

**Analizando el Experimento de Pensamiento de Iacono, Acerca de los
Estudios de Campo del Polígrafo:**

¿Razón o Fantasía?

Charles R. Honts

Boise State University

y

Steven Thurber

Departamento de Servicios Humanos de Minnesota

Traductor

rodolfo@poligrafia.com.mx

Analizando el Experimento de Pensamiento de Iacono, Acerca de los Estudios de Campo del Polígrafo: ¿Razón o Fantasía?

Charles R. Honts

Boise State University

y

Steven Thurber

Departamento de Servicios Humanos de Minnesota

Abstract

Revisamos y analizamos un experimento de pensamiento publicado por primera vez en Iacono (1991) y reintroducido en Iacono y Ben-Shakhar (2019). El Experimento de Pensamiento de Iacono, *Iacono Thought Experiment (ITE)* parece haber utilizado el método de *Backcasting* para generar una serie de supuestos y condiciones previas que podrían hacer posible tener una prueba de polígrafo con una precisión al azar, que produce un estudio de campo con alta precisión bajo el criterio de confesión. A partir de este experimento de pensamiento, Iacono promulgó una hipótesis de que todos los estudios de polígrafo utilizando criterios de confesión producen estimaciones de precisión exageradamente altas. Nuestro análisis de los supuestos y condiciones previas del ITE encontró que no son representativos y es muy poco probable que se cumplan en condiciones del mundo real. Utilizamos un enfoque de evidencia convergente que aplicó resultados meta analíticos, estudios de campo que no utilizaron un criterio de confesión, y datos de casos que involucraron exámenes de polígrafo con condena injusta para probar la hipótesis de Iacono. Encontramos una fuerte evidencia de falsificación de la hipótesis de Iacono y concluimos que debe ser abandonada como una descripción del significado de la investigación del polígrafo de campo.

Palabras clave: ocular-motor, detección de engaño, seguimiento ocular, lectura

Nota del autor

La correspondencia debe dirigirse a Charles R. Honts, Ph. D., Departamento de Ciencias Psicológicas, Boise State University, 1910 University Drive MS-1715, Boise ID 83725-1715. Los autores desean agradecer a Adela Stephanescu por su ayuda en la edición del manuscrito completo

Analizando el Experimento de Pensamiento de Iacono Sobre los Estudios de Campo del Polígrafo: ¿Razón o fantasía?

Las pruebas de polígrafo representan una aplicación importante y generalizada de una prueba psicológica en la procuración de la ley, seguridad nacional y el empleo en todo el mundo. Internacionalmente, la American Polygraph Association muestra miembros de 62 países. Zhang (2011) estimó que había hasta 8000 examinadores de polígrafo operando solo en China. A pesar de la naturaleza ubicua de las pruebas de polígrafo, ha recibido relativamente poca atención de la psicología académica y, a menudo, esa atención ha sido en forma de comentarios negativos.

La prueba de polígrafo más comúnmente utilizada - y criticada - es la Prueba de Preguntas de Comparación (CQT). El CQT viene en diferentes variantes, pero en todos los casos, monitorea la fisiología autonómica del sujeto (generalmente respiración, actividad electrodérmica, presión arterial relativa y a menudo, actividad vasomotora periférica) mientras el sujeto responde una serie de preguntas. La serie de preguntas contiene dos categorías críticas de preguntas (generalmente dos o tres de cada una). Las preguntas relevantes abordan directamente el asunto bajo investigación. Las preguntas de comparación están diseñadas y presentadas de manera tal que cada sujeto mienta, o se asume que mienta, en su respuesta a ellas durante la prueba. Se espera que las respuestas fisiológicas del sujeto muestren una interacción, de modo que los sujetos que presentan engaño ante las preguntas relevantes muestren respuestas fisiológicas más grandes a las preguntas relevantes en comparación con las preguntas de comparación. Se espera que los sujetos que son veraces ante las preguntas relevantes muestren un patrón opuesto, y que las respuestas fisiológicas a las preguntas de comparación sean mayores que a las de las preguntas relevantes.

Hay una serie de revisiones de la investigación del CQT, tipificadas pero no limitadas a los siguientes ejemplos: Raskin, Honts y Kircher (1997), Iacono y Lykken (1997), The National Research Council (NRC, 2003), Honts (2004), Vrij (2008); Asociación Americana de Poligrafía (2011); Raskin, Honts y Kircher (2014), y Iacono y Ben-Shakhar (2018). Existe una variación entre las revisiones, sin embargo, generalmente produjeron estimaciones de precisión en general superiores al 80%.

