



## **Una Breve Historia de las Revisiones Científicas de las Investigaciones acerca de la Precisión del Polígrafo.**

**Raymond Nelson y Mark Handler**

¿Qué tan preciso es el polígrafo? Una pregunta engañosamente simple. Simple hasta el punto de ser simplista. Sobre simplificada. Hay que decir que la respuesta no es tan simple. Parte de la complejidad es que no hay un solo índice numérico que pueda representar adecuadamente un resultado de prueba con una solución que represente una probabilidad con un grado de incertidumbre o rango de error – y para el cual las discusiones acerca de la precisión de la prueba puedan tomar diferentes perspectivas de sus dimensiones, incluyendo las de la sensibilidad, especificidad, tasas, de error, y otras formas de describirla. Y aun cuando el asunto de la precisión no es tan simple, la pregunta tampoco es tan compleja al grado que no pueda ser estudiada y respondida, aunque no de manera simplista.

Para reducir la complejidad de las discusiones científicas, las discusiones acerca de la

precisión dentro del contexto de la investigación y las pruebas a menudo pueden hacerse mediante discusiones acerca de la tasa o probabilidad de error (también conocido como un valor  $p$ , y comúnmente expresado como decimal). La expectativa de cuantificar la posibilidad o probabilidad de un resultado erróneo es uno de los distintivos o estándares de la búsqueda científica.

Las probabilidades de error sirven las necesidades de conveniencia de los pensadores científicos, pero para las personas sin entrenamiento en metodologías científicas les es más fácil trabajar con un concepto diferente – lo contrario de la tasa de error, convenientemente referido como nivel de confianza (CL) con el que nuestra estimación de la probabilidad de un resultado correcto se expresa comúnmente en términos de porcentaje. Ya que la investigación en poligrafía involucra la precisión de resultados

tanto de veracidad como de engaño, el promedio simple no ponderado es uno de los números más útiles que podemos utilizar, con un solo número, para intentar describir los complejos fenómenos de la precisión del polígrafo. Esto puede pensarse como una precisión no ponderada, para la cual promediamos la precisión de los resultados veraces con los de engaño, al tiempo que asumimos tasas-base uniformes de personas culpables e inocentes.

Al igual que no esperamos lograr una estimación generalizable de la precisión del polígrafo a partir de un solo caso, tampoco esperamos lograr estimaciones generalizables de la precisión del polígrafo a partir de un solo estudio. En su lugar, es más probable que se logren estimaciones generalizables de la precisión del polígrafo mediante la revisión sistemática de los resultados de múltiples estudios. Se podría obtener una comprensión aún más amplia de la precisión del polígrafo mediante una evaluación histórica de los resultados de las revisiones científicas publicadas.

### **Una Cronología de las Revisiones Científicas acerca de la Precisión del Polígrafo**

Dentro de las primeras revisiones científicas publicadas acerca de la precisión del polígrafo está la de Abrams (1973), que revisó la literatura publicada que data de la primera parte del siglo XX, e informó una tasa de precisión agregada impresionante de 0,980.

Pocos años después, Abrams (1977) publicó los resultados de otra revisión de la precisión del polígrafo, y reportó una estimación de

.910 utilizando estudios más recientes y criterios más cuidadosamente estructurados para la selección de los estudios.

Podlesny y Raskin (1978) también revisaron la literatura disponible en un intento de cuantificar una respuesta a las preguntas relacionadas con la exactitud del polígrafo y reportaron en ese momento una tasa de precisión estimada de 0,890.

Ansley (1983) también intentó cuantificar la literatura disponible acerca de la exactitud del polígrafo y reportó los resultados de 1,964 casos de laboratorio y 1,113 casos de campo. Él describió un nivel de precisión en la decisión de .968 excluyendo los resultados inconclusos. Los resultados con el uso de la técnica relevante-irrelevante fueron más precisos (.960) que los que utilizaron la técnica con preguntas de comparación (.952). Los resultados de las pruebas de información oculta se reportaron con un índice de precisión de .912.

Por las mismas fechas, la Office of Technology Assessment (OTA, 1983) informó que los resultados de los estudios de laboratorio tenían una tasa de precisión agregada de .832, mientras que la precisión de los estudios de campo se reportó con un .847.

Abrams (1989) informó los resultados de un tercer estudio acerca de la precisión del polígrafo y describió una tasa de precisión agregada de .880 - algo más conservador que sus cálculos anteriores.

Ansley (1990) publicó otra revisión científica de la literatura relacionada con la precisión del polígrafo y resumió los resultados de 10 estudios, que incluyeron 2,042 polígrafos de investigación criminal desde 1980. Reportó una tasa de precisión de .980 para los casos de culpables y .970 para los casos de inocentes.

