

# Lecciones de Cinco Minutos de Ciencia: Teoría de la Medición y Detección del Engaño

Raymond Nelson



Traductor

Rodolfo Prado Pelayo

rodolfo@poligrafia.com.mx

**APA Magazine** 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

# Lecciones de Cinco Minutos de Ciencia: Teoría de la Medición y Detección del Engaño

## Raymond Nelson<sup>1</sup>



Raymond Nelson es un psicoterapeuta y examinador poligráfico que ha realizado varios miles de exámenes poligráficos. Tiene experiencia trabajando con autores y víctimas de delitos sexuales y de otros tipos de abusos y violencia. El Sr. Nelson es experto en estadística y análisis de datos y es uno de los desarrolladores del algoritmo de puntuación OSS-3 y del Sistema de Puntuación Empírico. Es investigador de Lafayette Instrument Company (LIC), desarrollador y fabricante de tecnologías poligráficas y ciencias de la vida, y es también expresidente de la American Polygraph Association (APA), que actualmente funge como Director electo.

El Sr. Nelson frecuentemente instruye y dicta conferencias en los Estados Unidos e internacionalmente, y ha publicado numerosos estudios y documentos sobre todos los aspectos de la prueba poligráfica, incluyendo bases de psicología y fisiología, análisis de datos de prueba, fingimiento/contramedidas, entrevista y formación de preguntas y selección de objetivos de prueba. El Sr. Nelson ha participado en el desarrollo de políticas a nivel local, estatal, nacional e internacional tanto en la poligrafía como en la psicología, y ha testificado como testigo experto en casos de cortes municipales, distritales, de apelaciones, superiores y suprema corte. El Sr. Nelson es también el director académico de la International Polygraph Training Center (IPTC). No hay intereses de propiedad o comerciales y no hay conflictos de interés asociados con el contenido de esta publicación. Las opiniones y opiniones expresadas en esta publicación son las del autor y no necesariamente las de la APA, LIC o IPTC. Sr. Nelson puede ser contactado en [raymond.nelson@gmail.com](mailto:raymond.nelson@gmail.com)

**APA Magazine** 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

A veces se dice que no es realmente posible medir una mentira. Personas con pensamiento simplista y concreto, y aquellos que se oponen a la prueba poligráfica se contentan con terminar la discusión en este punto y ofrecen la conclusión impulsiva y errónea de que la prueba científica de la detección de mentira y evaluación de credibilidad no es posible. Esta conclusión es errónea, un non-sequitur, ya que muchas áreas de ciencia están involucradas en la cuantificación de fenómenos para los cuales la medición física directa no es posible. La teoría de la prueba poligráfica, y de la detección del engaño y la evaluación de la credibilidad en general, de hecho, no implican la medición del engaño o la veracidad. Ni tampoco involucran la medición o registro del miedo o de cualquier otra emoción específica. Esta publicación intenta introducir y orientar al lector en la teoría de la medición y en su aplicación en el problema del polígrafo y la detección científica del engaño o de la prueba de la evaluación de la credibilidad.

La teoría analítica del polígrafo es que los cambios más fuertes de la actividad fisiológica se cargan hacia diferentes tipos de estímulos de prueba en función del engaño o veracidad, en respuesta al estímulo relevante objetivo (Nelson, 2015a, 2016; Senter, Weatherman, Krapohl & Horvath, 2010). En ausencia de una teoría analítica o hipótesis de la prueba poligráfica, las teorías poligráficas

previamente fueron explicadas con términos que intentaban describir el proceso psicológico o mecanismo responsable de las reacciones ante los estímulos de la prueba poligráfica. Aunque se ha aprendido mucho acerca de la fisiología registrable que está asociada con el engaño y la prueba poligráfica, se ha trabajado poco en la investigación acerca de la hipótesis psicológica del engaño. En general, la base psicológica del polígrafo asume actualmente, que se involucran una combinación de factores emocionales, cognitivos y de condicionamiento conductual (Handler, Shaw & Gougler, 2010; Handler, Deitchman, Kuczek, Hoffman, & Nelson, 2013; Kahn, Nelson & Handler, 2009). La teoría analítica de la prueba poligráfica presupone que hay cambios fisiológicos asociados con el engaño y veracidad, y que estos cambios pueden ser registrados, analizados y cuantificados mediante la comparación de las respuestas ante los diferentes tipos de estímulos de prueba. La comparación y cuantificación son objetivos centrales de la teoría de la medición. La aplicación de la teoría de la medición en la prueba poligráfica requiere de un entendimiento básico de la teoría de la medición.

