.28 para su análisis estadístico, con un intervalo de confianza de 95% y una tasa de error de 5%. Se utilizaron medidas de tendencia central y dispersión para describir las variables sociodemográficas, académicas y entorno asistenciales.

Se realizaron comparaciones entre las puntuaciones obtenidas por los internos de medicina humana y las respuestas generadas por ChatGPT-4o, ambas en relación con el panel de expertos. Se evaluó la confiabilidad del SCT mediante el coeficiente alfa de Cronbach, aplicado tanto al panel de expertos como a los estudiantes. Además, se aplicó el test de normalidad de Kolmogorov-Smirnov a ambos grupos, obteniéndose una distribución no normal en ambos casos. Por este motivo, se empleó la prueba de Levene para verificar la homogeneidad de las varianzas y, posteriormente, la prueba U de Mann-Whitney para comparar el desempeño académico mediante el SCT entre el panel de expertos y los estudiantes, ya que esta es adecuada para distribuciones no normales. Adicionalmente, se calculó el tamaño del efecto  $r$  basado en el estadístico  $Z$  de la U de Mann-Whitney para evaluar la magnitud de las diferencias entre los grupos.

Se realizó la prueba de hipótesis utilizando el estadístico de chi cuadrado para analizar si existía asociación entre la variable 1 (Desempeño académico en el Script Concordance Test) y la variable 2 (Factores asociados al razonamiento clínico), considerando un valor significativo de  $p < 0.05$ .

### 3.9. Aspectos éticos

La presente investigación tuvo validez tanto científica como ética, ya que se respetaron los principios bioéticos fundamentales de beneficencia, no maleficencia, autonomía y justicia<sup>206</sup> durante todo el proceso de diseño y ejecución del estudio. Estos principios fueron implementados siguiendo las referencias científicas relevantes en el contexto del Script Concordance Test, la evaluación del razonamiento clínico en internos de medicina humana y el uso de herramientas de inteligencia artificial, como ChatGPT, en la creación de los casos clínicos.

1. **Beneficencia:** La investigación se orienta a contribuir al fortalecimiento de la educación médica en el internado, enfocado específicamente en el desarrollo de habilidades de razonamiento clínico entre los internos de medicina. Al evaluar estas competencias mediante el SCT, se espera que los resultados proporcionen información valiosa para mejorar las metodologías de enseñanza y evaluación en la etapa avanzada de formación médica. Teniendo un impacto directo en la preparación de los internos para su futura práctica profesional, especialmente en el contexto del SERUMS, optimizando su capacidad de respuesta ante situaciones clínicas habituales en atención primaria. Además, el uso de ChatGPT-4o como herramienta para generar casos clínicos permitirá incrementar la complejidad y diversidad de los escenarios evaluativos.

2. **No maleficencia:** El estudio se diseñó para evitar cualquier tipo de daño físico o psicológico a los participantes. La implementación del Script Concordance Test y el uso de ChatGPT-4o no representan un riesgo para los sujetos, ya que ambos fueron utilizados con fines educativos y de investigación, en un entorno controlado y validado científicamente.
3. **Autonomía:** Los participantes recibieron información completa sobre los objetivos, métodos y beneficios del estudio, garantizando su derecho a decidir voluntariamente su participación.
4. **Justicia:** La selección de los participantes fue equitativa, asegurando que todos los internos de medicina humana que cumplan con los criterios de inclusión puedan formar parte del estudio. No se realizarán exclusiones arbitrarias y se garantizó que los resultados beneficiarán por igual a todos los participantes.

Para la ejecución de la presente investigación, se solicitó la aprobación del Comité de Ética de la Universidad Norbert Wiener, con el fin de cumplir con los estándares éticos antes de proceder con la recolección de datos.

### **Confidencialidad de los Datos**

La recolección de datos mediante el Script Concordance Test respetó estrictamente la confidencialidad de los participantes. A cada interno se le asignó un código único generado aleatoriamente para proteger su identidad. Además, los casos clínicos generados por ChatGPT-4o fueron revisados y validados por expertos antes de ser utilizados en el estudio. Los datos fueron almacenados de manera segura en Microsoft Excel, protegido con contraseña, lo que garantizará que solo los investigadores autorizados tengan acceso a la información.

## CAPÍTULO IV: RESULTADOS Y DISCUSIÓN

### 4.1. Resultados

#### 4.1.1. Presentación de resultados descriptivos

##### 4.1.1.1 Análisis sociodemográfico

**Tabla 8.** Características sociodemográficas de estudiantes internos de una escuela de medicina humana, Universidad Privada Norbert Wiener, Lima 2024.

Características sociodemográficas	N (%)
Sexo	
Masculino	28 (32.1)
Femenino	59 (67.8)
Edad *	27.48 ± 4.75

\* Media ± desviación estándar

La tabla 8 presenta la distribución de los 87 participantes evaluados, donde el 67.8% pertenecía al sexo femenino y el 32.1% al sexo masculino, evidenciando una mayor representación de mujeres en la población estudiada. La edad promedio de los participantes fue de 27.48 años, con una desviación estándar de 4.75 años, reflejando una ligera variabilidad en los rangos etarios dentro del grupo analizado.

#### 4.1.1.2. Análisis académicos

**Tabla 9.** Características académicas de estudiantes internos de una escuela de medicina humana, Universidad Privada Norbert Wiener, Lima 2024.

Características académicas	N (%)
Estudios o carrera previa	
Ninguno	58 (66.6)
Enfermería	3 (3.4)
Obstetricia	1 (1.1)
Nutrición	1 (1.1)
Odontología	0 (0.0)
Terapia Física y Rehabilitación	0 (0.0)
Psicología	0 (0.0)
Otros	24 (27.5)
Experiencia en investigación	
No tengo	69 (78.3)
Ponencia en congreso	4 (4.6)
Participación administrativa	3 (3.4)
Artículo en revista	9 (10.3)
Varios artículos	2 (2.3)
Participación en mentoría académica	
Estudiante	42 (48.2)
Mentor	5 (5.7)
Ninguno	40 (45.9)
Prácticas clínicas extracurriculares	
Nula	33 (37.9)
Escasa	23 (26.4)
Moderada	26 (29.8)
Considerable	5 (5.7)

La tabla 9 presenta las características académicas de los estudiantes evaluados. Se observó que la mayoría no contaba con estudios o carreras previas (66.6%), mientras que un pequeño porcentaje indicó haber cursado Enfermería (3.4%), Obstetricia (1.1%) y Nutrición (1.1%), mientras que un 27.5% reportó haber estudiado otras

carreras diferentes a las mencionadas. Respecto a la experiencia en investigación, el 78.3% afirmó no tener experiencia, mientras que un 10.3% reportó haber publicado artículos en revistas científicas, un 4.6% participó en congresos, un 3.4% en roles administrativos y un 2.3% indicó contar con varios artículos publicados. En términos de participación en mentoría académica, el 48.2% se identificó como estudiantes, el 5.7% como mentores y el 45.9% no participó en este tipo de actividades. Finalmente, en relación con las prácticas clínicas extracurriculares, el 37.9% reportó no tener experiencia, el 26.4% indicó escasa participación, el 29.8% señaló un nivel moderado y el 5.7% describió una participación considerable.

#### 4.1.1.3. Análisis entorno asistenciales

**Tabla 10.** Características entorno asistenciales de estudiantes internos de una escuela de medicina humana, Universidad Privada Norbert Wiener, Lima 2024.

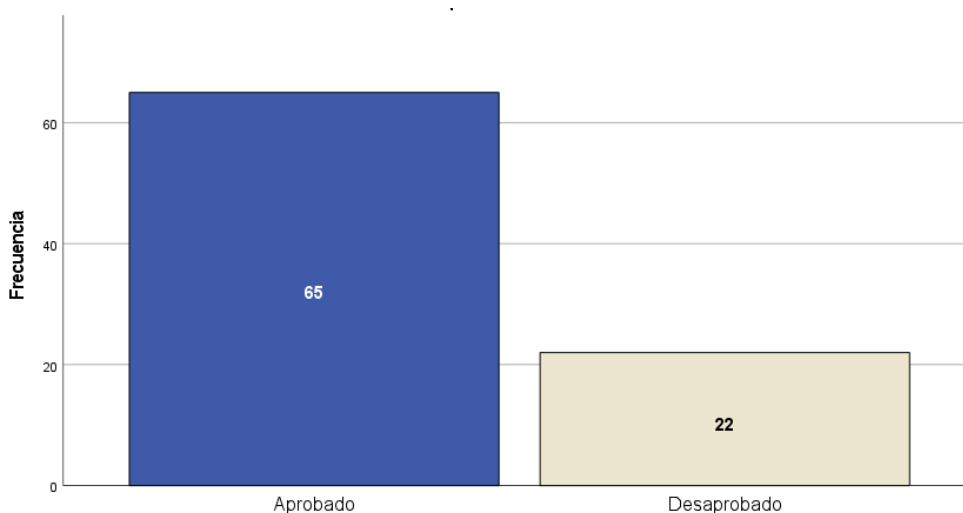
Características entorno asistenciales	N (%)		
Supervisión médica en el internado			
Nunca	12 (13.7)		
Raramente	7 (8.0)		
A veces	25 (28.7)		
Frecuentemente	29 (33.3)		
Siempre	14 (16.0)		
Retroalimentación del personal médico en el internado			
	Baja	Moderada	Alta
Cirugía	33 (37.9)	36 (41.3)	18 (20.6)
Gineco-obstetricia	10 (37.9)	43 (41.3)	34 (20.6)
Pediatría	14 (16.0)	36 (41.3)	37 (42.5)
Medicina interna	23 (16.0)	36 (42.5)	28 (41.3)
Nivel de atención del establecimiento de salud en el internado			
Primer nivel de atención	45 (51.7)		
Segundo nivel de atención	9 (10.3)		
Tercer nivel de atención	33 (37.9)		
Ubicación del establecimiento de salud en el internado			
Lima sur	13 (14.9)		
Lima norte	15 (17.2)		
Lima centro	44 (50.5)		
Lima oeste	0 (0.0)		
Lima este	12 (13.79)		
Otro departamento	3 (3.45)		

La tabla 10 presenta los datos sobre la supervisión médica y la retroalimentación recibida durante el internado. El 33.3% de los participantes señaló haber recibido supervisión médica frecuentemente, el 28.7% indicó que esta ocurría a veces, el 16.0% reportó que siempre estuvo supervisado, el 13.7% manifestó nunca haber sido supervisado y el 8.0% mencionó que esta ocurría raramente. En cuanto a la retroalimentación del personal médico, se observaron diferencias según la especialidad: en Cirugía, el 37.9% calificó la retroalimentación como baja, el 41.3%

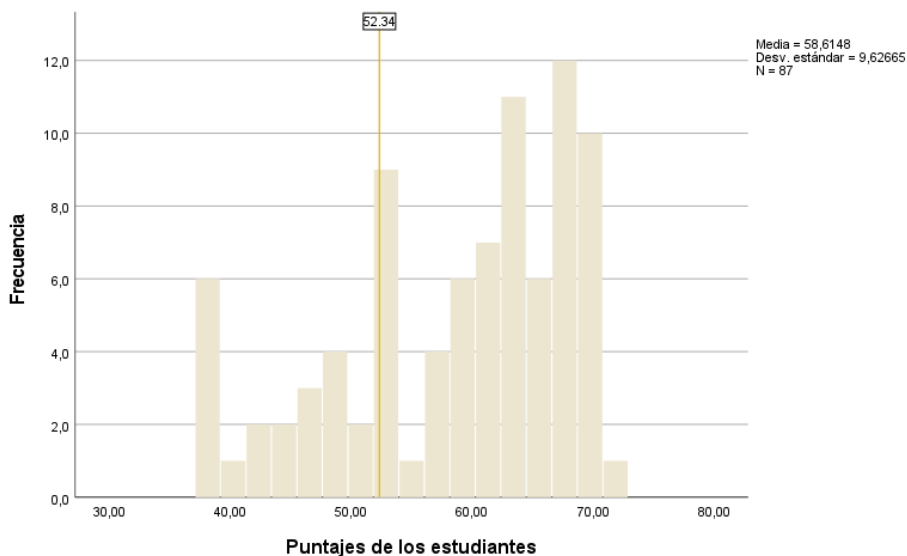
como moderada y el 20.6% como alta; en Gineco-obstetricia, los valores fueron 37.9% baja, 41.3% moderada y 20.6% alta; en Pediatría, el 16.0% reportó una retroalimentación baja, el 41.3% moderada y el 42.5% alta; mientras que en Medicina Interna, los porcentajes fueron 16.0% baja, 42.5% moderada y 41.3% alta. Respecto al nivel de atención del establecimiento de salud donde se desarrolló el internado, el 51.7% de los participantes realizó sus prácticas en el primer nivel de atención, el 10.3% en el segundo nivel y el 37.9% en el tercer nivel. Finalmente, en cuanto a la ubicación geográfica de los establecimientos de salud, la mayoría estuvo en Lima Centro (50.5%), seguido de Lima Norte (17.2%), Lima Sur (14.9%), Lima Este (13.79%), mientras que un 3.45% realizó su internado en otro departamento y ninguno reportó haber estado en Lima Oeste.

#### 4.1.1.4. Cohorte aprobado/no aprobado del Script Concordance Test

**Figura 4.** Cohorte aprobado/no aprobado total.



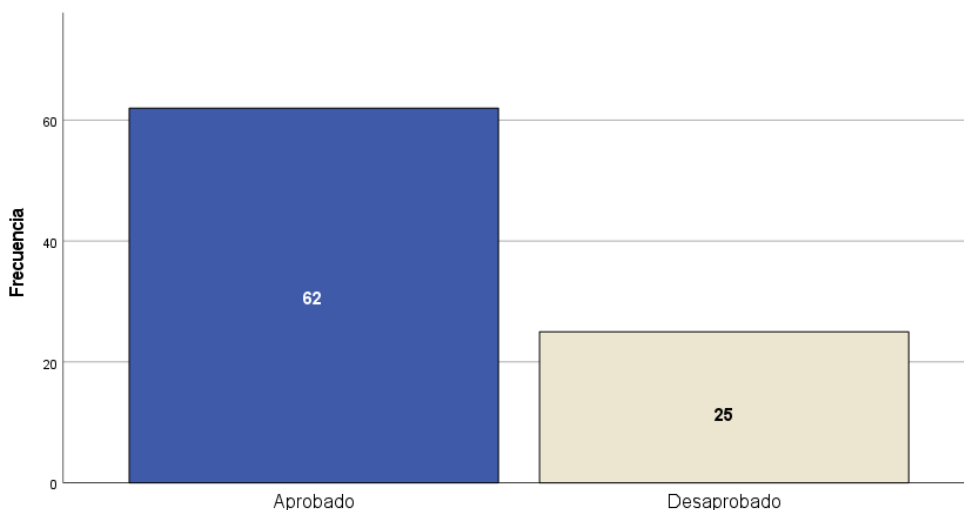
**Figura 5.** Distribución de frecuencias de las puntuaciones de los estudiantes.



La figura 4 y la figura 5 presentan el análisis de los puntajes del Script Concordance Test, con una cohorte de 52.34. Se observó que de los 87 estudiantes, el 74.7% (n=65) fueron clasificados como "Aprobado", mientras que el 25.3% (n=22) correspondieron a la categoría de "Desaprobado". Las puntuaciones oscilaron entre un valor mínimo de 37.61 y un máximo de 71.74, reflejando la distribución de los participantes en relación con el desempeño obtenido en el total de la prueba.

**4.1.1.4.1 Dimensión clínico-médica**

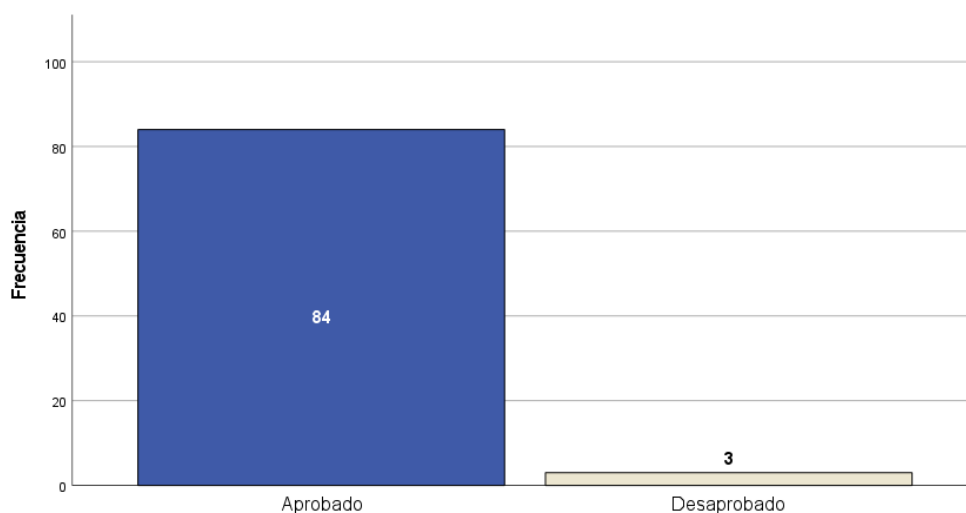
**Figura 6.** Dimensión 1: clínico-médica



La figura 6 presenta el análisis de los puntajes correspondientes a la dimensión clínico-médica del Script Concordance Test. Se observó que, de los 87 estudiantes evaluados, el 71.3% (n=62) fueron clasificados como "Aprobado", mientras que el 28.7% (n=25) correspondieron a la categoría de "Desaprobado". Estos resultados reflejan la distribución de los participantes en relación con su desempeño en esta dimensión específica de la prueba.

#### 4.1.1.4.2 Dimensión clínico-quirúrgica

**Figura 7.** Dimensión 2: clínico-quirúrgica



La figura 7 presenta el análisis de los puntajes correspondientes a la dimensión clínico-quirúrgica del SCT. Se observó que, de los 87 estudiantes evaluados, el 96.6% (n=84) fueron clasificados como "Aprobado", mientras que el 3.4% (n=3) correspondieron a la categoría de "Desaprobado".