Sin embargo, todas esas revisiones pueden ser criticadas por las elecciones selectivas de estudio y la falta de escrutinio meta analítico. Ninguno intentó probar las variables moderadoras, aunque a veces llegaron a conclusiones que hipotetizaban o incluso

asumían efectos moderadores poderosos. El informe de la NRC (2003) fue particularmente notorio al respecto. NRC encontró objetivamente una alta discriminación al estimar el área bajo la curva (AUC) de las Características Operativas del Receptor (ROC) en 0.89. El uso del análisis ROC y el AUC como tamaño del efecto ha sido criticado como una aplicación inapropiada de una tecnología desarrollada para examinar el desempeño de la detección de señal con tecnología (específicamente la capacidad del operador para ver pantallas RADAR y distinguir barcos enemigos, barcos amigos y ruido; Tape, 2019) con la psicología en general (Balakrishnan, 1999), y con las pruebas de polígrafo en particular (Honts y Schweinle, 2009).

Sin embargo, el NRC utilizó el AUC como su índice de tamaño del efecto, pero ¿qué significa realmente el AUC y qué implica un valor de AUC de 0.89 con respecto al desempeño del polígrafo? El valor de AUC puede variar de 0.50, que representa un desempeño al azar, a 1.00 que representa un desempeño perfecto (100% de precisión; Tape, 2019). Tape (2019) caracteriza cualitativamente los valores AUC entre 0.80 y 0.90 como indicativos de un buen discriminador y los valores de AUC superiores a 0.90 como excelentes. Las tablas (Rice y Harris, 2005) y el software (DeCoster, 2012) para convertir entre AUC y otras mediciones el tamaño del efecto, están fácilmente disponibles. La referencia de esas tablas y el software muestra que un valor de AUC de 0.89 correspondió a un valor *d de Cohen* de 1.74 y un valor de r_{pb} de 0.66. Cohen (1969, 1988, 1992) describió a los grandes efectos en la psicología como aquellos con valores de *d* superiores a 0.80 (r_{pb} correspondiente > 0.49). De manera memorable, Cohen dijo que, en psicología aplicada, los tamaños de efecto de $d = .8$ son "casi tan altos como llegan" (Cohen, 1988, p. 81). Por lo tanto, el tamaño del efecto AUC informado por NRC (2003) indica un desempeño extremadamente alto para el CQT en comparación con otras pruebas y medidas psicológicas.

A pesar de la poderosa evidencia empírica de la utilidad del CQT como discriminador entre la veracidad y el engaño, la NRC descontó esos hallazgos y dijo que los métodos de investigación eran deficientes. Para los presentes autores, esto parece ser una conclusión arrogante, ya que el NRC sustituyó su juicio sobre las cualidades de la investigación publicada en revistas de ciencia psicológica revisadas por pares de primer nivel. Tal posición es insultante para los editores de esas revistas de primer nivel y para los científicos que trabajan para ellos en la revisión por pares. La opinión de la NRC es peor por el hecho de que ninguno de los miembros del comité de la NRC que escribieron el informe había publicado un estudio sobre detección del engaño.

Adicionalmente, el NRC y otros fueron especialmente críticos con el uso de estudios experimentales (de laboratorio) para evaluar la validez de CQT. Iacono y Lykken (1997) descartaron por completo la investigación experimental sobre el CQT, argumentando que los contextos motivacionales del mundo real no podían modelarse experimentalmente y, por lo tanto, los resultados de laboratorio eran cualitativamente diferentes de aquellos en casos reales. Sin embargo, las conclusiones de la NRC y Iacono y Lykken (1997) sobre la generalización de los métodos de investigación en revistas archivadas de ciencia psicológica y la generalización de la investigación experimental acerca del CQT deben considerarse como una opinión y no como un hecho, ya que ninguna de esas opiniones estaba o está basada en datos.

La historia de desacuerdos académicos sobre la precisión (validez de criterio) del CQT es larga y en ocasiones ha sido polémica. Esos desacuerdos se caracterizan, y no se limitan a, los intercambios publicados entre investigadores de la Universidad de Utah y la Universidad de Minnesota, a partir de 1978 en las revistas *Psychophysiology* (Raskin y Hare, 1978; Lykken, 1978; Raskin, 1978) y *Psychological Bulletin* (Lykken, 1979; Raskin y Podlesny, 1979). Los intercambios directos en la literatura entre estos grupos continuaron hasta 2002 (Honts, Raskin y Kircher, 2002; Iacono y Lykken, 2002). Esos desacuerdos se discutieron en varios niveles sobre diversos temas. A lo largo de los desacuerdos, el grupo de Minnesota rechazó radicalmente la noción de que la detección del engaño podría modelarse de manera válida en laboratorio y sostuvo que los resultados de los estudios de laboratorio no eran útiles para estimar la validez del criterio en aplicaciones de campo porque carecían de validez externa (generalizabilidad). El grupo de Minnesota mantiene esa posición hasta hoy a pesar del rechazo general de tales críticas en toda la Ciencia Psicológica y específicamente para la detección del engaño. Hartwig y Bond (2014) proporcionaron una discusión general sobre la generalizabilidad de los estudios de laboratorio y proporcionaron un rechazo empírico específico de las diferencias entre los entornos experimentales y de campo, dentro de un meta análisis de la literatura de investigación de la detección del engaño interpersonal.

