Directrices sobre la seguridad de los tests, exámenes y otras evaluaciones

Comisión Internacional de Tests

2014

Traducido por: Vicente Ponsoda

Adopción formal

El Consejo de la Comisión Internacional de Tests (ITC) adoptó formalmente las directrices en su reunión de julio de 2014 en San Sebastian, España.

Publicación online

Este documento se publicó online oficialmente después de la Reunión General de la ITC en Julio de 2014 en San Sebastián, España, y desde entonces se puede encontrar en la página web de la ITC: http://www.intestcom.org.

Publicación en papel

Este documento no ha sido todavía publicado en papel.

Por favor, cite este documento así:

International Test Commission (2014). International Guidelines on the Security of Tests, Examinations, and Other Assessments. [www.intestcom.org]

AGRADECIMIENTOS

Estas directrices han sido preparadas bajo la dirección del Dr. David Foster, Kryterion, Inc. y Caveon Test Security (USA), con la colaboración de Eugene Burke, SHL (UK), y Casey Marks, Cambridge Assessments (USA). Los estándares se basan en varios artículos y publicaciones en el campo de la seguridad de los tests. Los autores de las directrices agradecen sus contribuciones a los autores de dichas publicaciones y a los que han participado personalmente en el desarrollo de los estándares contenidos en este documento. Específicamente, querríamos agradecer las aportaciones de

David Bartram (Reino Unido)
Ian Coyne (Reino Unido)
Dragos Iliescu (Rumania)
Tom Oakland (Estados Unidos de América)

El autor quiere agradecer también el esfuerzo, valiosos comentarios y sugerencias de diversos miembros que han colaborado en la fase de revisión: Sara Gutiérrez (CEB SHL Talent Measurement), William G. Harris (Asociación de Editores de Tests), John Hattie, John Kleeman (Questionmark), Fredi Lang (Consejo Asesor de Tests y Diagnósticos de la Asociación Alemana de Psicólogos), Peter Macqueen (Grupo de Referencia sobre Tests y Evaluación con Tests de la Sociedad Australiana de Psicología), Marcus Scott (Caveon Test Security) y Richard Smith (Sociedad Británica de Psicología).

También se deja constancia de los estándares y directrices más importantes que han ayudado a desarrollar los contenidos que va a encontrar en este documento:

- American National Standards Institute (ANSI) (2006). *Guidance for Conformity to ANSI/ISO/IEC 17024: Requirement for Certification Program Security*.

- American Educational Research Association (AERA), American Psychological Association (APA), y National Council on Measurement in Education (NCME) (1999). *Standards for Educational and Psychological Testing*.

- Caveon Test Security (2009). *Test Security Standards*.

- Association of Test Publishers (ATP) (2002). *Guidelines for Computer-Based Testing*.

- International Test Commission (ITC) (2005). *International Guidelines on Computer-Based and Internet Delivered Testing*.

- National Council on Measurement in Education (NCME) (2012). *Testing and Data Integrity in the Administration of Statewide Student Assessment Programs*.

- National Organization for Competency Assurance (NOCA) (2001). *Certification Testing on the Internet.*

Además, se reconoce la contribución del siguiente texto de referencia dedicado enteramente a la protección de los tests y las evaluaciones:

- Wollack, J. A. y Fremer, J. J. (2013). *Handbook of Test Security*. New York: Routledge.

RESUMEN

La cantidad y gravedad de las amenazas a la seguridad han aumentado considerablemente durante las pasadas dos décadas, lo que cuestiona la validez de las evaluaciones aplicadas en todo el mundo. Estas amenazas han aumentado por diversas razones, como el uso frecuente de las tecnologías online e informatizadas para la aplicación de tests y el uso casi indetectable de tecnologías para sustraer el contenido del test y compartirlo ilegalmente al instante a través de fronteras y culturas. Ningún programa de evaluación, por grande o pequeño que sea, es inmune a este daño potencial.

La Comisión Internacional de Tests ha reconocido la necesidad de que cualquier organización que tenga un programa importante de evaluación sea consciente de estas amenazas y esté preparada para hacerles frente. Esta es la razón por la que se han elaborado estas directrices. Conocer las amenazas y estas directrices permitirá adoptar medidas efectivas para proteger el programa de evaluación y lo conseguido con él, y mantener el valor que los tests y las evaluaciones tienen para la comunidad internacional.

Las directrices contenidas en este documento proporcionan recomendaciones sobre cómo planificar un sistema de seguridad, cómo mantener la seguridad durante el desarrollo y aplicación de los tests, y cómo responder correctamente cuando se produce un fallo en dicho sistema de seguridad. Al aplicar estas directrices se creará una importante barrera contra el fraude en el uso de los tests que proteja los logros conseguidos por el programa tras dedicar tiempo y dinero.

CONTENIDO

INTRODUCCIÓN

Propósito de las Directrices de la Comisión Internacional de Tests sobre Seguridad de los Tests y Otras Evaluaciones

La necesidad de hacer seguros los tests, exámenes y otras formas de evaluación ha aumentado en importancia en paralelo con el incremento del uso de los tests y la mayor contribución de la tecnología en la implementación, aplicación y puntuación de los tests, especialmente mediante Internet.

Todos las partes interesadas en el desarrollo y uso de los tests estarían de acuerdo en que el valor de una puntuación obtenida en un test u otra evaluación estructurada disminuye cuando se ha hecho trampa al responder o se ha robado el contenido del test. Hacer trampa se define como el intento de mejorar la puntuación en un test, examen o evaluación, utilizando medios fraudulentos. Robo se define como el intento de sustraer el contenido del test, antes, durante o después de la aplicación del test.

El principal propósito de estas directrices es compartir las mejores prácticas a través de las cuales los creadores de los tests, los que los encargan, los que proporcionan los servicios de evaluación y los usuarios pueden impulsar la seguridad de sus programas de evaluación con tests y defender el valor de la información que proporcionan las puntuaciones obtenidas con estos programas.

En el mejor de los programas se puede hacer trampas al responder, robar el contenido del test o se puede producir algún otro fallo de seguridad. No obstante, un programa de gestión activa de la seguridad ayudará a que se produzcan menos fallos de seguridad y que sus daños sean limitados.

Destinatarios de la Directrices

Son muchas las partes interesadas que participan en los procesos de evaluación con tests y de evaluación en general. Cada una de estas partes interesadas puede resultar afectada por los fallos de seguridad y puede beneficiarse del conocimiento y aplicación de estas directrices. Se describen a continuación siete grupos de partes interesadas.

- **Los evaluados.** Son las personas que responden personalmente a un test o reciben otro tipo de evaluación. Habrán de registrarse para hacer el test, pagar por ello y acordar cuándo ha de ser la aplicación.

- **Los creadores del test**. Son los individuos u organizaciones responsables del diseño y creación del test o de la evaluación. Pueden formar parte de un servicio proporcionado por otros.

- **Los proveedores del servicio de aplicación de tests.** Son organizaciones, como los centros de evaluación con tests, que disponen de tecnología y canales de distribución para asegurar que un test publicado esté disponible cuando y donde convenga a las personas que van a ser evaluadas.

- **Los proveedores del servicio de seguridad de los tests.** Ofrecen servicios específicos, como el análisis forense, para mejorar la seguridad. Pueden formar parte o no de una organización que sea responsable además de otros servicios.

- **Los proveedores de servicios tecnológicos.** Son organizaciones que proporcionan a otras partes interesadas servicios de bases de datos, tecnología de bancos de ítems, servicios de comunicación, soporte de almacenamiento, etc.

- **Los editores o propietarios del test.** Son las organizaciones o personas propietarias del contenido del test, que autorizan su uso para propósitos concretos. Contratan con los proveedores de distintos servicios cuando lo necesitan.