#### 4.1.1.5. Presentación de análisis estadístico del Script Concordance Test

##### 4.1.1.5.1. Confiabilidad del SCT en la muestra

**Tabla 11.** Cálculo del Alfa de Cronbach para los estudiantes

<b>Estadísticas de fiabilidad</b>	
Alfa de Cronbach	N de elementos
0.907	53

La tabla 11 presenta la evaluación de la confiabilidad para los estudiantes mediante el coeficiente Alfa de Cronbach, obteniéndose un valor de 0.907 para los 53 ítems del test. Este resultado indica una alta consistencia interna, lo que sugiere que los ítems presentan una adecuada correlación interna y miden de manera consistente los constructos evaluados.

#### 4.1.1.5.2. Test de normalidad Kolmogorov-Smirnov

**Tabla 12.** Test de normalidad Kolmogorov-Smirnov para el panel experto y los estudiantes.

<b>Resultados</b>	<b>Estadístico</b>	<b>Kolmogorov-Smirnov</b>	
		<b>gl</b>	<b>Sig.</b>
<b>Total del test:</b>			
Panel experto	0.252	12	0.033
Estudiantes	0.118	87	0.004
<b>Clínico médico:</b>			
Panel experto	0.123	12	0.200
Estudiantes	0.138	87	<0.001
<b>Clínico quirúrgico:</b>			
Panel experto	0.182	12	0.200
Estudiantes	0.125	87	0.002

gl: Grados de libertad (tamaño de la muestra)

La tabla 12 presenta la evaluación de la normalidad de los puntajes del Script Concordance Test mediante la prueba de Kolmogorov-Smirnov, aplicada tanto al panel de expertos como a los estudiantes en las diferentes dimensiones evaluadas. En el total del test, el estadístico fue de 0.252 para el panel experto (gl=12, p=0.033) y de 0.118 para los estudiantes (gl=87, p=0.004), indicando que los puntajes no siguen una distribución normal en ambas cohortes. En la dimensión clínico-médica, el panel

experto presentó un estadístico de 0.123 (gl=12, p=0.200), sugiriendo una distribución normal, mientras que los estudiantes obtuvieron un estadístico de 0.138 (gl=87, p<0.001), indicando desviación de la normalidad. Finalmente, en la dimensión clínico-quirúrgica, el panel experto registró un estadístico de 0.182 (gl=12, p=0.200), compatible con una distribución normal, mientras que los estudiantes alcanzaron un estadístico de 0.125 (gl=87, p=0.002), confirmando la ausencia de normalidad en sus puntajes.

#### 4.1.1.5.3. Análisis descriptivo de la distribución de las puntuaciones SCT

**Tabla 13.** Análisis descriptivo de la distribución de las puntuaciones SCT para el panel experto y los estudiantes.

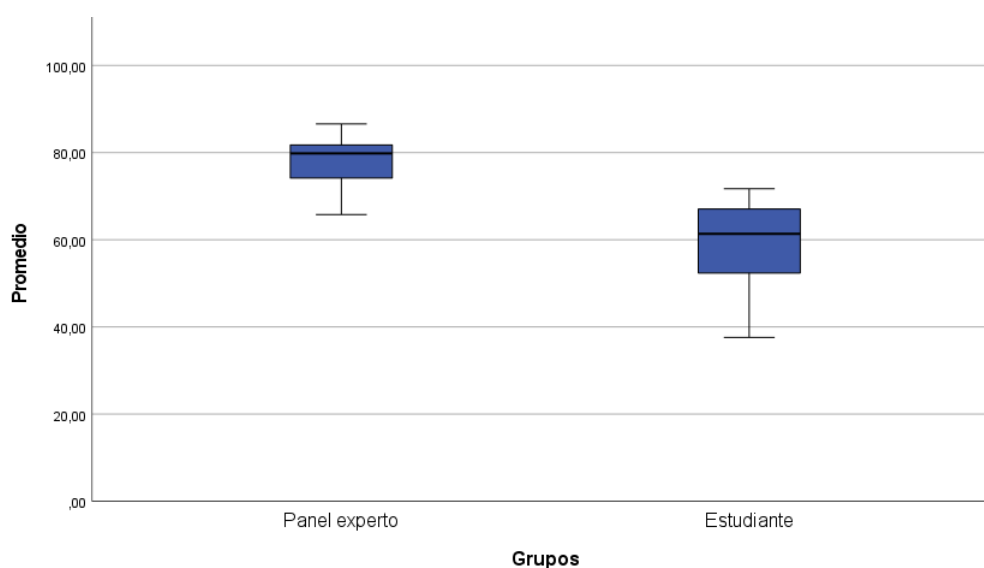
Resultados	n	Media	DE	Test de Levene		U de Mann-Whitney	Efecto r
				F	Sig.	Sig.	r
<b>Total del test:</b>							
Panel experto	12	77.98	6.41	3.815	.054	<.001	-0.521
Estudiantes	87	58.61	9.62				
<b>Clínico médico:</b>							
Panel experto	12	77.40	6.58	3.424	.067	<.001	-0.541
Estudiantes	87	57.17	9.15				
<b>Clínico quirúrgico:</b>							
Panel experto	12	79.76	10.50	1.119	.293	<.001	0.369
Estudiantes	87	63.03	14.52				

La tabla 13 presenta el análisis descriptivo de las puntuaciones obtenidas en el Script Concordance Test por el panel experto y los estudiantes, considerando tanto el puntaje total como el de las dimensiones evaluadas. En el puntaje total del test, el panel experto presentó una media de 77.98 (DE=6.41), mientras que los estudiantes obtuvieron una media de 58.61 (DE=9.62). La prueba de Levene mostró homogeneidad marginal de varianzas (F=3.815, p=0.054), y la U de Mann-Whitney evidenció diferencias estadísticamente significativas entre los grupos (p<0.001), con un efecto moderado (r=-0.521). En la dimensión clínico-médica, el panel experto alcanzó una media de

77.40 (DE=6.58), en comparación con la media de 57.17 (DE=9.15) obtenida por los estudiantes. Aunque la prueba de Levene mostró homogeneidad marginal de varianzas ( $F=3.424$ ,  $p=0.067$ ), la U de Mann-Whitney confirmó diferencias significativas entre ambos grupos ( $p<0.001$ ), con un efecto moderado ( $r=-0.541$ ).

En la dimensión clínico-quirúrgica, el panel experto obtuvo una media de 79.76 (DE=10.50), frente a la media de 63.03 (DE=14.52) de los estudiantes. La prueba de Levene indicó homogeneidad de varianzas ( $F=1.119$ ,  $p=0.293$ ), y la U de Mann-Whitney mostró diferencias significativas ( $p<0.001$ ) con un efecto moderado ( $r=0.369$ ).

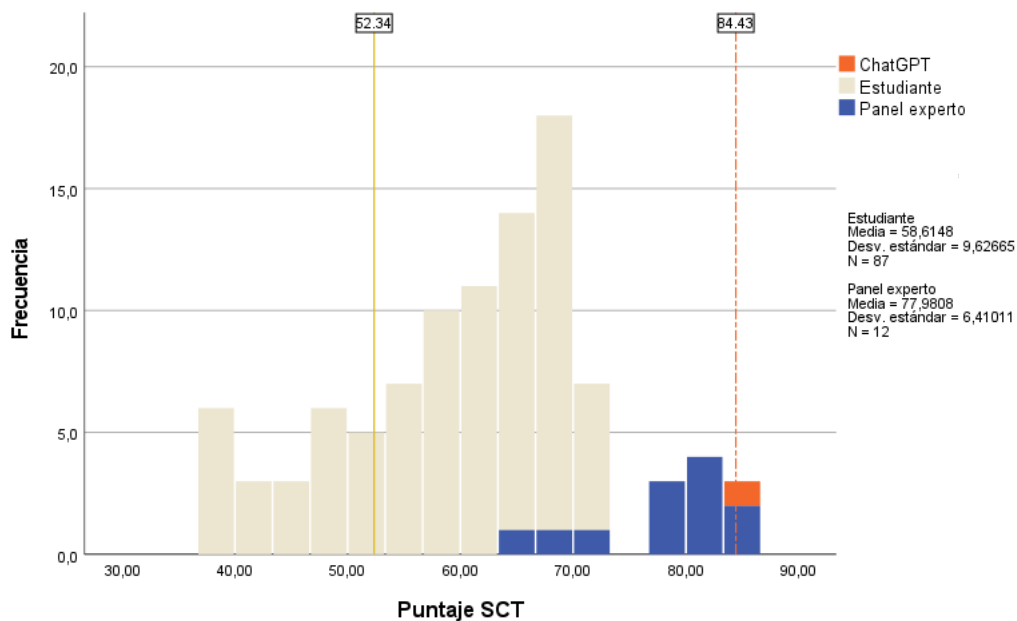
**Figura 8.** Distribución de las puntuaciones de los miembros del panel experto y de los estudiantes.



La figura 8 muestra la distribución de las puntuaciones promedio obtenidas por el panel experto y los estudiantes en el SCT. Se observa que el panel experto presenta un promedio más alto y una menor dispersión de las puntuaciones, mientras que el grupo de estudiantes muestra un promedio inferior con un rango de dispersión más amplio.

#### 4.1.1.6 Desempeño de ChatGPT-4o en el Script Concordance Test

**Figura 9.** Distribución de frecuencias de las puntuaciones del panel experto vs estudiantes vs ChatGPT-4o. La línea amarilla representa el umbral de aprobación (52.34 puntos), y la línea naranja punteada indica la puntuación de ChatGPT (84.43 puntos).



La figura 9 muestra que ChatGPT-4o obtuvo una puntuación de 84.43, superando el umbral de aprobación establecido en 52.34 puntos y posicionándose entre los puntajes más altos del panel experto. Además, alcanzó un puntaje superior a la media del panel experto (77.98, DE = 6.41) y significativamente mayor que la media de los estudiantes (58.61, DE = 9.63). Al analizar específicamente las puntuaciones del panel experto, se observa que solo un participante obtuvo un puntaje superior (86.59) al de ChatGPT-4o, mientras que los demás quedaron por debajo del modelo de inteligencia artificial.

## 4.1.2. Prueba de hipótesis

### 4.1.2.1. Hipótesis general

**H1:** Existen factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

**H0:** No existen factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

Nivel de significancia: 0.05

Estadístico de prueba: Prueba de chi cuadrado

**Tabla 14.** Análisis bivariado de los factores sociodemográficos asociados al desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

Características sociodemográficas	Calificación		p*
	Aprobado (n=65) 74.71 (%)	Desaprobado (n=22) 25.29 (%)	
Sexo			0.627
Masculino	20 (71.4)	8 (28.6)	
Femenino	45 (76.3)	14 (23.7)	
Edad**			0.317

\*Valores p calculados con la prueba Chi Cuadrado de independencia

\*\*Valor p calculado con la prueba T de Student

El análisis estadístico descrito en la tabla 14 muestra que no se encontró una asociación significativa entre el sexo y el desempeño académico en el Script Concordance Test (aprobado o desaprobado), según el valor de  $p=0.627$  obtenido con la prueba Chi Cuadrado

de independencia. Este resultado indica que el sexo de los estudiantes no influyó significativamente en la probabilidad de aprobar o desaprobar el test.

De manera similar, la prueba T de Student, utilizada para analizar la relación entre la edad y el desempeño académico en el SCT, arrojó un valor de  $p=0.317$ , lo que sugiere que tampoco se identificó una diferencia significativa en el desempeño en función de la edad.

**Tabla 15.** Análisis bivariado de los factores académicos asociados al desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

Características académicas	Calificación		p*
	Aprobado (n=65) 74.71 (%)	Desaprobado (n=22) 25.29 (%)	
Estudios o carrera previa			0.505
Ninguno	46(79.3)	12(20.7)	
Enfermería	2(66.7)	1(33.3)	
Obstetricia	1(100)	0(0.0)	
Nutrición	1(100)	0(0.0)	
Otros	15(62.5)	9(37.5)	
Experiencia en investigación			0.119
No tengo	52 (75.4)	17 (24.6)	
Ponencia en congreso	1 (25.0)	3 (75.0)	
Participación administrativa	2 (66.7)	1 (33.3)	
Artículo en revista	8 (88.9)	1 (11.1)	
Varios artículos	2 (100)	0 (0.0)	
Participación en mentoría académica			0.137
Estudiante	28 (66.7)	14 (33.3)	
Mentor	3 (60.0)	2 (40.0)	
Ninguno	34 (85.0)	6 (15.0)	
Prácticas clínicas extracurriculares			<b>0.030</b>
Nula	20 (60.6)	13 (39.4)	
Escasa	22 (95.7)	1 (4.3)	
Moderada	19 (73.1)	7 (26.9)	
Considerable	4 (80.0)	1 (20.0)	

\*Valores p calculados con la prueba Chi Cuadrado de independencia

En cuanto a los estudios o carreras previas, evidenciado en la tabla 15, no se observó una asociación significativa con el desempeño académico en el SCT ( $p=0.505$ ). Esto indica que haber cursado una carrera como Enfermería, Obstetricia, Nutrición u otros estudios previos

no influyó de manera relevante en los resultados del test, en comparación con aquellos que no tenían formación académica previa. En relación con la experiencia en investigación, aunque los participantes con múltiples artículos publicados obtuvieron mejores resultados, no se encontró una diferencia significativa entre los distintos niveles de experiencia ( $p=0.119$ ). De manera similar, la participación en mentorías académicas, ya sea como estudiante, mentor o sin participación, tampoco mostró una asociación estadísticamente significativa con el desempeño académico en el SCT ( $p=0.137$ ). Por otro lado, la participación en prácticas clínicas extracurriculares sí presentó una relación significativa ( $p=0.030$ ).

**Tabla 16.** Análisis bivariado de los factores entorno asistenciales asociados al desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

Características entorno asistenciales	Calificación		p*
	Aprobado (n=65) 74.71 (%)	Desaprobado (n=22) 25.29 (%)	
Supervisión médica en el internado			0.077
Nunca	9 (75.0)	3 (25.0)	
Raramente	3 (42.9)	4 (57.1)	
A veces	20 (80.0)	5 (20.0)	
Frecuentemente	25 (86.2)	4 (13.8)	
Siempre	8 (57.1)	6 (42.9)	
Retroalimentación del personal médico en el internado			
Cirugía			0.858
Baja	24 (72.7)	9 (27.3)	
Moderada	28 (77.8)	8 (22.2)	
Alta	13 (72.2)	5 (27.8)	
Gineco-obstetricia			0.301
Baja	8 (80.0)	2 (20.0)	
Moderada	29 (67.4)	14 (32.6)	
Alta	28 (82.4)	6 (17.6)	
Pediatria			0.884
Baja	11 (78.6)	3 (21.4)	
Moderada	26 (72.2)	10 (27.8)	
Alta	28 (75.7)	9 (24.3)	
Medicina interna			0.568
Baja	16 (69.6)	7 (30.4)	
Moderada	29 (80.6)	7 (19.4)	
Alta	20 (71.4)	8 (28.6)	
Nivel de atención del establecimiento de salud en el internado			0.725
Primer nivel de atención	33 (73.3)	12 (26.7)	
Segundo nivel de atención	6 (66.7)	3 (33.3)	
Tercer nivel de atención	26 (78.8)	7 (21.2)	
Ubicación del establecimiento de salud en el internado			0.571
Lima sur	10 (76.9)	3 (23.1)	
Lima norte	11 (73.3)	4 (26.7)	
Lima centro	34 (77.3)	10 (22.7)	
Lima este	9 (75.0)	3 (25.0)	
Otro departamento	1 (33.3)	2 (66.7)	

\*Valores p calculados con la prueba Chi Cuadrado de independencia

En la tabla 16, se presenta el análisis bivariado de los factores a del entorno asistencial asociados al desempeño académico (aprobación o desaprobación) en el Script Concordance Test. En cuanto a la supervisión médica durante el internado, no se encontró una asociación significativa con el desempeño académico ( $p=0.077$ ). Aunque los estudiantes con supervisión frecuente presentaron mayores tasas de aprobación (86.2%), las diferencias entre los niveles de supervisión no alcanzaron significancia estadística. Para la retroalimentación del personal médico, ninguna de las áreas analizadas mostró asociaciones significativas con el desempeño: Cirugía ( $p=0.858$ ), Gineco-obstetricia ( $p=0.301$ ), Pediatría ( $p=0.884$ ) y Medicina interna ( $p=0.568$ ). Aunque se observaron ligeras variaciones entre los niveles de retroalimentación (baja, moderada o alta), estas no fueron estadísticamente relevantes. El nivel de atención del establecimiento de salud tampoco estuvo asociado significativamente con el desempeño académico ( $p=0.725$ ). Los estudiantes distribuidos entre el primer, segundo y tercer nivel de atención presentaron tasas de aprobación similares. Finalmente, la ubicación geográfica del establecimiento de salud no mostró una relación significativa con los resultados del test ( $p=0.571$ ). Aunque los estudiantes en Lima Centro y Lima Sur presentaron tasas de aprobación ligeramente superiores, las diferencias no fueron estadísticamente significativas.

#### **4.1.3. Discusión**

Luego de la presente investigación y considerando la importancia del razonamiento clínico en la educación médica, este estudio representa un aporte significativo en el Perú, al ser el primero en aplicar el Script Concordance Test como herramienta para evaluar esta habilidad fundamental el país. Este trabajo destaca, además, por el desarrollo de 53 ítems con el apoyo de inteligencia artificial, marcando un paso importante hacia la integración de tecnologías innovadoras en el diseño de evaluaciones educativas.