### **Tipos de medición**

Stevens (1946) intentó proporcionar un encuadre para el entendimiento de los tipos de medición. En ese tiempo, parte del interés era clarificar la selección de métodos estadísticos y

analíticos asociados con los diferentes tipos de medición de datos. Casi de inmediato fue evidente que la selección de la estadística era un esfuerzo mucho más complejo que podría ser caracterizada por la reducción de la colección de tipos de datos y de las preguntas científicas a un pequeño grupo de categorías.

Las escalas *nominales* no tienen significado alguno de orden de rango (ejemplo, gato, ratón, perro, avestruz, zombie, robot). La transformación matemática de los ítems nominales no es posible. Las mediciones *ordinales* tienen significado en el orden de rango, pero tienen un significado impreciso acerca de la distancia entre los ítems (ejemplo, saber el primer, segundo y tercer lugar de ganadores en una carrera de avestruces, no proporciona información acerca de la diferencia en los tiempos de carrera). Con las mediciones ordinales son posibles ciertas transformaciones matemáticas, bajo el requisito de que preserven la información y significado ordinal. Las mediciones de escalas *intervalares* tienen un significado de orden de rango y también proveen información con significado acerca de la diferencia entre los ítems. De cualquier manera, el punto cero de una escala ordinal es arbitrario y por lo tanto no tiene significado. Un ejemplo típico de enseñanza de la arbitrariedad del punto cero en una escala intervalar es la escala de la temperatura para la cual tenemos tanto escalas Fahrenheit y Celsius con diferentes puntos cero

arbitrarios, y no se espera que el cero signifique que no hay temperatura o calor que sea medido. Las mediciones de *proporción* incluyen la combinación del significado de orden de rangos y significado de intervalos de distancia, junto con la noción de un punto cero no arbitrario. En mediciones de escalas proporcionales, el cero significa nada (ejemplo, no hay diferencia). Posteriormente, Stevens (1951) ofreció un grupo de prescripciones y proscipciones acerca del tipo de estadística que es apropiada para cada tipo de datos.

La forma más común de crítica de Stevens se ha enfocado en el hecho de que es innecesariamente restrictivo (Valleman & Wilkinson, 1993), resultando en el sobre uso de los métodos no paramétricos, que se sabe son menos eficientes que los métodos paramétricos (Baker, Hardyck, & Petrinovich, 1966; Borgatta & Bohrnstedt, 1980), y que el tipo de análisis debería determinarse por la pregunta de investigación en cuestión (Guttman, 1977; Lord, 1953; Turkey, 1961). Luce (1997) aseveró directamente que hoy en día los teóricos de la medición no aceptan la definición demasiado abierta de Stevens acerca de la medición. Sin embargo, el trabajo de Stevens provee una introducción útil al lenguaje conceptual y a los problemas de la teoría de la medición.

## Teoría de la Medición

La teoría de la medición es un área de la ciencia preocupada por la investigación de lo medible y de lo que hace posible la medición. Helmholtz (1887) comenzó la tradición del cuestionamiento científico y filosófico dentro de la teoría de la medición al hacer la pregunta “¿Por qué los números pueden ser asignados a cosas?”, junto con otra pregunta como “¿Qué se puede entender de esos números?” De acuerdo con Campbell (1920/1957), *la medición es el proceso de utilizar números para representar cualidades*. En general, las propiedades de un fenómeno medible deben, en cierta forma, parecerse a las propiedades de los números. Un trabajo posterior de Suppes (1951) acerca de las diferencias entre los fenómenos medibles y los no medibles, comenzó a formalizar la tradición de la teoría de la medición al clarificar nuestro entendimiento de los requisitos de la medición y generó un crecimiento en la teoría de la representación moderna de la medición. (Diez, 1997; Suppes, 2002; Suppes & Zinnes, 1963; Suppes, Krantz, Luce, & Tversky, 1989; Niederee, 1992). Establecieron de manera simple, que la teoría representativa de la medición involucra la asignación de números a fenómenos físicos de manera tal que se preservaran las relaciones empíricas u observables.

La existencia de relaciones de orden (orden de rangos) entre los objetos observados es central para los requerimientos de mensurabilidad de un fenómeno. Debemos de ser capaces de cuantificar una instancia del fenómeno, cuando tiene mayor magnitud que otro. Otro requerimiento central del fenómeno medible es que debe de existir una forma de combinar los objetos medibles de manera análoga a una adición matemática. Estas son unas de las principales diferencias entre los fenómenos medibles y los no medibles.