- **Los usuarios del test.** Son las partes interesadas que utilizan la información del test y sus puntuaciones para la decisión individual o de grupo o para la elaboración de políticas educativas o de otro tipo.

Cómo se han estructurado las Directrices de Seguridad

Las directrices se han estructurado alrededor de las acciones claves que sustentan la seguridad efectiva de los tests y de las evaluaciones. Estas acciones se han clasificado como sigue:

- **Desarrollo e implementación de un plan de seguridad** que detalle la preparación necesaria, incluya la creación de un plan de respuesta a incidentes de seguridad, y establezca las políticas y procedimientos para gestionar activamente la seguridad.

- **Implementación de un sistema de seguridad en el proceso de evaluación con tests** que cubra tanto el diseño y desarrollo del test/evaluación, como los procedimientos administrativos para la realización de los tests y evaluaciones.

- **Respuesta ante un fallo en el sistema de seguridad** cuando se descubra que ha habido trampa en las respuestas o robo del contenido.

Cómo aplicar estas directrices

Las directrices se han pensado para que sean aplicables internacionalmente. Son muchas las condiciones del contexto que pueden afectar a la gestión y puesta en práctica de las directrices. Estas condiciones deben ser tenidas en cuenta a nivel local cuando sean implementadas. Algunas de ellas se exponen a continuación:

- Diferencias sociales, políticas, institucionales, lingüísticas y culturales entre los entornos en los que se realizan las evaluaciones.

- Leyes, estatutos, políticas, estándares internacionales y otros documentos legales relacionados con la evaluación con tests.

- Leyes relativas al tratamiento que se ha de dar a los datos procedentes de los tests en los distintos países.

- Directrices nacionales y estándares de desempeño establecidos por sociedades profesionales y asociaciones.

LAS DIRECTRICES

Objetivo de las Directrices

La incidencia de fraude en los tests es mayor en los escenarios en los que un test, examen o evaluación estructurada proporciona una puntuación que tiene consecuencias significativas para la persona evaluada y/u otras partes interesadas. Uno de estos escenarios sería la aplicación de tests para la admisión o no a un programa educativo, o los tests realizados durante y al final de un programa para obtener la acreditación. En el contexto clínico, estamos ante escenarios de consecuencias importantes cuando se toman decisiones relativas al tratamiento clínico a seguir o los procedimientos legales que se derivan del diagnóstico de la persona evaluada. En contextos organizacionales, son escenarios de consecuencias importantes la obtención de empleo, la promoción en una organización, así como la evaluación de destrezas que permite a las personas evaluadas obtener acreditaciones profesionales, como certificaciones y carnets. En contextos forenses, estaríamos también ante un escenario de consecuencias importantes para el evaluado cuando se ha de determinar la capacidad de un individuo para ser juzgado por un delito y, en caso de ser condenado, la severidad de la sentencia.

Dada la madurez de la investigación y la práctica en la seguridad de los tests y aunque estas directrices se centran en el uso de los tests, la seguridad es un asunto que puede surgir en cualquier forma de evaluación estructurada de conocimientos, habilidades, capacidades y atributos psicológicos de los individuos, como muestran los ejemplos proporcionados en el párrafo anterior. Por ejemplo, en un entorno laboral, una persona puede ser evaluada a través de una entrevista, haber sido preparada por un instructor y haber tenido acceso a las preguntas típicas de la entrevista antes de la evaluación. Las observaciones del comportamiento en el lugar de trabajo o en el aula constituyen otra forma de evaluación que puede resultar afectada por las amenazas de seguridad, sobre todo si el observador pudiera resultar beneficiado de producirse en la observación algún resultado concreto. Aunque los términos test y examen se utilizan con más frecuencia, el lector encontrará en estas directrices información útil para mejorar la seguridad de todos los tipos de evaluaciones.

Las evaluaciones planteadas originalmente como de consecuencias poco importantes para el evaluado (por ejemplo, una evaluación de 360 grados para identificar las necesidades de formación y desarrollo de los empleados) pueden ganar importancia cuando los evaluados se dan cuenta de sus repercusiones, como la posibilidad de acceder a programas de formación y desarrollo y de recibir posteriormente recompensas, como aumentos de sueldo y/o promociones. Las partes involucradas en la

evaluación probablemente encontrarán útiles estas directrices independientemente de que se considere que las consecuencias de la evaluación son inicialmente muy o poco importantes. Reconocido lo anterior, hay que añadir que estas directrices no son de aplicación a los contextos que no requieren seguridad, como las autoevaluaciones y los tests de prácticas.

Estas directrices hacen referencia con frecuencia a la tecnología para facilitar la prevención o la detección del fraude en los tests. Aunque la aplicación de los tests se hace cada vez más mediante ordenadores e Internet en muchos lugares, la seguridad del test es un asunto relevante para cualquier forma de aplicación o evaluación. En consecuencia, los principios descritos en estas directrices atañen a los tests de lápiz y papel, a los tests informatizados y a las situaciones en las que se combinan varios modos de aplicación.

En resumen, el objetivo de estas directrices es potenciar la seguridad de todos los tests y evaluaciones con independencia de que sean aplicados en escenarios de consecuencias de mucha o poca importancia para el evaluado, así como fomentar las mejores prácticas tanto si las pruebas o evaluaciones se aplican de forma manual como informatizada. No obstante, los métodos y niveles de seguridad que se establezcan pueden variar dependiendo del tipo y/o condiciones de la evaluación.

La seguridad no es un asunto todo o nada. Por lo general, se ha de alcanzar un equilibrio entre los riesgos de que se pueda hacer trampa al responder y de robo del contenido del test, por un lado, y los costes de la prevención de dichos riesgos, por otro. Este equilibrio depende de los intereses en juego. Estas directrices pretenden mostrar todo lo que se puede hacer para maximizar la seguridad, pero se reconoce que no todas las directrices han de ser necesariamente implementadas en todos los casos. Así pues, es necesario llevar a cabo un análisis de los riesgos en cada nueva situación e implementar medidas de seguridad que permitan abordar y mitigar esos riesgos. También pone de relieve la necesidad de que el plan de seguridad tenga un enfoque global que abarque todo el proceso de evaluación.

Las directrices se dividen en tres partes: (1) Desarrollo e implementación de un plan de seguridad, (2) Implementación de un sistema de seguridad en el proceso de evaluación con tests, y (3) Respuesta ante un fallo en el sistema de seguridad. Cada una de estas partes se presenta en este orden a continuación.

Parte 1: Desarrollo e implementación de un plan de seguridad

La terminología básica de un buen trabajo sobre seguridad incluye los conceptos de amenazas, riesgos, vulnerabilidades y fallos de seguridad. La preparación de un buen programa requiere conocer que existen amenazas específicas de seguridad y que se puede estimar cuanto riesgo suponen. A modo de ejemplo, las vulnerabilidades o debilidades en la seguridad del programa, así como la inadecuada preparación del personal aumentan el nivel de riesgo. Teniendo en cuenta las circunstancias que se dan en un momento concreto, el riesgo puede ser calculado de manera informal a partir de:

- la probabilidad de que una amenaza se haga realidad,

- la facilidad con la que las vulnerabilidades del programa pueden ser aprovechadas para realizar acciones fraudulentas,
- la cantidad de daño que una amenaza puede causar si acaba produciendo un fallo en la seguridad,
- el grado de preparación del programa para detectar/detener el fallo de seguridad y reparar el daño.