En línea con este esfuerzo por implementar este test, el análisis de la confiabilidad interna del SCT refuerza su solidez como método de evaluación. Los resultados mostraron valores elevados del Alfa de Cronbach tanto en expertos ( $\alpha = 0.839$ ) como en estudiantes ( $\alpha = 0.907$ ), lo que refleja una excelente consistencia interna del test. Este nivel de confiabilidad respalda no solo la calidad del diseño del instrumento, sino también su validez estadística, demostrando que el SCT es una herramienta robusta, incluso cuando ha sido desarrollado parcialmente con el apoyo de inteligencia artificial. Este alto nivel de confiabilidad interna también permite observar diferencias clave entre los grupos evaluados.

Por ejemplo, la menor consistencia interna observada en los expertos podría atribuirse a la diversidad en sus enfoques y experiencia clínica, mientras que el mayor Alfa de Cronbach en los estudiantes refleja una mayor homogeneidad en sus respuestas, característica de su nivel de formación. Esto evidencia que el SCT generado con ChatGPT-4o no solo es confiable, sino también capaz de discriminar efectivamente entre niveles de experiencia. Este hallazgo puede explicarse desde la teoría de los guiones de enfermedad, que establece que los expertos construyen esquemas complejos y flexibles basados en experiencia previa, mientras que los estudiantes aún dependen de estructuras más rígidas y generalizadas para abordar los casos clínicos.<sup>207</sup> Este contraste refuerza la capacidad del SCT para captar y evaluar diferencias significativas en el razonamiento clínico a través de distintas etapas de formación profesional.

Asimismo, un estudio reportado por Kiyak YS y Emekli E. (2024) evaluó la calidad de los ítems de SCT generados por ChatGPT, específicamente con el Script Concordance Test Generator, el cual fue usado en este estudio.<sup>19</sup> Encontrando que estos ítems generados demostraban evaluar predominantemente el razonamiento clínico en lugar del recuerdo fáctico (92.50 %). Además, evidenció que el 73.33 % de los ítems generados por ChatGPT podían considerarse aceptables para su uso educativo.<sup>19</sup> El estudio de Kiyak también destacó

que el uso de modelos de inteligencia artificial para generar ítems de SCT puede reducir significativamente el tiempo y esfuerzo necesarios para su desarrollo.<sup>19</sup> No obstante, subrayó la importancia de revisar y modificar los ítems generados por inteligencia artificial antes de su implementación educativa, como fue realizado en esta tesis, en la que los SCT generados por ChatGPT fueron revisados y ajustados por un panel de expertos. Este proceso de revisión aseguró la calidad y adecuación de los ítems para evaluar el razonamiento clínico en contextos reales.

Después de analizar la confiabilidad y capacidad discriminativa de este test, es relevante discutir los hallazgos relacionados con el análisis bivariado entre los factores asociados al razonamiento clínico y su desempeño académico con el SCT. El presente estudio no encontró asociaciones significativas entre el género, la edad, la formación académica previa, la experiencia en investigación, la participación en mentorías académicas, la supervisión médica, la retroalimentación del personal médico, el nivel de atención y la ubicación del establecimiento de salud con el desempeño académico en el SCT. En cuanto al género ( $p=0.627$ ), los resultados coinciden con Iglesias Gómez et al. (2022) ( $p=0.331$ ) y Ross et al. (2023), aunque Hege et al. (2022,  $p<0.001$ ) encontró diferencias en contextos específicos, posiblemente debido a factores relacionados con los instrumentos de evaluación.<sup>22,24,66</sup> Respecto a la edad ( $p=0.317$ ), los resultados son similares a los de Iglesias Gómez et al. (2022) y Ross et al. (2023), mientras que otros estudios han reportado ventajas para estudiantes mayores debido a su experiencia previa.<sup>22,24,68</sup> La formación académica previa tampoco mostró un impacto significativo ( $p=0.505$ ), aunque investigaciones previas como Groves et al. (2003) indican ventajas iniciales que se nivelan con la formación.<sup>67</sup> En relación con la experiencia en investigación, este análisis obtuvo un valor de  $p=0.119$ , indicando una tendencia positiva sin significancia estadística, en contraste con Javedani Masroor et al.

(2024), quienes reportaron una relación significativa con habilidades prácticas del razonamiento clínico ( $p < 0.05$ ).<sup>68</sup>

La participación en mentorías académicas no mostró un impacto significativo ( $p = 0.137$ ), posiblemente porque estas no cuentan con un diseño específico centrado en el razonamiento clínico, aunque estudios como Royan et al. (2018) y Reifenrath et al. (2022) destacan su potencial para desarrollar habilidades críticas si están estructuradas adecuadamente.<sup>4,208</sup> La supervisión médica durante el internado ( $p = 0.077$ ) tampoco mostró una relación significativa, aunque podría ser más relevante en términos cualitativos, como refuerzan Wang et al. (2016).<sup>71</sup> La retroalimentación del personal médico no se asoció significativamente con el desempeño, lo cual podría depender del enfoque y estructura del proceso. Finalmente, ni la ubicación del establecimiento ( $p = 0.571$ ) ni el nivel de atención ( $p = 0.725$ ) influyeron significativamente, aunque estudios como el de Young et al. (2023) destacan que entornos rurales y complejos pueden promover habilidades avanzadas debido a la exposición a casos desafiantes.<sup>72</sup>

Estos resultados sugieren que las diferencias iniciales relacionadas con género, edad, carrera previa, experiencia en investigación, mentorías, supervisión, retroalimentación, ubicación y nivel de atención del internado tienden a diluirse en programas médicos estandarizados. Sin embargo, factores contextuales como el diseño de actividades formativas, la estructura de las mentorías y el entorno clínico podrían ser determinantes en el desarrollo del razonamiento clínico, destacando la necesidad de estrategias educativas que prioricen la práctica variada, el aprendizaje colaborativo y metodologías bien estructuradas.

En cambio, con las prácticas clínicas extracurriculares, se identificó una asociación estadísticamente significativa entre la participación en estas actividades y un mejor desempeño en razonamiento clínico, medido mediante el SCT ( $p = 0.030$ ). Este resultado

enfatisa la relevancia de estas experiencias para fortalecer las habilidades analíticas y la toma de decisiones en escenarios clínicos complejos. De manera análoga, Javedani Masroor et al. (2024) reportaron que la experiencia en medicina general se correlaciona con mayores puntuaciones en la práctica del razonamiento clínico (CR-P) ( $p < 0.001$ ).<sup>68</sup> Asimismo, Royan et al. (2018) señalaron que la participación en sesiones del programa CRE, una iniciativa educativa que expone a los estudiantes a escenarios clínicos reales bajo supervisión mentorizada, mostró una relación dosis-respuesta, con mejores resultados clínicos en áreas como la historia clínica, el examen físico y la comunicación proporcional al número de sesiones asistidas ( $p < 0.05$ ).<sup>70</sup> Permitiendo a los estudiantes desarrollar habilidades de razonamiento clínico mediante la integración de datos obtenidos en la historia clínica y el examen físico, junto con la generación de planes diagnósticos y terapéuticos.

La similitud entre estos hallazgos radica en que las prácticas clínicas permiten un aprendizaje situado, facilitando la integración de conocimientos teóricos y prácticos en contextos que simulan la práctica médica real. Este enfoque no solo favorece el desarrollo de competencias técnicas y analíticas, sino que también fortalece la confianza y autonomía necesarias para el ejercicio clínico. En el estudio de Royan et al. (2018), el programa CRE proporcionó un entorno seguro y estructurado donde los estudiantes podían interactuar con pacientes y recibir retroalimentación directa, lo cual mejoró significativamente sus habilidades en la toma de historia clínica, el examen físico y la comunicación clínica.<sup>70</sup>

Por otro lado, los resultados también evidencian que el razonamiento clínico es una habilidad que se desarrolla a través de la experiencia repetida en entornos supervisados. Según los datos de Javedani Masroor et al. (2024), el conocimiento y la práctica del razonamiento clínico están influenciados positivamente por factores como la experiencia previa en medicina general y la participación en investigaciones académicas.<sup>68</sup> Esto sugiere que los programas que integran investigación y práctica clínica temprana pueden maximizar el

impacto en el desarrollo de habilidades analíticas. Además, estudios internacionales como el de Kononowicz et al. (2020) destacan la necesidad de implementar currículos longitudinales que enseñen y evalúen de manera explícita el razonamiento clínico a lo largo de toda la formación. Este enfoque no solo permitiría abordar déficits en las habilidades analíticas y prevenir errores cognitivos, sino también facilitaría una enseñanza más contextualizada y orientada a las necesidades actuales de la práctica médica.<sup>2</sup>

En otro aspecto de la evaluación de los estudiantes, la aplicación de este test mostró que el 74.7% (n=65) de los estudiantes fueron clasificados como "Aprobado" y el 25.3% (n=22) como "Desaprobado", con un puntaje promedio de 58.61 (DE = 9.62) y un rango de 37.61 a 71.74. En comparación, el estudio de Paul Duggan y Bernard Charlin (2012), realizado con 132 estudiantes de quinto año de medicina, reportando una tasa de aprobación significativamente mayor a nuestro estudio (74.7% frente al 97%), lo que podría atribuirse a una mayor variabilidad en los desempeños individuales de los estudiantes, reflejada en un rango de puntajes más amplio (37.61 a 71.74 frente a 54.5 a 75.5).<sup>205</sup>

Del mismo modo, la media de nuestro estudio fue moderadamente menor a la reportada (58.61 frente 68.6) por estos mismos autores. Además, el puntaje de corte para la aprobación en este estudio (52.34) fue más bajo que el puntaje límite reportado por Paul Duggan y Bernard Charlin (56.5), aunque ambos utilizaron el mismo criterio metodológico de 4 desviaciones estándar por debajo de la media del panel experto.<sup>204</sup> Por otro lado, el estudio de Nazim et al. (2019) reportó en internos de medicina humana un puntaje promedio de 54.48 (DE = 12.68), mostrando una tendencia similar en la variabilidad de los puntajes observada en el presente análisis.<sup>209</sup> También se describen medias mucho mayores a nuestro estudio, por ejemplo, la descrita por van den Broek WES et al. (2012) en estudiantes de medicina de último año.<sup>210</sup> En ese estudio, los participantes obtuvieron una puntuación

promedio de 82.94 (DE = 5.01), con rangos claramente superiores a los de nuestro estudio (70.12–92.08).<sup>210</sup>

Estas diferencias podrían deberse a variaciones en la composición de las poblaciones estudiadas, como estudiantes provenientes de un contexto académico más estandarizado. Además, factores como la dificultad de los escenarios clínicos presentados y las condiciones de evaluación podrían haber influido en la distribución de los puntajes. Así como la diferencia en la composición del panel experto y número de caso-ítems.

Así también en nuestro análisis descriptivo se mostraron diferencias significativas entre el panel experto y los estudiantes. El panel experto obtuvo una media de 77.98 (DE = 6.41) en el puntaje total, en comparación con los estudiantes, quienes obtuvieron una media de 58.61 (DE = 9.62). Estos resultados contrastan con los reportados en el estudio de Mzoughi et al. (2018), donde el panel experto alcanzó una media superior de 80.8 (DE = 4.09), mientras que los estudiantes obtuvieron una media de 54.78 (DE = 9.14), reflejando una mayor brecha entre ambos grupos.<sup>193</sup> De igual manera, el estudio de Mathieu et al. (2013) que incluyó a 15 estudiantes de quinto año de medicina y con un número de ítems similares a nuestro estudio, mostraron que la puntuación media de los estudiantes en el SCT fue significativamente inferior a la de los expertos ( $61,5 \pm 11,4$  vs  $76,6 \pm 15,7$ ,  $p=0,001$ ).<sup>1</sup>

Otro aspecto a discutir fue la comparación directa entre ChatGPT-4o, el panel de expertos y los estudiantes, siendo esta la primera vez que un modelo de lenguaje como ChatGPT-4o es evaluado en un contexto de razonamiento clínico utilizando un SCT. En este análisis, ChatGPT-4o obtuvo una puntuación de 84.43, superando tanto la media del panel experto (77.98) por un margen del 8.3% como la de los estudiantes (58.61) por un 44.1%. Solo un experto alcanzó un puntaje superior al modelo, posicionándolo entre los niveles más altos de desempeño registrados. Este resultado es consistente con estudios realizados en diferentes

países, incluido Perú, en el cual Flores-Cohaila JA et al. (2023) evaluaron el Examen Nacional de Medicina (ENAM) con ChatGPT. En este caso, GPT-4 alcanzó un puntaje promedio del 82%, situándose por encima del promedio nacional de los estudiantes (70%) y mostrando un rendimiento equivalente al del decil superior de participantes humanos.<sup>138</sup>

En el ámbito internacional, ChatGPT también ha demostrado resultados destacados. En Irán, logró tasas de precisión entre el 32% y el 84%, dependiendo del idioma y del formato de las preguntas, mientras que en el USMLE alcanzó una tasa de aprobación del 86% en el Step 2 CK, superando holgadamente el umbral requerido del 60% para aprobar.<sup>211,212</sup> Esta concordancia de resultados puede explicarse por la capacidad de ChatGPT para emular razonamientos complejos a partir de patrones entrenados en grandes bases de datos médicas y científicas. Dichas bases incluyen guías clínicas, artículos científicos y bancos de preguntas, lo que le permite ofrecer respuestas precisas y consistentes en diversos escenarios clínicos. Sin embargo, aunque su desempeño es notable, es fundamental reconocer que el modelo carece de juicio clínico y experiencia contextual, lo que refuerza la necesidad de emplearlo como una herramienta complementaria bajo supervisión experta y dentro de un marco ético y metodológicamente riguroso.

En el contexto de esta investigación, el Test de Concordancia de Guiones reafirma su relevancia como herramienta clave para evaluar el razonamiento clínico en condiciones de incertidumbre, alineándose con los hallazgos presentados en esta tesis. Como se discute, el TCS no solo permite discriminar eficazmente entre niveles de experiencia clínica, sino que también refleja la integración de razonamiento analítico e intuitivo, fundamentales para el desempeño clínico.<sup>4,5,12</sup> Además, el uso de inteligencia artificial en la construcción de ítems, validado por este estudio, optimiza los recursos y produce resultados equiparables a los obtenidos por expertos, consolidando su viabilidad y eficiencia.<sup>39</sup> Estas evidencias refuerzan la posición del TCS como una herramienta indispensable en estrategias educativas

innovadoras, como las planteadas en esta tesis, para preparar a los estudiantes frente a los retos de la práctica médica actual.

## **CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES**

### **5.1. Conclusiones**

**Primero**, se identificaron los factores del razonamiento clínico asociados al desempeño académico en el Script Concordance Test, evidenciando que solo uno de ellos mostró una asociación significativa. Resaltando la importancia de fortalecer la enseñanza del razonamiento clínico mediante estrategias que integren la práctica y el uso de enfoques innovadores en la educación médica.

**Segundo**, no se halló una relación estadísticamente significativa entre los factores sociodemográficos del razonamiento clínico y el desempeño en el Script Concordance Test, dado que todas las variables evaluadas presentaron valores de  $p > 0.05$ . Esto indica que características como la edad y el sexo no influyeron de manera determinante en los resultados de la prueba.

**Tercero**, no se encontró una asociación estadísticamente significativa entre los factores académicos del razonamiento clínico y el desempeño académico en el Script Concordance Test, ya que todas las variables analizadas presentaron valores de  $p > 0.05$ .

**Cuarto**, dentro los factores del entorno asistencial del razonamiento clínico solo se identificó una asociación significativa entre la participación en prácticas clínicas extracurriculares y el desempeño en el Script Concordance Test ( $p = 0.030$ ). Este hallazgo resalta el impacto positivo de la exposición a escenarios clínicos adicionales en el desarrollo del razonamiento clínico y las competencias diagnósticas de los estudiantes.

## 5.2. Recomendaciones

**Primero**, se sugiere que los directivos de la EAP Medicina Humana y de la Facultad de Ciencias de la Salud de la Universidad Privada Norbert Wiener, prioritariamente trabajen de la mano con la Sociedad Científica de Estudiantes de Medicina de la Universidad Norbert Wiener (SOCIEM UNW), dado su rol en la promoción de actividades extracurriculares como intercambios, talleres y cursos prácticos en colaboración con diversas instituciones. Fortalecer esta relación permitirá que los estudiantes accedan a más oportunidades de aprendizaje en escenarios reales, mejorando sus habilidades clínicas y su razonamiento médico. Asimismo, se recomienda que la universidad incremente la organización de cursos, congresos y talleres enfocados en la práctica clínica, proporcionando a los estudiantes experiencias que complementen su formación teórica y refuercen su preparación para enfrentar desafíos médicos.

**Segundo**, se recomienda que la EAP Medicina Humana de la Universidad Privada Norbert Wiener incorpore evaluaciones y métodos educativos basados en evidencia, integrando el razonamiento clínico como un eje transversal desde las fases preclínicas. Integrar métodos de aprendizaje como la reflexión estructurada, la autoexplicación, las listas y mnemotécnicos para diagnósticos diferenciales, SNAPPS, guiones de enfermedades y el aprendizaje en pares. Ejecutar distintos tipos de evaluaciones como el uso del SCT, simulaciones clínicas, el ECOE y exámenes de opción múltiple rigurosos. Para finalmente implementar el Entrustable Professional Activities (EPAs), un enfoque avanzado y desafiante que permitirá evaluar competencias específicas en actividades clave que los futuros médicos deberán realizar de manera autónoma.

**Tercero**, se recomienda a los líderes del ámbito educativo en salud en Perú la realización de futuras investigaciones que evalúen el desempeño en el Script Concordance Test según los

diferentes niveles de formación académica dentro de la carrera de medicina. La literatura ha evidenciado que el razonamiento clínico varía entre los años de estudio y niveles de preparación académica, pero sería relevante validar estos hallazgos en el contexto peruano. Asimismo, dado que en el presente estudio no encontró una asociación significativa en la mayoría de los factores analizados, se sugiere ampliar la investigación explorando distintas variables. Esto permitiría identificar en qué etapas se presentan mayores brechas en la toma de decisiones clínicas y orientar estrategias pedagógicas más efectivas.