Los problemas científicos que rodean el contraste entre los entornos experimentales y de campo para la investigación de la detección del engaño interpersonal son casi idénticos a los del CQT. Recientemente, Honts y Thurber (2019) informaron un metanálisis integral del CQT que siguió el enfoque analítico de Hartwig y Bond (2019). Honts y Thurber (2019) informaron efectos estadísticamente no detectables para los moderadores de motivación, población de sujetos o entorno (experimental versus campo) en su meta análisis integral del CQT.

El grupo de Minnesota inicialmente apoyó los estudios de campo que se ajustaban a su criterio para estudios útiles de campo. Sin embargo, a partir de la década de 1980, se publicaron estudios de campo que produjeron altos niveles de precisión con el CQT (Honts y Raskin, 1988; Raskin, Kircher, Honts y Horowitz, 1988). Esos estudios fueron diseñados específicamente para cumplir con los criterios del grupo de Minnesota. Posteriormente, el grupo de Minnesota rechazó todos los estudios de campo y tomó la posición radical de que no se podían realizar investigaciones válidas sobre el CQT. Una piedra angular de esa posición fue un experimento de pensamiento reportado por primera vez por Iacono (1991) y luego reintroducida con alguna modificación en Iacono y Ben-Shakhar (2018)¹. El experimento de pensamiento de Iacono (1991) se presentó originalmente de la siguiente manera:

Supongamos que se investigan 800 delitos utilizando una técnica de polígrafo que funciona con precisión al azar exacta; es decir, la mitad de los sospechosos culpables e inocentes fracasarán y la otra mitad pasará. Debido a que el polígrafo se usa a menudo en crímenes por los cuales hay múltiples sospechosos, supongamos, sin pérdida de generalidad, que estamos tratando con 800 crímenes de dos sospechosos, y que para cada uno de ellos uno es culpable y el otro inocente. Supongamos además que (1) el sospechoso culpable es examinado el 50% de las veces, (2) el segundo sospechoso no será evaluado si el resultado de la primera prueba indica engaño, (3) ni los sospechosos inocentes ni los sospechosos culpables quienes aprueben el examen confesarán y (4) el 20% de los culpables que no pasen el examen y posteriormente son interrogados confiesan. (Iacono, 1991, págs. 202-203).

Los experimentos de pensamiento son bien conocidos en la filosofía y la ciencia. Los experimentos de pensamiento pueden definirse como "dispositivos de la imaginación utilizados para investigar la naturaleza de las cosas" (Brown, 2014). Uno de los experimentos de pensamiento científico más famosos fue el razonamiento de Galileo de que dos objetos de diferente peso deben caer a la misma velocidad.

¹En Iacono y Ben-Shakhar proporcionaron dos versiones simplificadas del Experimento de Pensamiento de Iacono con sujetos individuales y con sujetos apareados. Los supuestos 3, 4 y 5 no se aplican ni al sujeto individual ni a las pruebas pareadas, ya que todos los sujetos se evalúan independientemente de los resultados. Las tasas de confesión no se especifican para ninguno de los análisis, por lo tanto, la suposición 7 no es específica. Los otros supuestos son explícitos o implícitos en la última versión del Experimento de Pensamiento de Iacono.

El experimento de pensamiento de Galileo se validó fácilmente a partir de observaciones, como en la demostración concluyente de Neil Armstrong en la luna cuando dejó caer una pluma y un martillo simultáneamente y aterrizaron en la superficie al mismo tiempo (Pigllucci, 2006). Pigllucci señala además que los experimentos de pensamiento también pueden estar equivocados y ser falsificados por los datos. Si el experimento de pensamiento de Galileo hubiera sido invalidado por los datos, se habría perdido en la historia y olvidado.

Los experimentos de pensamiento pueden tomar varias formas o tipos. Si bien una discusión sobre los múltiples tipos de experimentos de pensamiento está más allá del alcance de este documento, vale la pena señalar que el experimento de pensamiento de Iacono parece ser de un tipo conocido como Backcasting (Robinson, 1982). En Backcasting uno imagina un estado deseado o posible del mundo y luego razona hacia atrás desde ese estado final hacia los precursores necesarios. Por definición, dicha lógica no necesariamente proporciona una descripción de la realidad, solo proporciona una cadena de precursores que podrían producir el estado final deseado. Tales experimentos mentales, como todos los experimentos mentales, son útiles en el mundo real solo en la medida en que puedan ser probados y validados o falsificados con datos. Comenzamos nuestro análisis del Iacono Thought Experiment (ITE), Experimento de pensamiento de Iacono, definiendo los precursores hipotéticos que él inventó o seleccionó para alcanzar el estado final deseado donde las pruebas de polígrafo con precisión al azar podrían producir un estudio de campo con altas tasas de precisión.