Un ejemplo real puede ser útil para ilustrar estos conceptos. Dado que los tests estatales aplicados en USA tienen consecuencias importantes para el evaluado, se han producido y se han investigado en este país un gran número de incidentes de seguridad. En ellos estuvieron implicados directores y profesores que supuestamente manipularon las puntuaciones de los tests (la amenaza) cambiando las hojas de respuestas, entrenando a los estudiantes en las respuestas a los tests, filtrando el test a los estudiantes antes de la aplicación, y de otras maneras (mediante los distintos métodos de hacer trampas que se exponen más adelante). La probabilidad de que se haga trampa y se cause el daño esperado (el riesgo) se puede estimar de antemano evaluando cuantos fallos de seguridad se han producido en otros estados, el daño que han causado, y siendo conscientes de que los profesores y los directores son las personas realmente responsables de la aplicación del test (una vulnerabilidad). Se produce un fallo de seguridad cuando alguien ha hecho trampas al responder al test y se ha detectado.

Mediante un proceso de análisis de riesgos, un programa puede priorizar sus limitados recursos y asignarlos a eliminar o reducir las amenazas, reforzar las vulnerabilidades, instalar mecanismos de detección para descubrir rápidamente los fallos de seguridad, y a la preparación para minimizar y remediar los efectos de dichos fallos.

El establecimiento de un plan de seguridad eficaz requiere que uno comprenda la naturaleza de las amenazas actuales de seguridad del programa y sus riesgos asociados. Una situación que facilite que se haga trampas o se robe el contenido del test es una amenaza a la seguridad. Por ejemplo, existe amenaza de que se haga trampas cuando se puede utilizar un teléfono móvil para recibir mensajes de texto mientras se responde al test. Existe amenaza de robo cuando alguien puede acceder a un dispositivo o lugar de almacenamiento del test y puede conseguir una parte o todo su contenido. La necesidad de establecer y revisar un plan eficaz de seguridad aumenta a medida que se conocen mejor las amenazas particulares y los riesgos de un programa de evaluación con tests. Un plan de seguridad correctamente desarrollado y gestionado reducirá las amenazas y los daños causados por los fallos de seguridad.

Las tablas 1 y 2 exponen las distintas amenazas relacionadas con hacer trampas y robo, respectivamente (Foster y Miller, 2012).

Tabla 1. Tipos de amenazas relacionadas con hacer trampas

Amenaza	Descripción
Conocer el contenido del test antes de su aplicación	La persona evaluada obtiene preguntas del test de una fuente de confianza antes de que le sea aplicado.
Recibir ayuda de expertos cuando está respondiendo al test	La persona evaluada recibe ayuda del profesor o de otra persona durante el test.

Uso de ayudas no autorizadas	La persona evaluada usa ayudas no autorizadas, como chuletas, teléfonos móviles, auriculares, calculadoras programables, etc.
Un suplantador hace el test por la persona evaluada	La persona evaluada se vale de un servicio que proporciona suplantadores o pide a un amigo o colega que haga el test en su lugar.
Manipulación de las hojas de respuestas o de los resultados almacenados	Hecho el test, una persona (por ejemplo, un profesor) puede alterar las hojas de respuesta, cambiando las respuestas erróneas por correctas, o puede cambiar directamente la puntuación asignada al evaluado.
Copia de las respuestas de otra persona	La persona evaluada copia las respuestas de otra persona que está también respondiendo al test.

Tabla 2. Tipos de amenazas relacionadas con el robo del contenido

Amenaza	Descripción
Robo de los archivos que contienen el test o los cuadernillos	El contenido del examen es especialmente vulnerable al robo en algunas etapas de la distribución del test (por ejemplo, cuando los archivos se almacenan en el servidor o los cuadernillos están guardados en un despacho o almacén). Los ladrones pueden conseguir todo el contenido del test y las respuestas correctas si los controles del acceso son inadecuados.
Robo de las preguntas del test mediante fotografía digital o dispositivos de copia	Las preguntas del examen se pueden conseguir durante la aplicación del test. Un ladrón puede utilizar una cámara digital oculta e indetectable u otros dispositivos de copia (por ejemplo, bolígrafos que escanean).
Robo de las preguntas mediante la grabación electrónica del contenido de la prueba	En el caso de los tests informatizados, se puede grabar una sesión completa de la aplicación del test, incluyendo todas las preguntas de la prueba, con un procedimiento de registro digital conectado a uno de los puertos de salida del ordenador.
Memorización del contenido de la prueba	La persona evaluada memoriza preguntas que son recordadas y grabadas en un momento posterior. Dado que hace falta un esfuerzo colectivo organizado para memorizar todas las preguntas, a este tipo de robo se le denomina "robo organizado".
Transcripción verbal de las preguntas	El contenido oral o escrito puede obtenerse durante la aplicación del test utilizando aparatos de grabación de audio o de texto, como teléfonos móviles, radios bidireccionales o tomando notas con un dispositivo electrónico o en papel.
La obtención de material del test a partir de alguien que trabaja para el programa	Un empleado o responsable de elaborar un programa de evaluación con tests puede robar el contenido de la prueba durante su desarrollo, publicación o distribución.

Un análisis de riesgos evalúa la probabilidad de que las amenazas de las tablas 1 y 2 se hagan realidad y la cantidad de daño que podrían causar si el fallo de seguridad llegara a producirse. Se muestran a continuación dos ejemplos.

Que una persona sola, con sus propios medios, haga trampas al responder al test es probable e incluso frecuente en cualquier programa de evaluación con tests. El daño ocasionado generalmente se limita a una sola decisión errónea pues solo hay una puntuación incorrecta. Por otro lado, un cuadernillo robado y distribuido online puede aumentar indebidamente las puntuaciones de miles o decenas de miles de personas evaluadas. Este evento es menos probable pero produce un daño mucho mayor. Una organización debe decidir cuántos de sus limitados recursos deben aplicarse a detectar, disuadir y hacer frente a los que hacen trampa individualmente y cuántos a establecer procedimientos que dificulten el robo y la distribución de cuadernillos.

Las tablas 1 y 2 presentan una taxonomía de los pocos tipos de amenazas que se conocen hoy en día. Sin embargo, para cada tipo, los métodos que la gente realmente utiliza para hacer trampa o robar se cuentan por cientos. Siguiendo el ejemplo de la industria bancaria, un esfuerzo integral de seguridad debe utilizar múltiples capas de procedimientos de seguridad, dado el supuesto bien establecido de que varios procedimientos, trabajando a la vez, tendrán más éxito que uno solo. Estas directrices se han pensado para que se apliquen de forma combinada e indiquen cómo se ha de preparar un programa para enfrentarse de forma eficaz a los riesgos de seguridad.

Directrices específicas sobre el desarrollo e implementación de un plan de seguridad

Para gestionar la integridad del test y de todos los materiales de la evaluación, así como las puntuaciones de los tests y las decisiones basadas en ellas, es necesario un documento detallado que describa el plan de seguridad.

1. Este documento debe identificar los roles y responsabilidades relacionadas con la seguridad en todas las etapas clave del proceso, desde el diseño y desarrollo, hasta la implementación y aplicación del test, y la recogida, almacenamiento y análisis de los resultados. El documento puede incluir todas o algunas de las funciones que se detallan a continuación:

   a. Director de Seguridad. Cuando sea posible, un programa debe nombrar a un director de seguridad que sea responsable de todos los aspectos de la seguridad del programa.

   b. Comité de Seguridad. El programa debe nombrar un comité de seguridad, presidido por el Director de Seguridad, e integrado por los responsables de la creación/mantenimiento del plan de seguridad, de la evaluación de la severidad de los incidentes de seguridad y de las políticas a aplicar cuando se produzcan dichos incidentes, de la supervisión de las respuestas a los fallos de seguridad, y por quienes realizan otras funciones relacionadas con crear y mantener un plan de seguridad viable.

   c. Gestores. Las personas con responsabilidad en el desarrollo y aplicación de los tests, y en la recogida y almacenamiento de los resultados deben conocer bien y

estar de acuerdo con las políticas y procedimientos de seguridad establecidos en el plan de seguridad.

d. Supervisor, vigilante, o aplicador del test. Como parte del proceso de seguridad, son responsables de la aplicación segura de la prueba, incluyendo la autenticación y vigilancia de la persona evaluada a lo largo de la sesión de evaluación. Los supervisores y/o aplicadores de la prueba no deberían participar como instructores, expertos, formadores, o en otras funciones que les proporcionen acceso a los contenidos de la prueba o en tareas que puedan dar lugar a posibles conflictos de intereses que pudieran afectar al rendimiento de la persona evaluada.

e. Proveedores de servicios de seguridad de los tests. Colaboran con el programa en la identificación de vulnerabilidades, ayudan a prevenir problemas de seguridad, detectan fallos de seguridad cuando se producen, determinan la magnitud de los daños, recomiendan las acciones a tomar y eventualmente las llevan a cabo. Entre estos profesionales de la seguridad se encuentran consultores, investigadores, analistas forenses de datos, especialistas en la supervisión de la web, expertos legales y algunos otros.

