

Enfoque Utah para Pruebas Poligráficas de Preguntas Comparativas

Mark Handler y Raymond Nelson

Rodolfo Prado Pelayo
“traductor”

Enfoque Utah para Pruebas Poligráficas de Preguntas Comparativas

Mark Handler y Raymond Nelson

ABSTRACT

En 1970, David Raskin, un psicólogo e investigador de la Universidad de Utah, inició un estudio de la técnica poligráfica de preguntas de comparación de mentira probable. Raskin y sus colegas refinaron sistemáticamente los elementos de la poligrafía mediante la determinación de qué aspectos de la técnica podrían ser probados científicamente para incrementar su validez y confiabilidad (Raskin & Honts, 2002). Sus esfuerzos culminaron en la creación de lo que hoy se conoce como el enfoque Utah para la Prueba de Preguntas de Comparación (CQT), un enfoque de poligrafía empíricamente consistente y unificado. La Utah-CQT, tradicionalmente es empleada como una Prueba de Zonas de Comparación de asunto único (ZCT). Es susceptible a otros usos, incluyendo; pruebas multi-faceta de un solo asunto criminal, similar al formato de la Técnica de Preguntas Generales Modificado (MGQT), o como una Técnica de Preguntas Generales (GQT) de asuntos-múltiples (asuntos mixtos). La Utah-CQT y su correspondiente Sistema de Calificación Numérica Utah (Bell, Raskin, Honts y Kircher, 1999; Handler, 2006) fue el resultado de más de 30 años de investigación científica y de su revisión-por-pares científicos. La técnica resultante ofrece algunas de las tasas más altas de precisión de criterio de y confiabilidad inter-evaluador de cualquiera de los protocolos de examinación poligráfica (Senter, Dollins & Krapohl, 2004; Krapohl, 2006) cuando se aplica en una situación de prueba de evento-específico. Los autores describen la Utah-CQT utilizando la Prueba de Mentira Probable (PLT), así como la menos conocida Prueba de Mentira Dirigida (DLT) y examinan algunos de los beneficios posibles de cada método.

ESTRUCTURA Y ADMINISTRACIÓN DE LA PRUEBA

El Utah-CQT al igual que otros procedimientos de prueba, comienza con la entrevista de pre-test, llevada a cabo de una manera no acusatoria. El examinador debe obtener la liberación legal necesaria de la prueba que incluye una breve declaración de los alegatos o asuntos por resolver, y si aplica, una renuncia a los derechos legales y después la recolección de información biográfica general y médica del sujeto por evaluar. El fomento de la construcción de rapport le da al examinador la oportunidad de evaluar la idoneidad del sujeto evaluado para el examen. La interacción con el sujeto evaluado también da al examinador la oportunidad de hacer una valoración aproximada de las habilidades verbales y mentales del sujeto evaluado que posteriormente serán utilizadas para ayudar a la redacción de las preguntas del examen. En la versión PLT, el examinador utiliza este período de conversación para desarrollar el material de las preguntas de comparación que se utilizarán durante la fase de prueba del examen, aunque por lo general la naturaleza de los asuntos por resolver dicta el contenido general de las

preguntas de comparación. Sin embargo, el examinador no habla con el sujeto evaluado con respecto a transgresiones pasadas durante esta revisión del material de las preguntas de comparación. Esta parte de la entrevista se conduce con preguntas abiertas y con el uso cuidadoso de sugerencias, opuesto a un interrogatorio sobre acciones pasadas.

El examinador señala los dispositivos de monitoreo o grabación en la sala de examinación y explica el propósito de tener ese examen supervisado y/o grabado. En el enfoque Utah-CQT, todos los exámenes deben ser grabados en su totalidad. En una época en la que la tecnología de vídeo y audio grabación está disponible de forma sencilla y plenamente integrada en todos los sistemas poligráficos modernos en el campo, no hay razón alguna para renunciar a las ventajas de una grabación completa de video y audio de todos los exámenes de polígrafo. Es sólo a través de las grabaciones completas que es posible el tan importante control de calidad. La franqueza con respecto a los dispositivos de monitoreo ayuda a asegurarle al sujeto evaluado que la prueba se llevará a cabo de manera profesional y podría ayudar a convencer al sujeto evaluado de que el examinador está siendo abierto y veraz. Una breve explicación de cualquier programa de control de calidad también ayuda a establecer un ambiente profesional y de confianza.

El examinador avisa al sujeto evaluado acerca de la naturaleza general de los alegatos y de los temas específicos a ser resueltos mediante el examen. Se le da entonces al sujeto evaluado la oportunidad de ofrecer una "narrativa libre" para discutir su conocimiento y/o papel en el incidente. El objetivo de la discusión mediante narrativa libre es la de obtener información del sujeto evaluado sin confrontación ni estrés innecesario.

En general, el examinador debe permitir que el sujeto evaluado cuente su historia sin interrupción. El examinador informa al sujeto evaluado de los hechos del caso con un enfoque de bajo-perfil y debe avisar al sujeto evaluado de que solo son acusaciones y debe asegurarse que el sujeto entiende la diferencia entre acusaciones y hechos conocidos que son verdad.

El examinador debe notar inconsistencias u otros asuntos a los que podría querer regresar una vez que el sujeto evaluado termine la narración. El examinador no discute con el sujeto evaluado ni desafía la versión de los hechos del caso del sujeto evaluado. El examinador anima al sujeto evaluado a ser sincero con el fin de formular las preguntas de prueba de una manera sucinta y clara.

En el polígrafo exploratorio o en los programas de monitoreo (es decir, LEPET, seguridad, PCSOT), el Utah-CQT podría ser utilizado como una examinación de asunto-mixto (asuntos-múltiples), similar al AFMGQT, en ausencia de un alegato conocido o incidente conocido. En estos programas, la discusión de la acusación conocida o incidente conocido se reemplaza por un protocolo de entrevista estructurada, que se dirige hacia áreas de contenido pertinentes con los asuntos de riesgo o de cumplimiento bajo investigación. Cabe señalar que estas aplicaciones de la prueba de polígrafo no se han investigado tan a fondo como los otros usos, y la investigación científica y la verificación de tales usos están más limitadas.

Este enfoque de bajo perfil, no-acusatorio presenta al examinador como un buscador neutral de la verdad y ayuda a disipar temores de culpa preconcebida. Si hay inconsistencias u otros asuntos que requieran seguimiento o aclaración antes del examen, se discuten en este momento de una manera no confrontacional.

Después de la narrativa y de la discusión de otros asuntos, se colocan los componentes en el sujeto evaluado. Durante este proceso, se discuten las funciones de los diferentes sensores que componen el polígrafo, y se da una explicación general de la psicofisiología que subyace a la prueba del polígrafo. Esto puede hacerse a través de una discusión general de anécdotas que ilustren la respuesta psicofisiológica y las diferentes causas posibles de una reacción (Handler y Honts, 2007). El objetivo de esta parte de la entrevista es garantizar en el sujeto evaluado la comprensión de que la mentira inevitablemente se asocia con una respuesta fisiológica.

Una vez que los componentes se colocan en el sujeto evaluado, el examinador conduce una prueba de familiarización. La prueba de familiarización es conocida generalmente como una prueba de pico de tensión de solución conocida que se utiliza para demostrar la eficacia del examen de polígrafo. No se prohíben otros enfoques para la prueba de familiarización ni invalidarían una examinación. En la prueba de familiarización de solución-conocida, se le dice al sujeto evaluado que seleccione un número tal que estará entre algunas preguntas adicionales o de relleno antes y después del número seleccionado. Esto se puede lograr dirigiendo al sujeto evaluado para que seleccione un número entre el 3 y el 6 y para que escriba ese número en un trozo de papel. El papel podría entonces mostrarse en frente del sujeto evaluado y se le instruye para que niegue haber escogido cualquiera de los números entre 1 y 7 mientras que el polígrafo registre sus reacciones fisiológicas. La prueba de familiarización permite al examinador asegurarse de la producción de registros de calidad adecuada y de tomar las acciones correctivas para remediar cualquier falta de ella.

El examinador puede usar la prueba de familiarización durante la revisión de preguntas para demostrarle al sujeto evaluado que es un candidato adecuado para el polígrafo, y ofrecer garantías de que se puede conseguir una finalización exitosa del examen cuando se responde a todas las preguntas de prueba con

la verdad (en la versión PLC del examen) o de que se encontró indicador claro cuando el sujeto evaluado no estaba respondiendo con la verdad (en la versión DLC).

