Empleo de obstáculos sucesivos en la evaluación con polígrafo – AIIP Mérida, MX

Un marco para la estandarización y la mejora de la precisión

Mark Handler – Gerente de Productos de Polígrafo de Lafayette Instrument Company

Chip Morgan – Director de la Escuela del Centro de Capacitación en Evaluación de Credibilidad Peak

13 de octubre de 2026

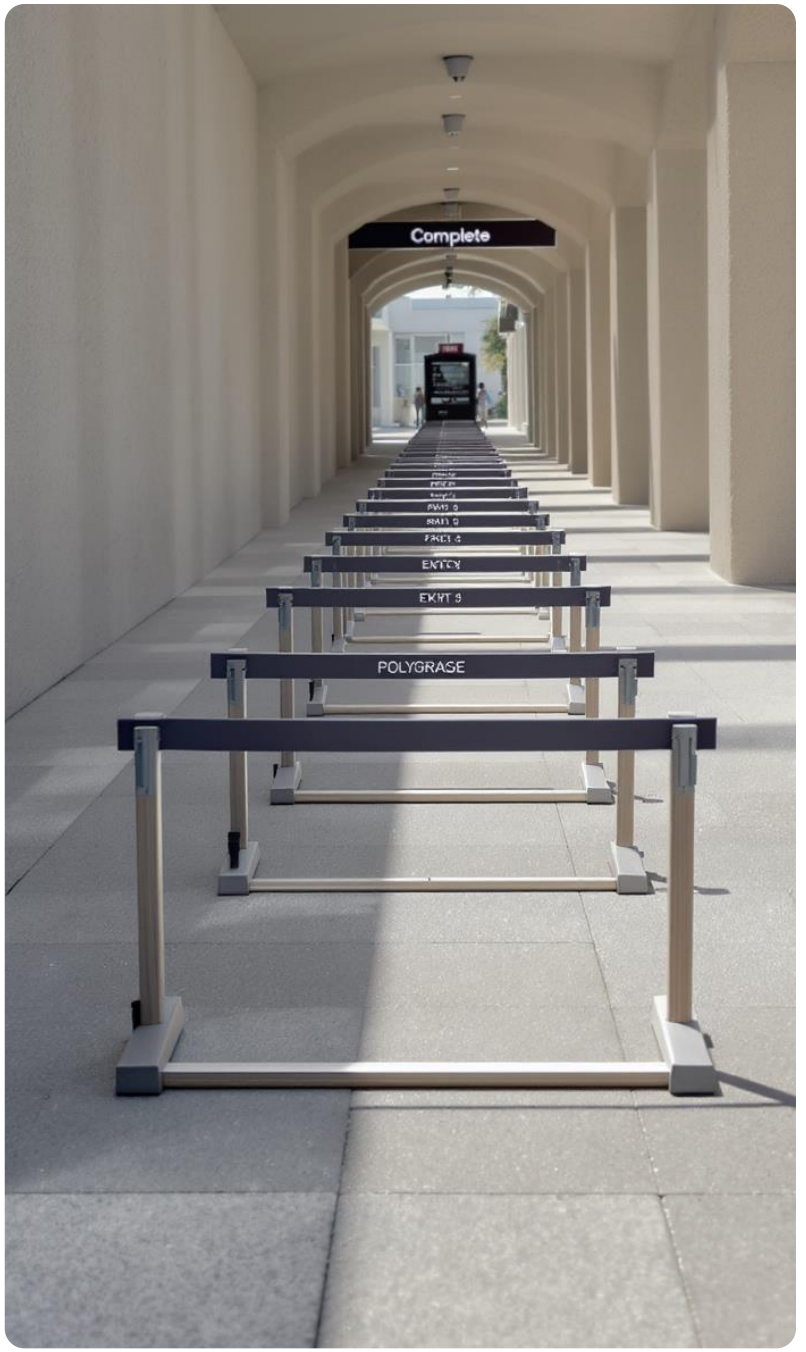


Agenda

1. Introducción a los obstáculos sucesivos. Limitaciones de las pruebas de múltiples temas. Estándares de implementación del marco de obstáculos sucesivos. Consideraciones científicas y éticas. Eficacia operativa. Aportaciones de los autores. Síntesis asistida por IA. Contexto institucional. Conclusión y perspectivas futuras