**Cuarto**, se sugiere a CONAREME, ASPEFAM y al Ministerio de Salud (MINS) reconsideren la estructura y objetividad de sus evaluaciones (ENAM, ENARM, ENCAPS), ya que estas pruebas se centran principalmente en medir el conocimiento teórico de manera directa, lo que limita la evaluación de habilidades prácticas y el razonamiento clínico en situaciones reales. Sería beneficioso explorar la implementación de una combinación de pruebas teóricas y evaluaciones basadas en escenarios clínicos simulados, que permitan evaluar no solo el conocimiento, sino también la capacidad de los estudiantes para tomar decisiones y resolver problemas en contextos dinámicos. Esta integración de diferentes tipos de pruebas ofrecería una visión más completa de las competencias de los futuros profesionales de la salud, mejorando así la preparación para el ejercicio profesional y asegurando una evaluación más alineada con los estándares internacionales de la educación médica.

**REFERENCIAS**

1. Mohd Tambéh SN, Yaman MN. Clinical reasoning training sessions for health educators—A scoping review. *Journal of Taibah University Medical Sciences*. diciembre de 2023;18(6):1480-92.
2. Kononowicz AA, Hege I, Edelbring S, Sobocan M, Huwendiek S, Durning SJ. The need for longitudinal clinical reasoning teaching and assessment: Results of an international survey. *Medical Teacher*. 2 de abril de 2020;42(4):457-62.
3. Hoffert MM, Passalacqua KD, Haftka-George A, Abreu Lanfranco O, Martin RA. A Curriculum for Enhancing Physician Teaching Skills: The Value of Physician-Educator Partnerships. *Journal of Medical Education and Curricular Development*. enero de 2021;8:238212052110320.
4. de-Sousa MR, Aguiar TRXD. Dedução, Indução e a Arte do Raciocínio Clínico na Educação Médica: Revisão Sistemática e Proposta Bayesiana. *Arquivos Brasileiros de Cardiologia*. 9 de noviembre de 2022;119(5 suppl 1):27-34.
5. Young ME, Thomas A, Lubarsky S, Gordon D, Gruppen LD, Rencic J, et al. Mapping clinical reasoning literature across the health professions: a scoping review. *BMC Med Educ*. diciembre de 2020;20(1):107.
6. Araújo B, Gomes SF, Ribeiro L. Critical thinking pedagogical practices in medical education: a systematic review. *Front Med*. 14 de junio de 2024;11:1358444.
7. Xu H, Ang BWG, Soh JY, Ponnampereuma GG. Methods to Improve Diagnostic Reasoning in Undergraduate Medical Education in the Clinical Setting: a Systematic Review. *J GEN INTERN MED*. septiembre de 2021;36(9):2745-54.
8. Kün-Darbois JD, Annweiler C, Lerolle N, Lebdaï S. Script concordance test acceptability and utility for assessing medical students' clinical reasoning: a user's survey and an institutional prospective evaluation of students' scores. *BMC Medical Education*. 13 de abril de 2022;22(1):277.
9. Karila L, François H, Monnet X, Noel N, Roupret M, Gajdos V, et al. Le test de concordance de script : un outil pédagogique multimodal. *La Revue de Médecine Interne*. julio de 2018;39(7):566-73.
10. Thiessen N, Fischer MR, Huwendiek S. Assessment methods in medical specialist assessments in the DACH region – overview, critical examination and recommendations for further development. *GMS Journal for Medical Education*; 36(6):Doc78 [Internet]. 15 de noviembre de 2019 [citado 16 de julio de 2024]; Disponible en: <https://www.egms.de/en/journals/zma/2019-36/zma001286.shtml>
11. Atayee RS, Lockman K, Brock C, Abazia DT, Brooks TL, Pawasauskas J, et al. Multicentered Study Evaluating Pharmacy Students' Perception of Palliative Care and Clinical Reasoning Using Script Concordance Testing. *Am J Hosp Palliat Care*. noviembre de 2018;35(11):1394-401.

12. Nouh T, Boutros M, Gagnon R, Reid S, Leslie K, Pace D, et al. The script concordance test as a measure of clinical reasoning: a national validation study. *The American Journal of Surgery*. abril de 2012;203(4):530-4.
13. Nomura O, Itoh T, Mori T, Ihara T, Tsuji S, Inoue N, et al. Creating Clinical Reasoning Assessment Tools in Different Languages: Adaptation of the Pediatric Emergency Medicine Script Concordance Test to Japanese. *Front Med*. 7 de diciembre de 2021;8:765489.
14. Hamui M. Prueba de concordancia de script: primera experiencia en pediatría de alcance nacional. *Arch Argent Pediatr* [Internet]. 1 de febrero de 2018 [citado 19 de junio de 2024];116(1). Disponible en: <http://www.sap.org.ar/docs/publicaciones/archivosarg/2018/v116n1a43.pdf>
15. Roberti A, Roberti MDRF, Pereira ERS, Costa NMDSC. Script concordance test in medical schools in Brazil: possibilities and limitations. *Sao Paulo Med J*. 19 de enero de 2016;134(2):116-20.
16. Pleguezuelos EM, Hornos E, Dory V, Gagnon R, Malagrino P, Brailovsky CA, et al. Script concordance testing in continuing professional development: local or international reference panels? *Adv in Health Sci Educ*. octubre de 2013;18(4):779-85.
17. Kiyak YS, Emekli E. A Prompt for Generating Script Concordance Test Using ChatGPT, Claude, and Llama Large Language Model Chatbots. *Rev Esp Edu Med* [Internet]. 15 de mayo de 2024 [citado 18 de junio de 2024];5(3). Disponible en: <https://revistas.um.es/edumed/article/view/612381>
18. Hudon A, Kieपुरa B, Pelletier M, Phan V. Using ChatGPT in Psychiatry to Design Script Concordance Tests in Undergraduate Medical Education: Mixed Methods Study. *JMIR Med Educ*. 4 de abril de 2024;10:e54067-e54067.
19. Kiyak YS, Emekli E. Uso de modelos de lenguaje de gran tamaño para generar pruebas de concordancia de guiones en la educación médica: ChatGPT y Claude. *Rev Esp Edu Med* [Internet]. 3 de diciembre de 2024 [citado 17 de enero de 2025];6(1). Disponible en: <https://revistas.um.es/edumed/article/view/636331>
20. Nalini Y, Manivasakan S, Pai DR. Comparison between MCQ, Extended matching questions (EMQ) and Script concordance test (SCT) for assessment among first-year medical students – A pilot study. *Journal of Education and Health Promotion* [Internet]. febrero de 2024 [citado 16 de julio de 2024];13(1). Disponible en: [https://journals.lww.com/10.4103/jehp.jehp\\_839\\_23](https://journals.lww.com/10.4103/jehp.jehp_839_23)
21. Naylor K, Hislop J, Torres K, Mani ZA, Goniewicz K. The Impact of Script Concordance Testing on Clinical Decision-Making in Paramedic Education. *Healthcare*. 22 de enero de 2024;12(2):282.
22. Ross L, Semaan E, Gosling CM, Fisk B, Shannon B. Clinical reasoning in undergraduate paramedicine: utilisation of a script concordance test. *BMC Med Educ*. 19 de enero de 2023;23(1):39.

23. Redmond C, Jayanth A, Beresford S, Carroll L, Johnston ANB. Development and validation of a script concordance test to assess biosciences clinical reasoning skills: A cross-sectional study of 1st year undergraduate nursing students. *Nurse Education Today*. diciembre de 2022;119:105615.
24. Iglesias Gómez C, González Sequeros O, Salmerón Martínez D. Clinical reasoning evaluation using script concordance test in primary care residents. *Anales de Pediatría (English Edition)*. agosto de 2022;97(2):87-94.
25. Moghadami M, Amini M, Moghadami M, Dalal B, Charlin B. Teaching clinical reasoning to undergraduate medical students by illness script method: a randomized controlled trial. *BMC Med Educ*. diciembre de 2021;21(1):87.
26. Cohen Aubart F, Papo T, Hertig A, Renaud MC, Steichen O, Amoura Z, et al. Are script concordance tests suitable for the assessment of undergraduate students? A multicenter comparative study. *La Revue de Médecine Interne*. abril de 2021;42(4):243-50.
27. Peyrony O, Hutin A, Truchot J, Borie R, Calvet D, Albaladejo A, et al. Impact of panelists' experience on script concordance test scores of medical students. *BMC Med Educ*. diciembre de 2020;20(1):313.
28. Ottolini MC, Chua I, Campbell J, Ottolini M, Goldman E. Pediatric Hospitalists' Performance and Perceptions of Script Concordance Testing for Self-Assessment. *Academic Pediatrics*. marzo de 2021;21(2):252-8.
29. SADEGHI A, ALI ASGARI A, MOULAEI N, MOHAMMADKARIMI V, DELAVARI S, AMINI M, et al. Combination of different clinical reasoning tests in a national exam. *Journal of Advances in Medical Education & Professionalism [Internet]*. octubre de 2019 [citado 16 de julio de 2024];7(4). Disponible en: <https://doi.org/10.30476/jamp.2019.83101.1083>
30. Lineberry M, Hornos E, Pleguezuelos E, Mella J, Brailovsky C, Bordage G. Experts' responses in script concordance tests: a response process validity investigation. *Medical Education*. julio de 2019;53(7):710-22.
31. Subra J, Chicoulaa B, Stillmunkès A, Mesthé P, Oustric S, Rougé Bugat ME. Reliability and validity of the script concordance test for postgraduate students of general practice. *European Journal of General Practice*. 2 de octubre de 2017;23(1):209-14.
32. Sibert L, Charlin B, Corcos J, Gagnon R, Grise P, Vleuten CVD. Stability of clinical reasoning assessment results with the Script Concordance test across two different linguistic, cultural and learning environments. *Medical Teacher*. enero de 2002;24(5):522-7.
33. Echeverry-Díaz T, Restrepo Escobar JA. Prueba de razonamiento clínico script: Una mirada desde la práctica educativa en medicina interna. *Acta Med Col [Internet]*. 1 de septiembre de 2022 [citado 26 de agosto de 2024];48(1). Disponible en: <http://www.actamedicacolombiana.com/ojs/index.php/actamed/article/view/2569>

34. Higgs J, Jensen GM, Loftus S, Christensen N, editores. Clinical reasoning in the health professions. Fourth edition. Edinburgh London New York: Elsevier; 2019. 511 p.
35. Mattingly C. What is Clinical Reasoning? The American Journal of Occupational Therapy. 1 de noviembre de 1991;45(11):979-86.
36. García-Sánchez PC. Creación y Validación de un cuestionario para medir el Razonamiento Clínico en los Fisioterapeutas [Internet]. [España]: Universidad de Salamanca; 2024 [citado 1 de septiembre de 2024]. Disponible en: [https://gredos.usal.es/bitstream/handle/10366/139835/REDUCIDA\\_Cuestionariofisioterapeutas.pdf?sequence=1&isAllowed=y](https://gredos.usal.es/bitstream/handle/10366/139835/REDUCIDA_Cuestionariofisioterapeutas.pdf?sequence=1&isAllowed=y)
37. Gormaz C, Brailovsky C. Desarrollo del Razonamiento Clínico en Medicina. Clinical Reasoning Development in Medicine. Revista de Docencia Universitaria. 15 de septiembre de 2012;Vol.10:177-99.
38. Vargas Bianchi L. Kahneman, D. (2011). Thinking, Fast and Slow. Revista de Comunicación. mayo de 2022;11(1):251-3.
39. Schwartzstein RM. CLINICAL REASONING AND ARTIFICIAL INTELLIGENCE: CAN AI REALLY THINK? Trans Am Clin Climatol Assoc. 2024;134:133-45.
40. Melo M, Gusso GDF, Levites M, Amaro E, Massad E, Lotufo PA, et al. How doctors diagnose diseases and prescribe treatments: an fMRI study of diagnostic salience. Sci Rep. 2 de mayo de 2017;7(1):1304.
41. Eva KW. What every teacher needs to know about clinical reasoning. Med Educ. enero de 2005;39(1):98-106.
42. Fish M, de Cossart L. Cultivating a thinking surgeon: using a clinical thinking pathway as a learning and assessment process. En: Clinical reasoning in the health professions. Sydney: Butterworth-Heinemann; 2008. p. 451-60.
43. Charlin B, Lubarsky S, Millette B, Crevier F, Audétat M, Charbonneau A, et al. Clinical reasoning processes: unravelling complexity through graphical representation. Medical Education. mayo de 2012;46(5):454-63.
44. Charlin B, Tardif J, Boshuizen HPA. Scripts and Medical Diagnostic Knowledge: Theory and Applications for Clinical Reasoning Instruction and Research. Academic Medicine. febrero de 2000;75(2):182-90.
45. Bordage G, Connell KJ, Chang RW, Gecht MR, Sinacore JM. Assessing the semantic content of clinical case presentations: studies of reliability and concurrent validity. Academic Medicine. octubre de 1997;72(Supplement 1):S37-9.
46. Durning SJ, Artino AR, Schuwirth L, Van Der Vleuten C. Clarifying Assumptions to Enhance Our Understanding and Assessment of Clinical Reasoning: Academic Medicine. abril de 2013;88(4):442-8.

47. Yeung E. Enhancing Rigour in Orthopaedic Manual Physical Therapy Certification: Development of an Assessment Rubric for a Test of Clinical Reasoning [Internet]. [Canada]: University of Toronto; 2015. Disponible en: <https://api.semanticscholar.org/CorpusID:80061550>
48. Daniel M, Rencic J, Durning SJ, Holmboe E, Santen SA, Lang V, et al. Clinical Reasoning Assessment Methods: A Scoping Review and Practical Guidance. *Academic Medicine*. junio de 2019;94(6):902-12.
49. Gruppen LD, Wolf FM, Billi JE. Information Gathering and Integration as Sources of Error in Diagnostic Decision Making. *Med Decis Making*. diciembre de 1991;11(4):233-9.
50. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. *Med Educ*. octubre de 2015;49(10):961-73.
51. Krupat E, Wormwood J, Schwartzstein RM, Richards JB. Avoiding premature closure and reaching diagnostic accuracy: some key predictive factors. *Med Educ*. noviembre de 2017;51(11):1127-37.
52. Pelaccia T, Tardif J, Tribby E, Ammirati C, Bertrand C, Charlin B, et al. Insights into emergency physicians' minds in the seconds before and into a patient encounter. *Intern Emerg Med*. octubre de 2015;10(7):865-73.
53. Cutrer WB, Sullivan WM, Fleming AE. Educational Strategies for Improving Clinical Reasoning. *Current Problems in Pediatric and Adolescent Health Care*. octubre de 2013;43(9):248-57.
54. Monteiro SD, Sherbino JD, Ilgen JS, Dore KL, Wood TJ, Young ME, et al. Disrupting Diagnostic Reasoning: Do Interruptions, Instructions, and Experience Affect the Diagnostic Accuracy and Response Time of Residents and Emergency Physicians? *Academic Medicine*. abril de 2015;90(4):511-7.
55. Graber ML, Tompkins D, Holland JJ. Resources medical students use to derive a differential diagnosis. *Medical Teacher*. enero de 2009;31(6):522-7.
56. Stojan JN, Daniel M, Morgan HK, Whitman L, Gruppen LD. A Randomized Cohort Study of Diagnostic and Therapeutic Thresholds in Medical Student Clinical Reasoning. *Academic Medicine*. noviembre de 2017;92(11S):S43-7.
57. Williams RG, Klamen DL, Markwell SJ, Cianciolo AT, Colliver JA, Verhulst SJ. Variations in Senior Medical Student Diagnostic Justification Ability: *Academic Medicine*. mayo de 2014;89(5):790-8.
58. Goldszmidt M, Minda JP, Bordage G. Developing a Unified List of Physicians' Reasoning Tasks During Clinical Encounters: *Academic Medicine*. marzo de 2013;88(3):390-4.
59. Norman G. Research in clinical reasoning: past history and current trends. *Med Educ*. abril de 2005;39(4):418-27.

60. Feltovich PJ, Johnson PE, Moller JH. LCS: The role and development of medical knowledge in diagnostic expertise. En 1984. p. 275-319. Disponible en: <https://api.semanticscholar.org/CorpusID:18706223>
61. Barrows HS, Feightner J, Neufeld V, Norman G. Analysis of the clinical methods of medical students and physicians. Hamilton, Ontario: McMaster University. 1978;
62. Kim TH, Alraek T, Bian ZX, Birch S, Bovey M, Lee J, et al. Clinical reasoning in traditional medicine exemplified by the clinical encounter of Korean medicine. *Integrative Medicine Research*. junio de 2021;10(2):100641.
63. Schuwirth L. Can clinical reasoning be taught or can it only be learned?: Can clinical reasoning be taught or can it only be learned? *Medical Education*. agosto de 2002;36(8):695-6.
64. Groen GJ, Patel VL. Medical problem-solving: some questionable assumptions. *Medical Education*. marzo de 1985;19(2):95-100.
65. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication [published erratum appears in *Acad Med* 1992 Apr;67(4):287]. *Academic Medicine*. octubre de 1990;65(10):611-21.
66. Hege I, Hiedl M, Huth KC, Kiesewetter J. Differences in clinical reasoning between female and male medical students. *Diagnosis*. 15 de mayo de 2023;10(2):100-4.
67. Groves M, O'rourke P, Alexander H. The association between student characteristics and the development of clinical reasoning in a graduate-entry, PBL medical programme. *Medical Teacher*. noviembre de 2003;25(6):626-31.
68. Javedani Masroor M, Rezaei S, Hashemi Kiapay SA, Hakiminezhad M, Ahmadi SAY. Knowledge, Attitude, and Practice (KAP) Status toward Clinical Reasoning and Evidence-Based Medicine among the Medical Interns and Gynecology Residents of Iran University of Medical Sciences. Nakhostin Ansari N, editor. *The Scientific World Journal*. 13 de marzo de 2024;2024:1-9.
69. Humphrey KN, Daulton RS, Weber D, Sall D, Kelleher M. Coproducing the Learning Environment: Lessons Learned from a Year of Near-Peer Teaching. *Journal of Medical Education and Curricular Development*. enero de 2022;9:23821205221096288.
70. Royan R, Wu C, Theyyanni N, Montas S, Cranford J, House J, et al. Anything but Shadowing! Early Clinical Reasoning in Emergency Department Improves Clinical Skills. *WestJEM*. 18 de enero de 2018;177-84.
71. Wang YA, Chen CF, Chen CH, Wang GL, Huang AT. A clinical clerkship collaborative program in Taiwan: Acquiring core clinical competencies through patient care responsibility. *Journal of the Formosan Medical Association*. junio de 2016;115(6):418-25.
72. Young L, Anderson E, Gurney T, McArthur L, McGrail M, O'Sullivan B, et al. A comparative evaluation of quality and depth of learning by trainee doctors in regional, rural, and remote locations. *BMC Med Educ*. 5 de abril de 2023;23(1):215.