Por ejemplo: las mediciones pueden ser aplicadas a fenómenos físicos como la altura, peso y presión sanguínea de una persona. Esto es posible debido a que estas cosas implican un fenómeno físico: la distancia lineal o unificada entre la cabeza y los dedos de los pies. La fuerza gravitacional en la masa física de una persona y la presión única requerida para sobrellevar y bloquear la presión arterial relativa con respecto a un punto de referencia como podría ser la presión atmosférica promedio al nivel del mar (ejemplo, 29.92 inHG o 760 mmHG). Estos fenómenos pueden combinarse de manera tal que son de alguna forma análogos a una adición numérica. Esto es, que existe una forma de interpretación física coherente para adicionar las combinaciones de diferentes instancias de estos fenómenos físicos. Los eventos de tiempo limitado también pueden ser medidos. Por ejemplo: si una persona

brinca en el aire dos veces y marcamos el peso físico de cada salto, y después combinamos ambas distancias, entonces esto será también una adición numérica análoga. Sin embargo, los intentos de registrar cambios fisiológicos ante estímulos poligráficos, no necesariamente cumplen con esos requerimientos de relaciones de orden de rangos y adición. Los detalles de cómo los datos registrados poligráficamente pueden resultar en la cuantificación del engaño y la veracidad son presentados en la parte restante de esta publicación.

En primer lugar, se ha establecido desde hace tiempo, que las respuestas ante los estímulos poligráficos no pueden ser tomados o interpretados directamente como una medición del engaño. Ni tampoco las respuestas ante los estímulos poligráficos deben de ser interpretados como registro o medición del miedo o de cualquier otra emoción específica. Las respuestas ante los estímulos poligráficos son una forma de proxi o data substituta para la que existe una relación o correlación

con el engaño y la veracidad. Las reacciones y datos registrados, por sí mismos no son ni engaño ni veracidad per se. En segundo lugar, aun cuando podría ser posible interpretar las relaciones de orden de rango entre los estímulos de prueba de acuerdo con la magnitud de respuesta, la instrumentación de registro poligráfico hoy en día no fue diseñada para proporcionar datos que cumplan el requerimiento de aditividad para la medición de datos. En otras palabras, los intentos de realizar cualquier combinación aditiva sensible de los datos de respuesta reales dentro de cada uno de los sensores respiratorios, cardiacos, electro dérmicos y vasomotores tampoco se ha intentado o establecido. En lugar de esto, los datos poligráficos deben de ser transformados a una forma más abstracta para darles significado, antes de que puedan ser analizados e interpretados. La calificación poligráfica y algoritmos de análisis, ya sean manuales o automatizados, intentan lograr y facilitar dicha transformación, análisis e interpretación. 1

1. Una diferencia importante entre los algoritmos manuales y automáticos del análisis poligráfico es que los protocolos de puntaje manual se desarrollaron durante un tiempo en el que los practicantes de campo no tenían acceso y no estaban familiarizados con el uso de potentes microcomputadoras. Por tanto, los algoritmos de calificación manual se basan en transformaciones matemáticas que son, por necesidad, muy simples, si no es que hasta obtusas. Las versiones previas de los protocolos de puntuación manual no hicieron uso de distribuciones de referencia normativas, correcciones estadísticas o intervalos de confianza. Otra diferencia importante es que los protocolos de puntuación manual realizan tareas de extracción de características - la extracción de la información de la señal dentro de otra información registrada y ruido - usando métodos de visión subjetiva. Los algoritmos de análisis automatizado harán uso de métodos estadísticos más avanzados, y se basará en Extracción de características más objetivos y automatizados que son menos vulnerables a la interferencia subjetiva.

**APA Magazine** 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

## **Mediciones fundamentales y derivadas**

Algunas mediciones pueden ser referidas como *fundamentales* y no requieren de un fenómeno previamente medido para lograr su determinación. El requisito fundamental para una medición fundamental es que existan algún fenómeno físico para el que exista alguna cantidad que pueda ser entendida ya sea como mayor o menor que (ejemplo, *es pesado*) en oposición a un fenómeno que se entiende mejor como un algo contra nada (ejemplo, *es un avestruz*). Si tenemos dos avestruces, tiene cierto sentido hacerse una pregunta tal como, cuál avestruz es la más pesada, ya que hay una intuición con significado en torno a la idea de que algunas avestruces son más pesadas. Pero no tiene sentido hacerse la pregunta de cuál es más avestruz, ya que no existe una intuición con sentido de que se va a obtener algo en la respuesta. Ser una avestruz es una propiedad, no una cantidad. El peso de una avestruz también es una propiedad, y esto ilustra que algunas propiedades también pueden ser cantidades. El fenómeno físico del peso o la pesadez pueden ser cuantificados para lograr una mayor precisión en lugar de simplemente decir, *muy pesado* o *muy muy pesado* cuando se intenta comparar el peso de dos avestruces. Sin el uso de cantidades numéricas, dos observadores distintos podrían llegar a dos conclusiones diferentes

acerca de cuál avestruz es la más pesada sin importar el cómo intentemos utilizar sus adjetivos descriptivos. Es más probable que observadores diferentes logren conclusiones similares cuando utilizamos mediciones vs. la alternativa de no utilizarlas. El uso de mediciones nos permite pensar, entender, describir y planear el mundo que nos rodea con mayor precisión, lo que puede presentarse como *reproductibilidad*.