Introducción a los «obstáculos sucesivos»

- Los «obstáculos sucesivos» se refieren a un proceso de pruebas con polígrafo de múltiples etapas en el que los examinados pasan por pruebas o evaluaciones secuenciales. El enfoque tiene como objetivo mejorar la precisión de la evaluación mediante el filtrado progresivo de falsos positivos y la confirmación del engaño a través de pasos independientes. No existe una definición operativa clara ni un protocolo estandarizado por parte de la Asociación Estadounidense de Poligrafía para las pruebas de obstáculos sucesivos. Esta falta de claridad da lugar a prácticas inconsistentes y limita la eficacia potencial de la metodología de obstáculos sucesivos. El desarrollo de un marco basado en la evidencia es crucial para brindar orientación práctica y mejorar el rigor científico y procedimental en la evaluación con polígrafo.



Limitaciones de las pruebas de detección de múltiples temas

Heterogeneidad de la tasa base y falsos positivos

Los distintos problemas de conducta presentan tasas de prevalencia (probabilidad a priori) variables, lo que provoca un aumento de las tasas de falsos positivos en las pruebas de múltiples temas. Esto reduce la precisión predictiva y conduce a la clasificación errónea de personas inocentes.

Relevancia: efectos ponderados

Las respuestas fisiológicas intensas a las preguntas relevantes iniciales pueden suprimir las reacciones a las preguntas posteriores. Esto da lugar a posibles falsos negativos en cuestiones posteriores dentro de la misma prueba debido a la saturación fisiológica.

Ambiguidad diagnóstica

Las pruebas con múltiples temas producen resultados compuestos que carecen de claridad sobre qué tema específico desencadenó una respuesta, lo que requiere entrevistas subjetivas posteriores a la prueba y complica la interpretación objetiva.

Complicaciones estadísticas derivadas de la multiplicidad

Las múltiples preguntas independientes aumentan la tasa acumulada de falsos positivos. Las estrategias de pruebas en serie no mitigan esto por completo, lo que da como resultado una menor precisión general y mayores tasas de error entre los examinados que responden con veracidad.



Limitaciones

Heterogeneidad de la tasa base y falsos positivos

- Cada tema de comportamiento evaluado tiene una tasa de prevalencia única en la población evaluada, conocida como tasa base o probabilidad a priori. Las bajas tasas base para algunos comportamientos hacen que incluso las pruebas de alta precisión produzcan muchos falsos positivos y una baja confianza en los resultados. Por ejemplo, con una tasa base del 5 % y una validez de la prueba del 95 %, casi la mitad de los resultados positivos pueden ser falsos positivos. Las pruebas que abordan múltiples problemas agravan este problema, ya que las respuestas no permiten aislar qué problema provocó los indicios de engaño. Esta heterogeneidad de la tasa base dificulta la interpretación probabilística y reduce la confiabilidad general de la evaluación.

Respuesta fisiológica ponderada por la relevancia

- La respuesta fisiológica ponderada por la relevancia describe la reducción de la intensidad de la respuesta fisiológica tras una reacción inicial fuerte a una pregunta relevante. Este efecto se produce porque los examinados concentran su mayor reactividad fisiológica en el estímulo más relevante o amenazante, lo que hace que las respuestas posteriores parezcan más débiles. La saturación fisiológica o habituación significa que la respuesta autonómica del cuerpo disminuye con estímulos repetidos o muy cercanos en el tiempo durante las pruebas de múltiples temas. Las observaciones de campo confirman que los examinados pueden reaccionar de manera notable solo ante un tema, incluso cuando mienten sobre varios temas, lo que genera desafíos diagnósticos. Este fenómeno puede causar resultados falsos negativos en preguntas posteriores, lo que socava la validez de las pruebas de polígrafo de múltiples temas.