73. Ilgen JS, Humbert AJ, Kuhn G, Hansen ML, Norman GR, Eva KW, et al. Assessing Diagnostic Reasoning: A Consensus Statement Summarizing Theory, Practice, and Future Needs. Kowalenko T, editor. *Academic Emergency Medicine*. diciembre de 2012;19(12):1454-61.
74. Case SM, Swanson DB, Ripkey DR. Comparison of items in five-option and extended-matching formats for assessment of diagnostic skills. *Academic Medicine* [Internet]. 1994;69(10). Disponible en: [https://journals.lww.com/academicmedicine/fulltext/1994/10000/comparison\\_of\\_items\\_in\\_five\\_option\\_and.23.aspx](https://journals.lww.com/academicmedicine/fulltext/1994/10000/comparison_of_items_in_five_option_and.23.aspx)
75. Beullens J, Struyf E, Van Damme B. Diagnostic ability in relation to clinical seminars and extended-matching questions examinations. *Med Educ*. diciembre de 2006;40(12):1173-9.
76. Day SC, Norcini JJ, Diserens D, Cebul RD, Schwartz JS, Beck LH, et al. The validity of an essay test of clinical judgment. *Academic Medicine* [Internet]. 1990;65(9). Disponible en: [https://journals.lww.com/academicmedicine/fulltext/1990/09000/the\\_validity\\_of\\_an\\_essay\\_test\\_of\\_clinical\\_judgment.34.aspx](https://journals.lww.com/academicmedicine/fulltext/1990/09000/the_validity_of_an_essay_test_of_clinical_judgment.34.aspx)
77. Graaff E, Post GJ, Drop MJ. Validation of a new measure of clinical problem-solving. *Medical Education*. mayo de 1987;21(3):213-8.
78. Feletti GI. Reliability and validity studies on modified essay questions. *Academic Medicine* [Internet]. 1980;55(11). Disponible en: [https://journals.lww.com/academicmedicine/fulltext/1980/11000/reliability\\_and\\_validity\\_studies\\_on\\_modified\\_essay.6.aspx](https://journals.lww.com/academicmedicine/fulltext/1980/11000/reliability_and_validity_studies_on_modified_essay.6.aspx)
79. Palmer EJ, Devitt PG. Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Med Educ*. diciembre de 2007;7(1):49.
80. McCARTHY WH, Gonnella JS. The Simulated Patient Management Problem: a Technique for Evaluating and Teaching Clinical Competence. *Medical Education*. 29 de enero de 2009;1(5):348-52.
81. Newble DI, Hoare J, Baxter A. Patient management problems Issues of validity. *Medical Education*. mayo de 1982;16(3):137-42.
82. Page G, Bordage G. The Medical Council of Canada's key features project: a more valid written examination of clinical decision-making skills. *Academic Medicine* [Internet]. 1995;70(2). Disponible en: [https://journals.lww.com/academicmedicine/fulltext/1995/02000/the\\_medical\\_council\\_of\\_canada\\_s\\_key\\_features.12.aspx](https://journals.lww.com/academicmedicine/fulltext/1995/02000/the_medical_council_of_canada_s_key_features.12.aspx)
83. Hrynchak P, Glover Takahashi S, Nayer M. Key-feature questions for assessment of clinical reasoning: a literature review. *Med Educ*. septiembre de 2014;48(9):870-83.
84. Lineberry M, Kreiter CD, Bordage G. Threats to validity in the use and interpretation of script concordance test scores. *Med Educ*. diciembre de 2013;47(12):1175-83.

85. Charlin B, Van Der Vleuten C. Standardized Assessment of Reasoning in Contexts of Uncertainty: The Script Concordance Approach. *Eval Health Prof.* septiembre de 2004;27(3):304-19.
86. Ber R. The CIP (comprehensive integrative puzzle) assessment method. *Medical Teacher.* enero de 2003;25(2):171-6.
87. Capaldi VF, Durning SJ, Pangaro LN, Ber R. The Clinical Integrative Puzzle for Teaching and Assessing Clinical Reasoning: Preliminary Feasibility, Reliability, and Validity Evidence. *Military Medicine.* 1 de abril de 2015;180(suppl\_4):54-60.
88. Pottier P, Hardouin JB, Hodges BD, Pistorius MA, Connault J, Durant C, et al. Exploring how students think: a new method combining think-aloud and concept mapping protocols: A new method for exploring how students think. *Medical Education.* septiembre de 2010;44(9):926-35.
89. Daley BJ, Torre DM. Concept maps in medical education: an analytical literature review. *Medical Education.* mayo de 2010;44(5):440-8.
90. Anastakis DJ, Cohen R, Reznick RK. The structured oral examination as a method for assessing surgical residents. *The American Journal of Surgery.* julio de 1991;162(1):67-70.
91. Wass V, Wakeford R, Neighbour R, Van Der Vleuten C. Achieving acceptable reliability in oral examinations: an analysis of the Royal College of General Practitioners membership examination's oral component. *Med Educ.* febrero de 2003;37(2):126-31.
92. Khan KZ, Ramachandran S, Gaunt K, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective. *Medical Teacher.* septiembre de 2013;35(9):e1437-46.
93. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: Organisation & Administration. *Medical Teacher.* septiembre de 2013;35(9):e1447-63.
94. Ilgen JS, Sherbino J, Cook DA. Technology-enhanced Simulation in Emergency Medicine: A Systematic Review and Meta-Analysis. Zehtabchi S, editor. *Academic Emergency Medicine.* febrero de 2013;20(2):117-27.
95. Cook DA, Brydges R, Zendejas B, Hamstra SJ, Hatala R. Technology-Enhanced Simulation to Assess Health Professionals: A Systematic Review of Validity Evidence, Research Methods, and Reporting Quality. *Academic Medicine.* junio de 2013;88(6):872-83.
96. Kogan JR, Holmboe ES, Hauer KE. Tools for Direct Observation and Assessment of Clinical Skills of Medical Trainees: A Systematic Review. *JAMA.* 23 de septiembre de 2009;302(12):1316.
97. J J Norcini, Blank LL, Duffy FD, Fortna GS. The Mini-CEX: A Method for Assessing Clinical Skills. *Annals of Internal Medicine.* 2003;138(6):476-81.

98. Crossley J, Johnson G, Booth J, Wade W. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales: Construct alignment improves workplace-based assessment scales. *Medical Education*. junio de 2011;45(6):560-9.
99. Shea JA, Norcini JJ, Kimball HR. Relationships of ratings of clinical competence and ABIM scores to certification status. *Academic Medicine* [Internet]. 1993;68(10). Disponible en: [https://journals.lww.com/academicmedicine/fulltext/1993/10000/relationships\\_of\\_ratings\\_of\\_clinical\\_competence.34.aspx](https://journals.lww.com/academicmedicine/fulltext/1993/10000/relationships_of_ratings_of_clinical_competence.34.aspx)
100. Lewin LO, Beraho L, Dolan S, Millstein L, Bowman D. Interrater Reliability of an Oral Case Presentation Rating Tool in a Pediatric Clerkship. *Teaching and Learning in Medicine*. enero de 2013;25(1):31-8.
101. Durning SJ, Artino A, Boulet J, La Rochelle J, Van Der Vleuten C, Arze B, et al. The feasibility, reliability, and validity of a post-encounter form for evaluating clinical reasoning. *Medical Teacher*. enero de 2012;34(1):30-7.
102. Goulet F, Jacques A, Gagnon R, Racette P, Sieber W. Assessment of Family Physicians' Performance Using Patient Charts: Interrater Reliability and Concordance With Chart-Stimulated Recall Interview. *Eval Health Prof*. diciembre de 2007;30(4):376-92.
103. Schipper S, Ross S. Structured teaching and assessment. *Canadian Family Physician*. 2010;56(9):958-9.
104. Chatterjee S, Ng J, Kwan K, Matsumoto ED. Assessing the Surgical Decision Making Abilities of Novice and Proficient Urologists. *Journal of Urology*. mayo de 2009;181(5):2251-6.
105. Sibbald M, De Bruin ABH. Feasibility of self-reflection as a tool to balance clinical reasoning strategies. *Adv in Health Sci Educ*. agosto de 2012;17(3):419-29.
106. Cleary TJ, Callan GL, Zimmerman BJ. Assessing Self-Regulation as a Cyclical, Context-Specific Phenomenon: Overview and Analysis of SRL Microanalytic Protocols. *Education Research International*. 2012;2012:1-19.
107. Artino AR, Cleary TJ, Dong T, Hemmer PA, Durning SJ. Exploring clinical reasoning in novices: a self-regulated learning microanalytic assessment approach. *Med Educ*. marzo de 2014;48(3):280-91.
108. Hornberger J. INTRODUCTION TO BAYESIAN REASONING. *International Journal of Technology Assessment in Health Care*. 2001/05/25 ed. 2001;17(1):9-16.
109. Dauchet L, Bentégeac R, Ghauss H, Hazzan M, Truffert P, Amouyel P, et al. Le panel d'experts pour les tests de concordance de script : une référence réellement adéquate ? *La Revue de Médecine Interne*. julio de 2024;S0248866324006301.
110. Clancey WJ. The epistemology of a rule-based expert system —a framework for explanation. *Artificial Intelligence*. mayo de 1983;20(3):215-51.

111. Schank RC, Abelson RP. Scripts, Plans, Goals, and Understanding [Internet]. 0 ed. Psychology Press; 1977 [citado 1 de septiembre de 2024]. Disponible en: <https://www.taylorfrancis.com/books/9781134919666>
112. Feltovich PJ, Barrows HS. Issues of generality in medical problem solving. *Tutorials in problem-based learning*. 1984;128-42.
113. Bowen JL. Educational Strategies to Promote Clinical Diagnostic Reasoning. Cox M, Irby DM, editores. *N Engl J Med*. 23 de noviembre de 2006;355(21):2217-25.
114. Custers EJFM, Boshuizen HPA, Schmidt HG. The influence of medical expertise, case typicality, and illness script component on case processing and disease probability estimates. *Memory & Cognition*. mayo de 1996;24(3):384-99.
115. Charlin B, Brailovsky C, Leduc C, Blouin D. The Diagnosis Script Questionnaire: A New Tool to Assess a Specific Dimension of Clinical Competence. *Advances in Health Sciences Education*. 1998;3(1):51-8.
116. Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: From theory to practice: AMEE Guide No. 75. *Medical Teacher*. marzo de 2013;35(3):184-93.
117. Charlin B, Roy L, Brailovsky C, Goulet F, Van Der Vleuten C. The Script Concordance Test: A Tool to Assess the Reflective Clinician. *Teaching and Learning in Medicine*. octubre de 2000;12(4):189-95.
118. Lubarsky S, Charlin B, Cook DA, Chalk C, Van Der Vleuten CPM. Script concordance testing: a review of published validity evidence: Validity evidence for script concordance tests. *Medical Education*. abril de 2011;45(4):329-38.
119. Dory V, Gagnon R, Vanpee D, Charlin B. How to construct and implement script concordance tests: insights from a systematic review. *Medical Education*. junio de 2012;46(6):552-63.
120. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology*. enero de 2006;3(2):77-101.
121. Carrière B, Gagnon R, Charlin B, Downing S, Bordage G. Assessing Clinical Reasoning in Pediatric Emergency Medicine: Validity Evidence for a Script Concordance Test. *Annals of Emergency Medicine*. mayo de 2009;53(5):647-52.
122. Cooke S, Lemay JF, Beran T. Evolutions in clinical reasoning assessment: The Evolving Script Concordance Test. *Medical Teacher*. 3 de agosto de 2017;39(8):828-35.
123. Fournier JP, Demeester A, Charlin B. Script Concordance Tests: Guidelines for Construction. *BMC Med Inform Decis Mak*. diciembre de 2008;8(1):18.
124. Ramaekers S, Kremer W, Pilot A, Beukelen PV, Keulen HV. Assessment of competence in clinical reasoning and decision-making under uncertainty: the script concordance test method. *Assessment & Evaluation in Higher Education*. octubre de 2010;35(6):661-73.

125. Gagnon R, Charlin B, Lambert C, Carrière B, Van Der Vleuten C. Script concordance testing: more cases or more questions? *Adv in Health Sci Educ.* agosto de 2009;14(3):367-75.
126. Stanford JS. To what extent is the Script Concordance Test a valid measure of clinical reasoning for Advanced Paediatric Life Support Training [Internet]. [Australia]: The University of Melbourne; 2018 [citado 25 de julio de 2024]. Disponible en: <https://api.semanticscholar.org/CorpusID:149961436>
127. Gagnon R, Charlin B, Coletti M, Sauve E, Van Der Vleuten C. Assessment in the context of uncertainty: how many members are needed on the panel of reference of a script concordance test? *Med Educ.* marzo de 2005;39(3):284-91.
128. Aniort J, Trefond J, Tanguy G, Bataille S, Burtey S, Pereira B, et al. Impact of reference panel composition on scores of script concordance test assessing basic nephrology knowledge in undergraduate medical education. *Medical Teacher.* 2 de enero de 2024;46(1):110-6.
129. Fabry G. *Medizindidaktik: ein Handbuch für die Praxis.* Huber; 2008.
130. Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores: Reliability of objective structured clinical examination scores. *Medical Education.* diciembre de 2011;45(12):1181-9.
131. Wolpaw T, Papp KK, Bordage G. Using SNAPPS to Facilitate the Expression of Clinical Reasoning and Uncertainties: A Randomized Comparison Group Trial: *Academic Medicine.* abril de 2009;84(4):517-24.
132. Chamberland M, Mamede S, St-Onge C, Setrakian J, Bergeron L, Schmidt H. Self-explanation in learning clinical reasoning: the added value of examples and prompts. *Med Educ.* febrero de 2015;49(2):193-202.
133. Mamede S, Van Gog T, Moura AS, De Faria RMD, Peixoto JM, Rikers RMJP, et al. Reflection as a strategy to foster medical students' acquisition of diagnostic competence. *Medical Education.* mayo de 2012;46(5):464-72.
134. Matheny M, Israni ST, Ahmed M, Whicher D. *Artificial intelligence in health care: The hope, the hype, the promise, the peril.* Washington, DC: National Academy of Medicine. 2019;10.
135. Kanjee Z, Crowe B, Rodman A. Accuracy of a Generative Artificial Intelligence Model in a Complex Diagnostic Challenge. *JAMA.* 3 de julio de 2023;330(1):78.
136. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. Dagan A, editor. *PLOS Digit Health.* 9 de febrero de 2023;2(2):e0000198.
137. Nori H, King N, McKinney SM, Carignan D, Horvitz E. Capabilities of GPT-4 on Medical Challenge Problems [Internet]. arXiv; 2023 [citado 2 de septiembre de 2024]. Disponible en: <https://arxiv.org/abs/2303.13375>

138. Flores-Cohaila JA, García-Vicente A, Vizcarra-Jiménez SF, De La Cruz-Galán JP, Gutiérrez-Arratia JD, Quiroga Torres BG, et al. Performance of ChatGPT on the Peruvian National Licensing Medical Examination: Cross-Sectional Study. *JMIR Med Educ.* 28 de septiembre de 2023;9:e48039.
139. Sun L, Yin C, Xu Q, Zhao W. Artificial intelligence for healthcare and medical education: a systematic review. *Am J Transl Res.* 2023;15(7):4820-8.
140. Dziurka M, Machul M, Ozdoba P, Obuchowska A, Kotowski M, Grzegorzczak A, et al. Clinical Training during the COVID-19 Pandemic: Experiences of Nursing Students and Implications for Education. *IJERPH.* 23 de mayo de 2022;19(10):6352.
141. Huo CC, Zheng Y, Lu WW, Zhang TY, Wang DF, Xu DS, et al. Prospects for intelligent rehabilitation techniques to treat motor dysfunction. *Neural Regen Res.* 2021;16(2):264.
142. Li H, Chen B, Li J, Liu W, Ren T, Zheng S, et al. Status, problems and countermeasures of artificial intelligence application in medical education. *Chin J Evid Based Med.* 2020;20:1092-7.
143. Agarwal M, Sharma P, Goswami A. Analysing the Applicability of ChatGPT, Bard, and Bing to Generate Reasoning-Based Multiple-Choice Questions in Medical Physiology. *Cureus [Internet].* 26 de junio de 2023 [citado 2 de septiembre de 2024]; Disponible en: <https://www.cureus.com/articles/165118-analysing-the-applicability-of-chatgpt-bard-and-bing-to-generate-reasoning-based-multiple-choice-questions-in-medical-physiology>
144. Ayub I, Hamann D, Hamann CR, Davis MJ. Exploring the Potential and Limitations of Chat Generative Pre-trained Transformer (ChatGPT) in Generating Board-Style Dermatology Questions: A Qualitative Analysis. *Cureus [Internet].* 18 de agosto de 2023 [citado 2 de septiembre de 2024]; Disponible en: <https://www.cureus.com/articles/179695-exploring-the-potential-and-limitations-of-chat-generative-pre-trained-transformer-chatgpt-in-generating-board-style-dermatology-questions-a-qualitative-analysis>
145. Kıyak YS, Coşkun Ö, Budakoğlu İİ, Uluoğlu C. ChatGPT for generating multiple-choice questions: Evidence on the use of artificial intelligence in automatic item generation for a rational pharmacotherapy exam. *Eur J Clin Pharmacol.* mayo de 2024;80(5):729-35.
146. Kıyak YS, Emekli E. ChatGPT prompts for generating multiple-choice questions in medical education and evidence on their validity: a literature review. *Postgraduate Medical Journal.* 6 de junio de 2024;qgae065.
147. Flores-Cohaila JA, Calderón CDR, Castro-Blancas CI, Vela-Pandal E, Cruz-Galán JPDL, Bustamante-Ordoñez MA, et al. Using chatbots to develop multiple-choice questions. We got evidence, but we ain't there yet! [Internet]. 2023 [citado 2 de septiembre de 2024]. Disponible en: <https://www.researchsquare.com/article/rs-3742281/v1>