2. El documento del plan de seguridad debe especificar los derechos y responsabilidades relativas a la evaluación de las personas evaluadas y cómo se deja constancia de que el evaluado conoce esos derechos y responsabilidades.

   a. Los evaluados tienen el derecho de que se les apliquen tests seguros, especialmente si las consecuencias de la evaluación son importantes para ellos. Ningún evaluado debiera obtener una ventaja injusta por hacer trampas o cometer algún otro tipo de fraude cuando responde al test.

   b. Las personas evaluadas sospechosas o acusadas de haber cometido algún tipo de fraude en el test tienen derecho a un juicio justo.

   c. Los evaluados tienen la responsabilidad de no divulgar el contenido del test y de informar de dicha actividad si la descubren.

3. El documento del plan de seguridad debe estar a disposición de las partes interesadas que lo soliciten.

4. El plan de seguridad debe incluir un plan de acción ante los fallos de seguridad, que describa qué hacer cuando se produzca uno. El plan de acción debe incluir objetivos, plazos, personal clave, sistemas de elaboración de informes, la cadena de notificación, las reglas sobre cómo comunicar al público el incidente, las relaciones con los medios y las acciones correctivas concretas a adoptar en función de la naturaleza del incidente o fallo de seguridad. Algunas de las posibles acciones correctivas son las sanciones para los infractores, la anulación de puntuaciones, la repetición del test, la sustitución de los bancos de ítems o del test y el emprendimiento de acciones legales.

5. Las normas de seguridad deben estar claramente indicadas en el plan de seguridad y comunicadas a todas las partes interesadas. Las consecuencias de la violación de

estas normas deben ser claras.

6. El plan de seguridad debe ser aprobado por un conjunto representativo de las partes interesadas y revisado al menos anualmente.

7. El plan de seguridad debe detallar los requisitos de seguridad de las políticas y procedimientos relacionados con las tecnologías de la información y comunicación que han de cumplir los empleados, responsables y todos los proveedores de servicios. Entre los requisitos a detallar están los relacionados con el almacenamiento seguro y el acceso a los contenidos de la prueba, resultados y posible información adicional de la evaluación e información de la persona evaluada, y la protección de dicha información durante los procesos de comunicación y transferencia de datos.

8. El plan de seguridad debe incluir referencias a las leyes de protección de datos de los diferentes países y regiones donde se aplica el test. El plan debe indicar cómo se deben modificar las políticas y procedimientos para dar cabida a estas diferencias. Los esfuerzos para proteger los datos de los individuos y las organizaciones deben ser compatibles con las leyes y políticas vigentes.

9. Debe haber una reserva suficiente de recursos que permita el funcionamiento de las medidas de seguridad preventivas y la supervisión de las actividades descritas en el plan de seguridad. Además, debiera haber un fondo de reserva suficiente para responder al más grave de los posibles fallos de seguridad. El presupuesto de seguridad debe revisarse periódicamente, ajustarse según sea necesario, y estar en consonancia con la identificación de nuevas amenazas.

10. Se deben elaborar materiales de formación en seguridad que cubran las funciones y responsabilidades descritas en el documento del plan de seguridad y repartirlos a todos los individuos involucrados en el programa de evaluación con tests.

11. El compromiso de no divulgar el contenido del test y otros acuerdos debe cumplirse de forma rutinaria por todas las partes implicadas, y en particular por las personas evaluadas, proveedores de servicios y empleados del programa. Estos acuerdos requerirán el reconocimiento del copyright, de la propiedad de la prueba y del contenido de la evaluación, y el reconocimiento de los actos que se consideran fraudulentos y las posibles consecuencias de tales actos. Los acuerdos exigirán a los individuos que se comprometan a no divulgar los contenidos no públicos aludidos.

12. El propietario de la prueba debe establecer legalmente que tiene el copyright y es el propietario del test, para proteger su contenido en los países en los que el test va a ser aplicado.

13. Los protocolos de seguridad de todos los proveedores de servicios han de ser supervisados y auditados periódicamente por expertos en seguridad internos o externos para evaluar su eficacia.

Parte 2: Implementación de un sistema de seguridad en el proceso de evaluación con tests

Después de haber desarrollado y aprobado un plan de seguridad, se puede diseñar, crear, implementar y gestionar la seguridad de los procesos que ocurren antes, durante o después de la aplicación de un test o evaluación. Algunos procesos que tienen implicaciones para la seguridad son:

- Registro de la persona evaluada
- Autenticación o identificación de la persona evaluada
- Diseño del test y del ítem
- Desarrollo del test
- Publicación y distribución del test
- Aplicación del test
- Puntuación del test
- Resultados del test, recopilación de datos sobre el candidato y almacenamiento prolongado de la información.

Muchos de estos procesos requieren la gestión y distribución de materiales confidenciales entre las partes interesadas.

Directrices específicas sobre Implementación de un sistema de seguridad en el proceso de evaluación con tests

1. Las personas evaluadas han de registrarse formalmente para la evaluación. El registro para la realización de un test o evaluación debe requerir, como mínimo, la asignación de un nombre de usuario y clave de acceso o contraseña específicos y únicos para cada persona evaluada.

2. Previamente a la realización del test, las personas que van a ser evaluadas han de ser informadas de que deberán seguir los procedimientos de autenticación establecidos. La organización responsable de la evaluación puede conocer a las personas que van a ser evaluadas o pueden disponer de carnets de identificación emitidos por la organización. Si no es el caso, se debe pedir a las personas que van a ser evaluadas que proporcionen información verificable (por ejemplo, mediante documentos oficiales con fotografías) o habrán de participar en procedimientos de autenticación biométrica. En algunas evaluaciones, como cuando se aplican pruebas de cribado en selección de personal, se habrá de informar a las personas a evaluar de la necesidad de someterse al proceso de autenticación en un momento posterior a la realización del test.

3. Los procedimientos de registro ayudan a garantizar que accedan al registro y programación de la realización de la evaluación solo las personas cualificadas para ello. Entre los requisitos para una adecuada cualificación se encuentran los siguientes: haber completado ciertos estudios, haber respondido y aprobado un test que asegure que se cumplen ciertos prerrequisitos, o haber realizado un pago. En ocasiones, para los que realizan la prueba en repetidas ocasiones, otro requisito es el tiempo que ha de transcurrir antes de poder repetir el test.

a. Si está permitido y es compatible con las leyes sobre protección de datos, el programa de evaluación puede crear y tener actualizada una lista de personas de alto riesgo. Con ello se indicaría a los sistemas de registro y programación de la fecha de aplicación del test la relación de personas a las que se les ha de prohibir o limitar la aplicación del test, de acuerdo con las reglas establecidas en el programa.

4. Las políticas relativas a la repetición del test que se establezcan deben procurar reducir las oportunidades de robo organizado de ítems y otras formas de fraude. Por ejemplo, no se debe permitir a un evaluado repetir el test si lo ha aprobado o si no ha transcurrido un período de tiempo prefijado.