Para una mayor claridad, después de la prueba de familiarización se revisan con el sujeto evaluado las preguntas del examen. Algunas agencias o algunos protocolos de prueba locales podrían especificar que las preguntas de prueba serán totalmente revisadas antes de conectar cualquier componente al sujeto evaluado. No hay ninguna razón teórica que sugiera que esta diferencia podría invalidar el resultado del examen. La colocación previa de los sensores puede permitir que estos se estabilicen, especialmente los electrodos de grabación electrodérmica.

El examinador comienza con la pregunta relevante-de-sacrificio seguido de las preguntas relevantes. La pregunta relevante-de-sacrificio se utiliza para introducir el asunto relevante bajo investigación durante la prueba y no se califica. En las pruebas de polígrafo de investigación, los objetivos de las preguntas relevantes son dictados por las circunstancias de la investigación y comúnmente se formulan alrededor de los aspectos más destacados o intensos de la acusación. En los programas exploratorios, las preguntas relevantes deben describir la participación del sujeto evaluado en las preocupaciones de conductas posibles para los administradores de riesgos o árbitros y deben ser diseñados para agregar validez incremental a su programa particular.

Los objetivos del polígrafo exploratorio idealmente deberían ser seleccionados para investigar áreas de contenido pertinentes a los protocolos actuariales o derivados empíricamente para la evaluación y administración de riesgos. El requisito fundamental para la selección del objetivo de la pregunta relevante es que el asunto conductual de preocupación ofrezca información útil a la autoridad remitente.

Las preguntas relevantes formuladas efectivamente evaluarán directamente el involucramiento conductual del sujeto evaluado en el tema de interés. Las preguntas relevantes no deben generar confusión por el uso de lenguaje o de conceptos relacionados con la motivación psicológica o la intención, ya que se piensa que introducen dimensiones de excusa o racionalización por parte de los sujetos evaluados o de mentirosos hábiles. Por el contrario, los sujetos evaluados veraces podrían producir reacciones falsas debido a la ambigüedad y la falta de concreción de tales preguntas. Las preguntas directas con una estructura gramatical sencilla son el mejor enfoque. Las preguntas relevantes deben estar libres de jerga idiomática y legal que es desconocida para el sujeto evaluado y no deben incluir asuntos de evaluación psicológica o de inferencia. Las preguntas relevantes son preguntas simples que pueden ser contestadas fácilmente con un "Sí" o "No". La renuencia, por parte del sujeto evaluado

para proporcionar una respuesta sencilla a una pregunta sencilla puede ser un indicador de que el asunto no es evaluable o de que el sujeto evaluado es incapaz de eliminar la ambigüedad del asunto. La discusión y resolución de este, no debe ser acusatoria, pero lo suficientemente persistente para obtener una respuesta evaluable sencilla a una pregunta que es comportamentalmente descriptiva del posible involucramiento del sujeto evaluado en el tema de preocupación. Prevalece la preferencia en la práctica, de que las preguntas relevantes por lo general sean respondidas con un "No", aunque se han sugerido algunas excepciones como con las presuntas víctimas de agresiones sexuales graves (Hardy y Murphy, 1996). Los autores del presente estudio no encontraron soporte de que el uso de la respuesta "Sí" en las preguntas relevantes pudiera invalidar la prueba.

A continuación, el examinador introduce las preguntas de comparación. Las preguntas PLC se presentan al sujeto evaluado como necesarias para evaluar más profundamente el carácter del sujeto evaluado y el asunto bajo investigación. Las preguntas PLC se basan en transgresiones cuya materia temática está genérica o conceptualmente relacionada con las acusaciones del examen y que prácticamente todas las personas han cometido, pero que probablemente van a negarse en el contexto de la examinación. Las preguntas PLC son de amplio alcance y por lo general se basan en acciones categóricamente similares a las del asunto bajo investigación. Es decir, las preguntas relevantes acerca de robo, normalmente se asocian con preguntas de comparación de robo o de honestidad en general. Las preguntas relevantes acerca de actos violentos son típicamente asociadas con preguntas de comparación acerca de causar daño. La construcción estándar de preguntas de comparación, como se enseña en las escuelas de poligrafía acreditadas por la Asociación Americana de Poligrafía y la Asociación Americana de Policías Poligrafistas, está recomendada para asegurar su saliencia. Sin embargo, no hay razón para prohibir el uso de preguntas de comparación estándar de "mentira" en casi cualquier contexto de prueba.

Las preguntas de comparación en la Utah-CQT son tradicionalmente "exclusivas" en el sentido de que se separan de asunto relevante por tiempo, lugar o categoría. Las preguntas de comparación que no se separan del asunto relevante a veces se denominan de tipo no-excluyente. Tres estudios (Horvath, 1988; Amsel, 1999; Palmatier, 1991) no lograron establecer alguna ventaja clara y consistente de las preguntas de comparación excluyentes sobre de las preguntas no-excluyentes (Krapohl, Stern & Ryan, 2003). Podlesny y Raskin, (1978) mostraron cierta superioridad en las preguntas excluyentes, en que el tiempo de media recuperación de la Respuesta en la Conductancia de la Piel (RCP), el ancho del tiempo de media recuperación en RCP y la amplitud de la Respuesta Potencial de la Piel (RPP) fueron significativamente más efectivos con las preguntas de comparación excluyentes.

Podlesny y Raskin (1978) también reportaron que los dos tipos de preguntas de comparación producen una identificación significativa de los sujetos evaluados inocentes, pero que sólo las preguntas de

comparación excluyentes producen una identificación significativa de los sujetos evaluados culpables usando calificaciones numéricas. En conjunto, estos reportes sugieren que las preguntas de comparación excluyentes podrían no tener ventaja sobre las preguntas de comparación no-excluyentes cuando los datos son evaluados utilizando los criterios de reacción típicamente empleados en las pruebas de campo (para obtener descripciones de los criterios de reacción, véase: Bell et al, 1999; Handler, 2006; Raskin y Honts, 2002; Kircher et al, 2005). El uso de preguntas de comparación excluyentes puede evitar posibles críticas de que las preguntas PLC también son relevantes y pueden causar un resultado falso negativo. Los autores del presente estudio no encontraron algo que sugiera que una prueba sería inválida si el examinador elige emplear el tipo de preguntas de comparación no-excluyentes.

Al igual que en otras técnicas CQT, el sujeto evaluado es fuerte, pero indirectamente desalentado a realizar admisiones en las preguntas PLC. Si el sujeto evaluado hace una admisión en una pregunta PLC, el examinador toma nota de esa admisión con cierta consternación, "*De veras?, ¿Usted hizo algo que hace pensar que usted es un ladrón?*", y/o también minimiza la admisión, "No, yo sólo estoy preocupado acerca de cosas serias", o modifica la pregunta de comparación. Un ejemplo de esto último es: "*Además de lo que me dijiste, antes de este año ¿Alguna vez mentiste a alguien que confiaba en tí?*"

Tenga en cuenta el modificador en cursiva que precede a la pregunta de comparación. El objetivo final es desalentar admisiones a las preguntas PLC para asegurar que el sujeto evaluado las perciba como ambiguas y amplias en naturaleza. También es importante que el examinador insinúe al sujeto evaluado que mentir en cualquiera de las preguntas relevantes o PLC tendrá como resultado el reprobar la prueba de polígrafo y en una conclusión de engaño en el asunto relevante bajo investigación.

Entonces el examinador introduce y revisa las preguntas neutrales que proporcionan un tiempo para volver a una línea base cuando hay distorsión o una reacción fisiológica a una pregunta específica. Kircher, Kristjansson, Gardner, y Webb (2005) sugieren que los intervalos entre-preguntas se deben incrementar a un mínimo de 35 segundos después de una respuesta cardiovascular fuerte para permitir la recuperación, o insertar una pregunta neutral. En general, el enfoque preferido es esperar para permitir un retorno hacia, o por lo menos dirigido hacia, los niveles de línea base. Las preguntas neutrales deben ser no-emocionales en naturaleza y se responden generalmente con "Sí" para garantizar que el sujeto evaluado esté prestando atención a las preguntas de prueba. Sin embargo, no hay nada que sugiera, que un examen en el que se responde "No" a cualquier pregunta neutral sería inválido. El examinador puede revisar preguntas neutrales adicionales en caso de que se necesiten durante la prueba para re-establecer el trazo de línea base.

Después, el examinador revisa la pregunta introductoria que está redactada de manera similar a las preguntas "sintomáticas" que se utilizan en otros formatos CQT. La pregunta introductoria intenta asegurarle al sujeto evaluado que no se harán preguntas no-revisadas durante el examen y podrían permitir una respuesta de orientación al inicio del examen.

La investigación realizada por Honts, Amato y Gordon (2004) falló en demostrar las funciones de la pregunta sintomática como están descritas y que en realidad podría producir una precisión más pobre, especialmente en los sujetos evaluados inocentes.