Aclaración y Análisis de las Precondiciones Hipotéticas y Suposición del Experimento de Pensamiento de Iacono

Suposiciones Explícitas del Experimento de Pensamiento de Iacono (1991).

Iacono (1991) hace una serie de suposiciones explícitas que se utilizaron para crear una posible ruta hacia el estado final deseado.

1. Ochocientos sujetos son evaluados donde 400 son Inocentes y 400 son Culpables.
2. El polígrafo se desempeña exactamente con una probabilidad de 50% correcta, 50% incorrecta y sin resultados inconclusos. Esta suposición es parte del estado final general deseado donde una prueba de polígrafo al azar podría

producir resultados de alta precisión. Todos los demás supuestos también sirven para el establecimiento de ese estado final.

3. Cada delito tiene solo dos sospechosos. (Iacono hace esta suposición y afirma que se hace "sin pérdida de generalidad" (p. 203).

4. El sospechoso culpable se evalúa primero en la mitad de los casos

5. Si el primer sospechoso falla la prueba del polígrafo, el segundo sospechoso no será evaluado.

6. Ni los inocentes ni los culpables que pasan la prueba confesarán.

7. Solo el 20% de los sospechosos culpables que fallan y son interrogados confesarán.

Suposiciones implícitas del experimento de pensamiento de Iacono. Los siguientes supuestos implícitos son también necesarios para que las matemáticas y la lógica del experimento de pensamiento de Iacono alcancen el estado final deseado.

8. El polígrafo es la única fuente de información sobre quién es culpable en un caso criminal.

9. Las personas culpables solo confiesan después de los exámenes poligráficos.

Análisis de los Supuestos Explícitos del Pensamiento de Iacono

Experimento

La suposición 1 es que la tasa base de sujetos culpables e inocentes es igual. La tasa base de culpables en un caso criminal variará mucho dependiendo de cuándo se use el polígrafo. Si se utiliza al principio de una investigación, es probable que haya mucho más inocentes que sujetos culpables; si se utiliza muy tarde en una investigación, puede haber muchos más culpables que inocentes. La suposición de tasas base iguales es aceptable para un experimento mental siempre que se reconozca que las variaciones en la tasa base podrían alterar drásticamente los resultados del estado final y que una tasa base del 50% será inusual en la práctica real.

La suposición 2 es que el polígrafo funciona exactamente como el lanzamiento de una moneda. Esta suposición se realiza como una premisa del experimento de pensamiento y es un componente necesario del estado final deseado. Sin embargo, esta premisa carece de apoyo empírico en el mundo real. Hasta donde sabemos, no hay estudios que muestren ninguna versión de que el CQT se desempeña a niveles de azar.

La suposición 3 es que cada caso tiene solo dos sospechosos. Esta premisa simplifica las matemáticas necesarias para lograr el estado final deseado del experimento de pensamiento de Iacono, pero es una premisa que rara vez se cumple en el mundo real, y no es en absoluto representativa del campo en el momento en que se escribió Iacono (1991) (por ejemplo, Honts y Raskin, 1988, Raskin, Kircher, Honts y Horowitz, 1988 contienen muchos casos de sospechosos únicos y múltiples al igual que los estudios de campo más recientes). La afirmación de Iacono de que esta suposición se hace sin una pérdida de generalidad (para el experimento de pensamiento Backcasting) claramente no está respaldada por los datos.

La suposición 4 es que un sospechoso culpable es examinado primero en la mitad de los casos. Esta suposición es sostenible solo si únicamente hay dos sospechosos y el examinador no tiene ningún motivo para evaluar primero a uno u otro sospechoso. Es una suposición conveniente para el experimento de pensamiento, pero es poco probable que sea ampliamente representativo de las pruebas de polígrafo de campo.

La suposición 5 establece que, si la primera persona es examinada y falla, el segundo sospechoso no será evaluado. Sin embargo, este no es el caso en investigaciones reales. Si el primer sujeto se prueba y falla, pero no confiesa, entonces el sospechoso o sospechosos restantes probablemente serán evaluados para evaluar su participación en el crimen. La lógica probable de los investigadores sería que los sospechosos adicionales no serían sospechosos a menos que hubiera una razón para sospechar de ellos, y bien podrían estar involucrados. En nuestra experiencia, de hecho, es una práctica común evaluar a todos los sospechosos en un caso durante una investigación.