148. Kiyak YS, Kononowicz AA. Case-based MCQ generator: A custom ChatGPT based on published prompts in the literature for automatic item generation. *Medical Teacher*. 2 de agosto de 2024;46(8):1018-20.
149. Lee J, Wu AS, Li D, Kulasegaram K (Mahan). Artificial Intelligence in Undergraduate Medical Education: A Scoping Review. *Academic Medicine*. noviembre de 2021;96(11S):S62-70.
150. Fischetti C, Bhattar P, Frisch E, Sidhu A, Helmy M, Lungren M, et al. The Evolving Importance of Artificial Intelligence and Radiology in Medical Trainee Education. *Academic Radiology*. mayo de 2022;29:S70-5.
151. Valikodath NG, Cole E, Ting DSW, Campbell JP, Pasquale LR, Chiang MF, et al. Impact of Artificial Intelligence on Medical Education in Ophthalmology. *Trans Vis Sci Tech*. 14 de junio de 2021;10(7):14.
152. Blease C, Kharko A, Bernstein M, Bradley C, Houston M, Walsh I, et al. Machine learning in medical education: a survey of the experiences and opinions of medical students in Ireland. *BMJ Health Care Inform*. febrero de 2022;29(1):e100480.
153. Wartman SA, Combs CD. Medical Education Must Move From the Information Age to the Age of Artificial Intelligence. *Academic Medicine*. agosto de 2018;93(8):1107-9.
154. LIU D lu, LI J. Cultivating practical literacy of machine learning for medical students. *Basic & Clinical Medicine*. 2021;41(7):1071.
155. Simpson SA, Cook TS. Artificial Intelligence and the Trainee Experience in Radiology. *Journal of the American College of Radiology*. noviembre de 2020;17(11):1388-93.
156. Duong MT, Rauschecker AM, Rudie JD, Chen PH, Cook TS, Bryan RN, et al. Artificial intelligence for precision education in radiology. *BJR*. noviembre de 2019;92(1103):20190389.
157. Han ER, Yeo S, Kim MJ, Lee YH, Park KH, Roh H. Medical education trends for future physicians in the era of advanced technology and artificial intelligence: an integrative review. *BMC Med Educ*. diciembre de 2019;19(1):460.
158. Attia ZI, Harmon DM, Behr ER, Friedman PA. Application of artificial intelligence to the electrocardiogram. *European Heart Journal*. 7 de diciembre de 2021;42(46):4717-30.
159. Nagi F, Salih R, Alzubaidi M, Shah H, Alam T, Shah Z, et al. Applications of Artificial Intelligence (AI) in Medical Education: A Scoping Review. En: Mantas J, Gallos P, Zoulias E, Hasman A, Househ MS, Charalampidou M, et al., editores. *Studies in Health Technology and Informatics [Internet]*. IOS Press; 2023 [citado 2 de septiembre de 2024]. Disponible en: <https://ebooks.iospress.nl/doi/10.3233/SHTI230581>

160. Preiksaitis C, Rose C. Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review. *JMIR Med Educ*. 20 de octubre de 2023;9:e48785.
161. Kiyak YS. A ChatGPT Prompt for Writing Case-Based Multiple-Choice Questions. *Rev Esp Edu Med [Internet]*. 20 de octubre de 2023 [citado 2 de septiembre de 2024];4(3). Disponible en: <https://revistas.um.es/edumed/article/view/587451>
162. Abero L, Berardi L, Capocasale A. Investigación educativa. [Internet]. Buenos Aires: CLACSO; 2015 [citado el 9 de julio de 2021]. Disponible en: <https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=4721826>
163. Fernández P, Díaz P. Investigación cuantitativa y cualitativa. *Cad Aten Primaria [Internet]*. 2002 [citado el 9 de julio de 2021];9:76–8. Disponible en: [https://fisterra.com/mbe/investiga/cuanti\\_cuali/cuanti\\_cuali2.pdf](https://fisterra.com/mbe/investiga/cuanti_cuali/cuanti_cuali2.pdf)
164. Maldonado JJC, Macho LKG, Casallas EC. La investigación aplicada y el desarrollo experimental en el fortalecimiento de las competencias de la sociedad del siglo XXI. *Tecnura [Internet]*. el 1 de enero de 2023 [citado el 13 de diciembre de 2024];27(75):140–74. Disponible en: <https://revistas.udistrital.edu.co/index.php/Tecnura/article/view/19171>
165. Müggenburg Rodríguez V. MC, Pérez Cabrera I. Tipos de estudio en el enfoque de investigación cuantitativa. *Enferm Univ [Internet]*. 2007 [citado el 9 de julio de 2021];4(1):35–8. Disponible en: <http://revista-enfermeria.unam.mx:80/ojs/index.php/enfermeriauniversitaria/article/view/469> 54.
166. Ávila AJM, Suarez AKB, Martínez ZKP, Gonzaga JAR, Calderón JEZ, Suárez CEC. Diseños de Investigación. *Educ Salud Bol Científico Inst Cienc Salud Univ Autónoma Estado Hidalgo [Internet]*. el 5 de diciembre de 2019 [citado el 9 de julio de 2021];8(15):119
167. Vílchez Gutarra AAP, Ramos Muñoz WC, Guerrero Ramírez NN, Oyola García AE. Análisis de Situación de Salud (ASIS) 2021 [Internet]. Lima: Ministerio de Salud. Centro Nacional de Epidemiología, Prevención y Control de Enfermedades; 2023 [citado 16 de septiembre de 2024] p. 144. Disponible en: <https://www.gob.pe/institucion/ensap/informes-publicaciones/4509305-analisis-de-situacion-de-salud-asis-2021>
168. Mejía Álvarez CR, Inga Berrospi F, Contreras P, Requena T, Ruiz D. Serums: Recomendaciones basadas en la experiencia [Internet]. Universidad Continental. Universidad Continental; 2018 [citado 16 de septiembre de 2024]. Disponible en: <https://repositorio.continental.edu.pe/handle/20.500.12394/4624>
169. Contenidos según bloque temático de la evaluación para el Serums - Medicina [Internet]. Ministerio de Salud; 2024 ago [citado 16 de septiembre de 2024]. Disponible en: <https://www.gob.pe/institucion/minsa/informes-publicaciones/5911827-contenidos-segun-bloque-tematico-de-la-evaluacion-para-el-serums>
170. Programa-de-internado-medico | EsSalud [Internet]. [citado 16 de septiembre de 2024]. Disponible en: <https://www.essalud.gob.pe/programa-de-internado-medico/>

171. Tabla de especificaciones ENAM [Internet]. ASPEFAM; 2020 [citado 16 de septiembre de 2024] p. 17. Disponible en: <https://www.aspefam.org.pe/enam/>
172. Giuston Mendoza Chuctaya, Milagros Calla Torres, Kevin Ramos Chuctaya, Christian Mejía Álvarez. National Medicine Exam: an Analysis of the last decade of theoretical exams in peruvian last year medical students. *Acta Med Peru* [Internet]. 24 de octubre de 2021 [citado 16 de septiembre de 2024];38(3). Disponible en: <https://amp.cmp.org.pe/index.php/AMP/article/view/2164>
173. Flores-Cohaila J, Rivarola-Hidalgo M. El desempeño académico previo como predictor del examen nacional de medicina: un estudio transversal en Perú. *FEM: Revista de la Fundación Educación Médica*. 2022;25:243-7.
174. Tsai TC. Twelve tips for the construction of ethical dilemma case-based assessment. *Medical Teacher*. 3 de abril de 2017;39(4):341-6.
175. Brailovsky C, Charlin B, Beausoleil S, Cote S, Van Der Vleuten C. Measurement of clinical reflective capacity early in training as a predictor of clinical reasoning performance at the end of residency: an experimental study on the script concordance test. *Med Educ*. 13 de mayo de 2001;35(5):430-6.
176. Charlin B, Desaulniers M, Gagnon R, Blouin D, Van Der Vleuten C. Comparison of an Aggregate Scoring Method With a Consensus Scoring Method in a Measure of Clinical Reasoning Capacity. *Teaching and Learning in Medicine*. julio de 2002;14(3):150-6.
177. Brazeau-Lamontagne L, Charlin B, Gagnon R, Samson L, Van Der Vleuten C. Measurement of perception and interpretation skills during radiology training: utility of the script concordance approach. *Medical Teacher*. junio de 2004;26(4):326-32.
178. Llorca G, Roy P, Riche B. Evaluation de résolution de problèmes mal définis en éthique clinique: variation des scores selon les méthodes de correction et les caractéristiques des jurys. *Pédagogie médicale*. 2003;4(2):80-8.
179. Kassirer JP. Teaching Clinical Reasoning: Case-Based and Coached: *Academic Medicine*. julio de 2010;85(7):1118-24.
180. Wilson AB, Pike GR, Humbert AJ. Analyzing Script Concordance Test Scoring Methods and Items by Difficulty and Type. *Teaching and Learning in Medicine*. abril de 2014;26(2):135-45.
181. Sibert L, Darmoni SJ, Dahamna B, Hellot MF, Weber J, Charlin B. On line clinical reasoning assessment with Script Concordance test in urology: results of a French pilot study. *BMC Med Educ*. diciembre de 2006;6(1):45.
182. Charlin B, Gagnon R, Kazi-Tani D, Thivierge R. Le test de concordance comme outil d'évaluation en ligne du raisonnement des professionnels en situation d'incertitude. *Revue internationale des technologies en pédagogie universitaire*. 2005;2(2):22.
183. Kania RE. Online Script Concordance Test for Clinical Reasoning Assessment in Otorhinolaryngology The Association Between Performance and Clinical Experience. *Arch Otolaryngol Head Neck Surg*. 1 de agosto de 2011;137(8):751.

184. Hornos EH, Pleguezuelos EM, Brailovsky CA, Harillo LD, Dory V, Charlin B. The practicum script concordance test: an online continuing professional development format to foster reflection on clinical practice. *J Contin Educ Health Prof.* 2013;33(1):59-66.
185. Sibert L, Darmoni SJ, Dahamna B, Weber J, Charlin B. Online clinical reasoning assessment with the Script Concordance test: a feasibility study. *BMC Med Inform Decis Mak.* diciembre de 2005;5(1):18.
186. Stiglich Valenzuela EA. Test de concordancia script para la evaluación del razonamiento clínico en estudiantes de la carrera de tecnología médica mención oftalmología en la Universidad de Valparaíso [Internet]. [Santiago, Chile]: Universidad de Chile - Facultad de Medicina; 2021 [citado 5 de septiembre de 2024]. Disponible en: <https://repositorio.uchile.cl/handle/2250/183063>
187. Lubarsky S, Dory V, Meterissian S, Lambert C, Gagnon RG. Examining the effects of gaming and guessing on script concordance test scores. *Perspect Med Educ.* 12 de junio de 2018;7(3):174-81.
188. Wan MS, Tor E, Hudson JN. Improving the validity of script concordance testing by optimising and balancing items. *Med Educ.* marzo de 2018;52(3):336-46.
189. Ben Hamida E, Ayadi I, Marrakchi Z, Quinton A. The script concordance test as a tool to evaluate clinical reasoning in neonatology. *Tunis Med.* mayo de 2017;95(5):326-30.
190. Caire F, Sol JC, Charlin B, Isodiri P, Moreau JJ. Le test de concordance de script (TCS) comme outild'évaluation formative des internes en neurochirurgie :implantation du test sur Internet à l'échelle nationale. *Pédagogie Médicale.* mayo de 2004;5(2):87-94.
191. Gantelet M, Demeester A, Pauly V, Gagnon R, Charlin B. Impact du panel de référence sur les qualités psychométriques d'un test de concordance de script développé en formation initiale des sages-femmes. *Pédagogie Médicale.* agosto de 2013;14(3):157-68.
192. Duggan P, Charlin B. Introducing recent medical graduates as members of Script Concordance Test expert reference panels: what impact? *MedEdPublish.* 8 de agosto de 2016;5:60.
193. Mzoughi K, Zairi I, Kedous MA, El Mhamdi S, Ben Dhiab M, Mghaieth F, et al. Script concordance test as a sanctionnal evaluation in cardiology. *Tunis Med.* junio de 2018;96(6):330-4.
194. Kazour F, Richa S, Zoghbi M, El-Hage W, Haddad FG. Using the Script Concordance Test to Evaluate Clinical Reasoning Skills in Psychiatry. *Acad Psychiatry.* febrero de 2017;41(1):86-90.
195. Humbert AJ, Besinger B, Miech EJ. Assessing Clinical Reasoning Skills in Scenarios of Uncertainty: Convergent Validity for a Script Concordance Test in an Emergency Medicine Clerkship and Residency: ASSESSING CLINICAL REASONING USING THE SCT-EM. *Academic Emergency Medicine.* junio de 2011;18(6):627-34.

196. Mok SF, Tan TMD, Seow CJ. Modified endocrinology script concordance test: evaluating the reliability and construct validity for assessing clinical reasoning. *Singapore Medical Journal* [Internet]. 22 de agosto de 2023 [citado 19 de junio de 2024]; Disponible en: <https://journals.lww.com/10.4103/singaporemedj.SMJ-2021-230>
197. Registro Nacional de Identificación y Estado Civil (RENIEC). Portal Institucional. Lima: RENIEC; [2025]. Disponible en: <https://www.reniec.gov.pe>. Organización
198. Mundial de la Salud (OMS). Clasificación Internacional de Enfermedades, 11ª revisión (CIE-11). Ginebra: OMS; 2019. Disponible en: <https://www.who.int/es/publications/i/item/9789241565042>.
199. Mejia CR, Valladares-Garrido MJ, Quintana-Gomez S, Heredia P. Carrera previa como factor asociado al uso de buscadores científicos entre estudiantes de medicina latinoamericanos: cuando la experiencia no cuenta. *Educación Médica*. marzo de 2019;20:131-5.
200. Romaní-Romaní F, Gutiérrez C. Experiencia, actitudes y percepciones hacia la investigación científica en estudiantes de Medicina en el contexto de una estrategia curricular de formación de competencias para investigación. *Educación Médica*. mayo de 2022;23(3):100745.
201. Maturana A, Montenegro C, Nahuelhual P, Pérez C, Pizarro M, Riquelme A, et al. Manual del Tutor Clínico. Santiago: Universidad del Desarrollo, Centro de Desarrollo Educativo - Facultad de Medicina CAS-UDD; 2018. p. 88 [consultado 14 Feb 2025]. Disponible en: <https://medicina.udd.cl/cde/files/2018/09/Manual-tutor-2018.pdf>.
202. Gordillo Espinace EL, Dölz-Torres P, Palma-Bifani P, Rojas-Pino M. Impacto de actividades extracurriculares en la autorrealización de estudiantes de medicina: una revisión sistemática. *Rev Esp Edu Med* [Internet]. 22 de enero de 2024 [citado 14 de febrero de 2025];5(1). Disponible en: <https://revistas.um.es/edumed/article/view/592721>
203. Ministerio de Salud del Perú. Norma Técnica de Salud N° 021-MINSA/DGSP-V.03: .Categorías de Establecimientos del Sector Salud. Lima: MINSA; 2011. Disponible en: <https://www.gob.pe/institucion/minsa/normas-legales/243402-546-2011-minsa>
204. Universidad Nacional Mayor de San Marcos, Facultad de Medicina. Salud .Comunidad. Lima: UNMSM; [2000]. Disponible en: [https://sisbib.unmsm.edu.pe/bibvirtualdata/libros/medicina/salud\\_comun/sacomu.pdf](https://sisbib.unmsm.edu.pe/bibvirtualdata/libros/medicina/salud_comun/sacomu.pdf)
205. Duggan P, Charlin B. Summative assessment of 5th year medical students' clinical reasoning by script concordance test: requirements and challenges. *BMC Med Educ*. diciembre de 2012;12(1):29.
206. Protections (OHRP) O for HR. The Belmont Report [Internet]. 2010 [citado 7 de septiembre de 2024]. Disponible en: <https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html>

207. Custers EJFM. Thirty years of illness scripts: Theoretical origins and practical..applications. *Medical Teacher*. 4 de mayo de 2015;37(5):457-62.
208. Reifenrath J, Seiferth N, Wilhelm T, Holzmann-Littig C, Phillip V, Wijnen-Meijer M. Integrated clinical case discussions – a fully student-organized peer-teaching program on internal medicine. *BMC Med Educ*. 1 de diciembre de 2022;22(1):828.
209. NAZIM S, TALATI J, PINJANI S, BIYABANI S SR, ATHER M, NORCINI J. Assessing clinical reasoning skills using Script Concordance Test (SCT) and extended matching questions (EMQs): A pilot for urology trainees. *Journal of Advances in Medical Education & Professionalism [Internet]*. enero de 2019 [citado 17 de enero de 2025];7(1). Disponible en: <https://doi.org/10.30476/jamp.2019.41038>
210. van den Broek WES, van Asperen MV, Custers E, Valk GD, Ten Cate OTJ. Effects of two different instructional formats on scores and reliability of a script concordance test. *Perspect Med Educ*. agosto de 2012;1(3):119-28.
211. Keshtkar A, Atighi F, Reihani H. Systematic review of ChatGPT accuracy and performance in Iran's medical licensing exams: A brief report. *Journal of Education and Health Promotion [Internet]*. noviembre de 2024 [citado 22 de enero de 2025];13(1). Disponible en: [https://journals.lww.com/10.4103/jehp.jehp\\_1210\\_24](https://journals.lww.com/10.4103/jehp.jehp_1210_24)
212. Liu M, Okuhara T, Chang X, Okada H, Kiuchi T. Performance of ChatGPT in medical licensing examinations in countries worldwide: A systematic review and meta-analysis protocol. Khlaif Z, editor. *PLoS ONE*. 28 de octubre de 2024;19(10):e0312771.