Cuando no se intenta realizar una medición o no se espera una cantidad precisa o exacta, eso se conoce como una *estimación*. Debido a que no se espera que las probabilidades sean exactas, son estimaciones. Aunque algunos podrían usar o expresar subjetivamente la noción de probabilidad, la reproductibilidad de estimados probabilísticos computacionales es una diferencia importante entre el uso científico y no científico del concepto de probabilidad.

Algunas mediciones pueden ser pensadas como *derivadas*, ya que estas no se logran mediante una cuantificación directa del fenómeno físico, sino mediante la comparación de un fenómeno físico no cuantificable mediante otro fenómeno físico conocido. En principio, podemos medir una distancia *desconocida* si es que tenemos otras distancias y ángulos ya conocidos. Por ejemplo, si colocamos una serie de satélites en órbita alrededor de la tierra, podemos

calcular y conocer la localización de dichos satélites en relación a un grupo de objetos para los cuales es conocida su localización en la tierra. Entonces, si tenemos medios para recibir información de esos satélites con localización conocida, podemos utilizar la información de los satélites para calcular y medir nuestra propia ubicación si es que esta fuera desconocida. Esto es parecido a las viejas prácticas en las que podemos calcular la ubicación de objetos en el sistema solar de acuerdo a un sistema de conteo y cuantificación del número de días desde el evento observado previamente, entonces podemos utilizar la localización del objeto en el sistema solar. Y la localización de objetos en el sistema solar se podría utilizar, junto con un sistema definido de reglas científicas y matemáticas, para medir o cuantificar nuestra localización real en la tierra. Otro ejemplo de una medición derivada es la medición de la presión sanguínea, para lo cual utilizamos nuestro conocimiento acerca de la presión atmosférica, para cuantificar nuestra evaluación de las presiones cardiacas durante las fases sistólicas y diastólicas del ciclo cardiaco.

### **La prueba científica como una forma de medición probabilística.**

Como ocurre, muchos fenómenos interesantes e importantes no pueden ni ser observados directamente ni pueden ser sujetos a mediciones físicas. Esto en ocasiones es así porque

el fenómeno de interés es amorfo (sin sustancia física), y a veces porque la información no se ajusta a los requerimientos de orden y de aditividad de la medición. Si queremos mejorar la precisión de nuestras evaluaciones y de nuestras decisiones para estos fenómenos, no necesitamos basarnos en nuestras mediciones, pero sí en las pruebas científicas que cuantifican el fenómeno de interés utilizando estadísticas y teoría de la probabilidad. Nelson (2015b) proporcionó una descripción de cómo la prueba poligráfica, y las pruebas en general, pueden ser pensadas como un experimento de ciencia en un solo sujeto. Las pruebas científicas pueden también ser pensadas como una forma de medición probabilística, en la que se usan las teorías estadísticas y de probabilidad para cuantificar un fenómeno que no se sujeta a una medición real.

Un ejemplo de una prueba científica a manera de medición probabilística es la de prueba de medición un fenómeno psicológico amorfo y no medible como lo es la *personalidad* y el *funcionamiento intelectual*, durante el cual una cantidad observada de datos de un individuo es comparada matemáticamente contra una cantidad conocida en forma de distribución de referencia normativa, o modelo de referencia probabilística que caracteriza nuestro conocimiento de lo que esperamos observar. Los modelos de referencia pueden ser calculados empíricamente, mediante métodos de

muestreo estadísticos, y que también pueden tomar forma de distribuciones de referencia teóricas que caracterizan nuestras teorías de trabajo acerca del universo, o de una pequeña parte, y que funcionan al basarse solamente en hechos e información que se sujetan a pruebas matemáticas y lógicas.