Ambiguidad diagnóstica y complicaciones estadísticas

Ambiguidad diagnóstica en las pruebas de múltiples temas

- Las pruebas de múltiples temas combinan varios ámbitos conductuales, lo que hace que no quede claro qué tema específico causó una indicación de engaño. Los resultados positivos requieren entrevistas subjetivas posteriores a la prueba para aclararlos, lo que introduce un sesgo interpretativo y reduce la objetividad. La ambigüedad socava el valor del polígrafo como medida fisiológica al depender de aclaraciones verbales. Determinar la responsabilidad de las respuestas fisiológicas se vuelve un reto cuando se evalúan múltiples temas simultáneamente.



Complicaciones estadísticas y efectos de multiplicidad

- Cada tema evaluado representa una decisión de clasificación independiente, lo que aumenta las tasas de error acumuladas. Las comparaciones múltiples inflan las tasas de falsos positivos entre los examinados que dicen la verdad, lo que degrada la precisión general de la prueba. Las simulaciones de Monte Carlo muestran que las pruebas en serie pueden, paradójicamente, aumentar los falsos positivos en lugar de reducirlos. Los metaanálisis revelan que las pruebas de un solo tema tienen una mayor precisión diagnóstica (0,90–0,93) en comparación con los formatos de múltiples temas (0,85–0,88).

Marco basado en la evidencia: obstáculos sucesivos

Etapa uno: Evaluación SIST

Las pruebas iniciales de la batería se enfocan en comportamientos específicos, como el consumo de sustancias o el robo, dando prioridad a la sensibilidad para señalar inquietudes que requieran una revisión más detallada.

- Protocolos SIST validados Preguntas estandarizadas Resultados: NDI, DI, INC

Etapa dos: Entrevista posterior a la prueba

Las entrevistas enfocadas aclaran las respuestas y el contexto de cualquier resultado «DI» o «INC», lo que orienta la necesidad de realizar pruebas adicionales sin sesgos.

- Protocolo de entrevista estandarizado Documentación de las respuestas Decisión sobre la etapa tres

Etapa tres: Pruebas de seguimiento

Las pruebas de confirmación utilizan diferentes formatos de preguntas para verificar los resultados de manera independiente, lo que reduce el sesgo y los falsos positivos.

- SIST de confirmación. Calificación independiente. Clasificación final

Etapa uno: Evaluación con la batería SIST

- La batería SIST evalúa ámbitos conductuales específicos, tales como el consumo de sustancias, el robo, la conducta sexual inapropiada, la falsificación y la actividad delictiva grave. Cada prueba sigue formatos validados de preguntas comparativas con una redacción estandarizada para garantizar la consistencia. Los resultados de la puntuación se clasifican como «Sin indicios de engaño» (NDI), «Con indicios de engaño» (DI) o «No concluyente» (INC). Esta etapa prioriza la sensibilidad para detectar problemas reales, aceptando un mayor número de falsos positivos que se filtrarán posteriormente. Las pruebas de tema único ofrecen una interpretación probabilística más clara al aislar un comportamiento por prueba, lo que mejora el valor científico y probatorio.

Marco

Etapa dos: Entrevista estructurada posterior a la prueba

- Se lleva a cabo después de que los resultados de cualquier prueba de detección de un solo tema (SIST) indiquen «Con indicios de engaño» (DI) o «No concluyente» (INC). El propósito es aclarar la comprensión del examinado sobre el tema evaluado e identificar posibles causas no engañosas de las reacciones fisiológicas. Utiliza protocolos de entrevista estandarizados y basados en la evidencia para recabar información contextual y evitar el sesgo de confirmación. Determina si es necesario realizar pruebas adicionales; sin embargo, no decide por sí misma si se trata de la verdad o del engaño. Garantiza la independencia entre las etapas de la prueba al separar la medición fisiológica de la aclaración verbal.



Etapa tres: Pruebas de seguimiento específicas

- Se administra una Prueba de Cribado de Tema Único (SIST) confirmatoria para los temas no resueltos después de la Etapa Dos. Utilizar formatos de preguntas metodológicamente distintos; por ejemplo, si en la Etapa Uno se utilizó la comparación dirigida de mentiras, en la Etapa Tres se utiliza la estructura de mentira probable. Esto reduce el sesgo del examinador y del examinado al cambiar la tarea psicológica y la estrategia de respuesta. Mantener algoritmos de puntuación consistentes en todas las etapas para obtener resultados comparables. Los estudios de Monte Carlo muestran que este enfoque puede reducir los falsos positivos hasta en un 30 % sin perder sensibilidad.