La tendencia constante ilustrada por estos y otros investigadores, sugiere que la invención y la adición de nuevos tipos de preguntas no deben ser alentada en la era de las pruebas poligráficas científicas modernas, a menos que la investigación demuestra la eficacia de un nuevo enfoque (Hilliard, 1979).

Aunque es prudente que los examinadores de campo se adhieran a los principios y procedimientos generales que se imparten en la formación básica, es igualmente o incluso más sabio adaptar su práctica de campo para ajustarse a los enfoques modernos con validez comprobada. No creemos que desviaciones menores en las secuencias de preguntas mostradas abajo, o las diversas formulaciones del Utah-CQT podrían causar que una prueba sea inválida, ya que su surgimiento sugiere que los científicos que desarrollaron el método Utah-CQT no buscaron validez mediante la adhesión simplista a un "enfoque de pintura por números", sino que buscaron validez demostrable a través de la construcción de métodos CQT de acuerdo con principios de prueba válidos.

FORMATO DE TRES-PREGUNTAS

El Utah-CQT tiene dos versiones, una versión de tres-preguntas y una versión de cuatro-preguntas (Raskin y Honts, 2002).

La versión de tres-preguntas fue diseñada primero y fue usada principalmente para pruebas de asunto-único, pero también se puede utilizar para pruebas de facetas-múltiples de un alegato conocido individual. La versión de tres-preguntas del Utah-CQT permite un alto grado de flexibilidad en el formato de las preguntas relevantes.

A continuación se describe un ejemplo de la numeración de las preguntas y del tipo de preguntas que se utilizan en la versión de tres-preguntas del Utah-CQT. Para un examen de asunto-único, habrán tres preguntas relevantes, cada una ligeramente reformulada.

Ejemplo de redacción de 3-preguntas del Utah PLT.

Para una prueba de *asunto-único*, de evento-específico con respecto al robo a un banco que ocurrió el pasado jueves, uno podría hacer las siguientes preguntas:

Introductoria 1: ¿Entiende que sólo voy a hacerle las preguntas que discutimos?

Relevante de Sacrificio 2: con respecto a si usted robó o no ese banco, ¿Intenta responder a todas esas preguntas con la verdad?

Neutral 1: ¿Están en este momento encendidas las luces en el interior de esta sala?

Comparación 1: (Antes de que cumpliera X), ¿alguna vez hizo algo que fuera deshonesto o ilegal?

Relevante 1: ¿Robó usted ese banco ubicado en ____ en Austin?

Neutral 2: ¿Está usted ahora físicamente localizado en el Estado de Texas?

Comparación 2: (Entre sus edades de X e Y), ¿alguna vez tomó algo que no le pertenecía a usted?

Relevante 2: ¿Robó usted ese banco ubicado en ____ en Austin el pasado jueves?

Neutral 3: ¿A veces usted escucha música mientras viaja en un auto?

Comparación 3: ¿Alguna vez tomó algo de algún lugar donde usted trabajó, (antes de la edad de X)?

Relevante 3: ¿Robó usted ese banco en ____ el ____?

Los ejemplos anteriores muestran entre paréntesis la cláusula excluyente de la pregunta comparativa

Para un examen de facetas-múltiples, el examinador tiene la opción de hacer dos preguntas relevantes reformuladas con el mismo significado y otra pregunta relevante que se relacione directamente con asunto bajo investigación. Esta tercera pregunta relevante puede ser una pregunta conectora de evidencia, conocimiento culpable o de participación secundaria.

Una tercera alternativa es hacer tres preguntas relevantes separadas relacionadas con el mismo asunto específico bajo investigación. Se recuerda a los lectores que la investigación ha demostrado que los índices de precisión son más altos para las pruebas en las que el sujeto evaluado está siendo totalmente veraz o engañoso a todas las preguntas de la prueba en lugar de que lo esté siendo a sólo algunas de ellas (Honts, Kircher, y Raskin, 1988; Raskin, Kircher, Honts, y Horowitz, 1988; Barland, Honts y Barger, 1989).

Los autores de este estudio recomendarían fuertemente a los examinadores, de ser posible, tratar de limitar el examen a uno en el que el sujeto evaluado sea veraz o engañoso a todas las preguntas relevantes.

Si uno fuera a construir un examen de polígrafo de *facetas-múltiples* en torno a un evento de delito único que implique el robo a un banco, los ejemplos de preguntas relevantes alternativas relevantes podrían ser:

Introductoria 1: ¿Entiende que sólo voy a hacerle las preguntas que discutimos?

Relevante de Sacrificio 2: Con respecto a si usted robó o no ese banco ¿Tiene la intención de responder a todas las preguntas con la verdad?

Neutral 1: ¿Están encendidas en este momento las luces en el interior de esta sala?

Comparación 1: ¿Alguna vez robó cualquier cosa de alguien que confiaba en usted?

Relevante 1: ¿Robó usted ese banco en ____ el ____?

Neutral 2: ¿Está usted ahora físicamente localizado en el Estado de Texas?

Comparación 2: ¿Alguna vez robó cualquier cosa de un amigo o miembro de su familia?

Relevante 2: ¿Planeó o se organizó con alguien para robar ese banco en ____?

Neutral 3: ¿A veces escucha música mientras viaja en un auto?

Comparación 3: ¿Alguna vez robó cualquier cosa de un lugar donde usted trabajó?

Relevante 3: ¿Participó usted de cualquier manera en el robo de ese banco?

Tenga en cuenta que este ejemplo se proporciona con preguntas de comparación *no-excluyentes*.

FORMATO DE CUATRO-PREGUNTAS

El formato de cuatro preguntas es similar al diseño a una versión de la Técnica General de Preguntas Modificado de la Fuerza Aérea (DoDPI, 2006) que utiliza pares de preguntas relevantes que están en medio (entre corchetes) de dos preguntas de comparación. Esta permite mayor flexibilidad al examinador para cubrir más de un aspecto de los asuntos relevantes y el calificar mediante el uso de las preguntas de comparación que las rodean. Las preguntas relevantes pueden variar desde uno a cuatro aspectos o facetas conductuales distintas de un solo delito o alegato. Las reglas de construcción de preguntas son las mismas que las descritas anteriormente para la versión de facetas-múltiples en la versión de tres-preguntas.

A continuación se describe un ejemplo de la numeración de preguntas y del tipo de pregunta que se utiliza en la versión de cuatro preguntas.

- | | |
|-----|--|
| I1 | Introductoria |
| SR2 | Relevante de Sacrificio |
| N1 | Neutral |
| C1 | Comparación |
| R1 | Relevante |
| R2 | Relevante |
| C2 | Comparación |
| N3 | Neutral (opcional) Esta pregunta neutral se puede insertar a opción del examinador para permitir una cierta disminución de la tensión y recuperación de la línea de base. Si se inserta, el examinador saltará esta pregunta neutral durante la calificación de las preguntas. |
| R3 | Relevante |
| R4 | Relevante |
| C3 | Comparación |
| N2 | Neutral |

OPERACIÓN DE PRUEBA

Se instruye al sujeto evaluado que se siente estable y responda a cada pregunta con veracidad. Sin embargo, al abordarlo se evita hacer esto de una manera torpe. Como ejemplo, la siguiente admonición sería típica de este enfoque:

"Necesito que te sientes quieto durante el cuestionamiento de las preguntas. Cualquier movimiento creará distorsión y artefactos en las grabaciones, lo que requerirá que yo repita las preguntas, y hará más larga la prueba".

El examinador rota las preguntas neutrales, de comparación y relevantes (si lo desea) durante las presentaciones siguientes y posteriores. El examinador podría preferir dejar las preguntas relevantes siempre en la misma posición, y sólo rotar las preguntas comparativas y neutrales, haciendo más fácil la calificación de las gráficas al tener un orden fijo de preguntas relevantes. Mover las preguntas ayuda a prevenir el reconocimiento de patrones y a la anticipación de un orden específico de preguntas durante el examen.

Los siguientes son ejemplos de posicionamiento de las series en las cadenas de preguntas mostrando un ejemplo de rotación de preguntas.