En el caso de la prueba poligráfica – para la que la teoría analítica básica sostiene que los mayores cambios en la actividad fisiológica se cargarán en diferentes tipos de estímulos de prueba en función del engaño y veracidad, en respuesta al estímulo relevante – no es la comparación de las relevantes y otras preguntas de prueba lo que forma la base de nuestras conclusiones. En su lugar, es la comparación de las diferencias en las reacciones en las preguntas relevantes y en otras preguntas de prueba frente a distribuciones de referencia que anclan nuestro conocimiento acerca de las diferencias esperadas en la respuesta ante las relevantes y a otras preguntas de prueba entre las personas que engañan y las veraces. Idealmente, otras preguntas podrían tener el potencial de evocar actividad cognitiva y emocional de calidad similar, aunque quizá de diferente magnitud que ante el estímulo relevante objetivo. De cualquier manera, no es necesario que otras preguntas tengan un valor ecológico similar, comparado con el estímulo relevante para que sean una base útil y efectiva para la comparación

estadística. Un ejemplo de esto se puede ver en el uso de las comparativas de mentira dirigida (DLC), para las que Blalock, Nelson, Handler & Shaw (2011) proporcionaron un resumen de la investigación acerca de su efectividad (y del porque el nombre DLC no debería de ser utilizado para presuponer que la respuesta ante estas preguntas son realmente mentiras).

### **La prueba científica como una forma de predicción**

Si queremos cuantificar o mejorar la precisión asociada con nuestras evaluaciones y conclusiones acerca de eventos futuros que aún no han ocurrido – asumiendo que queremos cuantificar nuestras conclusiones ahora sin tener que esperar a que ocurra el evento – entonces de nuevo estamos intentando cuantificar un fenómeno que no es posible observar o medir directamente. Para esto, necesitamos una *prueba* con la que podamos realizar una conclusión probabilística acerca de un resultado futuro. Las pruebas que se utilizan de esta manera pueden ser pensadas como una forma de *predicción* científica. No es una forma de magia o adivinación. Es una forma de modelado probabilístico.

Un ejemplo de la cuantificación de un evento futuro es la medición o cuantificación del nivel de riesgo para un evento de peligro – del que está

APA Magazine 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

implícito que el evento futuro aún no ha ocurrido y por lo tanto no puede ser físicamente cuantificado u observado. Así, otro ejemplo que involucra la predicción de un evento futuro, sería la cuantificación del resultado de una elección que aún no ha ocurrido. Ambos ejemplos – resultados de riesgo y resultados de elecciones – pueden involucrar un evento futuro para el que el valor asociado es binario (ejemplo un evento que ocurrió o no, una elección que se ganó o no). En un punto único en el tiempo, el evento pudo ocurrir o no. Podríamos, en ocasiones, simplemente querer esperar para observar el resultado para obtener una conclusión determinista. La observación determinista de un resultado podría, por supuesto, obviar la necesidad de hacer una prueba y cuantificar.

Una diferencia notable entre la predicción de eventos de riesgo y resultados agendados, es que con los resultados de una elección se puede esperar que van a ocurrir en un punto de tiempo ya agendado, tiempo en el cual es posible observar el resultado. Después del evento agendado, el resultado es un hecho, no una probabilidad. Antes del evento agendado, el resultado podría ser pensado como una probabilidad, de manera que hay algunos factores que están asociados con los diferentes resultados posibles. Una meta de las predicciones científicas involucra la identificación de estos factores asociados de manera que puedan ser

caracterizados como variables aleatorias y ser utilizados para desarrollar una prueba o modelo predictivo.

Las probabilidades asociadas con los resultados de eventos agendados que aún no han ocurrido pueden pensarse como la proporción de resultados que podrían ocurrir de cierta manera, dadas las variables aleatorias que influyen en el resultado, si es que fuese posible observar el evento a través de numerosas repeticiones. La efectividad o precisión de una prueba como un modelo predictivo dependerá de nuestra habilidad para entender correctamente las variables aleatorias que están relacionadas con los posibles resultados. Ultimadamente, el resultado va a ser una certeza y no una probabilidad. Previo a la ocurrencia del resultado, se mantiene como una probabilidad o predicción. Cuando se presentan errores en la predicción, sus causas pueden deberse tanto a una variación aleatoria, o por malentender y caracterizar equivocadamente las variables aleatorias relacionadas con los posibles resultados.