Requisitos de implementación



Estandarización de los protocolos de evaluación

Adopte pruebas de cribado de tema único (SIST) validadas que utilicen algoritmos de puntuación respaldados empíricamente de manera consistente en todos los exámenes para mantener la confiabilidad y la uniformidad del procedimiento.

Elaboración de preguntas basada en la evidencia

Utilizar plantillas de preguntas estandarizadas que garanticen la claridad lingüística, la especificidad temporal y la adecuación cultural, adaptando la redacción a las diferencias de jurisdicción legal.

Competencia y capacitación de los evaluadores

Capacitar a los examinadores en razonamiento probabilístico, efectos de la tasa base, principios fisiológicos, entrevistas estandarizadas y prácticas éticas de pruebas en múltiples etapas.

Salvaguardas normativas y directrices éticas

Implementar políticas que prohíban las decisiones adversas basadas únicamente en los resultados iniciales de la evaluación, exigiendo evidencia corroborativa para obtener resultados justos y defendibles.

Garantía de calidad y revisión independiente

Establecer una revisión por pares de las interpretaciones de las pruebas, recalibrar los umbrales de puntuación periódicamente, documentar las decisiones en cada etapa y dar seguimiento a los indicadores de precisión de los examinadores.



Implementación

Estandarización de los protocolos de evaluación

- La adopción de formatos SIST validados garantiza que cada prueba se centre específicamente en un dominio conductual claramente definido, lo que mejora la claridad diagnóstica. La aplicación consistente de algoritmos de puntuación respaldados empíricamente, como ESS u OSS-3, mantiene la confiabilidad entre examinadores y reduce la interpretación subjetiva. Es fundamental el cumplimiento estricto de los procedimientos; las desviaciones de los protocolos estandarizados pueden comprometer la validez de la prueba y la comparabilidad entre casos. La transición a instrumentos poligráficos digitales facilita la estandarización, pero requiere disciplina por parte del examinador para mantener la fidelidad al protocolo. Los protocolos estandarizados permiten la acumulación de datos confiables para el control de calidad continuo y la mejora constante.