5. Las inscripciones para hacer un test o evaluación deben ser supervisadas cuidadosamente para evitar que se responda al test más a menudo de lo permitido, con el fin de reducir al mínimo las oportunidades para llevar a cabo un robo organizado de ítems.

6. Conviene prestar atención a los diseños de tests que limitan la exposición de los ítems o cambian su orden de presentación, sin deteriorar su calidad psicométrica. Un asunto relevante en estos diseños es cómo se seleccionan y se presentan los ítems. Entre las opciones posibles están los tests adaptativos informatizados, los tests lineales automatizados, los tests multietápicos y los conjuntos de tests equivalentes. Otros asuntos relevantes son si los evaluados pueden revisar y cambiar sus respuestas y las reglas de parada anticipada.

   a. El diseño del test puede interrumpir la presentación de ítems y evitar la exposición innecesaria de preguntas adicionales cuando el número de preguntas respondidas es suficiente para generar una puntuación con un nivel aceptable de fiabilidad y para proporcionar evidencias de validez (por ejemplo, sobre la cobertura de contenido).

   b. El diseño del test podría detener o modificar de algún modo la aplicación de los ítems del test si hay evidencia de que un evaluado carece de motivación, está haciendo trampas, robando preguntas, enfermo, fatigado, o no puede o no quiere proporcionar una estimación correcta del atributo que se está evaluando.

7. Los programas pueden considerar un diseño de presentación de ítems “solo hacia adelante” que no permita a los evaluados acumular ítems para su posible robo memorizándolos o grabándolos digitalmente. Los diseños de ítems y tests son más seguros cuando restringen o reducen la capacidad de marcar los ítems para su posterior revisión, como ocurre, por ejemplo, en los tests adaptativos informatizados, en los que no se permite por lo general la revisión de las respuestas.

8. La exposición del contenido del test o de la evaluación debe ser supervisada y controlada activamente. Por ejemplo, los creadores del test deben evitar que la selección de ítems del banco produzca una sobreexposición de los ítems imprevista y no controlada.

9. Los procedimientos de aplicación del test pueden conseguir un mejor control y gestión del uso y exposición de los ítems si el banco de ítems tiene muchos ítems.

10. Los ítems deben ser diseñados para gestionar y limitar la exposición de los contenidos. Hay muchas maneras de hacer esto. Se presentan algunos ejemplos a continuación. La elaboración y aplicación de los nuevos tipos y formatos de ítems, como son los formatos de elección-forzosa para ítems de auto-informe, las simulaciones, o los ítems que utilizan multimedia, etc., pueden requerir cambios en los sistemas disponibles para el desarrollo de tests, plataformas de distribución y sistemas de almacenamiento.

    a. Al utilizar el formato de opción múltiple, considere no mostrar todas las opciones. Por ejemplo, una posibilidad consiste en presentar las opciones de una en una hasta que la pregunta se responda correctamente, otra posibilidad sería extraer las opciones de un conjunto más grande; en ambos casos sólo se muestra al evaluado un subconjunto de las opciones disponibles.

    b. Cuando se utilicen ítems de opción múltiple, las opciones pueden presentarse en un orden aleatorio para confundir a los evaluados que puedan tener conocimiento previo del test.

    c. El uso de vídeo, audio, simulaciones y otros formatos multimedia puede hacer que sea más difícil filtrar el contenido de la prueba y evitar el uso de algunos métodos de hacer trampas.

11. Puede ser necesario emplear un test de verificación para confirmar las puntuaciones obtenidas en una prueba previa aplicada en condiciones menos seguras. Este proceso de verificación debe realizarse con el conocimiento y consentimiento de la persona evaluada.

12. El tiempo establecido para la aplicación del test debe basarse en un estudio estadístico, ser suficiente para completar la prueba y, a la vez, ser lo suficientemente ajustado para reducir las posibilidades de hacer trampas y robar el contenido de la prueba.

13. El contenido del test debe ser protegido cuidadosamente durante su desarrollo, ya que psicómetras, editores, expertos en la materia y otros profesionales han de tener acceso a los ítems y al test en las distintas fases de su desarrollo.

    a. Los ítems y tests deben ser protegidos limitando su exposición exclusivamente a los que han de elaborarlos o revisarlos, y por un tiempo limitado.

    b. Se deben aplicar procedimientos de acceso estrictos (por ejemplo, nombres de usuario y contraseñas o métodos biométricos).

    c. Se deben revisar los antecedentes de los que tienen acceso al contenido del test, que habrán de aceptar acuerdos estrictos de no divulgación.

    d. Los ítems y tests remitidos para su revisión a otros servidores, y temporalmente fuera de control directo, deben ser borrados de esos servidores o destruidos inmediatamente tras la revisión. Esa eliminación o destrucción debe ser verificada.

e. Se ha de establecer quién es el propietario de los ítems (por ejemplo, estableciendo su copyright) según las normas y políticas específicas del país.

f. Las personas involucradas en el proceso del desarrollo del test deben ser entrenadas en cómo reconocer e informar de los fallos en el sistema de seguridad.

14. Las pruebas deben ser protegidas durante las fases de publicación y distribución.

    a. Los servidores utilizados para almacenar el contenido de las pruebas se deben alojar en un centro de datos profesional certificado con los estándares internacionales (por ejemplo, ISO 27001 o SSAE 16) que aplique las medidas de seguridad de las tecnologías de la información y comunicación, como cortafuegos y detección de intrusos.

    b. Las personas que desarrollen los tests deben ser personas de confianza y han de suscribir acuerdos de no divulgación.

    c. El envío del test al centro en el que va a ser aplicado y su almacenamiento, ya sea en forma de cuadernillo o como archivo digital, deben hacerse de forma segura. Los proveedores del servicio de aplicación informatizada de tests deben mantenerse al tanto de la aparición de parches de seguridad del sistema operativo y software en uso, y aplicarlos con prontitud.

    d. El contenido digital debe ser protegido mediante métodos estrictos de encriptado, tanto si el contenido del test se envía en su totalidad a un servidor remoto para ser descargado, como si es enviado ítem a ítem, en tiempo real, en un test aplicado por Internet.

    e. El contenido del test que resida durante un periodo de tiempo determinado en un servidor de un centro de aplicación de tests debe ser protegido en todo momento mediante estrictos controles de acceso a usuarios (por ejemplo, nombres de usuario y contraseñas) y métodos estrictos de encriptado.

    f. Las pruebas deben permanecer en los centros de aplicación de tests el menor tiempo posible, según lo establecido en las políticas del programa y de la aplicación del test.

    g. Cuando un test ya no vaya a hacer falta en un centro de aplicación, su contenido debe ser borrado y/o destruido. La desaparición o destrucción debe ser verificada y el contenido no se debiera poder recuperar.

    h. Los creadores del test deben garantizar que quede claramente documentada la distribución de todos los materiales sensibles, de manera que pueda hacerse un seguimiento completo que incluya la devolución y/o destrucción de los materiales, si fuera el caso, cuando hayan sido utilizados. Su devolución o destrucción debe ser verificada.

15. Deben utilizarse métodos de seguimiento para registrar los períodos de control, acceso y cambios en los archivos, por ejemplo, mediante sistemas de registro en papel o digitales.

16. Los evaluados deben entender las reglas de seguridad y las consecuencias de su violación antes de que se registren y se programe la aplicación del test.

    a. Los evaluados deben conocer, con suficiente antelación a la aplicación del test, como parte de un acuerdo de código de honor o ético, que habrán de leer y ser conscientes de las reglas de seguridad y estar de acuerdo en respetarlas.

    b. Las consecuencias del incumplimiento de las reglas de seguridad deben ser claras.

    c. Los evaluados deben tener la oportunidad de expresar si están o no de acuerdo con esas reglas antes de la aplicación del test. No se debe permitir que respondan al test los evaluados que no estén de acuerdo con ellas.

    d. También se debe proporcionar y explicar una documentación relativa a los derechos del evaluado.