La suposición 6 establece que ni los sospechosos inocentes ni los sospechosos culpables que pasan la prueba confesarán. Esto es manifiestamente falso. Bajo ciertas circunstancias, como el de un resultado de prueba de polígrafo erróneamente reprobada o tergiversada deliberadamente, los sospechosos inocentes confesarán crímenes que no cometieron. El Documento Blanco de la Sociedad Americana de Derecho de Psicología (Kassin, Drizin, Grisso, Gudjonsson, Leo y Redlich, 2010) señala específicamente que los resultados del polígrafo fallados erróneamente o tergiversados intencionalmente son una estrategia poderosa de evidencia falsa que pone a los realmente inocentes en un mayor riesgo de una confesión falsa. Además, Bonpasse (2013) proporciona ejemplos y discute casos reales en los que resultados poligráficos incorrectos o tergiversados han contribuido a errores judiciales mediante su papel en la obtención de confesiones falsas. La suposición 6 también es incorrecta para los sospechosos culpables que aprueban polígrafos, ya que ignora el hecho de que las investigaciones rara vez se detienen solo porque se ha aprobado un polígrafo. Si la investigación subsecuente continúa y se obtiene información adicional, entonces el sospechoso probablemente será entrevistado por segunda vez, a pesar del polígrafo

aprobado, y puede proporcionar una confesión o confesarse más tarde como parte de un acuerdo de declaración de culpabilidad. Al menos uno de estos casos se incluyó en Honts y Raskin (1988).

La suposición 7 establece que solo el 20% de los sospechosos culpables que son interrogados confesarán. La opción del 20% es arbitraria y no tiene una base empírica. La tasa de confesión real dependerá de la situación en la que se realizaron las pruebas. Es poco probable que las pruebas de poligrafía realizadas por abogados defensores, o por la policía sobre sujetos que tienen un abogado defensor, sean seguidas de interrogatorios, independientemente del resultado del polígrafo. Por otro lado, el Departamento de Defensa de EE. UU. (2002) ha informado datos que indican que en un programa de polígrafo, más del 90% de los exámenes de polígrafo reprobados resultaron en admisiones relevantes. Claramente, la tasa elegida para este supuesto tendrá un impacto importante en los resultados resultantes del ITE. Además, dado que todos los valores de la tasa de confesión son específicos de la situación, no es sensato proporcionar un valor único para una tendencia central, ya que dicho valor no tendría sentido para ningún entorno específico aplicado.

Análisis de los Supuestos Implícitos del Experimento de Pensamiento de Iacono

La suposición 8 afirma que el polígrafo es la única fuente de información sobre quién es culpable en una investigación criminal. Esto ignora el hecho de que las personas pueden confesar en contextos distintos a los exámenes poligráficos, o que se puede obtener otra evidencia incontrovertible de culpabilidad o inocencia independiente a la prueba poligráfica. Esto fue examinado explícitamente en Honts (1996) y no se encontraron diferencias en los puntajes numéricos o las tasas de precisión entre los casos de confesión confirmada y de evidencia confirmada para sujetos culpables o inocentes. Curiosamente, en Honts (1996) ninguno de los sujetos inocentes fue confirmado por una confesión obtenida en el contexto de un examen poligráfico. El supuesto 8 es absurdo, pero el ITE no puede funcionar sin él.

La suposición 9 es que las personas culpables solo confiesan después de las pruebas de polígrafo. Según lo cubierto en nuestra discusión de la suposición 6, notamos que las pruebas de polígrafo inexactas y tergiversadas pueden dar lugar a confesiones falsas de sujetos inocentes. También notamos que incluso los sujetos culpables que pasan las pruebas de polígrafo a veces confiesan más tarde, cuando se enfrentan a evidencia nueva o abrumadora. Además, los sospechosos culpables, y ocasionalmente sujetos inocentes, confesarán como parte de un acuerdo de culpabilidad. Por lo tanto, es obviamente cierto que las personas culpables que no pasan una prueba de polígrafo pero que no son interrogadas, o que inicialmente se resisten a un interrogatorio,

pueden confesar más tarde. Al menos tres estudios de campo han tenido esto en cuenta explícitamente y buscaron información de confirmación en una muestra exhaustiva de casos dentro de un período de tiempo particular, y utilizaron toda la información disponible no solo en el archivo del examen poligráfico, sino en el registro policial completo del caso (Honts, 1996; Patrick e Iacono, 1991; Raskin et al., 2019).