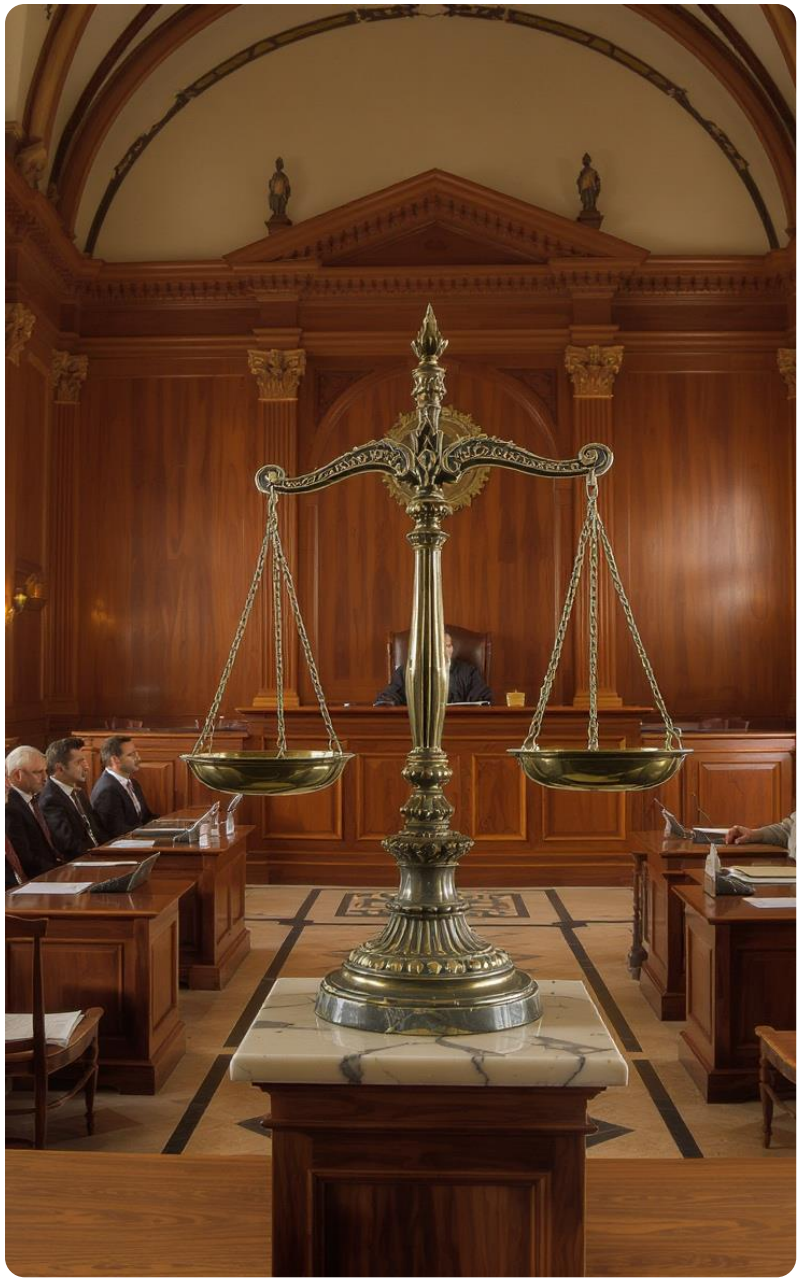
Implementación

Elaboración de preguntas basada en la evidencia

- Las plantillas de preguntas estandarizadas reducen la ambigüedad, asegurando que las preguntas sean claras y específicas. Las plantillas incorporan claridad lingüística para evitar malentendidos y mejorar la comprensión del examinado. La adaptación al contexto cultural es esencial para mantener la equivalencia semántica entre jurisdicciones con diferentes normas legales y sociales. Los protocolos de la Asociación Estadounidense de Poligrafía (American Polygraph Association) guían la elaboración estandarizada de preguntas de control, comparación y preguntas relevantes. La elaboración consistente de preguntas minimiza la variación interpretativa, lo que respalda la confiabilidad y la imparcialidad en los resultados de la prueba.

Competencia y capacitación de los evaluadores

- Los examinadores deben tener un sólido dominio del razonamiento probabilístico y la toma de decisiones bayesiana para interpretar con precisión los resultados de la evaluación inicial. Comprender principios fisiológicos como la respuesta fisiológica ponderada por la relevancia y la habituación es esencial para reconocer los patrones de respuesta y sus limitaciones. La capacitación debe abarcar técnicas estandarizadas de entrevista para garantizar la coherencia y reducir el sesgo del examinador durante las entrevistas posteriores a la prueba. Las directrices éticas exigen que los examinadores mantengan la equidad procesal, evitando decisiones adversas basadas únicamente en los resultados de la evaluación inicial. La educación continua debe hacer hincapié en la teoría de la decisión estadística, los efectos de la tasa base y los principios de las pruebas en múltiples etapas para mantener la competencia.



Ética

Salvaguardas normativas y directrices éticas

- Las decisiones nunca deben basarse únicamente en los resultados de la primera etapa de la evaluación para evitar medidas adversas injustas. Las medidas laborales, de seguridad o disciplinarias requieren evidencia corroborativa más allá de los resultados del polígrafo, como verificaciones de antecedentes o entrevistas. El marco de obstáculos sucesivos respalda la equidad procesal al exigir múltiples pasos de verificación antes de llegar a conclusiones definitivas. Las políticas deben prohibir explícitamente basarse en resultados inconclusos o de una sola prueba para tomar decisiones críticas, a fin de proteger los derechos de los examinados. El cumplimiento de estas pautas éticas mejora la defendibilidad legal y alinea las prácticas del polígrafo con los estándares reconocidos del debido proceso.