Se espera que ciertos tipos de resultados ocurrirán en un momento desconocido, o que podrían no ocurrir de ninguna forma por un periodo largo de tiempo. Podemos pensar en estos resultados como probabilidades. Por ejemplo: ¿cuál es la probabilidad de que un trasgresor criminal conocido, podría reincidir, o cual es la

probabilidad de un terremoto en la ciudad de México, o cual es la probabilidad de una inundación? Estos eventos también podrán ser considerados como certezas después de que ocurran, y también están sujetos a cierta relación con los factores relacionados que están asociados con su ocurrencia. Como en otros modelos predictivos, la identificación y caracterización de los factores asociados es un objetivo importante para el desarrollo de la evaluación del riesgo o de los modelos de predicción de riesgos. Las probabilidades asociadas con los resultados de la predicción de riesgos, pueden ser pensados en términos de frecuencias, como qué eventos de probabilidad alta ocurrirán con mucha frecuencia, mientras que los eventos de probabilidad baja ocurrirán con una menor frecuencia.

Casi nunca - incluyendo eventos para los que nuestra intuición nos dice que las posibilidades son muy bajas - se puede pensarse como una probabilidad. Esto puede en ocasiones, ser considerado un absurdo. Por ejemplo, ¿cuál es la probabilidad de que una horda de zombis nos ataque, o cual es la probabilidad de un apocalipsis robótico? Para estos ejemplos extremos nuestra intuición nos dice que la probabilidad es absolutamente cero o esencialmente cero, pero aun así podemos involucrar cierta imaginación ante los factores que podrían estar asociados con su ocurrencia. Si extendemos el periodo

bajo consideración, entonces las probabilidades asociadas con eventos muy raros podrían volverse considerablemente mayores. Por ejemplo: ¿Cuál es la probabilidad de que una avestruz caiga del cielo? Si expandimos nuestras dimensiones en tiempo y lugar hacia las nociones de cualquier momento y en cualquier lugar, podemos entender intuitivamente cierta probabilidad cero asociada con la avestruz cayendo del cielo, con el tipo de factores que podrían estar asociados con su posible ocurrencia (ejemplo, una emergencia en el aerotransporte de avestruces de una granja de avestruces que se inundó).

La cuantificación de eventos futuros como son algún riesgo o resultados de una elección, requieren que tratemos el resultado futuro de la misma manera que tratamos cualquier otro fenómeno amorfo que queremos cuantificar. Tratamos el resultado futuro como una probabilidad. La cuantificación de un resultado es útil, solamente cuando es un resultado futuro - un resultado que aún no ha ocurrido. Si existe información y está disponible para su observación o medición, entonces el resultado no es amorfo y es un fenómeno físico. La medición u observación directa de un resultado futuro requerirá que esperemos hasta ese punto futuro en el tiempo. Hasta entonces, si queremos tratar de predecir el resultado futuro que aún no ha ocurrido, necesitaremos confiar en las probabilidades para

describir ese evento futuro amorfo. De manera similar, la observación o medición de un evento pasado requerirá que estén disponibles ciertos fenómenos físicos del evento para su observación o medición. Si queremos cuantificar un evento pasado para el que no está disponible un fenómeno físico, entonces nuevamente vamos a necesitar confiar en la teoría de la probabilidad para cuantificar ese fenómeno amorfo.

Una cita famosa de un desconocido autor Danés durante los años 1937-1938 establece (en español) “*Es difícil hacer predicciones, especialmente acerca del futuro*”. Esta cita simple y humorosa nos recuerda que las predicciones de todo tipo son inherentemente imperfectas, incluyendo a las predicciones basadas en datos de prueba científicos. Las conclusiones probabilísticas son inherentemente imperfectas. De hecho, no se espera que sean perfectas. Se espera que las conclusiones probabilísticas solamente cuantifiquen el margen de incertidumbre asociado con una conclusión. La predicción estadística es una tarea inherentemente probabilística y estadística para la que una conclusión puede ser probablemente correcta y probablemente incorrecta. Las conclusiones acerca del engaño o la veracidad, a pesar del deseo de certidumbre e infalibilidad, será inherentemente probabilística e inherentemente imperfecta.

### **Conclusión: la prueba poligráfica científica como una forma de clasificación estadística.**

Los resultados de la prueba poligráfica deben de ser pensados como una forma de predicción de que existe otra forma de evidencia y que puede ser identificada como base de evidencia para confirmar o refutar un resultado de prueba. Una forma más simple y general de pensar en estas pruebas será en una forma de *clasificación* estadística. Como en otras pruebas científicas, las pruebas estadísticas que intentan clasificar, no se espera que sean perfectas, infalibles o deterministas. Tampoco se espera que las clasificaciones estadísticas proporcionen el mismo nivel de precisión que la medición real de un fenómeno físico. Como en otros esfuerzos probabilísticos, las pruebas científicas que pretenden una clasificación, se espera que cuantifiquen el margen de incertidumbre o nivel de confianza que puede ser atribuido a una conclusión. Más aún, el método de cuantificación estadística debe de ser explicable y los resultados deberían de ser reproducibles por otros. La medición extrema de la efectividad de una prueba estadística no es lograr la perfección o infalibilidad, sino es la observación correcta e incorrecta de las clasificaciones del mundo real que se ajusten a nuestras estimaciones de probabilidad calculadas.