Resumen del Análisis de Precondiciones y Supuestos

Nuestro análisis muestra que los supuestos del experimento de pensamiento de Iacono se eligieron generalmente sin referencia a los datos o la práctica profesional, al servicio de desarrollar lo que se convirtió en un conjunto altamente improbable de condiciones previas y supuestos que conducen a una solución específica que muestra que una prueba de polígrafo con precisión de azar podría producir un estudio de campo con tasas de alta precisión. El experimento de pensamiento de Iacono se transformó en una declaración normativa de que todos los estudios de campo del CQT eran, son y serán para siempre poco confiables y sobreestiman la precisión real. No creemos que esta conclusión normativa esté justificada a menos que pueda resistir el examen empírico y la falsificación. En el resto de este artículo, nos referiremos a la hipótesis derivada del experimento de pensamiento de Iacono, de que el CQT no es más preciso que el azar y que todos los estudios con criterios de confesión están sesgados para sobreestimar dramáticamente la precisión del CQT como la hipótesis del experimento de pensamiento de Iacono (ITEH)

Datos que Podrían Palsificar la Hipótesis del Experimento de Pensamiento de Iacono (ITEH)

Al igual que el experimento de pensamiento de Galileo sobre objetos que caen, el ITEH sobrevive a la prueba de la ciencia basada en la falta de datos de falsificación en la investigación científica. Esto lleva a la pregunta de qué datos falsificarían la ITEH. El resto de este documento aborda varias fuentes de datos convergentes que, de hecho, conducen a la conclusión de que la ITEH es falsa.

Convergencia de Datos Experimentales y de Campo sin Efectos Moderadores Detectables

Estudios recientes resumidos por Hartwig y Bond (2014) han indicado una fuerte convergencia entre los estudios experimentales y de campo en ciencias psicológicas y la detección del engaño interpersonal. Hartwig y Bond rechazaron explícitamente la

noción de que los experimentos y la investigación de campo sobre la detección del engaño interpersonal produjeron resultados significativamente diferentes. Si la ITEH fuera cierta, entonces las pruebas de polígrafo tendrían que ser mecanismos subyacentes cualitativamente diferentes al engaño interpersonal. En tales circunstancias, esperaríamos que los estudios de laboratorio del CQT produjeran precisiones dramáticamente más bajas que las precisiones exageradamente altas (según el ITEH) producidas por los efectos supuestamente inevitables de la ITEH en los estudios de campo del CQT. Las revisiones existentes simplemente no revelan resultados dramáticamente más precisos en el campo que en el laboratorio (NRC, 2003; Honts y Thurber, 2019).

Falta de Diferencias en la Precisión Entre los Estudios de Campo que se Basan en el Criterio de Confesión y los Que No.

Dado que la ITEH está críticamente vinculada al uso de confesiones como criterio de confirmación de Culpabilidad e Inocencia, y la ITEH predice que el criterio de confesión sesga críticamente los estudios de campo para mostrar una alta precisión, deberíamos esperar que los estudios de campo que incluyeron o utilizaron otros métodos de confirmación produciría precisiones que se acercan a niveles de precisión de azar. Empíricamente, este simplemente no es el caso. Honts (1996) probó directamente esta hipótesis, calificando la fuerza de la confirmación en una escala que iba desde confesiones con la generación de nueva evidencia en un extremo de la escala, hasta la no confirmación en el otro. Honts (1996) probó esa escala contra la precisión de la decisión y contra los puntajes numéricos. En oposición directa a las predicciones de la ITEH, Honts no encontró efectos para el nivel de confirmación. Es decir, los casos de confesión confirmada no tenían niveles de precisión más altos que los casos que fueron confirmados por métodos distintos de la confesión (evidencia física y/o declaraciones de testigos).

De igual manera, hay dos estudios de campo que utilizan pruebas apareadas y matemáticas para estimar la precisión (Ginton, 2013; Mao, Liang y Hu, 2015). Este enfoque de prueba apareada, aunque no sin problemas (Iacono y Ben-Shakhar, 2018), no depende de las confesiones y, por lo tanto, está fuera del alcance de la ITEH. Las tasas de precisión estimadas de los estudios de sujetos apareados convergen con los datos de los estudios de laboratorio y de campo y, por lo tanto, brindan apoyo para ambos.

Falta de Concurrencia entre Condenas Erróneas y Pruebas Fallidas de Polígrafo CQT.

Si la ITEH tiene razón en que los polígrafos CQT no son más precisos que el azar, esperaríamos que, en promedio, la mitad de los sujetos inocentes examinados en entornos de justicia penal produjeran errores falsos positivos. En los entornos de justicia penal, los sujetos inocentes que fallaron el polígrafo estarían expuestos a interrogatorios y, por lo tanto, correrían el riesgo de una confesión falsa. En tales circunstancias, esperaríamos que haya un número relativamente grande de resultados falsos positivos entre las filas de los condenados injustamente. Bonpasse (2013) revisó los archivos de casos del Registro Nacional de Exoneraciones², fundado en 1989 como un proyecto conjunto de la Facultad de Derecho de la Universidad de Michigan y la Universidad Northwestern. Bonpasse informó haber encontrado 215 casos de exoneración en los que participaron pruebas de polígrafo.