ANEXOS

Anexo 1: Matriz de consistencia

Objetivos	Objetivos	Hipótesis	Variable	Diseño metodológico
<p><b>Problema general</b></p> <p>PG: ¿Cuáles son los factores asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?</p>	<p><b>Objetivo general</b></p> <p>OG: Determinar factores asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.</p>	<p><b>Hipótesis alterna</b></p> <p>Existen factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.</p>	<p>Factores asociados al razonamiento clínico</p>	<ul style="list-style-type: none"> <li>• Investigación básica</li> <li>• Método hipotético-deductivo</li> <li>• Diseño no experimental, corte transversal, de alcance correlacional y prospectivo</li> </ul>
<p><b>Problemas específicos</b></p> <p>Pe1: ¿Cuáles son los factores sociodemográficos asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?</p> <p>Pe2: ¿Cuáles son los factores académicos asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?</p> <p>Pe3: ¿Cuáles son los factores del entorno clínico asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024?</p>	<p><b>Objetivos específicos</b></p> <p>Oe1: Determinar los factores sociodemográficos asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.</p> <p>Oe2: Determinar los factores académicos asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.</p> <p>Oe3: Determinar los factores del entorno clínico asociados al desempeño académico del razonamiento clínico mediante la aplicación del Script Concordance Test en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.</p>	<p><b>Hipótesis nula</b></p> <p>No existen factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.</p> <p><b>Hipótesis específicas</b></p> <p>Hipótesis específica 1</p> <ul style="list-style-type: none"> <li>· H1: Los factores sociodemográficos del razonamiento clínico están asociados con el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.</li> </ul> <p>Hipótesis específica 2</p> <ul style="list-style-type: none"> <li>· H1: Los factores académicos del razonamiento clínico están asociados con el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una</li> </ul>	<p><b>Dimensiones</b></p> <ol style="list-style-type: none"> <li>1. Factores sociodemográficos</li> <li>2. Factores académicos</li> <li>3. Factores del entorno clínico</li> </ol>	<p><b>Población</b></p> <p>La población del estudio estará compuesta por los internos de medicina 2024 de la Universidad Norbert Wiener, ubicada en Lima, Perú</p> <p><b>Muestra</b></p> <p>Se empleará un muestreo censal de esta población, considerando a todos los internos como participantes en el estudio.</p>
			<p><b>Variable</b></p> <p>Desempeño académico en el Script Concordance Test</p>	
			<p><b>Dimensiones</b></p> <ol style="list-style-type: none"> <li>1. Clínico-médica</li> <li>2. Clínico-quirúrgica</li> </ol>	

		<p>universidad privada de Lima, Perú, en el año 2024.</p> <p>Hipótesis específica 3</p> <p>·H1: Los factores del entorno asistencial del razonamiento clínico están asociados con el desempeño académico en la prueba Script Concordance Test en los internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, en el año 2024.</p>		
--	--	--	--	--

## Anexo 2: Instrumento

### Ficha de recopilación de datos

Script Concordance Test

SCT

El Script Concordance Test (SCT) es una herramienta innovadora diseñada para evaluar el razonamiento clínico en contextos de incertidumbre. Este formulario busca explorar los factores asociados al desempeño académico de los internos de medicina, contribuyendo a la comprensión y fortalecimiento de sus habilidades para la toma de decisiones en escenarios clínicos complejos.

Se solicita completar información básica y el SCT marcando con una "X" la opción correspondiente en cada pregunta.

**Edad**

 #

**Sexo** ▾

 1 Masculino

 2 Femenino

**Estudios o carrera previa**

 Ninguno

 Enfermería

 Obstetricia

 Nutrición

 Odontología

 Terapia Física y Rehabilitación

 Psicología

 Otros

**Participación en mentoría académica**

 Estudiante

 Mentor

 Ninguno

**Experiencia en investigación**

 No tengo

 Ponencia en congreso

 Participación administrativa

 Artículo en revista

 Varios artículos

**Prácticas clínicas extracurriculares**

 Nula

 Escasa

 Moderada

 Considerable

**Supervisión médica en el internado**

- Nunca
- Raramente
- A veces
- Frecuentemente
- Siempre

**Nivel de atención del establecimiento de salud en el internado**

- Lima sur
- Lima norte
- Lima centro
- Lima oeste
- Lima este
- Otro departamento

**Nivel de atención del establecimiento de salud en el internado**

- Primer nivel de atención
- Segundo nivel de atención
- Tercer nivel de atención

**Retroalimentación del personal médico en el internado**

	Baja	Moderada	Alta
<b>Cirugía</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Baja	Moderada	Alta
<b>Gineco-obstetricia</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Baja	Moderada	Alta
<b>Pediatría</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Baja	Moderada	Alta
<b>Medicina interna</b>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Script Concordance Test

Script Concordance Test

SCT

### SCT 1

Una mujer de 30 años, que trabaja como vendedora en un mercado local, acude a tu centro de salud por disuria y aumento de la frecuencia urinaria durante los últimos 3 días. No presenta fiebre ni dolor en la espalda baja. Refiere haber tenido un episodio similar hace 6 meses, que se resolvió con antibióticos. No ha tenido relaciones sexuales recientemente. En el examen físico, el abdomen está blando, sin signos de dolor a la palpación.

#### 1.-Si estas pensando en...

Cistitis aguda complicada

#### 2.-Y entonces aprendes que...

La paciente tiene antecedentes de diabetes mal controlada.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### 1.-Y si estas pensando en...

Pielonefritis

#### 2.-Y entonces aprendes que...

El examen físico no revela dolor en los ángulos costovertebrales.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 2

Un paciente de 35 años se presenta en tu consulta con dolor en el oído derecho, sensación de oído tapado, disminución de la audición desde hace 3 días, y sin fiebre. A la exploración otoscópica, observas un tapón de cerumen impactado en el oído derecho, que obstruye parcialmente el conducto auditivo. El paciente tiene antecedentes de infecciones recurrentes de oído, pero no recuerda cuándo fue su último tratamiento.

### 1.-Si estas pensando en...

Otitis externa

### 2.-Y entonces aprendes que...

El conducto auditivo está eritematoso y hay dolor a la palpación del trago.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Realizar lavado de oídos

### 2.-Y entonces aprendes que...

El paciente refiere dolor intenso en el oído afectado durante la exploración física.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 3

Una adolescente acude por primera vez al puesto de salud (I-2) para su control prenatal. Es su primer embarazo y no ha recibido controles prenatales previos. Según la fecha de su última menstruación (FUM), se encuentra en la semana 24 de gestación. Informa un aumento significativo de peso desde el inicio del embarazo. Además, refiere disuria leve y urgencia miccional ocasional en la última semana. Solo se cuenta al momento de la consulta con glucosa capilar de 145 mg/dL.

### 1.-Si estas pensando en...

Diabetes gestacional

### 2.-Y entonces aprendes que...

Su altura uterina se encuentra entre la apófisis xifoides y el ombligo.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Infección urinaria en el embarazo

### 2.-Y entonces aprendes que...

La paciente refiere un antecedente de ITU previo al embarazo.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Consejería integral (6 controles mínimos)

### 2.-Y entonces aprendes que

Tiene 14 años y no tiene factores de riesgo adicionales.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 4

Una joven de 19 años acude a la posta de salud tras haber sufrido abuso sexual hace 24 horas. No ha recibido atención médica desde el incidente. No presenta síntomas agudos, pero refiere cierta incomodidad general. Está preocupada por la posibilidad de infecciones de transmisión sexual y el riesgo de embarazo. No ha experimentado sangrado vaginal ni dolor pélvico significativo, pero refiere un leve dolor en la parte baja del abdomen, que describe como esporádico. También menciona una sensación de ardor al orinar, aunque sin dificultad significativa para miccionar. No ha usado anticonceptivos previamente.

### 1.-Si estas pensando en...

Ofrecer Anticonceptivo oral de emergencia

### 2.-Y entonces aprendes que...

Confiesa ser menor de edad y sus padres no tienen conocimiento del incidente.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Ofrecer tratamiento empírico para gonorrea y clamidia

### 2.-Y entonces aprendes que...

No se cuenta con los resultados de laboratorio ni posibilidad de realizar cultivos en el centro.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 5

Un hombre de 25 años, que colabora en actividades comunitarias, acude al centro de salud tras ser mordido por un perro callejero hace 8 horas, mientras se dirigía a una junta vecinal. La herida se localiza en su mano derecha y presenta sangrado moderado. El perro huyó después del incidente, por lo que no está disponible para observación. No hay acceso inmediato a inmunoglobulina antirrábica en el centro de salud, pero hay disponibilidad de vacunas antirrábicas.

### 1.-Si estas pensando en...

Limpiar la herida con abundante agua y jabón

### 2.-Y entonces aprendes que...

El paciente menciona que ha tenido contacto con agua no potable para limpiar la herida antes de llegar al centro de salud.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Administrar la vacuna antirrábica sin esperar observación del animal

### 2.-Y entonces aprendes que...

El paciente tiene antecedentes de vacunación completa contra la rabia hace 2 años tras otra mordedura.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 6

Una mujer de 60 años, trabajadora agrícola, acude al centro de salud por dolor lumbar crónico que ha empeorado en los últimos 3 meses. El dolor es más intenso por las mañanas y mejora levemente con el movimiento. También refiere dolor y rigidez en las manos al despertar, con episodios intermitentes de hinchazón en los dedos. Su trabajo diario implica largas horas parada cuidando su ganado.

### 1.-Si estas pensando en...

Artrosis lumbar

### 2.-Y entonces aprendes que...

La paciente menciona que el dolor mejora notablemente cuando toma descansos durante su trabajo físico.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Artritis reumatoide

### 2.-Y entonces aprendes que...

La paciente menciona que la hinchazón en las manos es ocasional y la rigidez dura menos de 30 minutos.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Lumbalgia mecánica

### 2.-Y entonces aprendes que...

El dolor lumbar disminuye en las noches y no irradia a las piernas.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 7

Estas en tu posta en Arequipa, en temporada de friaje, un hombre de 55 años se presenta en tu consultorio con una tos persistente de 4 semanas de evolución. Además, refiere pérdida de peso no intencionada y sudoración nocturna. Tiene antecedentes de tabaquismo crónico (30 años) y en su casa siempre ha cocinado con leña. No ha tenido acceso a radiografías o pruebas de esputo recientes debido a limitaciones en los recursos diagnósticos en la localidad. Los signos vitales: T 37.5°C, FR 20 rpm, SatO2 91% en aire ambiente. Estado general regular, con leve dificultad respiratoria al hablar, sin signos evidentes de angustia.

### 1.-Si estas pensando en...

Tuberculosis

### 2.-Y entonces aprendes que...

El paciente no ha tenido contacto con nadie que presente síntomas similares ni con personas diagnosticadas con tuberculosis.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Neumonía adquirida en la comunidad

### 2.-Y entonces aprendes que...

El paciente ha tenido varios episodios previos similares (tos, ↓ peso, sudoración nocturna, dificultad respiratoria leve) que se han resuelto sin tratamiento antibiótico específico.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

EPOC

### 2.-Y entonces aprendes que...

Se auscultan roncus y sibilancias generalizadas, sin focalización

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 8

Un hombre de 50 años llega a tu consulta con dolor abdominal progresivo en las últimas 48 horas. Refiere náuseas, vómitos y distensión abdominal. No ha tenido evacuaciones ni eliminación de gases en las últimas 24 horas. A la exploración, se observa distensión abdominal moderada, dolor a la palpación en todo el abdomen y timpanismo al percudir.

### 1.-Si estas pensando en...

Obstrucción intestinal alta

### 2.-Y entonces aprendes que...

El paciente ha estado vomitando grandes cantidades de contenido alimentario no digerido.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 9

Un hombre de 55 años acude por cefalea intensa de inicio súbito, asociada a mareos y sensación de presión en el pecho, que comenzó hace unas horas. Tiene antecedentes de hipertensión arterial no controlada y menciona que no ha estado tomando su medicación antihipertensiva de manera regular debido a la falta de acceso constante a los medicamentos. En el examen físico, su presión arterial es de 180/110 mmHg. No hay acceso inmediato a electrocardiograma ni a estudios de imagen.

### 1.-Si estas pensando en...

Emergencia hipertensiva

### 2.-Y entonces aprendes que...

El paciente refiere visión borrosa y tinnitus.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 10

Estás almorzando en el carnaval local del lugar donde estás realizando el SERUMS cuando un hombre de 45 años, que está comiendo cerca de ti, comienza a toser de forma violenta. La tos se vuelve menos efectiva y parece que tiene dificultad para respirar. Su respiración es superficial, jadea, emite ruidos respiratorios y va perdiendo la conciencia.

### 1.-Si estas pensando en...

Iniciar maniobra de Heimlich

### 2.-Y entonces aprendes que...

El hombre aún responde pidiendo ayuda con las manos y sigue tosiendo ruidosamente.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Iniciar maniobras de RCP

### 2.-Y entonces aprendes que...

Ves un cuerpo extraño en la vía aérea.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 11

Un niño de 6 meses es traído al centro de salud por una masa en la axila izquierda que ha aumentado de tamaño en las últimas dos semanas. Además, refiere que el niño está con calentura desde hace dos días. El niño está en buen estado general, se alimenta bien y no presenta otros síntomas sistémicos. A la exploración física se encuentra una masa de 2cm en la axila izquierda con signos de supuración.

### 1.-Si estas pensando en...

Linfadenitis supurativa

### 2.-Y entonces aprendes que...

Al revisar el historial, se confirma que el recibió la vacuna BCG al nacer.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Usar paracetamol para controlar la fiebre

### 2.-Y entonces aprendes que...

El niño presenta fiebre de 38.5°C y tuvo 1 episodio este año de convulsión febril.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 12

Una mujer de 35 años acude al centro de salud por diarrea de 2 días de evolución. Menciona que asistió a un carnaval local hace unos días y consumió alimentos de la calle antes del inicio de los síntomas. No ha presentado fiebre, pero refiere debilidad generalizada y calambres musculares. La paciente se encuentra en buen estado general, con signos de leve debilidad muscular y calambres, abdomen blando y no doloroso, ruidos intestinales ligeramente aumentados, piel algo pálida y seca.

### 1.-Si estas pensando en...

Rehidratación oral con sales de rehidratación oral (SRO)

### 2.-Y entonces aprendes que...

Encuentras mucosas secas y una presión arterial de 100/60 mmHg.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Suspender el consumo de alimentos sólidos

### 2.-Y entonces aprendes que...

La diarrea es acuosa y con 6 evacuaciones diarias

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 13

Un hombre de 30 años llega a la sala de urgencias tras haber caído del caballo en medio de una carrera. El paciente está consciente, pero refiere dolor abdominal intenso. En la inspección, tiene abrasiones en el abdomen y una pequeña herida puntiforme en el flanco derecho. Los signos vitales son: presión arterial de 100/60 mmHg, frecuencia cardíaca de 120 lpm, y frecuencia respiratoria de 22 rpm. A la palpación, el abdomen está doloroso de forma difusa con defensa leve.

### 1.-Si estas pensando en...

Traumatismo abdominal cerrado

### 2.-Y entonces aprendes que...

El paciente presenta una equimosis en forma de herradura en el abdomen y percusión mate en el cuadrante superior izquierdo.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 14

Un joven de 28 años, que trabaja en la cosecha de productos agrícolas, acude a tu centro de salud en Arequipa, con fiebre alta de 3 días de evolución, malestar general, dolor muscular generalizado y náuseas. También refiere dolor abdominal leve, más intenso en el cuadrante superior derecho. No ha presentado vómitos ni diarrea. En la exploración física, se observa hepatomegalía dolorosa. No tiene conductas sexuales de riesgo, ni viajes recientes fuera de su localidad.

### 1.-Si estas pensando en...

Dengue

### 2.-Y entonces aprendes que...

El paciente no presenta signos de rash ni petequias en la piel.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Hepatitis viral

### 2.-Y entonces aprendes que...

El paciente menciona que ha consumido alcohol (4-5 días x sem) en las últimas 3 semanas.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### SCT 15

Una mujer de 52 años con un índice de masa corporal de 32 kg/m<sup>2</sup> acude al centro de salud por fatiga persistente, hambre constante, sed excesiva y aumento de la frecuencia urinaria en las noches. Ha ganado peso progresivamente en los últimos dos años. No se dispone de acceso inmediato a pruebas de glucosa en ayunas ni a hemoglobina glicada, pero se sospecha diabetes mellitus tipo 2. Tiene de FC 82 lpm, PA 130/85 mmHg, abdomen blando y no doloroso, sin hepatomegalia ni masas palpables, ruidos intestinales normales y extremidades sin edemas. La paciente ha intentado seguir dietas, pero no ha tenido éxito en la pérdida de peso.