Versión de Tres Preguntas

Primera Gráfico I1, SR2, N1, C1, R1, N2, C2, R2, N3, C3, R3

Segundo Gráfico I1, SR2, N2, C3, R2, N3, C1, R3, N1, C2, R1

Tercer gráfico I1, SR2, N3, C2, R3, N1, C3, R1, N2, C1, R2

Versión de Cuatro preguntas

Primer Gráfico I1, SR2, N1, C1, R1, R2, C2, N3 (N3 es opcional), R3, R4, C3, N2

Segundo gráfico I1, SR2, N2, C2, R1, R2, C3, N3 (opcional), R3, R4, C1, N1

Tercer gráfico I1, SR2, N1, C3, R1, R2, C1, N3 (opcional), R3, R4, C2, N2

Como se puede ver arriba, cada pregunta relevante tiene la oportunidad de ser comparada con cada una de las preguntas de comparación a través de la serie de tres gráficos. Como se discutió anteriormente, si después de tres gráficas los resultados son inconclusos, se corren dos gráficas adicionales. El examinador simplemente podría utilizar la primera y segunda cadena de posicionamiento de preguntas de serie para la cuarta y quinta gráfica.

Después de la tercera gráfica, se califican numéricamente las gráficas. Sin embargo, se le dice al sujeto evaluado que el examinador siempre se detiene en este punto para corroborar cuidadosamente la calidad de los registros antes de recolectar más gráficas. Si los puntajes cumplen el umbral de los criterios de decisión, se ha completado la fase de recolección de datos. Si después de estas primeras tres gráficas el resultado de la prueba es inconcluso, se conducen dos gráficas adicionales siguiendo los mismos patrones de rotación descritos anteriormente. Después de la quinta gráfica, todos los puntajes se suman para hacer una determinación de veracidad. La Defense Academy for Credibility Assessment (DACA) Academia de la Defensa para la Evaluación de la Credibilidad, centro de entrenamiento Federal, (2006) permite que el examinador realice solo una cuarta gráfica, y si es necesario una quinta gráfica de prueba. Tenemos conocimiento de que no hay justificación teórica o evidencia que sugiera que esta diferencia en el procedimiento podría invalidar un resultado de prueba.

Las preguntas se presentan al sujeto evaluado por lo menos tres veces a través de las tres gráficas, con una breve discusión entre gráficas para aclarar y resolver cualquier problema percibido que le surja al sujeto evaluado y para reforzar el enfoque tanto en las preguntas relevantes como en las de comparación (Raskin y Honts, 2002). Después de cada presentación de las preguntas de prueba, el examinador debe preguntar al sujeto evaluado si tiene alguna duda con las preguntas de prueba. Honts (1999) revisó los datos de 19 estudios que involucraron 1,092 pruebas de polígrafo y encontró que la discusión entre gráficas (aunque se limite solamente a las preguntas de comparación) reduce el riesgo de error. En 9 de los 11 estudios, que incluyeron exámenes en los que fueron revisadas las preguntas entre gráficas, fueron revisadas tanto las preguntas relevantes como las de comparación. En 2 de esos estudios, sólo se revisaron las preguntas de comparación. Honts mostró que la estimulación y la revisión de preguntas entre gráficas redujo la tasa de falsos negativos (54%), tuvo una modesta reducción de la tasa de falsos positivos (2,9%) y una disminución sustancial de resultados inconclusos para los sujetos evaluados veraces (42%).

La siguiente es un tipo de conversación típica que podría tener lugar entre gráficas. Note cómo el examinador pone igual énfasis en cada grupo de preguntas durante la estimulación y la revisión.

Examinador: OK Roy, ¿Tuvo algún problema con alguna de estas preguntas en el examen?

Roy: No.

Examinador: ¿Viene algo a su mente cuando le hago esas preguntas?

Roy: No.

Examinador: ¿Qué hay de esas preguntas sobre la transacción de drogas? ¿Está claro lo que le estoy preguntando? ¿Las entiende?

Roy: Si.

Examinador: ¿Qué hay de esas preguntas acerca de mentir? ¿Algún problema con alguna de ellas?

Roy: No.

Ha existido controversia en torno a la revisión de preguntas entre gráficos de prueba. Abrams (1999) y Matte (2000) argumentaron que la revisión de preguntas comparativas entre gráficas es incorrecta y Offe y Offe (2007) no encontraron contribución alguna que mejorara o degradara los resultados en las decisiones como consecuencia de la revisión de las preguntas de prueba entre gráficas. En consideración de estos hallazgos, creemos prudente recomendar la revisión de preguntas entre cada gráfica, pero no se encuentra razón para apoyar el argumento de que la inclusión o exclusión de esta revisión podrían provocar un resultado de prueba que resultara inválido o erróneo. Honts (1999) no especuló sobre la causa psicológica de estos hallazgos. Él correctamente declaró "La esencia de la ciencia es el empirismo.

Esto es, el conocimiento científico se construye sobre datos, no por la especulación ni por autoridad". Honts optó por aceptar los datos por lo que es declarado en su propio mérito.

En el caso de que un sujeto evaluado hiciera admisiones adicionales a las preguntas de comparación o fuera necesario modificar una pregunta relevante, el examinador debe hacerlo y re-etiquetar la pregunta. Por ejemplo, si durante un Utah PLT el sujeto evaluado hace una admisión a la pregunta C1 "Antes de este año, ¿alguna vez robaste cualquier cosa de un negocio?", el examinador puede modificar esa pregunta con "Aparte de lo que me dijiste, antes de este año ¿alguna vez te robase cualquier cosa de un negocio?" y etiqueta esa pregunta como C1a. El examinador debe entonces revisar todas las preguntas del examen con el sujeto evaluado. El examinador conduce las siguientes dos gráficas y de nuevo comienza por dar instrucciones al sujeto evaluado para sentarse estable y responder a todas las preguntas con la verdad.

ANÁLISIS DE DATOS DE PRUEBA Y CRITERIOS DE DECISIÓN

El Sistema de Calificación de Utah (Bell, Raskin, Honts y Kircher, 1999) es una versión simplificada de las técnicas de calificación numérica introducidas por Backster en 1963 y modificadas por el US ARMY, Ejército de los Estados Unidos, alrededor de 1970 (Weaver, 1980; Swinford, 1999).

El sistema de calificación de Utah es un sistema de calificación simple y elegante diseñado para mejorar la precisión, reducir los resultados inconclusos, y mejorar la confiabilidad entre evaluadores. En comparación con otros sistemas de calificación que están actualmente en uso, tiene menos reglas que seguir y un menor número de criterios para calificar. El Sistema de Calificación de Evaluación Numérica de Utah fue diseñado, refinado y probado por Raskin y sus colegas.

El Sistema de calificación Utah se basa en datos de la respuesta fisiológica que se han demostrado que son indicadores válidos y confiables de la excitación simpática. Las correlaciones inter-evaluador de los resultados producidos utilizando el sistema de Calificación de Utah típicamente están alrededor de 0,90 (Bell et al., 1999). La precisión del sistema de calificación de Utah en varios estudios análogos fue del 90% cuando se promedia para en sujetos evaluados programados como inocentes y culpables, según lo informado por Bell et al. Los resultados de estudios de campo utilizando el sistema de calificación de Utah son consistentes con los resultados de estudios análogos (Bell et al., 1999)

La evaluación numérica de los datos de prueba se lleva a cabo mediante la comparación de la fuerza relativa de las respuestas a las preguntas de comparación y de las relevantes. El sistema de Utah utiliza un enfoque de puntuación numérica de 7-posiciones. Se comparan las fuerzas relativas de las reacciones fisiológicas de cada sensor y se asigna una puntuación. Las puntuaciones posibles varían de -3 a +3. La reacción de cada pregunta relevante se compara con la reacción de la pregunta de comparación previa en el formato CQT de 3-preguntas o contra la más fuerte de las dos preguntas de comparación circundantes en el formato MGQT de 4-preguntas. Si la fuerza relativa de la pregunta relevante es mayor que la de la pregunta de comparación, se le asigna un valor negativo.

Por el contrario, si la fuerza de la pregunta de comparación excede la fuerza de la pregunta relevante, se le asigna una puntuación positiva. Si no hay una diferencia observable, se le asigna un cero. En algunos componentes hay proporciones relativas mínimas que se deben alcanzar con el fin de asignar una puntuación.

Para la versión de tres-preguntas mostrada anteriormente, la pregunta relevante se compara normalmente contra la pregunta de comparación anterior para su evaluación. Si la pregunta de comparación previa está distorsionada por un artefacto, el examinador podría usar la pregunta de comparación libre-de-artefactos más cercana para su evaluación.

Para la versión de cuatro-preguntas mostrada previamente, el examinador compara la pregunta relevante contra las dos preguntas comparativas que la tienen “entre corchetes”, componente a componente. Por ejemplo, en el primer gráfico de la versión de cuatro-preguntas mostrada anteriormente, R1 se compara contra C1 y C2. El examinador encontrará el canal de reacción más fuerte por separado de cada canal para C1 y C2 y lo usará para compararlo contra el canal correspondiente de R1. Al usar la reacción de la pregunta de comparación entre corchetes más fuerte se ha demostrado que se producen resultados válidos de campo (Honts, 1996; Raskin et al., 1988).