APA Magazine 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

Si la teoría analítica básica de la prueba de polígrafo es incorrecta – si ningún cambio fisiológico está correlacionado con las diferencias entre el engaño y la veracidad – si toda la actividad fisiológica es meramente un caos aleatorio con respecto al engaño y la veracidad, entonces los humanos virtualmente no tendremos nunca la oportunidad de saber si alguien ha mentado, con una precisión superior a la posibilidad del azar. La única manera de protegerse a uno mismo del engaño sería mantenerse cínico, sospechar de todos, y no confiar en nadie. Aunque quizá suene tentador, esto no es realista y será insostenible después de un tiempo. Por otra parte, si es correcto que algunos cambios en la actividad fisiológica están asociados con el engaño y la veracidad en tasas significativamente superiores al azar, entonces es solamente un asunto de tiempo antes de que los tecnólogos, ingenieros, matemáticos, estadistas y analistas de datos diseñen algunas formas de incrementar la disponibilidad de información de señal útil, dentro del ruido caótico de otras actividades fisiológicas y exploten esas señales con una nueva forma de evaluación científica de la credibilidad o prueba de detección de mentiras.

Ultimadamente, si la prueba poligráfica es un interrogatorio y no una prueba científica, entonces la teoría de la medición no es de preocupación ni de consecuencias para la profesión poligráfica. Pero en

este caso, la gente comenzará a cambiar a otras metodologías científicas cuando deseen una prueba científica para la evaluación de la credibilidad, y la prueba poligráfica eventualmente podría ser reemplazada. Por otra parte, si la prueba poligráfica es una prueba científica, será de utilidad para todos los profesionales en poligrafía, familiarizarse con las bases de la teoría de la medición y con la discusión de resultados científicos de la prueba poligráfica, incluyendo las conclusiones categóricas acerca del engaño y veracidad y conclusiones acerca de contramedidas, utilizando el lenguaje conceptual de la teoría de la medición y de la probabilidad. Las conclusiones poligráficas no son mediciones físicas; son estimaciones probabilísticas.

En la ausencia de un pensamiento probabilístico aplicado a la prueba poligráfica, algunos tendrán el impulso de utilizar expectativas ingenuas y no realistas de una perfección determinista. Existirá también un deseo o impulso en algunos de fingir infalibilidad, debido a una magia o habilidades profesionales superiores, y esto podría por un tiempo parecer una estrategia efectiva de marketing. Pero fingir infalibilidad llevará a la confusión y frustración cuando inevitablemente sea observado que los errores de prueba pueden y seguramente ocurrirán. Una solución correctiva temporal para esta frustración será encontrar errores en

**APA Magazine** 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

el profesionalista, y no en la prueba – y por lo tanto será restaurada la suposición falsa de infalibilidad, mientras podamos evitar esos magos menos competentes, expertos menos competentes. Aunque será gratificante por un tiempo, este tipo de propuesta no es científica, y no podrá sostenerse en el contexto de la experiencia del mundo real y de la evidencia científica.

Los resultados de la prueba poligráfica deberían ser entendidos y descritos al igual que los resultados de otras pruebas científicas, utilizando el lenguaje conceptual de las probabilidades estadísticas. Expresiones de conclusiones supuestamente científicas, incluyendo las conclusiones acerca del engaño y la veracidad y conclusiones acerca del uso de contramedidas, sin el uso de métrica probabilística provocará acusaciones de que el polígrafo es meramente pseudociencia subjetiva cubierta por un exceso de confianza. Un enfoque científico de la prueba poligráfica reconocerá que la tarea de cualquier prueba es cuantificar un fenómeno probabilístico cuando la observación directa o medición física no son posibles, y de reconocer y realizar un uso explicable del potencial del error de la prueba cuando decidimos que valor colocarle y como usar o confiar en el resultado de prueba.

Como en otras pruebas científicas, la prueba poligráfica intenta realizar clasificaciones probabilísticas de

engaño y veracidad en ausencia de la capacidad de una observación directa o medición física del asunto en cuestión. Si el fenómeno físico estuviera disponible para su observación o medición, entonces no sería necesaria una prueba científica. Debido a que el engaño y la veracidad son constructos amorfos, la detección científica de la mentira y la evaluación de la credibilidad son finalmente, preocupaciones epistemológicas que a veces están sujetas a cuestiones filosóficas complejas e importantes tales como: *¿qué significa decir que algo es verdad, y que tipo de cosas pueden decirse para ser verdad?* Aunque es profundamente interesante, esto deberá ser tema de otra publicación.