Implementación

Garantía de calidad y revisión independiente

- La revisión por pares de las interpretaciones de los gráficos del polígrafo ayuda a detectar sesgos y a confirmar la precisión de la puntuación al involucrar a múltiples examinadores en el proceso de evaluación. La recalibración periódica de los umbrales de puntuación basada en datos de campo acumulados asegura que los criterios de decisión sigan siendo válidos y reflejen el desempeño en el mundo real (véase el artículo de PPO y NPO de Nelson, Handler y Thiel, 2021). La documentación exhaustiva de la justificación de las decisiones en cada obstáculo sucesivo mantiene la transparencia y respalda la rendición de cuentas procesal. El seguimiento de los indicadores de precisión entre los examinadores y los centros de pruebas facilita la identificación de tendencias, promueve la rendición de cuentas de los examinadores e impulsa la mejora continua de la calidad. La revisión independiente en la tercera etapa de las pruebas, en la que los examinadores desconocen los resultados previos, reduce el sesgo de confirmación y refuerza la equidad del procedimiento.



Validez

Mejora de la validez científica

- Utiliza pruebas de cribado de tema único (SIST) validadas con precisión y confiabilidad documentadas. Emplea protocolos estandarizados de elaboración de preguntas para reducir la ambigüedad y garantizar la consistencia. Integra algoritmos de puntuación empíricos como ESS y OSS-3 para mejorar la confiabilidad entre examinadores. Permite la acumulación de datos de calidad para la calibración y mejora continuas basadas en tasas de base y umbrales de decisión conocidos. Alinea la práctica de la detección con polígrafo con los estándares psicométricos y de diagnóstico contemporáneos, abordando las críticas de las principales revisiones científicas.

Ética

Equidad procesal y debido proceso

- Las pruebas en múltiples etapas evitan decisiones adversas basadas únicamente en un solo resultado de cribado, lo que respalda el debido proceso. Cada etapa sucesiva ofrece a los examinados la oportunidad de aclarar y verificar sus respuestas, lo que reduce el riesgo de falsos positivos. Los procedimientos estandarizados y los puntos de decisión documentados mejoran la transparencia y la rendición de cuentas. Este marco equilibra los objetivos de seguridad de las organizaciones con la protección de los derechos legales y éticos de los examinados. Las múltiples etapas de verificación garantizan que las decisiones se basen en evidencia corroborada, no en resultados ambiguos del polígrafo.



Eficacia operativa

- La reducción de los falsos positivos protege a los candidatos calificados de una exclusión injustificada, preservando el valioso capital humano y la confianza en la organización. La disminución de los falsos negativos mejora la seguridad al identificar riesgos reales que, de otro modo, podrían pasar desapercibidos. Los obstáculos sucesivos y estructurados permiten un uso específico de los recursos de pruebas intensivas, centrando los esfuerzos en casos con preocupaciones demostradas en lugar de una aplicación generalizada. La mayor precisión y equidad contribuyen a fortalecer la reputación de la organización y la defendibilidad legal en contextos de selección de alto riesgo. El enfoque se asemeja a los sistemas de triaje médico al asignar recursos de manera eficiente con base en una evaluación de riesgos basada en evidencia.