#### 1.-Si estas pensando en...

Modificación de la dieta y el estilo de vida con un plan de alimentación saludable

#### 2.-Y entonces aprendes que...

La paciente menciona que tiene acceso limitado a alimentos frescos debido a dificultades económicas.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### 1.-Y si estas pensando en...

Iniciar metformina de forma empírica

#### 2.-Y entonces aprendes que...

La paciente refiere visión borrosa reciente.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

#### 1.-Y si estas pensando en...

Derivar a un nutricionista para manejo especializado de la obesidad

#### 2.-Y entonces aprendes que...

El paciente tiene antecedentes familiares de obesidad y dificultades para controlar su peso a largo plazo.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**SCT 16**

Un niño de 6 años es llevado al centro de salud por presentar una erupción pruriginosa en varias partes del cuerpo. La madre menciona que las lesiones comenzaron hace aproximadamente 10 días, inicialmente pequeñas, pero han empeorado. También comenta que hace un mes el niño fue con sus compañeros a un bosque cerca del colegio y que recientemente han surgido casos de infecciones cutáneas en la escuela. En la exploración física se observan pápulas eritematosas dispersas, algunas excoriaciones, y costras superficiales debido al rascado.

**1.-Si estas pensando en...**

Escabiosis

**2.-Y entonces aprendes que...**

El prurito empeora notablemente por la noche, y el hermano del niño también ha comenzado a tener prurito recientemente.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**1.-Y si estas pensando en...**

Urticaria

**2.-Y entonces aprendes que...**

Las lesiones no son migratorias y persisten en las mismas áreas desde hace varios días.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**1.-Y si estas pensando en...**

Pediculosis

**2.-Y entonces aprendes que...**

La madre menciona que el prurito ha aumentado en el cuero cabelludo, en la nuca y detrás de las orejas.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 17

Un niño de 4 años es traído a tu centro de salud en Tumbes, por su madre debido a palidez progresiva e irritabilidad que inicio hace 10 días. No presenta fiebre ni otros síntomas agudos. La madre refiere que el niño no ha tenido infecciones recientes ni enfermedades crónicas. En el examen físico, se observa palidez en piel y mucosas, pero el niño está aparentemente estable.

### 1.-Si estas pensando en...

Iniciar tratamiento con sulfato ferroso y consejería integral

### 2.-Y entonces aprendes que...

A los 2 años, el niño interrumpió el tratamiento porque presentó deposiciones oscuras y nauseas.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Administrar hierro a 2 mg/kg/día durante 6 meses continuos

### 2.-Y entonces aprendes que...

La madre confirma que no ha habido antecedentes de anemia previos y la hemoglobina es 9 g/dL.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 18

Un niño de 6 años es traído al centro de salud por presentar fiebre alta (39.5 °C) desde hace 3 días, acompañada de un exantema que ha progresado de la cara hacia el tronco y las extremidades. La madre menciona que el niño no ha sido completamente vacunado. El niño también refiere dolor de garganta, malestar general y leve inflamación ocular. No hay antecedentes de enfermedades crónicas.

### 1.-Si estas pensando en...

Sarampión

### 2.-Y entonces aprendes que...

El exantema es maculopapular y en la cavidad oral se encuentran manchas de Koplik.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Rubéola

### 2.-Y entonces aprendes que...

No hay mejoría del cuadro clínico tras varios días y se observa artralgia en muñecas y tobillos.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 19

Un niño de 4 años presenta diarrea acuosa y maloliente desde hace 5 días, sin sangre ni vómitos. También ha tenido dolor abdominal intermitente y distensión abdominal. En la exploración física, se encuentra afebril, sin signos de deshidratación aparente, pero con distensión abdominal leve. El niño vive en una zona del pueblo sin acceso a agua potable.

### 1.-Si estas pensando en...

Hidratación oral con solución de rehidratación oral (SRO)

### 2.-Y entonces aprendes que...

El niño no ha vomitado ni presenta signo del pliegue.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Antibióticos sistémicos

### 2.-Y entonces aprendes que...

El niño ha perdido peso y presentó esteatorrea.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 20

Un hombre de 30 años, agricultor en una zona rural, es llevado al centro de salud por sus familiares tras presentar confusión, sudoración excesiva, vómitos y debilidad muscular que comenzaron aproximadamente 40 minutos antes de su llegada. Fue encontrado en su chacra, donde había una botella de insecticida derramada cerca. Se desconoce si inhaló o ingirió alguna sustancia tóxica mientras trabajaba. En el examen físico, presenta diaforesis intensa, pupilas mióticas y una frecuencia cardíaca de 45 lpm.

### 1.-Si estas pensando en...

Intoxicación por organofosforados

### 2.-Y entonces aprendes que...

El paciente presenta aumento de secreciones bronquiales y fasciculaciones musculares.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Envenenamiento por picadura de viuda negra

### 2.-Y entonces aprendes que...

El paciente presenta dolor abdominal severo y espasmos musculares en la pierna derecha.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 21

Un hombre de 45 años llega a urgencias tras caerse de una escalera de dos metros. Refiere dolor intenso en la muñeca y el hombro derecho, con dificultad para mover ambas articulaciones. La muñeca muestra deformidad, hinchazón y hematoma, y el hombro tiene movilidad limitada. Los signos vitales son estables y no presenta otras lesiones.

### 1.-Si estas pensando en...

Fractura distal de radio (fractura de Colles)

### 2.-Y entonces aprendes que...

El paciente presenta desplazamiento epifisiario ventral y dolor intenso a la palpación de la muñeca.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Luxación de hombro anterior

### 2.-Y entonces aprendes que...

El paciente tiene dolor en el hombro y se observa prominencia del acromion con limitación de la movilidad.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 22

Un ganadero de 40 años se presenta en urgencias con dolor abdominal que comenzó hace 12 horas. Inicialmente difuso, intensidad 6/10, luego localizado en el abdomen inferior derecho que no alivia con analgésicos orales. Refiera náuseas, pero no ha vomitado ni fiebre. Hace años le dijeron que tenía "piedras en la vesícula".

### 1.-Y si estas pensando en...

Colecistitis aguda

### 2.-Y entonces aprendes que...

Paciente menciona que también le duele su hombro derecho y al examen físico el dolor progresó al cuadrante superior derecho.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Coledocolitiasis de bajo riesgo

### 2.-Y entonces aprendes que...

Al referenciar al hospital, solo se observa una dilatación del conducto biliar común.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### SCT 23

Mujer de 32 años, en la semana 34 de gestación por FUM, altura uterina correspondiente a edad gestacional, controles prenatales irregulares, se presenta a consulta con cefalea, edema en extremidades inferiores y dolor epigástrico que evoluciona de moderado a intenso. Su presión arterial es de 160/110 mmHg, y la tira reactiva de orina muestra proteinuria +2.

**1.-Si estas pensando en...**

Iniciar metildopa y referenciar solicitando hemograma + perfil hepático

**2.-Y entonces aprendes que...**

El dolor no mejora con paracetamol, y la paciente tiene antecedentes de multiparidad y cesáreas.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**1.-Y si estas pensando en...**

Iniciar hidratación EV + sulfato de magnesio + metildopa

**2.-Y entonces aprendes que...**

La paciente reporta visión borrosa y su presión arterial sube a 180/110 mmHg.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 24

Un niño de 3 años es traído al centro de salud por su madre debido a fiebre leve, tos que ha empeorado en los últimos 2 días y congestión nasal, todo esto inició hace 4 días. No tiene antecedentes patológicos y cuenta con sus vacunas completas según su edad. En el examen físico, se observa temperatura de 37.9°C, congestión nasal, tos ronca, enrojecimiento leve en la faringe y linfadenopatías palpables. El niño está respirando normalmente en el momento de la consulta.

### 1.-Si estas pensando en...

Crup viral

### 2.-Y entonces aprendes que...

La madre informa que la tos se agrava durante la noche y suena como si se estuviera ahogando.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Resfrío común

### 2.-Y entonces aprendes que...

La temperatura subió a 40°C y la tos empeoró.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## SCT 25

Una joven de 22 años acude a la posta de salud por dolor menstrual incapacitante desde la adolescencia, el cual ha empeorado en los últimos meses. Además, en los últimos seis meses ha presentado episodios de sangrados intermenstruales y períodos de más de 10 días. Estos síntomas han afectado su capacidad para trabajar y ayudar a su familia. También refiere sentirse fatigada y le preocupa la posibilidad de anemia.

### 1.-Si estas pensando en...

Ofrecer anticonceptivos orales combinados (AOC)

### 2.-Y entonces aprendes que...

La madre murió por cáncer de ovario y útero.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu manejo es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Dismenorrea primaria

### 2.-Y entonces aprendes que...

Además, presenta náusea, diarrea y dolor de cabeza durante la menstruación.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

### 1.-Y si estas pensando en...

Endometriosis

### 2.-Y entonces aprendes que...

La paciente menciona que su hermana no puede quedar embarazada.

	(-2) mucho menos apropiado	(-1) menos apropiado	(0) sin cambios	(+1) más apropiado	(+2) mucho más apropiado
Con esta nueva información tu diagnóstico es:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Anexo 3: Validez del instrumento**

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. Lisdamys Morera Gonzalez

DNI: 48665295

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. Nailin Melany Inga Mayta

DNI: 72195627

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. CONNETH SUSY CCANTO AUCCAPUMA

DNI: 40299717

Especialidad del validador: Medicina familiar y comunitaria



Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. JACKELINEE FIORELLA QUISPE JARA

DNI: 44423971

Especialidad del validador: Medicina familiar y comunitaria



Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dr. Alex Maguiña Maldonado

DNI: 08169949

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

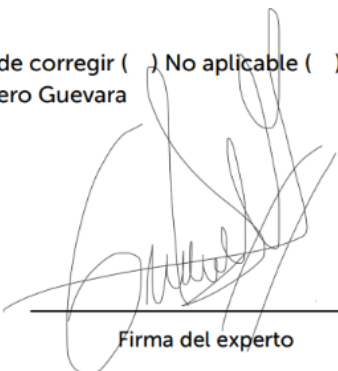
Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dr. José Miguel Guerrero Guevara

DNI: 70435101

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. YOVANA ALENA ARGOTE CASACHAGUA

DNI: 20722114

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. Jehny Almonte Laura

DNI: 29206080

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. Katherine Alessandra Hinojosa Bocanegra

DNI: 72929521

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. Nataly R. Espinoza Suarez

DNI: 45796511

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. JORGE EDUARDO ROJAS VILLANUEVA

DNI: 46054648

Especialidad del validador: Medicina familiar y comunitaria



Dr. Jorge E. Rojas Villanueva  
MÉDICO FAMILIAR  
C.M.P. 069293

---

Firma del experto

## Certificado de validez de contenido del instrumento

Observaciones (si las hubiera):

Opinión de aplicabilidad: Aplicable ( x ) Aplicable ( ) Después de corregir ( ) No aplicable ( )

Apellidos y nombre del juez validador: Dra. Alicia del Socorro Morales Covarrubias

DNI: 44971684

Especialidad del validador: Medicina familiar y comunitaria



---

Firma del experto

## Anexo 4: Aprobación del comité de ética



### COMITÉ INSTITUCIONAL DE ÉTICA E INTEGRIDAD CIENTÍFICA

#### CONSTANCIA DE APROBACIÓN

Lima, 30 de Diciembre de 2024

Investigador(a)  
**JIMMY ALEJANDRO DIAZ CARRILLO**  
**Exp. N°:1376-2024**

---

De mi consideración:

Es grato expresarle mi cordial saludo y a la vez informarle que el Comité Institucional de Ética e Integridad Científica de la Universidad Privada Norbert Wiener (CIEIC-UPNW) **evaluó y APROBÓ** los siguientes documentos:

- Protocolo titulado: **“Factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024” Versión 01 con fecha 12/12/2024.**
- Formulario de Consentimiento Informado Versión 01 con fecha **12/12/2024.**

El cual tiene como investigador principal al Sr(a) Jimmy Alejandro Diaz Carrillo.

La APROBACIÓN comprende el cumplimiento de las buenas prácticas éticas, el balance riesgo/beneficio, la calificación del equipo de investigación y la confidencialidad de los datos, entre otros.

El investigador deberá considerar los siguientes puntos detallados a continuación:

1. **La vigencia** de la aprobación es de **dos años** (24 meses) a partir de la emisión de este documento.
2. **El Informe de Avances** se presentará cada 6 meses, y el informe final una vez concluido el estudio.
3. **Toda enmienda o adenda** se deberá presentar al CIEIC-UPNW y no podrá implementarse sin la debida aprobación.
4. Si aplica, **la Renovación** de aprobación del proyecto de investigación deberá iniciarse treinta (30) días antes de la fecha de vencimiento, con su respectivo informe de avance.

Es cuanto informo a usted para su conocimiento y fines pertinentes.

Atentamente,

Raúl Antonio Rojas Ortega  
**Presidente**

**Comité Institucional de Ética e Integridad Científica**  
**UPNW**



## Anexo 5: Formato de consentimiento informado

**Institución:** Universidad Privada Norbert Wiener

**Investigador principal:** Jimmy Alejandro Díaz Carrillo

**Título del estudio:** Factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024.

---

### Propósito del estudio:

El presente estudio tiene como objetivo explorar los factores del razonamiento clínico asociados a su desempeño académico en internos de medicina humana, utilizando la aplicación del Script Concordance Test (SCT) en un contexto educativo específico de Lima, Perú. Esta herramienta, validada para la evaluación del razonamiento clínico, se utilizará para identificar factores sociodemográficos y académicos que influyen en el rendimiento de los estudiantes en la toma de decisiones clínicas. Además, este estudio considera la adaptación cultural y la inclusión de inteligencia artificial en la generación de casos clínicos, lo que podría fortalecer la práctica educativa en medicina y apoyar la evaluación de competencias clínicas esenciales en futuros médicos, especialmente en entornos complejos y de alta demanda asistencial.

### Procedimientos:

Si acepta participar en este estudio, se le solicitará completar una encuesta física que incluirá tanto preguntas generales sobre su perfil sociodemográfico y formación académica, como también una serie de ítems clínicos pertenecientes al SCT. Esta encuesta está diseñada para ser completada en aproximadamente 60 minutos. La recopilación de sus respuestas será completamente confidencial y los datos serán anónimos para proteger su identidad. Los resultados se analizarán de manera global, sin identificar a los participantes individualmente.

### Riesgos:

Este estudio no implica riesgos físicos, psicológicos o emocionales para los participantes. La participación consiste únicamente en la respuesta a un cuestionario, por lo que no se espera ningún tipo de impacto negativo en su salud o bienestar. Sin embargo, si en algún momento durante el estudio desea detener su participación, puede hacerlo sin consecuencias.

### Beneficios:

La participación en este estudio le permitirá contribuir al avance del conocimiento en el campo de la educación médica, específicamente en la evaluación y mejora del razonamiento clínico. Los hallazgos del estudio pueden ser utilizados para el desarrollo de mejores estrategias de enseñanza y evaluación en la formación de médicos, lo cual beneficiará a futuras generaciones de profesionales de la salud.

### Costos e incentivos:

No se requerirá ningún tipo de pago para participar en este estudio. Del mismo modo, los participantes no recibirán compensación económica ni incentivos materiales por su colaboración. La participación es voluntaria y libre de cualquier costo.

### Confidencialidad:

La información proporcionada será tratada bajo estrictas normas de confidencialidad. Sus datos personales serán protegidos mediante el anonimato, es decir, su identidad no estará vinculada a las respuestas proporcionadas en la encuesta. Los resultados del estudio se presentarán de forma colectiva y en ningún momento se revelará información que permita identificar a los participantes. Los archivos de datos estarán almacenados de forma segura y accesibles solo para el investigador principal.

### Derechos del participante:

Su participación en este estudio es completamente voluntaria. Si en algún momento durante el desarrollo de la encuesta decide retirarse o no continuar, podrá hacerlo sin ninguna penalización ni necesidad de justificar su decisión. Asimismo, si experimenta cualquier duda, inquietud o malestar relacionado con el estudio, puede comunicarse directamente con el investigador principal, Jimmy Alejandro Díaz Carrillo, a través del número de contacto 910854147 o al correo electrónico jimmy2104dc@gmail.com. Toda preocupación será atendida con la mayor celeridad y discreción. Además, podrá comunicarse con el comité de ética que validó el presente estudio, Dra. Yenny M. Bellido Fuentes, presidenta del Comité de Ética para la investigación de la Universidad Norbert Wiener, telf. 7065555 anexo 3285/comité.etica@uwiener.edu.pe

### Consentimiento:

Al firmar este documento, usted confirma que ha leído la información detallada acerca del estudio, ha tenido la oportunidad de hacer preguntas y ha recibido respuestas satisfactorias. Entiendo que mi participación es voluntaria y que puedo retirarme en cualquier momento. Acepto participar en este estudio titulado Factores del razonamiento clínico asociados a su desempeño académico mediante la aplicación del 'Script Concordance Test' en internos de la Escuela Académico Profesional de Medicina Humana de una universidad privada de Lima, Perú, 2024. Recibiré una copia firmada de este consentimiento.

Participante:  
DNI:



Investigador: Jimmy Alejandro Díaz Carrillo  
DNI: 74289105  
Fecha: 02/01/2024

## ● 10% de similitud general

Principales fuentes encontradas en las siguientes bases de datos:

- 8% Base de datos de Internet
- Base de datos de Crossref
- 7% Base de datos de trabajos entregados
- 3% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

### FUENTES PRINCIPALES

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	<b>repositorio.uwiener.edu.pe</b> Internet	2%
2	<b>uwiener on 2024-08-23</b> Submitted works	<1%
3	<b>uwiener on 2024-07-07</b> Submitted works	<1%
4	<b>admissiononline.uft.cl</b> Internet	<1%
5	<b>repositorio.upsjb.edu.pe</b> Internet	<1%
6	<b>hdl.handle.net</b> Internet	<1%
7	<b>sap.org.ar</b> Internet	<1%
8	<b>revistas.um.es</b> Internet	<1%