17. Siempre que lo permitan las leyes vigentes, los evaluados deben ser adecuadamente autenticados[1] , lo que puede hacerse antes, durante o después de haber respondido al test. Entre los procedimientos aceptables de autenticación se encuentran la presentación de un documento de identificación oficial con foto o el uso de dispositivos biométricos que permitan la lectura de la huella dactilar o de la palma de la mano, el escaneo del iris, analizar la dinámica de las pulsaciones en el teclado, o el reconocimiento facial.

18. El riesgo para la seguridad es máximo durante la aplicación del test y después de la autenticación. Es en estas fases cuando los ítems presentados pueden ser robados y se puede hacer trampas. Además de las medidas de seguridad acordadas durante las etapas de planificación y diseño, puede ser necesario un esfuerzo adicional para garantizar, en la medida de lo posible, que la prueba no sea robada y que se minimice[2] la probabilidad de que se haga trampas.

    a. En las aplicaciones informatizadas, el sistema de aplicación del test debe utilizar un programa de bloqueo o navegador seguro que impida el acceso a recursos externos que no sean necesarios para responder al test.

    b. Los supervisores podrían iniciar la aplicación del test utilizando una "clave" especial suministrada por el sistema de aplicación de la prueba. Una clave similar podría proporcionarse al evaluado, siendo necesarias ambas claves para que el test comience.

---

[1] La autenticación no es lo mismo que la identificación. En exámenes o evaluaciones que puedan conllevar consecuencias importantes para el evaluado no es necesario identificar realmente a la persona. Es suficiente con asegurarse de que la persona que va a hacer la prueba es la misma que se registró y se inscribió en el programa.

[2] Es obvio que es posible hacer trampas y que habrá evaluados que las hagan incluso cuando se hayan implementado las medidas de seguridad máximamente efectivas. El objetivo de un programa de seguridad es minimizar los efectos del robo y reducir a niveles razonables los casos en los que se hace trampas.

c. Los supervisores deben vigilar atentamente a los evaluados sin distraerles. Si lo permite la legislación vigente, la supervisión se puede hacer a distancia (mediante cámaras web), o localmente (en el lugar donde se aplica el test y/o con circuito cerrado de televisión).

d. Los supervisores no deben tener la posibilidad, o tenerla limitada, de ver la pantalla o las páginas del cuadernillo de la persona evaluada mientras está haciendo el test.

e. Los supervisores deben conocer los métodos más comunes de hacer trampas y robo del contenido del test, y deben saber qué hacer si se produce un fallo de seguridad y cómo elaborar el correspondiente informe cuando se produce un suceso irregular en la aplicación del test.

f. Los supervisores deben estar suficientemente motivados para estar atentos a los posibles problemas de seguridad y, si es necesario, enfrentarse a un evaluado cuando se produce una sospecha de fallo en el sistema de seguridad.

g. Los supervisores no deberían tener un interés especial en las puntuaciones que el test asigne. No deben ser instructores o profesores de los evaluados ni estar familiarizados con el contenido del test.

h. Si lo permiten las leyes vigentes, debe haber cámaras que faciliten la supervisión, grabación y registro de lo ocurrido durante la aplicación del test, así como de cualquier incidente de seguridad.

i. Si es posible y lo permiten las leyes vigentes, los dispositivos que pueden ayudar a hacer trampas o robar el contenido del test, como teléfonos móviles inteligentes, tabletas, cámaras, etc., deben recogerse antes del inicio del test y ser devueltos cuando acabe la aplicación.

j. La posibilidad de que haya descansos debe gestionarse con cuidado. Después de un descanso, los evaluados no deberían poder revisar las respuestas a las preguntas que contestaron antes del mismo.

k. Las hojas de papel disponibles o utilizadas durante la realización del test deben recogerse al acabar y ser tratadas de acuerdo con las políticas del programa.

l. Si durante la aplicación de un test se observa que alguien está haciendo trampas o grabando su contenido, se debe responder con rapidez y eficacia, de acuerdo con la orientación específica facilitada por el programa de evaluación. La respuesta a dar pudiera ser la suspensión temporal o permanente de una sesión del test, la confiscación de los equipos o materiales utilizados, y la elaboración de un informe oficial sobre la irregularidad de seguridad observada.

19. En los tests informatizados, cuando los resultados se han de obtener de servidores remotos, la transferencia de datos debe hacerse en cuanto acabe el test o tras cada ítem, en el caso de tests online aplicados por Internet. Los datos deben ser protegidos con procedimientos estrictos de control de acceso, mientras residan en un servidor remoto, y con una encriptación estricta durante la transmisión.

20. Los tests y los ítems deben ser evaluados periódicamente para descubrir si hay indicios de que se haya hecho trampas o de que haya ítems filtrados. El rendimiento en los ítems y en las pruebas será diferente si los ítems han sido robados y conocidos por los evaluados antes de la realización del test, y si se ha hecho trampas. Se muestran a continuación algunos ejemplos:

    a. La presencia de patrones de respuesta inusuales, como responder a preguntas sencillas de forma incorrecta y correctamente a preguntas difíciles, puede indicar que se ha hecho trampas o que ha habido robo.

    b. Los tiempos de respuesta inusualmente largos o cortos, tanto en el test como en los ítems, pueden indicar un fallo de seguridad o algún otro problema.

    c. Demasiados borrones en las hojas de respuestas, en particular las tachaduras que convierten fallos en aciertos, pueden indicar que habido manipulación de las hojas de respuesta o que alguien ha filtrado al evaluado las opciones que son correctas.

    d. Respuestas inusualmente parecidas entre pares o grupos de evaluados puede indicar que ha habido colusión o suplantación (se han compinchado algunos evaluados).

    e. Los patrones de respuesta y latencias que son similares en parejas o grupos de evaluados pueden indicar que se ha habido colusión o suplantación, o que alguien ha filtrado a los evaluados las respuestas correctas.

    f. Mejoras inusuales en los resultados de una sesión del test a otra, tanto para grupos de evaluados o para un individuo, pueden indicar que se ha hecho trampas.

    g. Cambios inusuales en el funcionamiento de un ítem (por ejemplo, en sus parámetros estadísticos) pueden indicar que ha sido filtrado. Estos ítems deben ser reemplazados inmediatamente.

    h. Un funcionamiento diferencial de unos tipos de ítems frente a otros puede indicar que ha habido conocimiento previo del contenido del test. Por ejemplo, el rendimiento en ítems tipo “Caballo de Troya”, que se puntúan adrede con una clave incorrecta, o en ítems nuevos, y que por tanto han sido menos expuestos, puede indicar el uso de pre-conocimiento cuando se compara con el rendimiento de los ítems operativos del test.

    i. Si lo permiten las leyes vigentes, los datos demográficos de la persona evaluada también se pueden analizar para detectar posibles fraudes (por ejemplo, suplantación). Por ejemplo, que un evaluado haga el test varias veces en países diferentes de su país de residencia, con intervalos cortos de tiempo entre aplicaciones, puede indicar que ha habido suplantación o colusión.

    j. Cuando las sesiones de aplicación del test tienen un horario regular, la comprobación de las horas de inicio y finalización del test nos informaría de si se ha aplicado durante el horario regular. Los tests aplicados fuera del horario normal pueden indicar que se ha intentado hacer trampas o robo organizado de ítems.

21. El software necesario para la creación y aplicación de exámenes o con fines de gestión del programa debe ser desarrollado usando procedimientos seguros que lo protejan contra las vulnerabilidades comunes de programación y debe ser evaluado periódicamente (por ejemplo, mediante pruebas de penetración de terceros).