Los trazos fisiológicos que se encuentran afectados por artefactos se excluyen para fines de evaluación. Si el sujeto evaluado respondió "Sí" a la pregunta de comparación durante la prueba, podría ser utilizada la respuesta de la pregunta de comparación en la calificación, siempre y cuando la respuesta revisada hubiera sido "No" (ver Honts, Raskin, y Kircher, 1992). El examinador podría insertar una pregunta neutral rutinariamente después de la segunda pregunta de comparación o en cualquier otro momento en que sea necesario para restablecer la estabilidad del trazo. Durante el análisis de los datos de prueba, el examinador se salta esa pregunta neutral.

El Sistema de Calificación de Utah utiliza un total de siete criterios primarios de puntuación en los canales de respiración, cardiógrafo, electrodérmico, y de actividad vasomotora periférica.

Se asignan valores de -3, -2, -1, 0, +1, +2, +3 por canal para cada pregunta relevante. Como se mencionó anteriormente, si la pregunta relevante es la mayor de las dos, la calificación será un número negativo. Si la pregunta de comparación es la más fuerte de las dos, la calificación será un número positivo, y si no hay diferencia se da una puntuación de cero.

Sólo se puede asignar una calificación de +/- 3 por gráfica, en el canal de cardio y electrodérmico, y sólo si la línea base para el canal está estable y la reacción es la más grande de ese canal en el gráfico. Los totales a las preguntas relevantes se suman después de tres gráficas, y si es inconcluso, después de cinco gráficas.

Para el canal respiratorio, existen cuatro características empíricamente confirmadas que se consideran diagnósticas (ASTM, 2005). Tres de esas características están contenidas por el fenómeno conocido como Longitud de la Línea Respiratoria (LLR) (Timm, 1982). El LLR es simplemente la medición de la longitud de la línea respiratoria por un periodo establecido de tiempo. Se compara la longitud total de la línea en el periodo de tiempo designado entre la pregunta relevante y la o las preguntas comparativas. Mientras más grande sea la supresión o más corta sea la longitud de la línea, más fuerte será la respuesta. Esas tres características son supresión de la amplitud de la respiración (Figura 2), reducción de la tasa de respiración (que incluye los cambios en la tasa de inhalación/exhalación si dan como resultado un decremento de tasa respiratoria, Figura 3) y la apnea cuando ocurre cerca del ciclo de exhalación (Figure 4). El cuarto criterio respiratorio es el incremento temporal en la línea base del trazo.

Un trazo respiratorio se considera como diagnóstico si se encuentran al menos tres ciclos sucesivos de una característica LLR o de incremento temporal de línea base. La excepción es la apnea, donde podría no haber ciclos respiratorios discernibles. Aunque las respiraciones torácicas y abdominales se registran por separado, se les asigna un solo valor. Ese valor se basa en la diferencia combinada observada entre las preguntas relevantes y de comparación.

Quienes desarrollaron el Utah-CQT han enseñado y practicado el ser conservador cuando se evalúa el canal respiratorio. Bell y sus colegas utilizaron una muestra de 50 exámenes de polígrafo para conducir una estudio que proporcionó 450 puntuaciones numéricas. Bell et al. (1999) anotaron esas calificaciones para determinar la distribución de las calificaciones, y reportaron que las puntuaciones de respiración de

0 se asignaron aproximadamente en un 75% de las ocasiones. Calificaciones de +/- 1 fueron asignadas aproximadamente el 20% de las ocasiones, y de +/- 2 o 3 menos del 5% de las ocasiones.

Para el canal electrodérmico, las calificaciones se basan principalmente en la comparación de la amplitud de pico (Figura 6), un criterio que ha demostrado empíricamente ser de diagnóstico. La amplitud se mide desde la línea base pre-estímulo y hasta el pico más alto alcanzado dentro de la ventana de calificación (Bell et al., 1999). Se calcula la proporción entre la pregunta relevante y la de comparación.

Se asigna una puntuación de +/-1 si la fuerza relativa es dos veces más grande, se le asigna una puntuación de +/-2 si la fuerza relativa es tres veces más grande y se asigna una puntuación de +/-3 si la fuerza relativa es cuatro veces más grande. Si el trazo electrodérmico es lábil, no se debe asignar una puntuación de 3.

Pueden ser consideradas la duración y la complejidad de la respuesta como criterios de reacción secundarias. Las reacciones que tienen una clara complejidad o una duración más larga podrían aumentar de un 0 a un +/-1 o de un +/- 1 a un +/-2 (Figuras 7 y 8). Si las proporciones de amplitud son de al menos 1.5:1 con complejidad contra no complejidad o aumento en la duración del tiempo de reacción, esto permite el aumento de una puntuación de 0 a +/- 1. De forma similar, una proporción de al menos 2.5:1 puede aumentar una puntuación de +/- 1 a +/- 2 siguiendo las mismas reglas con respecto al aumento de la complejidad o de la duración. Bell et al. señalaron que aproximadamente en el 50% de las ocasiones fueron asignadas calificaciones de 0 en el canal electrodérmico, puntuaciones de +/- 1 alrededor del 25%, +/- 2 aproximadamente el 20% y +/- 3 en menos del 10% de las ocasiones.

Krapohl y Handler (2006) demostraron que se puede obtener información diagnóstica adicional al interpretar diferencias más pequeñas en las respuestas de las preguntas relevantes y las de comparación usando exámenes ZCT federales. Si bien estos no le pertenecen al formato Utah-CQT, los autores de este trabajo no conocen razón alguna para creer que los valores de proporción mejorados de calificación serían susceptibles de uso en un formato Utah-CQT. Estas proporciones mejoradas están siendo investigadas al tiempo de escribir este artículo.

Para el canal de la presión sanguínea relativa, la fuerza relativa de las reacciones se evalúa con base en el movimiento hacia arriba de la línea base (excitación de línea base) como se muestra en la Figura 9. Se requiere una proporción mínima de 1.5:1 para una puntuación de +/-1; una proporción de 2:1 para una puntuación de +/-2; y de 3:1 para una puntuación de +/-3. Bell et al. (1999) informaron que la duración

de la respuesta puede ser considerada cuando se evalúa la fuerza relativa de la reacción y que una reacción con una mayor duración puede aumentar una puntuación de 0 a 1 o de 1 a 2 (ver Figura 10).

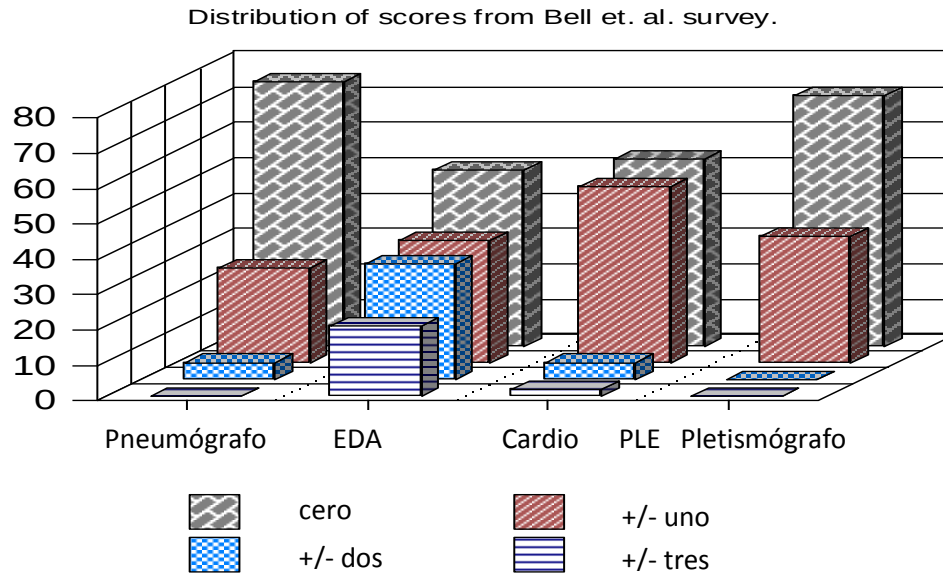
Sin embargo, ellos no discutieron el procedimiento para la aplicación de esta regla y los autores de este trabajo sugieren limitar la puntuación de este canal al incremento en la línea base según lo sugerido por Kircher et al (2005), Harris, Horner y McQuarrie (2000) y ASTM (2002). Bell et al. notaron que las calificaciones de la presión sanguínea relativa de 0 fueron asignadas aproximadamente el 50% de las ocasiones, las calificaciones de ± 1 alrededor del 45%, ± 2 menos de 5% de las ocasiones. Las calificaciones de ± 3 son raras y sólo una de esas calificaciones pueden asignarse por gráfica como se explica en la sección electrodérmica (Bell et al., 1999).