Honts y Peterson (1997) resumieron los resultados de estudios de polígrafos de alta calidad para el caso del escrito presentado en los EUA de Amicus vs Scherer, e informaron que la tasa de precisión excedía .900. Estos resultados también fueron reportados por Raskin y Honts (2002).

En un estudio independiente financiado por el U.S. Department of Defense, Crewson (2002) revisó los resultados de polígrafos diagnósticos y exploratorios y los comparó contra los resultados de pruebas exploratorias y diagnósticas utilizadas en medicina y psicología. Se reportó en ese momento que las pruebas de polígrafo diagnósticas tenían una tasa de precisión agregada de .880, y que las pruebas exploratorias tenían una tasa de precisión de .740.<sup>1</sup>

El National Research Council (NRC, 2003) realizó un estudio más reciente acerca de la precisión del polígrafo en el cual describió la precisión de la prueba en términos del Área Bajo la Curva (AUC) para la Característica Operativa del Receptor (ROC). La mediana del

AUC se reportó con .860 para estudios de laboratorio y .890 para los de campo.<sup>2</sup>

Utilizando un enfoque diferente para estimar la precisión de la prueba, Kokish, Levenson y Blasingame (2005) reportaron los resultados de una encuesta de opinión a delincuentes sexuales convictos que estaban sujetos a pruebas de polígrafo como requisito de libertad condicional. Reportaron que los encuestados describieron una alta tasa de acuerdo con los resultados de la prueba del polígrafo de más de .900.

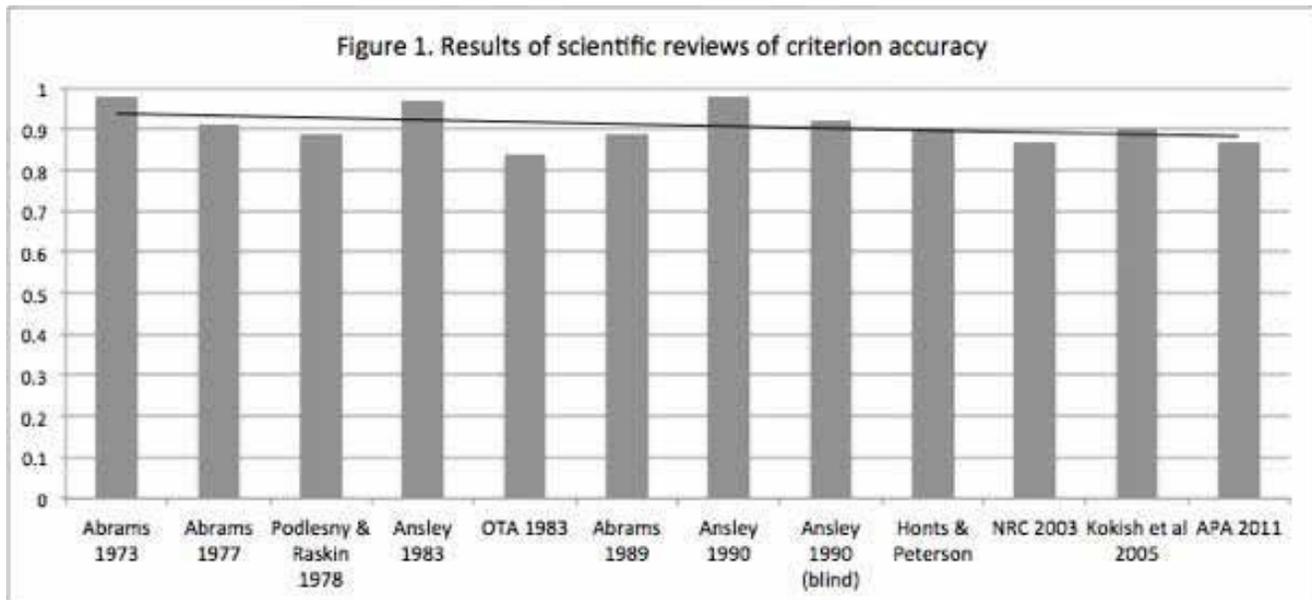
En la revisión meta-analítica del criterio de validez, la American Polygraph Association (APA, 2011) estimó la precisión de los polígrafos diagnósticos de evento específico en .890 con un rango de confianza del 95% de .829 a .951. Se informó que las técnicas poligráficas interpretadas bajo el supuesto de que los estados de criterio de las preguntas relevantes objetivo varían de manera independiente, tienen una tasa de precisión agregada de .850 con un rango de confianza del 95% de .773 a .926. La tasa de precisión agregada para todas las técnicas se reportó con .869 con un rango de confianza del 95% de .798 a .940.