22. Los tests informatizados por lo general son puntuados inmediatamente después de que el test haya acabado o mientras se responde al test, después cada respuesta, como ocurre en los tests adaptativos informatizados. Las amenazas y los riesgos de que se haga trampas durante el proceso de puntuación para este tipo de pruebas son mínimos. En los tests de lápiz y papel el proceso de puntuación es largo y consta de varias etapas, y se requieren más medidas de seguridad para evitar que las puntuaciones puedan alterarse.

    a. Las puntuaciones se pueden proporcionar con carácter provisional y se confirmarían solo después de que se haya determinado su validez. En este sentido, se puede establecer que las puntuaciones no sean distribuidas o hechas oficiales hasta que se hayan revisado los informes sobre irregularidades y se haya completado el análisis forense de los datos.

    b. La obtención de las puntuaciones de las pruebas informatizadas debe hacerse en servidores remotos bien protegidos y no en el ordenador de la persona evaluada. En los tests en papel, las hojas de respuestas pudieran ser modificadas cuando van a ser escaneadas o puntuadas. Se debe implementar un proceso de supervisión para seguir de cerca y proteger las hojas de respuestas hasta que sean procesadas para la obtención de las puntuaciones.

23. Los tests, ítems, resultados de los tests y otra información importante (por ejemplo, la información demográfica de los evaluados), tanto en las pruebas en papel como en las informatizadas, a menudo se almacenan durante largos periodos de tiempo, a veces durante varios años. Independientemente de dónde o cuándo se hayan producido y recogido los datos, deben establecerse procedimientos que aseguren que es extremadamente difícil el acceso no autorizado a esta información y que no se puede acceder a las puntuaciones y otros datos, modificarlos o suprimirlos sin la debida autorización. Los procedimientos y sistemas que utilicen las tecnologías de la información y comunicación deben ser auditados y actualizados periódicamente.

24. Antes, durante y después del periodo en el que la prueba puede aplicarse, un programa debe iniciar un proceso de supervisión para comprobar si se ha hecho público su contenido en Internet. En dicha supervisión se puede encontrar a individuos hablando informalmente de una prueba o de algunas de las preguntas, o la reproducción exacta de una pregunta o de todo un conjunto de preguntas. Cuando se descubra algo así, el programa debe enviar una solicitud al administrador del sitio web para que interrumpa la charla, advierta a los participantes, y elimine el contenido. Se deben considerar acciones más duras, incluso acciones legales, si el material no se retira rápidamente.

25. Antes, durante y después del periodo de aplicación del test, un programa debe evitar que personas no autorizadas puedan acceder al contenido de un test.

Parte 3: Respuesta ante un fallo en el sistema de seguridad

De vez en cuando las defensas del programa no consiguen controlar adecuadamente una amenaza y se produce un fallo de seguridad, como, por ejemplo, cuando alguien hace trampas o se produce robo del contenido del test. Las siguientes directrices proporcionan asesoramiento y apoyo para hacer frente a estos hechos. Cuando se descubre que se ha hecho trampas o que el test o los ítems han sido robados, el programa de evaluación tiene la responsabilidad de investigar a fondo, evitar que aumente la importancia del fallo de seguridad, reparar los daños y tomar otras medidas que se consideren adecuadas. Deben adoptarse medidas para evitar que el fallo de seguridad vuelva a ocurrir en el futuro.

El comité de seguridad debe tener toda la responsabilidad en la respuesta a un fallo de seguridad y disponer de los poderes necesarios para la toma de decisiones.

Un programa tendrá conocimiento de un fallo de seguridad, posible o real, de diversas maneras, que difieren en lo fáciles de tratar y útiles que resultan. Se exponen algunas a continuación:

- por un periodista u otro medio de comunicación
- por el informe de irregularidad elaborado por un supervisor
- por un aviso confidencial
- a partir de los informes del análisis forense de datos
- mediante los informes de la supervisión de la web
- a partir de los "sistemas" de seguridad automáticos implementados que pueden detectar, por ejemplo, los intentos de pirateo o las pulsaciones de teclas inapropiadas durante la realización del test.

Independientemente de qué cause el fallo de seguridad, el programa de evaluación tiene que actuar con rapidez para confirmar la validez de la información y el alcance del fallo. A continuación, deberá tomar las medidas adecuadas. Cuando se está aplicando un test, el sistema de supervisión o vigilancia debe responder inmediatamente cuando se está produciendo un fallo de seguridad. Antes y después de la aplicación del test, el comité de seguridad es responsable de revisar los detalles del fallo de seguridad y responder en consecuencia.

Directrices específicas sobre cómo responder ante un fallo en el sistema de seguridad

1. La aplicación del test a una persona debe ser interrumpida provisional o definitivamente si los supervisores (quienes vigilan a distancia online o quienes lo hacen donde se aplica el test) o los aplicadores de la prueba observan que hace trampas o está robando su contenido. Se debe proporcionar al evaluado una explicación de por qué se ha interrumpido la aplicación.

   a. Después de interrogar a un evaluado sospechoso de haber hecho trampas o de incumplimiento de las reglas establecidas en la sesión de evaluación, un supervisor

o aplicador del test puede permitir que la aplicación del test continúe o puede decidir que siga parada o sea interrumpida definitivamente.

b. Si lo permite la legislación vigente, los instrumentos no apropiados (cámaras, etc.) deben ser confiscados. Se debe explicar al evaluado las razones por las que se procede así.

c. El supervisor, vigilante o aplicador de la prueba debe elaborar un informe de la irregularidad observada y remitirlo a la comisión de seguridad del programa de evaluación.

d. Si se permite la continuación del test, se deberá elaborar un informe de irregularidad que habrá de ser revisado por el comité de seguridad del programa de evaluación.

2. Un fallo en el sistema de seguridad debe ser investigado a fondo para determinar su extensión y la magnitud del daño. Las investigaciones pueden incluir entrevistas con las personas sospechosas de estar involucradas en el fallo de seguridad o con testigos, análisis forenses de datos para comprobar los efectos en las puntuaciones, y/o la supervisión de la web para comprobar la cantidad de contenido de la prueba que se ha filtrado.

3. Un test o conjunto de ítems que hayan sido filtrados debe ser reemplazado lo más rápidamente posible.

   a. Algunos programas de evaluación utilizan un test creado anteriormente, conocido como "test fallo de seguridad", cuando han de reemplazar un test o conjunto de ítems robado.

4. Las puntuaciones que se sepa que son incorrectas, por haberse cometido algún tipo de fraude en el test, han de anularse.

   a. Este proceso se simplifica si se implementa la política de revisar las puntuaciones habitualmente y considerarlas "provisionales" hasta que sean confirmadas.

   b. Si las puntuaciones no válidas ya han sido entregadas a los evaluados, se les deberá comunicar inmediatamente que sus puntuaciones no son válidas y que todas las decisiones basadas en dichas puntuaciones habrán de ser revisadas.

5. En ocasiones conviene volver a obtener la puntuación en un test que ha sido filtrado, si al proceder de este modo se consigue que la toma de decisiones se base en puntuaciones más correctas, como en las que no intervienen los ítems que han sido filtrados.

6. En ocasiones se habrán de tomar medidas adicionales como acciones legales, civiles o penales, en función de las políticas establecidas en el programa de evaluación y de la importancia de los daños causados por el fallo en el sistema de seguridad.