Desafíos

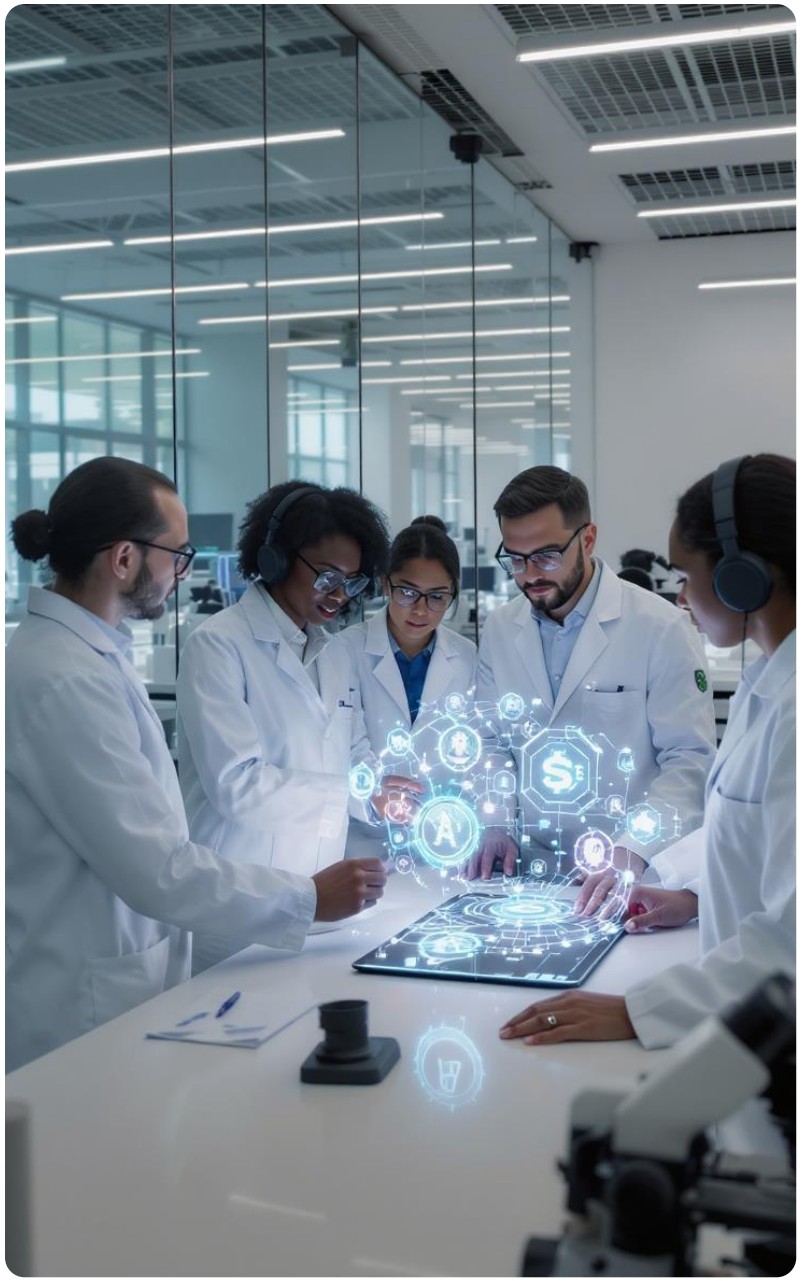
Limitaciones y contraindicaciones

Fortalezas del marco de «obstáculos sucesivos»

- Mejora la precisión diagnóstica mediante el uso de etapas secuenciales de pruebas de un solo aspecto. Reduce los falsos positivos al incorporar pruebas de confirmación metodológicamente distintas. Mejora la equidad procesal a través de múltiples oportunidades de verificación. Apoya la mejora continua de la calidad mediante la recopilación de datos empíricos. Mitiga algunos sesgos de respuesta fisiológica, como la atenuación anticlimática.

Desafíos y limitaciones actuales

- Persisten los problemas de tasa base, especialmente al detectar eventos muy raros, lo que conduce a altas tasas de falsos positivos. Las diferencias fisiológicas individuales y los estados psicológicos pueden afectar la validez y la confiabilidad de la prueba. Los falsos negativos siguen siendo una preocupación debido a la habituación, la respuesta fisiológica ponderada por la relevancia y la variabilidad de la respuesta entre pruebas. La susceptibilidad a las contramedidas físicas y mentales no se elimina por completo con los obstáculos sucesivos. Ningún método de polígrafo, incluido el de los obstáculos sucesivos, alcanza una precisión perfecta ni resultados libres de errores.