Para la actividad vasomotora periférica, la fuerza relativa de las reacciones se evalúa mediante la comparación de la reducción de la amplitud del pulso (Figura 11). La fuente de este canal es una reducción de monitoreo fotoplefetismográfico en la amplitud de pulso dedal. Las calificaciones numéricas se basan en la duración y el grado de reducción de la amplitud.

Se pueden asignar puntuaciones cuando no hay diferencia en la disminución de la amplitud pero sí una diferencia discernible en la duración de las reacciones (Figura 12). Bell et al. señalaron que en la calificación de la amplitud del pulso dedal, las puntuaciones de 0 fueron asignados aproximadamente el 70% de las veces y calificaciones de ± 1 aproximadamente el 30%, (Bell et al. 1999).

El siguiente gráfico muestra la distribución de las calificaciones numéricas obtenidas por la encuesta de Bell et al. al (1999). Como puede verse en la gráfica, la mayoría de las calificaciones numéricas asignadas son cero o ± 1 para la mayoría de los canales.

Figura 1. Distribución de las calificaciones de la encuesta de Bell et al.



CRITERIOS DE DECISIÓN

El examinador procede a través de las gráficas y totaliza las calificaciones para cada pregunta relevante en cada gráfico. Se determina entonces la calificación total para cada pregunta relevante para las tres primeras gráficas. Para las pruebas asunto-único donde el sujeto evaluado debería ser veraz o engañoso a todas las preguntas relevantes, el corte de calificación es de ± 6 . En otras palabras, cuando hay un total de $+6$ o mayor, el resultado es veraz. Un gran total de -6 o menos se traduciría en una determinación de no veracidad o indicador de engaño. Las calificaciones que caen entre -5 y $+5$ resultarían en una determinación de inconcluso y el examinador conducirá otras dos gráficas como se describió anteriormente.

Tras esas dos gráficas adicionales, las calificaciones a las preguntas relevantes son nuevamente totalizadas. Las cortes de calificación de ± 6 siguen siendo los mismos para cinco gráficas.

Los criterios de decisión son ligeramente diferentes para los exámenes de facetas-múltiples donde el sujeto evaluado podría ser veraz en alguna, pero no en todas las preguntas relevantes. Si los totales por spot (punto) para todas las preguntas relevantes son todos positivos o todos negativos (haciendo caso

omiso de las calificaciones por spot de cero), entonces utilice la regla de Gran Total +/-6 descrita anteriormente para las pruebas de asunto único.

Si cualquiera de los spots es opuesto (algunos positivos y algunos negativos, nuevamente haciendo caso omiso de las calificaciones de cero), entonces use una Regla de Calificación por Spot (RCS) por cada spot. El RCS significa que el total de cada spot debe ser de +3 para una conclusión de que no se indica engaño (NDI), y que cualquier total por spot de -3 o menos llama a una decisión de engaño indicado (DI) para el examen. Sin embargo, si se toman decisiones basándose en preguntas individuales, se recomienda precaución ya que las investigaciones indican que cuando los sujetos evaluados responden algunas preguntas con la verdad y otras con engaño, la precisión para resultados en preguntas individuales se reduce (véase la discusión en Raskin y Honts, 2002.) El problema puede exacerbarse cuando se intenta constatar la veracidad a una o más preguntas cuando las puntuaciones totales para cualquier pregunta relevante han indicado engaño (Raskin y Honts, 2002).

Aunque ha existido un esfuerzo constante para evaluar la validez empírica de diferentes puntos de corte, se ha hecho poco énfasis en la determinación de cortes de calificación determinados estadísticamente, la forma del modelo de detección de señales Gaussiana, como se describe por Barland (1985). Krapohl y McManus (1999), Krapohl (2002), y Nelson, Handler y Krapohl (2007) es una excepción a esta tendencia.

PRUEBA UTAH DE MENTIRA DIRIGIDA

Las preguntas DLC son aquellas en las que el examinador da instrucciones al sujeto evaluado para responder falsamente (Honts y Raskin, 1988; Raskin y Honts, 2002). Las preguntas DLC podrían ofrecer cierto alivio a los problemas potenciales identificados en las versiones de prueba poligráfica PLC. Los examinadores podrían experimentar dificultades en la estandarización de las preguntas de comparación en la versión PLC. Cada sujeto evaluado trae consigo sus propias experiencias de vida e idiosincrasia que pueden obstaculizar el mantenimiento del rapport cuando se intenta sentar las bases para establecer las preguntas PLC. Los sujetos evaluados que tienen experiencia poligráfica previa, o aquellos que han investigado acerca de técnicas de polígrafo, podrían no ser ingenuos a los principios del PLC. Esta sofisticación podría hacer desafiante el poder fundamentar las preguntas de comparación. Los sujetos evaluados no-ingenuos podrían aceptar el procedimiento con el fin de no parecer que protestan, en cuyo caso las preguntas PLC se vuelven similares a las preguntas DLC.

Una de las teorías detrás de la propuesta DLC es de que es similar a la propuesta PLC en el sentido que se asume que la cognición y atención del sujeto (posiblemente algo de emoción) se centrará más en las preguntas que plantean la mayor preocupación por reprobar la prueba. Por lo tanto, el sujeto veraz estará más preocupado en si será o no un sujeto idóneo y podrá producir respuestas adecuadas a las preguntas DLC para mostrar cuando está mintiendo. Aunque tienen permiso para mentir en estas preguntas, las preguntas aun así sirven para llamar la atención del examinado durante la prueba. La teoría además propone que el sujeto culpable permanecerá preocupado principalmente por las preguntas relevantes de la prueba y por lo tanto producirá en ellas las mayores reacciones. En este sentido, las preguntas DLC operan como un elemento distractor para los sujetos veraces, que son más propensos a distraerse lejos de las preguntas relevantes que los sujetos engañosos.

Las preguntas DLC son fáciles de estandarizar, requieren de poca manipulación psicológica y tienen mayor validez aparente. La estandarización y simplificación de cualquier técnica puede servir para incrementar la confiabilidad entre-evaluadores y de test-retest, y ambas dimensiones limitan la validez potencial de una técnica. La variabilidad excesiva en la administración o la interpretación de una prueba necesariamente compromete la confiabilidad y validez de cualquier método de prueba. La confiabilidad entre-evaluadores es una preocupación que se mantiene con suma importancia ante las preguntas acerca de la validez del polígrafo. Cuando las prácticas estandarizadas se basan en principios que son consistentes con constructos validados y datos obtenidos mediante el estudio objetivo de los datos, podemos anticipar razonablemente que esas mejoras contribuirán de manera significativa al objetivo de diseño de prueba, de la validez de criterio y de la precisión en la decisión.

Hay mucho menos estudios de campo y de laboratorio que hablan de la validez del DLC en comparación con los de PLC. Sin embargo, los resultados de los estudios existentes (Barland, 1981; Barland *et al.*, 1989; DoDPI Research Division Staff, 1997; DoDPI Research Division Staff, 1998; Honts & Raskin, 1988; Horowitz, Kircher, Honts & Raskin, 1997; Kircher, Packard, Bell & Bernhardt, 2001; Reed, 1994; Raskin & Kircher, 1990) sugieren que las preguntas DLC se desempeñan tan bien o mejor que las preguntas PLC. Las DLCs requieren prácticas de administración menos complejas que aquellas asociadas a la propuesta del PLC y ofrecen un mayor potencial de estandarización. Los estudios que utilizan técnicas de DLC (DoDPI Research Division Staff, 1997; Research Division Staff, 1998) sugieren que la propuesta DLC y otras mejoras en la estructura de la administración de la prueba y en las políticas de toma de decisiones, contribuyen significativamente a los objetivos del programa de pruebas de polígrafo que son la de sensibilidad a la decepción y especificidad a la veracidad.

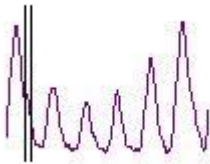
Hay ciertas advertencias que atender en el uso de pruebas DLC. En primer lugar, los examinadores no familiarizados con DLC deben buscar instrucción acerca de su correcto desarrollo e implementación. En

segundo lugar, hay algunos indicios en los datos de investigación de que algunos sujetos de prueba muestran respuestas respiratorias inusuales con el DLC (véase Horowitz et al 1997; Kircher 2001). Sin embargo, los procedimientos estándar de puntuación numérica en Horowitz et al. (1997) se comportaron bien con el DLC, aunque esto podría no ser óptimo y están actualmente en curso las investigaciones que exploran este tema. Más aún, actualmente no existen algoritmos computarizados disponibles que hayan sido especializados en datos DLT. Por lo tanto, los resultados de esos modelos deben ser vistos con cautela.