## Referencias

Baker, B.O., Hardyck, C.D., & Petrinovich, L.F. (1966). Weak measurements vs. strong statistics: An empirical critique of S.S. Stevens' proscriptions on statistics. *Educational and Psychological Measurement*, 26, 291-309.

Blalock, B., Nelson, R., Handler, M. & Shaw, P. (2011). A position paper on the use of directed lie comparison questions in diagnostic and screening polygraphs. *Police Polygraph Digest*, 2011, 2-5.

Borgatta, E.F., & Bohrnstedt, G.W. (1980). Level of measurement - once

APA Magazine 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

over again. *Sociological Methods and Research*, 9, 147- 160.

Campbell, N. (1920/1957). *Foundations of science: the philosophy of theory and experiment*. New York: Dover.

Diez, J.A. (1997). A hundred years of numbers. An historical introduction to measurement theory 1887–1990. *Studies in History and Philosophy of Science Part A* 28 (2), 237-265.

Guttman, L. (1977). What is not what in statistics. *The Statistician*, 26, 81-107.

Handler, M., Shaw, P. & Gougler, M. (2010). Some thoughts about feelings: a study of the role of emotion and cognition in polygraph testing. *Polygraph*, 39(3), 139-154.

Handler, M., Deitchman, G., Kuczek, T., Hoffman, S., & Nelson, R. (2013). Bridging emotion and cognition: a role for the prefrontal cortex in polygraph testing. *Polygraph* 42(1), 1-17.

Helmholtz, H.V. (1887). Zahlen und messen, erkenntnistheoretisch betrachtet, in H.v. Helmholtz, *Schriften zur Erkenntnistheorie*, 70-108. Engl. Trans., Numbering and measuring from an epistemological viewpoint”, in H.V. Helmholtz, (1977). *Epistemological Writings*, 72-114, Dordrecht: Reidel.

Kahn, J., Nelson, R. & Handler, M. (2009). An exploration of emotion and cognition during polygraph testing. *Polygraph*, 38, 184-197.

Lord, F. M. (1953). On the statistical treatment of football numbers. *American Psychologist*, 8(12), 750-751.

Luce, R. D. (1997). Quantification and symmetry: commentary on Michell'Quantitative science and the definition of measurement in psychology'. *British Journal of Psychology*, 88, 395–398.

Nelson, R. (2015a). Scientific basis for polygraph testing. *Polygraph*, 44(1), 28-61.

Nelson, R. (2015b). Polygraph as a single-subject scientific experiment. *APA Magazine*, 48(5) 101-106.

Nelson, R. (2016). Scientific (analytic) theory of polygraph testing. *APA Magazine*, 49(5), 69-82.

Niederee, R. (1992). What do numbers measure? A new approach to fundamental measurement. *Mathematical Social Sciences* 24. 237-276.

Senter, S., Weatherman, D., Krapohl, D., & Horvath, F. (2010). Psychological set or differential salience: a proposal for reconciling theory and terminology in polygraph testing. *Polygraph* 39(2), 109-117.

**APA Magazine** 2016, 49(6)

*This article is copyrighted by the American Polygraph Association (APA), and appears here with the permission of the APA. La American Polygraph Association (APA) tiene los derechos de autor de éste artículo, y aparece aquí con el permiso de la APA.*

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684). 677–680.

Stevens, S.S. (1951). Mathematics, measurement, and psychophysics. In S.S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley.

Suppes, P. (1951). A set of independent axioms for extensive quantities. *Portugaliae Mathematica* 10, 163-172.

Patrick Suppes (2002). Representational measurement theory. In J. Wixted & H. Pashler (eds.), *Stevens' Handbook of Experimental Psychology*. Wiley.

Suppes, P. and J. Zinnes (1963). Basic measurement theory. In *Handbook of Mathematical Psychology* vol. 1 (ed. D. Luce, R.R. Bush and E. Galanter). 3-76.

Suppes, P., Krantz, D., Luce, D. and A. Tversky (1989). *Foundations of Measurement* vol. 2. New York: Academic Press.

Tukey, J.W. (1961). Data analysis and behavioral science or learning to bear the quantitative man's burden by shunning badmandments, in *The Collected Works of John W. Tukey*, vol. III (1986), Lyle V. Jones (ed), Belmont, CA, Wadsworth, Inc. 391–484.

Velleman, P. F., & Wilkinson, L. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician* (1993), 47(1), 65-72.