De esos 215 casos, solo 23 (10.7%) contenían información de que un sujeto inocente había sido evaluado antes del juicio y había fallado el polígrafo. Sin embargo, hubo 44 (20.5%) sujetos inocentes que habían sido probados con el polígrafo antes del juicio que produjeron resultados veraces, pero esos resultados favorables no los ayudaron a evitar una condena injusta. Aunque la ITEH predice que los errores falsos positivos deberían ser comunes entre los condenados injustamente, ocurrieron a solo la mitad de la tasa de resultados verdaderos negativos. Bonpasse también informó que en todas las pruebas, antes y después del juicio, incluidas las pruebas del sospechoso inmediato y otros (coacusados y testigos), 135 (62.9%) del resultado de la prueba del polígrafo fueron favorables para la persona condenada injustamente, mientras que solo 31 (14.4%) produjo resultados desfavorables. Los datos de los condenados injustamente contradicen fuertemente y por lo tanto falsifican la ITEH.

Discusión

Hasta donde sabemos, no hay un solo estudio del CQT, ya sea de laboratorio o de campo, que produzca tasas de precisión de azar. Si bien existe una gran cantidad de variabilidad entre los estudios del CQT, no se ha encontrado ninguna revisión de que los estudios de laboratorio son dramáticamente menos precisos que los estudios de campo (NRC, 2003; Honts y Thurber, 2019). Por lo tanto, el ITEH carece por completo de fundamentación empírica.

²<http://www.law.umich.edu/special/exoneration/Pages/about.aspx>

Más aún, los datos del meta análisis de Honts & Thurber (2019), los estudios de campo que no utilizan el criterio de confesión y los condenados injustamente proporcionan todos evidencia de que la ITEH es falsa. Los resultados del ITE son entonces vistos como un experimento de pensamiento fallido que carece completamente de soporte empírico, y que debería ser relegado al bote de basura de las ideas históricamente fallidas.

Referencias

- American Polygraph Association (2011). Meta-analytic survey of criterion accuracy of validated polygraph techniques. *Polygraph*, 40, 194-305.
- Balakrishnan, J. D. (1999). Decision processes in discrimination: Fundamental misrepresentations of signal detection theory. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1189-1206.
- Bonpasse, M. (2013). Polygraph and 215 wrongful conviction exonerations. *Polygraph*, 42, 112-127.
- Brown, J. R. (2014). Thought experiments. *Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/thought-experiment/>
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Edition). Hillsdale NJ:Erlbaum.
- Cohen, J., (1992). A power primer. *Psychological Bulletin*, 122, 155-159.
- DeCoster, J. (2012). Converting effect sizes 2012-06-19-4.xls Retrieved from <http://www.stat-help.com/spreadsheets.html>
- Ginton, A. (2013). A non-standard method for estimating accuracy of lie detection techniques demonstrated on a self-validating set of field polygraph examinations. *Psychology, Crime, & Law*, 19, 577-594. DOI: <http://dx.doi.org/10.1080/1068316X.2012.656118>
- Hartwig, M., & Bond, C. F. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology*, 28, 661-676.
- Honts, C. R. (1996). Criterion development and validity of the control question test in field application. *The Journal of General Psychology*, 123, 309-324.

Honts, C. R. (2004). The psychophysiological detection of deception, in P. Granhag and L. Strömwall (Eds.) *Detection of deception in forensic contexts*. London: Cambridge University Press 103- 123.

Honts, C. R., & Raskin, D. C. (1988). A field study of the validity of the directed lie control question.

Journal of Police Science and Administration, 16, 56-61.

Honts, C. R., & Schweinle, W., (2009). Information gain of psychophysiological detection of deception in forensic and screening settings. *Applied Psychophysiology and Biofeedback*, 34, 161-172. (Available online July 2009)

Honts, C. R., & Thurber, S. (2019, March). A Comprehensive Meta-Analysis of the Comparison *Question Polygraph Test*. Paper presented at the annual meeting of the American Psychology Law Society, Portland, Oregon.

Honts, C. R., Raskin, D. C., & Kircher, J. C. (2002). The scientific status of research on polygraph techniques: The case for polygraph tests. In, D. L. Faigman, D. Kaye, M. J. Saks, & J. Sanders (Eds.) *Science in the Law: Social and Behavioral Sciences Issue*, American Casebook Series (pp. 598-634). West Group: St. Paul, Minnesota.