---

<sup>1</sup> Todos los estudios exploratorios en Crewson (2002) se realizaron utilizando la técnica relevante-irrelevante. Esta cifra no puede generalizarse para los métodos exploratorios que utilizan técnicas de preguntas comparativas.

<sup>2</sup> En varias ocasiones se ha observado una precisión ligeramente mayor en los estudios de campo en comparación con los estudios de laboratorio. Las diferencias no son estadísticamente significativas, pero son inconsistentes con la tendencia de las observaciones en otros campos de la ciencia para los cuales los estudios de campo con condiciones de control menos estrictas a menudo producen una precisión ligeramente menor que los estudios de laboratorio. Una interpretación de este sentido podría ser la superioridad de los estudios de campo sobre los de laboratorio. Una interpretación más conservadora sería que los estudios de campo están sujetos a sesgos inherentes y procesos de selección no aleatorios.

La Figura 1 muestra los resultados de la serie de revisiones científicas de estudios publicados sobre los criterios de precisión de las pruebas de polígrafo, junto con la línea de tendencia.



## Discusión

La precisión media de estas revisiones científicas fue de .900, y la mediana fue de .890. Excluyendo los tres estudios con una precisión extremadamente alta, tanto la precisión media como la mediana de las revisiones científicas de las investigaciones sobre la precisión del polígrafo fueron de .890.

Aunque las revisiones científicas incluidas en este documento tal vez no sean exhaustivas o inclusivas de todos los estudios del polígrafo, son razonablemente incluyentes de la mayoría de los intentos útiles y visibles de un recuento sistemático de la evidencia y la varianza que está inevitablemente presente en los resultados de décadas de estudios de calidad diversa. La mayoría de los resultados de estos estudios parecen estar dentro de tolerancias comprensibles de error entre sí. Aunque tres revisiones (Abrams, 1973; Ansley, 1983, 1990) son más optimistas que otras, con resultados arriba de los .90s, los resultados de la mayoría de estos estudios (Abrams; 1977, 1989; Raskin y Podlesney, 1979; la OTA, 1983; Honts y Peterson, 1997); Raskin y Honts, 2002; la NRC, 2002; Kokish, Levenson y

Blasingame, 2005; y la APA, 2011) sugieren que la precisión del polígrafo está dentro de un rango predecible entre los .80s medios y altos, que se extiende hasta los .90s bajos.

Una de las tendencias más obvias en estos datos es que los estudios realizados por académicos sin interés particular en el polígrafo tendieron a producir resultados que son ligeramente menores que las estimaciones históricas promedio producidas por investigadores dentro de la profesión del polígrafo. Esta tendencia también parece corregirse con el tiempo, ya que los investigadores dentro de la profesión han comenzado a esforzarse en aplicar una metodología y escrutinios más rigurosos al cuestionar la precisión de los criterios.

Una segunda tendencia obvia es que los resultados de los estudios de campo y de laboratorio han tendido a coincidir de manera más estrecha. En este momento no hay evidencia disponible que sugiera que una forma de investigación sea superior a la otra. De hecho, todos los distintos tipos de diseños de investigación tienen ventajas y desventajas. Es inherentemente benéfico para la profesión utilizar distintos tipos de metodologías de investigación.

Otra tendencia es que los resultados producidos por investigadores dentro de la profesión del polígrafo y los proporcionados por académicos e investigadores fuera de la profesión, han tendido a estar más de acuerdo que en desacuerdo y con el tiempo, han comenzado a converger de manera más cercana.

Una cuarta tendencia obvia es que las estimaciones acerca de la precisión del polígrafo han tendido a disminuir con el tiempo. Si bien es posible que esto sea un reflejo de una disminución en la precisión de la prueba de polígrafo, probablemente no es posible que el polígrafo se haya vuelto menos preciso con el tiempo. Una explicación más plausible es que las estimaciones de la precisión del polígrafo simplemente se han vuelto más realistas o más precisas. Una mirada a los detalles del estudio mostrará que los primeros estudios se realizaron con gran énfasis en el uso de casos de campo confirmados, lo que puede introducir una vulnerabilidad por la sobreestimación de la precisión de la prueba como resultado de la exclusión sistemática de casos de error falsos negativos y falsos positivos para los que no se pudo obtener ninguna confesión o evidencia de confirmación<sup>3</sup>. Es improbable que los investigadores estuvieran explotando intencionalmente este potencial; en cambio, es más probable que algunas investigaciones y revisiones de investigación tempranas sobrestimaran la precisión del polígrafo como resultado de una vulnerabilidad de muestreo no reconocida. Independientemente, aunque las afirmaciones de precisión extremadamente altas

o casi perfectas no están bien respaldadas, los datos que están disponibles en este momento respaldan la hipótesis de que la prueba del polígrafo puede proporcionar niveles de precisión de prueba usualmente altos.