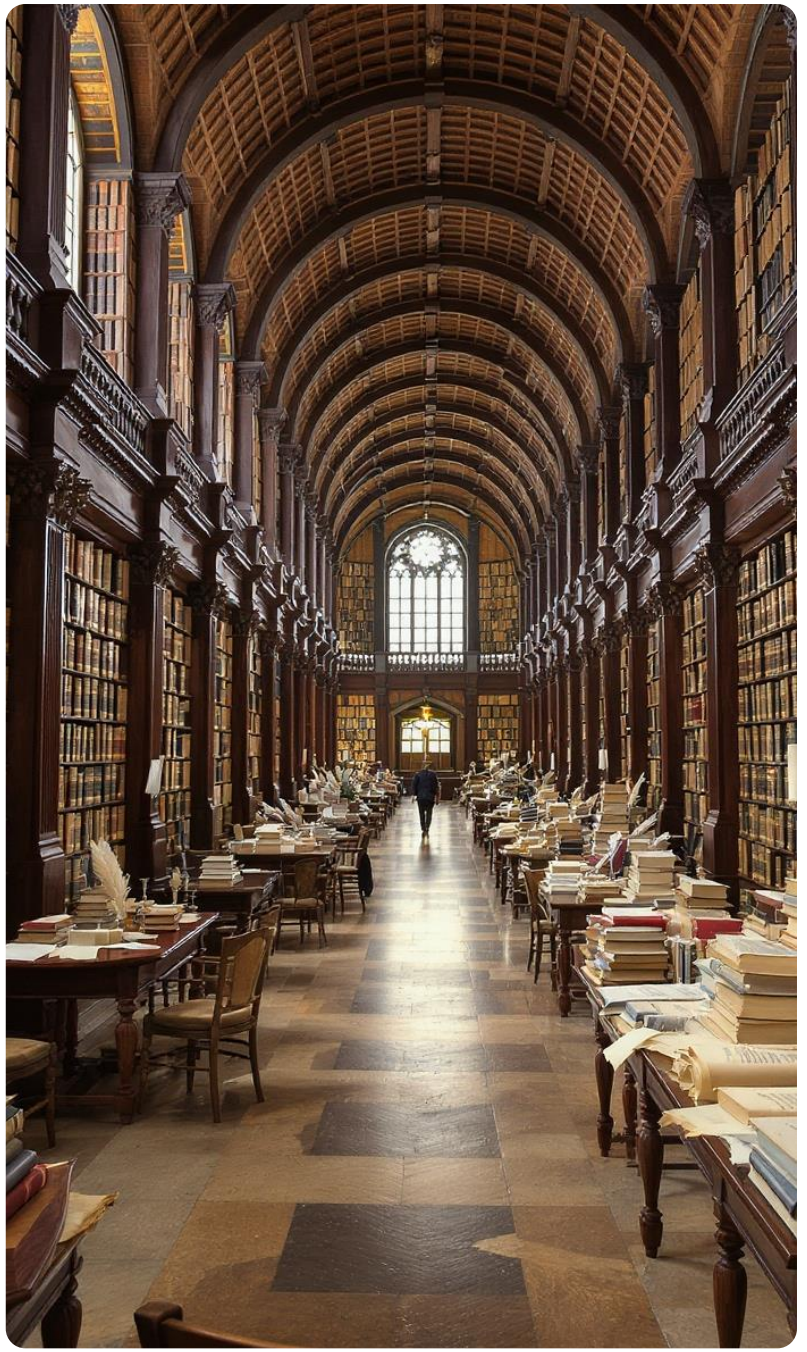


Conclusión y perspectivas futuras

- El marco de obstáculos sucesivos pone en práctica la evaluación con polígrafo mediante etapas claras, lo que mejora la precisión y la imparcialidad. Este modelo aborda desafíos científicos y éticos de larga data al dar prioridad a las pruebas de un solo tema y a los protocolos estandarizados. La síntesis ejemplifica la integración efectiva de la experiencia humana con la investigación, la edición y la verificación de datos asistidas por IA. La adopción de este marco puede impulsar la estandarización profesional y elevar la credibilidad de la evaluación con polígrafo. La investigación y la capacitación continuas son esenciales para perfeccionar la implementación y adaptarse a los contextos legales y operativos en constante evolución.

Referencias

- Normas de práctica de la Asociación Estadounidense de Poligrafía (2011, 2012, 2019) Informe del Consejo Nacional de Investigación (2003) sobre el polígrafo y la detección de mentiras Revisión de la validez científica de la Oficina de Evaluación Tecnológica (1983) Honts et al. (2021): metaanálisis de pruebas de polígrafo con preguntas comparativas; Nelson (2015, 2023): estudios sobre la precisión y las metodologías de las pruebas de polígrafo



Gracias