Iacono, W. G. (1991). Can we determine the accuracy of polygraph tests? In J. R. Jennings, P.K.

Ackles & M. G. H. Coles (Eds.) *Advances in psychophysiology* (pp. 201-207). London, UK: Jessica Kingsley Publishers.

Iacono, W. G., & Ben-Shakhar, G. (2019). Current Status of forensic lie detection with the comparison question test: An Update of the 2003 National Academy of Sciences report on polygraph testing. *Law and Human Behavior*, 43, 86-98.

Iacono, W. G., & Lykken, D. T. (1997). The scientific status of research on polygraph techniques: The case against polygraph tests. In, D. L. Faigman, D. Kaye, M. J. Saks, & J. Sanders (Eds.) *Science in the Law: Social and Behavioral Sciences Issue*, American Casebook Series (pp. 582-618-). West Group: St. Paul, Minnesota.

Iacono, W. G., & Lykken, D. T. (2002). The scientific status of research on polygraph techniques: The case against polygraph tests. In, D. L. Faigman, D. Kaye, M. J. Saks, & J. Sanders (Eds.) *Science in the Law: Social and Behavioral Sciences Issue, American Casebook Series* (pp. 634-688). West Group: St. Paul, Minnesota.

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, 34, 3-38.

Lykken, D. T. (1978). The psychopath and the lie detector. *Psychophysiology*, 15, 137-142. [http:// dx.doi.org/10.1111/j.1469-8986.1978.tb01349.x](http://dx.doi.org/10.1111/j.1469-8986.1978.tb01349.x)

Lykken, D. T. (1979). The detection of deception. *Psychological Bulletin*, 86, 47-53. <http://dx.doi.org/10.1037/0033-2909.86.1.47>

Mao, Y., Liang, Y., & Hu, Z. (2015). Accuracy rate of lie-detection in China: Estimate the validity of CQT on field cases. *Physiology & Behavior*, 140, 104-110.

National Research Council (2003). *The Polygraph and Lie Detection*. Washington, DC: The National Academies Press.

Patrick, C. J., & Iacono, W. G. (1991). Validity of the control question polygraph test: The problem of sampling bias. *Journal of Applied Psychology*, 76, 229-238.

Pigliucci, M. (2006). What is a thought experiment, anyhow? *Philosophy Now: A magazine of Ideas*, 58, Retrieved from

https://philosophynow.org/issues/58/What_is_a_Thought_Experiment_Anyhow

Raskin, D. C. (1978). Scientific assessment of the accuracy of deception of detection: A reply to Lykken. *Psychophysiology*, 15, 143-147.

Raskin, D. C., & Hare, R. D. (1978). Psychopathy and detection of deception in a prison population. *Psychophysiology*, 15, 126-136.

- Raskin, D. C., & Podlesny, J. A. (1979). Truth and deception: A reply to Lykken. *Psychological Bulletin*, 86, 54–59. <http://dx.doi.org/10.1037/0033-2909.86.1.54>
- Raskin, D. C., Honts, C. R., & Kircher, J. C. (1997). The scientific status of research on polygraph techniques: The case for polygraph tests. Chapter in, D. L. Faigman, D. Kaye, M. J. Saks, & J. Sanders (Eds.) *Modern scientific evidence: The law and science of expert testimony* (pp. 565–582).
- Raskin, D. C., Honts, C. R., & Kircher, J. C. (2014). *Credibility assessment: Scientific research and applications*. Oxford, UK: Academic Press. ISBN: 978-0-12-394433-7 (ebook version available online 17 December 2013).
- Raskin, D. C., Kircher, J. C., Honts, C. R., & Horowitz, S. W. (2019). A study of the validity of polygraph examinations in criminal investigations. Final report to the National Institute of Justice, Grant Number 85-IJ-CX-0400. *Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice*, 48, 10-39.
- Rice, M. E. and Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, 29, 615-620.
- Robinson, J. B. (1982) Energy backcasting: A proposed method of policy analysis. *Energy Policy*, December.
- Tape, T. G. (2019). Interpreting Diagnostic Tests. University of Nebraska Medical Center. Retrieved from, <http://gim.unmc.edu/dxtests/Default.htm> also referencing subordinate web pages on the same topic at this URL.
- U. S. Department of Defense (2002). *Department of Defense Polygraph Program Annual Report to Congress, Fiscal Year 2002*. Office of the Assistant Secretary of Defense (Command, Control, Communications, and Intelligence).
- Vrij, A. (2008). *Detecting Lies and Deceit: Pitfalls and Opportunities*, Second Edition. Chichester, UK: Wiley.

Zhang, X. (2011). The evolution of polygraph testing in the People's Republic of China. *Polygraph*, 40, 181-193.