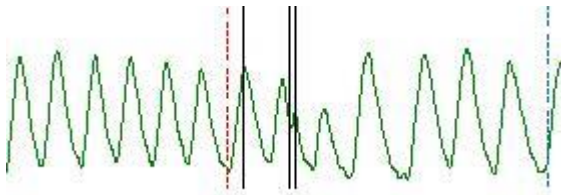
EJEMPLOS DE CRITERIOS DE CALIFICACIÓN

RESPIRACIÓN

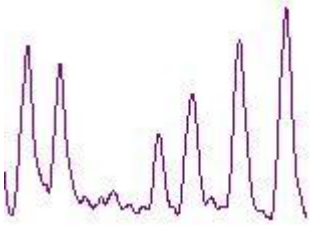
La Figura 2 abajo, muestra un ejemplo de supresión en la amplitud de respiración.



La Figura 3 abajo, muestra un ejemplo de reducción en la tasa de respiración.

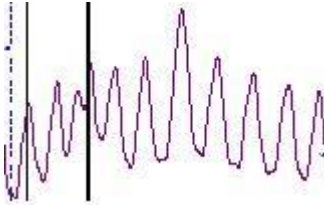


La Figura 4 abajo, muestra un ejemplo de apnea que ocurre en o cerca de la exhalación.



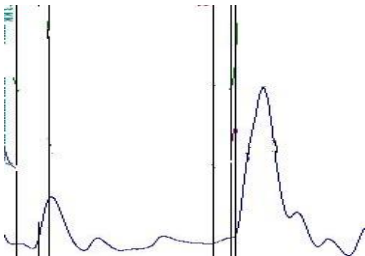
Nota: Los tres criterios de reacción anteriores son aquellos que son capturados por el fenómeno conocido como RLL.

La Figura 5 abajo, muestra un ejemplo del incremento temporal de línea de base.



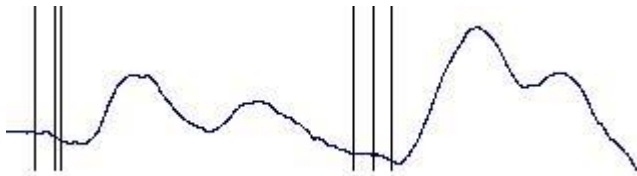
ACTIVIDAD ELECTRODÉRMICA

La Figura 6 abajo, muestra un ejemplo de Incremento de amplitud.



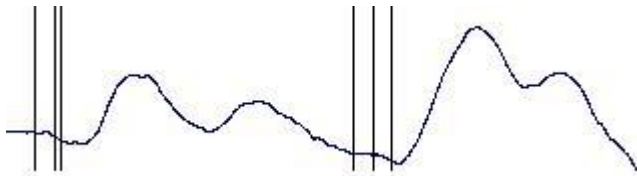
Tenga en cuenta que la proporción entre los trazos anteriores es de aproximadamente 2.6:1 (26 mm contra 10 mm de amplitud) con una duración obviamente más larga y complejidad en la porción de recuperación del trazo. Esta proporción podría calificar para una puntuación de +/- 2 con base en el aumento de la duración y la complejidad.

La Figura 7 muestra un ejemplo de incremento en la duración.



Tenga en cuenta que la proporción entre estos trazos es de aproximadamente 1.8:1 (18 mm contra 10 mm de amplitud) con una duración obviamente más larga y complejidad en la porción de recuperación del trazo. La proporción de amplitud no supera el 2:1 requerido normalmente para una puntuación de +/- 1. Esto calificaría para una puntuación de +/- 1 basado en la proporción de al menos 1.5:1 con una mayor duración y complejidad.

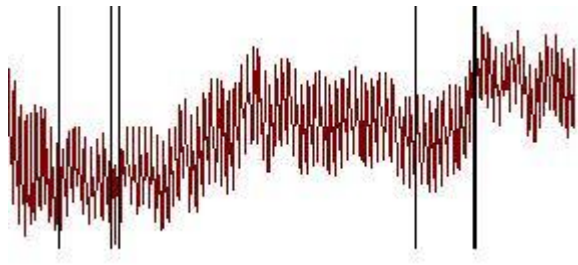
La Figura 8 abajo, muestra un ejemplo de complejidad.



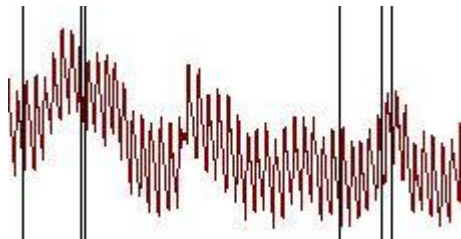
Tenga en cuenta que la proporción de los trazos en la Figura 8 es de aproximadamente 1.8:1 (18 mm contra 10 mm de amplitud) con una duración obviamente más larga y complejidad en la porción de recuperación del trazo. La proporción de amplitud no supera el 2:1 requerido normalmente para una puntuación de +/- 1. Esto calificaría para una puntuación de +/- 1 basado en la proporción de al menos 1.5:1 con una mayor duración y complejidad.

CARDÍÓGRAFO

La Figura 9 abajo, muestra un ejemplo de incremento de línea de base.



La Figura 10 abajo, muestra un ejemplo de aumento en la duración de la respuesta.

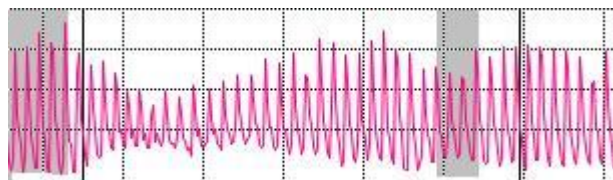


Tenga en cuenta que la proporción de los trazos en la figura 10 es de aproximadamente 1.3:1 (8 mm contra 6 mm de amplitud) con una duración obviamente más largo y complejidad en la porción de recuperación del trazo.

La proporción de la amplitud en la figura 10 no supera el 1.5:1 que se requiere normalmente para una puntuación de ± 1 . Esta proporción podría calificar para una puntuación de ± 1 con base en la mayor duración observada en la primera reacción.

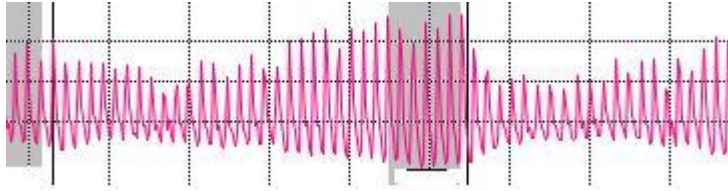
FOTOPLETISMÓGRAFO

La Figura 11 abajo, muestra un ejemplo de reducción de amplitud



La Figura 12 abajo, muestra un ejemplo de incremento en la duración de la reducción en la amplitud

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 este artículo, y aparece aquí con el permiso de la APA



CONCLUSIÓN

El Utah CQT fue creado por psicólogos/examinadores y fundada sobre principios conocidos y probados por la psicología y la psicofisiología. La confiabilidad y validez del Utah CQT se ha demostrado en muchos estudios científicos y publicados revisados-por-pares (véase la revisión en Raskin y Honts, 2002). Pueden encontrarse una serie de escritos en las revistas científicas y textos que discuten el Utah-CQT (por ejemplo, Raskin y Honts, 2002;.. Bell et al 1999; Handler, 2006). Esperamos haber consolidado algunos de esos escritos en una descripción básica de cómo administrar y evaluar adecuadamente el examen.

Aquellos científicos que crearon y perfeccionaron la técnica se esmeraron en la investigación y evaluación exhaustiva de la confiabilidad y la validez del examen. Esto incluyó numerosos estudios de campo y análogos realizados durante tres décadas. El sistema de puntuación Utah (Bell et al., 1999; Handler, 2006) adopta un enfoque algo conservador para la asignación de valores. Esto asegura que las puntuaciones sean asignadas a las reacciones que son claramente diferentes en la comparación y no sean asignadas arbitrariamente. Algunos sostienen este enfoque conservador puede dar lugar investigaciones inconclusas después de tres gráficas y por lo tanto se requieran que sean conducidas dos gráficas adicionales. Desde un punto de vista científico, mientras más datos mejor y las dos gráficas adicionales debe servir para incrementar la confianza en los resultados.