Una observación final es también interesante. Aunque se cree que los enfoques más rigurosos y científicos de la prueba de polígrafo aumentan su precisión, el grado de varianza en la evidencia publicada ha hecho difícil establecer si las intenciones de mejorar se han manifestado realmente en aumentos observables en la precisión de la prueba. Se puede ver que los resultados reportados por Honts y Raskin (1997) y Raskin y Honts (2002) han excedido marginalmente el nivel .900, pero la mayoría de las otras revisiones científicas de la literatura no exceden este nivel.

Esto puede ser un reflejo de que en la práctica se han utilizado métodos de polígrafo basados más en la tradición y menos en modelos científicos rigurosos. Pero también puede comenzar a ser un reflejo de que la precisión de la prueba del polígrafo se está acercando a cierto techo dentro del cual los aumentos observables y medibles en la precisión pueden ser difíciles de obtener sin una rigurosa adherencia a los protocolos científicos y las prácticas basadas en la evidencia.

---

<sup>3</sup> Las confesiones por sí mismas podrían ser el resultado de una decisión inherentemente no aleatoria de la búsqueda de información, lo que genera una advertencia importante sobre la necesidad de replicar los resultados observados mediante el estudio de datos de muestreo adquiridos a través del muestreo por confesiones. Además, si la decisión de buscar confesiones se basa en los resultados de una prueba poligráfica, puede haber una no-independencia inherente del criterio externo que provocará la posibilidad de que los resultados de la investigación sean auto complementarios.

La investigación del polígrafo ha estado en curso durante décadas y existe una gran cantidad de información en la literatura publicada. Es casi inevitable que los datos contengan información valiosa y al mismo tiempo han sido algo escandalosos. Al igual que los resultados de pruebas individuales tienen cierto potencial de error, los resultados de una investigación también tienen cierto potencial de error. Dado el tiempo y la evolución de los métodos estadísticos y de investigación, no es sorprendente que también pueda haber alguna variabilidad observable en los

resultados de las revisiones científicas de la evidencia publicada. La tendencia habitual de los datos agregados en la mayoría de los campos de la ciencia, medición y predicción estadística es que los resultados de las pruebas agregados tienden a superar a los resultados de las pruebas individuales y los resultados agregados de las investigaciones tienden a superar los resultados de los estudios individuales.

Por esta razón, se espera que se pueda encontrar algún valor en esta breve descripción general de los datos históricos.

## Referencias

---

- Abrams, S. (1973). Polygraph validity and reliability: A review. *Journal of Forensic Sciences*, 18, 313-326.
- Abrams, S. (1977). *A polygraph handbook for attorneys*. Lexington, MA: Lexington Books.
- Abrams, S. (1989). *Me complete polygraph handbook*. Lexington, MA: Lexington Books.
- American Polygraph Association (2011). Report of the Ad Hoc Committee on validated techniques. *Polygraph*, 40(4), 196-305.
- Ansley, N. (1983). A compendium on polygraph validity. *Polygraph*, 12, 53-61.
- Ansley, N. (1990). The validity and reliability of polygraph decisions in real cases. *Polygraph*, 19, 169-181.
- Crewson, P. E. (2001). A comparative analysis of polygraph with other screening and diagnostic tools. Research Support Service. Report No. DoDPI01-R-0003. Reprinted in *Polygraph* 32, (57-85).
- Honts, C. R. & Peterson, C.F. (1997). Brief of the Committee of Concerned Social Scientists as Amicus Curiae *United States v Scheffler*. Available from the author.
- National Research Council (2003). *The Polygraph and Lie Detection*. Washington, D.C.: National Academy of Sciences.
- Office of Technology Assessment (1983). *The validity of polygraph testing: A research review and evaluation*. Washington, D.C.: U.S. Congress, Office of Technology Assessment.
- Kokish, R., Levenson, J. S., & Blasingame, G. D. (2005). Post-conviction sexoffender polygraph examination: client-reported perceptions of utility and accuracy. *Sexual Abuse : A Journal of Research and Treatment*, 17, 211-21.
- Podlesny, J. A. & Raskin, D. C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. *Psychophysiology*, 15, 344-359.
- Raskin, D. C. & Honts, C. R. (2002). Handbook of polygraph testing. In M. Kleiner (Ed.), *Handbook of Polygraph Testing*. San Diego: Academic Press.