Es la esperanza sincera de los autores de que otros en el campo de la poligrafía consideren el aprendizaje y el uso del Utah CQT. Mientras más acerquemos a nuestra profesión hacia las técnicas que emplean principios validados científicamente, más respeto vamos a ganar de otras personas fuera de la profesión del polígrafo. Como en cualquier campo científico, se puede lograr el progreso a través del perfeccionamiento de técnicas probadas ya existentes.

AGRADECIMIENTOS

Una versión anterior de este artículo se publicó previamente por la American Polygraph Association, Asociación Americana de Poligrafía en la revista, *Polygraph* 2006, 35, (3) y se reimprime aquí con permiso del editor.

Los autores agradecen a David Raskin, Charles Honts, Don Krapohl, John Kircher y Frank Horvath por las revisiones exhaustivas y comentarios a los primeros borradores y por las revisiones de este artículo. Los autores desean conceder el uso ilimitado y derechos de reimpresión a cualquier escuela poligráfica acreditada por la American Polygraph Association o la American Association of Police Polygraphists, Asociación Americana de Poligrafía o la Asociación Americana de la Policías Poligrafistas. Son bienvenidas preguntas y comentarios a polygraphmark@gmail.com o raymondnelson@gmail.com

REFERENCES

- Abrams, S. (1999). The role of polygraphy with the professional board. *Polygraph*, 28, 300-303.
- Amsel, T. T. (1999). Exclusive or nonexclusive comparison questions: A comparative field study. *Polygraph*, 28, 273-283.
- ASTM (2002) E2229-02 Standard Practices for Interpretation of Psychophysiological Detection of Deception (Polygraph) Data. ASTM International.
- ASTM (2005). E2324-04 Standard Guide for PDD Paired Testing. ASTM International.
- Barland, G. H. (1981). *A validation and reliability study of counterintelligence screening test*. Unpublished manuscript, Security Support Battalion, 902nd Military Intelligence Group, Fort Meade, Maryland.
- Barland, G. H. (1985). A method of estimating the accuracy of individual control question polygraph tests. *Proceedings of IDENTA*, 1985, 142-147.
- Barland, G.H., Honts C.R. and Barger, S.D. (1989) Studies of the Accuracy of Security Screening Polygraph Examinations. Department of Defense Polygraph Institute, Fort McClellan, Alabama.
- Bell, B.G., Raskin, D.C., Honts, C.R., & Kircher, J.C. (1999). The Utah numerical scoring system. *Polygraph*, 28(1), 1-9.
- Department of Defense. (2006). Federal Psychophysiological Detection of Deception Examiner Handbook. Retrieved on 03/13/2008 from; <http://antipolygraph.org/documents/federal-polygraph-handbook-02-10-2006.pdf>
- DoDPI (2006) *Air Force Modified General Question Technique*. Presentation at the Senior Polygraph Examiner Course, Austin, TX.
- DoDPI Research Division Staff. (1998). Psychophysiological detection of deception accuracy rates obtained using the test for espionage and sabotage. *Polygraph*, 27, 68-73.
- DoDPI Research Division Staff, (1997). A Comparison of psychophysiological detection of deception accuracy rates obtained using the counterintelligence scope polygraph and the test for espionage and sabotage question formats. *Polygraph*, 26, 79-106.
- Handler, M. (2006). Utah Probable Lie Comparison Test, *Polygraph*, 35, (3) 139-148.
- Handler, M. and Honts, C. R. (2007). Psychophysiological mechanisms in deception detection: A theoretical overview. *Polygraph*, 36 (4), 221-232.
- Hardy, W.L. and Murphy, V.T. (1996). Victims and Polygraph Examinations., *Polygraph*, 25(4).

Harris, J, Horner, A & McQuarrie, A (2000), An Evaluation of the Criteria Taught by the Department of Defense Polygraph Institute for Interpreting Polygraph Examinations Prepared under contract DABT02-96-C-0012 – CLIN0004-Phase 11 for: The Department of Defense Polygraph Institute

Hilliard, D. L. (1979). A cross analysis between relevant questions and a generalized intent to answer truthfully question. *Polygraph*, 8, 73-77.

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. (1999) The discussion of comparison questions between list repetitions (charts) is associated with increased test accuracy. *Polygraph*, 28, 117-123.

Honts, C., Amato, S. & Gordon, A. (2004). Effects of outside issues on the Control Question Technique. *Journal of General Psychology*, 151, 53-74.

Honts, C.R., Kircher, J.C., & Raskin, D.C., (1988) Patterns of activation and deception. *Psychophysiology*, 25, 455.

Honts, C. & Raskin, D. (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., & Raskin, D. C. , & Kircher, J. C. (1992). Effectiveness of control questions answered "Yes": Dispelling a polygraph myth. *Forensic Reports*, 5, 265-272

Horowitz, S. W., Kircher, J. C., Honts, C. R. & Raskin, D.C. (1997). The role of comparison questions in physiological detection of deception. *Psychophysiology*, 34, 108-115.

Horvath, F. S. (1988). The utility of control questions and the effects of two control question types in field polygraph techniques. *Journal of Police Science and Administration*, 16, 198-209.

Kircher, J. C., Packard, T., Bell, B. G. & Bernhardt, P.C. (2001). *Effects of prior demonstration of polygraph accuracies on outcomes of probable-lie and directed-lie polygraph tests*. DoDPI02-R-0002.

Kircher, J.C., Kristjansson, S.D., Gardner, M.G., Webb, A. (2005) *Human and computer decision-making in psychophysiological detection of deception*. Submitted to the Department of Defense Polygraph Institute.

Krapohl, D. (2002). Short Report: An Update for the Objective Scoring System. *Polygraph*, 31, 298-302.

Krapohl, D. (2006). Validated Polygraph Techniques. *Polygraph*, 35(3), 149-155.

Krapohl, D. & Handler, M. (2006). New Ratios for Scoring Electrodermal Responses *Polygraph*, 35, (2) 65-70.

- Krapohl, D. & McManus, B. (1999). An objective method for manually scoring polygraph data. *Polygraph*, 28, 209-222.
- Krapohl, D., Stern, B. & Ryan, A. (2003). Exclusionary vs. nonexclusionary: A review of the evidence. *Polygraph*, 32 (4), 245-250.
- Matte, J. (2000). A critical analysis of Hont's study: The discussion (stimulation) of Comparison Questions. *Polygraph*, 29(2), 146-149.
- Nelson, R., Handler, M. & Krapohl, D. (2008). Brute-Force Comparison: A Monte Carlo Study of the Objective Scoring System version 3 (OSS-3) and Human Polygraph Scorers. *Polygraph*, 37(1), 185-215.
- Offe, H. & Offe, S. (2007). The comparison question test: does it work and if so how? *Law and human behavior*, 31, 291-303.
- Palmatier, J. J. (1991). Analysis of two variations of control question polygraph testing utilizing exclusive and nonexclusive controls. *Masters Abstracts*, 30, 380.
- Podlesny, J. A. & Raskin, D.C. (1978). Effectiveness of techniques and physiological measures in the detection of deception. *Psychophysiology*, 15, 344-58.
- Raskin, D.C. & Honts, C.R. (2002). The comparison question test. In M. Kleiner (Ed.), *Handbook of polygraph testing*. London: Academic (1-49).
- Raskin, D.C. and Kircher, J.C. (1990). Development of a Computerized Polygraph System and Physiological Measures for Detection of Deception and Countermeasures: A Pilot Study. A preliminary report under contract 88-L655300-000, Scientific Assessment Technologies, Inc. Salt Lake City, Utah.
- Raskin, D.C., Kircher, J.C., Honts, C.R. and Horowitz, S.W. (1988) *A Study of Validity of Polygraph Examinations in Criminal Investigation*, Grant number 85-IJ-CX-0040. Salt Lake City: Department of Psychology, University of Utah.
- Reed, S. (1994). Briefing on the Test of Espionage and Sabotage (TES), Expansion and Validation Study: Retrieved on 03/13/2008 from: <http://antipolygraph.org/documents/tesexpansion-validation.shtml>
- Senter, S., Dollins, A. & Krapohl, D. (2004). A comparison of polygraph data evaluation conventions used at the University of Utah and the Department of Defense Polygraph Institute. *Polygraph*, 33, 214-222.
- Swinford, J. (1999). Manually scoring polygraph charts utilizing the seven-position numerical analysis scale at the Department of Defense Polygraph Institute. *Polygraph*, 28(1), 10-27.
- Timm, H. W. (1982) Analyzing deception from respiration patterns. *Journal of Police Science and Administration*, 10(1), 47-51.

Weaver, R.S. (1980) The numerical evaluation of polygraph charts: Evolution and comparison of three major systems. *Polygraph*, 9, 94-108.

*Para cualquier corrección o consulta, el traductor puede ser contactado
en **rodolfo@poligrafia.com.mx***