7. A las páginas web que venden sin permiso preguntas con derechos de autor se les debe exigir que retiren dichos contenidos de la página web y de todas las posibles

ubicaciones. Se podrían tomar además otras medidas de mayor alcance. La página web debe ser estrechamente supervisada para comprobar que los contenidos han sido retirados.

a. Un responsable del programa podría enviar al administrador de la página web cartas de testigos del intento de venta para informar del problema y solicitar que se elimine el material. Este proceder ha demostrado ser un primer paso efectivo, pues la mayoría de los administradores responden positiva y rápidamente.

b. Si los contenidos no son retirados, se puede enviar una carta más formal de vulneración de derechos para exigir al administrador de la página web la eliminación de los contenidos, si no quiere enfrentarse a acciones legales. Estos avisos pueden citar estatutos regionales o nacionales relevantes (por ejemplo, el Acta sobre el Copyright del Milenio Digital de USA o la Directiva Europea sobre el Copyright).

8. En función de las políticas y la gravedad de la infracción, se puede solicitar y autorizar a los evaluados que respondan de nuevo a la misma u otra forma del test. Si lo permiten las leyes vigentes, se podría denegar la repetición del test a los evaluados que han hecho trampas o están involucrados en el robo del test.

   a. Los documentos que describen la política del programa deben establecer claramente las reglas a aplicar en relación a la repetición del test, según lo acordado por todas las partes interesadas, incluidos los evaluados.

   b. Si fuera posible, debiera utilizarse un test distinto para la repetición del test.

   c. Se puede requerir el cumplimiento de condiciones adicionales, como un pago o un período de espera, para que el evaluado pueda repetir el test.

9. Un fallo de seguridad puede despertar gran interés en las partes directamente implicadas y las que no lo están directamente, como los medios de comunicación o el público en general. Conviene elaborar comunicados y distribuirlos cuanto antes. Una empresa de relaciones públicas y/o portavoz pueden ser útiles. Conviene tener preparados borradores de comunicados que anuncien que se ha producido un fallo de seguridad, su gravedad y las medidas adoptadas para hacerle frente.

10. Tras un fallo de seguridad, puede ser necesaria una revisión del plan de seguridad existente para decidir si conviene cambiar las políticas y procedimientos de protección existentes y añadir otras nuevas. Estos cambios, si los hay, deben ser comunicados a todas las partes interesadas, y habrá de prepararse y aprobarse una nueva revisión del plan de seguridad.

TÉRMINOS Y DEFINICIONES

Amenaza. Un individuo o método que puede ocasionar que se haga trampas en una evaluación o que se robe con éxito el contenido del test.

Análisis de la latencia. Análisis forense de las latencias de respuesta. La latencia es el tiempo transcurrido desde que el contenido de un ítem se muestra al evaluado hasta que el ítem ha sido contestado y la respuesta enviada. Latencias inusualmente cortas o largas pueden indicar que se ha hecho trampas o algún otro problema de seguridad.

Análisis de riesgos. Análisis de las diversas amenazas de seguridad que tiene un programa con el fin de estimar la probabilidad de cada riesgo y su daño potencial, y distribuir los recursos de seguridad de manera adecuada.

Análisis forense de datos. Métodos que analizan los resultados de un test para detectar patrones que podrían sugerir que se ha hecho trampas o ha habido robo de su contenido.

Autenticación. Proceso para determinar que la persona que hace un test, lo hizo o se está preparando para hacerlo, es la persona que debe hacerlo. La autenticación no es el mismo proceso que la identificación, cuyo fin es conocer la identidad del evaluado.

Ayuda ilícita (coaching). Un individuo proporciona al evaluado las respuestas correctas, pistas, etc., durante la realización del test.

Ayudas del test (test aids). Dispositivos o documentos utilizados por un evaluado mientras responde al test. El uso de ciertas ayudas (por ejemplo, calculadoras) puede estar permitido.

Banco de ítems. Conjunto de ítems del que se extraen los que integran el test, antes o durante su aplicación.

Bloqueo. Programa de ordenador que se pone en marcha cuando comienza la aplicación de un test informatizado y que limita las funciones del teclado y del ordenador exclusivamente a las necesarias para navegar y responder a las preguntas del test. Con el bloqueo se prohíbe el acceso a otros recursos, tales como el disco duro e Internet, y se inhabilitan ciertas combinaciones de teclas.

Colusión. Personas que se ponen de acuerdo para hacer trampas en un examen o robar su contenido.

Descanso. Tiempo de descanso entre las secciones de un examen largo.

Dinámica de las pulsaciones en el teclado. Método biométrico que compara los patrones de escritura en el teclado de un evaluado cuando se registró en el programa y justo antes de comenzar el test.

Evaluaciones con consecuencias importantes. Los resultados obtenidos en estas evaluaciones tienen consecuencias importantes para un individuo u organización.

Evaluado. Individuo que responde a un test.

Exposición de un ítem. La exposición a los evaluados de un ítem del test, o la exposición del ítem tras haber sido robado y compartido a través de Internet o por alguna otra vía.

Fallo de seguridad. Un ataque realizado por amenazas conocidas o desconocidas que consigue romper las defensas del programa de evaluación.

Filtrado. Un ítem o test se dice que ha sido filtrado cuando se comprueba que se ha hecho público, habiendo dejado de ser apropiado su uso en el programa de evaluación.

Hacer trampas. Cualquier comportamiento que intenta o logra mejorar una puntuación en el test de manera inapropiada.

Identificación. El proceso de identificar realmente a la persona que hace o va a hacer el test. La identificación no es el mismo proceso que la autenticación, que, en general, no pretende identificar al evaluado.

Informes de irregularidad. Informes, elaborados por los vigilantes u otras personas, que describen que se ha hecho trampas u otro hecho inusual acaecido durante la aplicación del test.

Investigación. Proceso por el que se determinan las causas y el alcance de un fallo de seguridad. Las investigaciones pueden incluir entrevistas, análisis forense de datos, análisis de los procedimientos, la revisión de los informes, etc.

Ítems insertados. Son ítems no puntuables que se añaden al test intencionadamente para detectar individuos que han tenido acceso a los ítems del test que se les va a aplicar.

Ítems tipo "Caballo de Troya". Ítems insertados en un examen a los que intencionadamente se aplica una clave de corrección incorrecta. El propósito de estos ítems es detectar al evaluado que use ítems y clave de corrección robados para hacer trampas en el examen.

Manipulación de las hojas de respuestas. Forma de hacer trampas en la que las respuestas incorrectas de una hoja de respuestas son sustituidas por las correctas.

Mejoras en la puntuación. Análisis forense de datos sobre las mejoras (o disminuciones) en las puntuaciones para detectar cambios inusuales que pueden indicar que se ha hecho trampas.

Métodos biométricos. Métodos para la recogida de información corporal específica del evaluado que permite su autenticación o identificación.

Plan de seguridad. Un documento que describe las políticas y procedimientos de seguridad de una organización.

Presentación de ítems solo hacia adelante. Los ítems se presentan en el test de forma que no sea posible volver a ver los ya presentados.

Puntuaciones provisionales. Puntuaciones no oficiales proporcionadas a los evaluados una vez hecho el examen, que han de ser revisadas por un comité de seguridad.

Re-aplicación. Proceso por el que se permite a una persona volver a hacer el examen.

Reconocimiento facial. Método biométrico que utiliza las imágenes de la cámara web para comparar los rasgos faciales de un evaluado justo antes de comenzar el test con los obtenidos cuando hizo el registro.

Rendimiento diferencial del ítem. Análisis forense de datos sobre el rendimiento de un ítem en situaciones diferentes (por ejemplo, cuando el test se aplicó por primera vez y transcurridos seis meses) que podría indicar que el ítem ha sido filtrado.

Repetición de un examen. El mismo examen o una forma equivalente se vuelve aplicar a la persona a evaluar.

Repetición de la puntuación. Proceso por el que se vuelve a obtener la puntuación en un test, tras eliminar por ejemplo la influencia en la puntuación de los ítems filtrados.

Revisión de antecedentes. Proceso en el que se revisa el historial de un individuo y se determina si está cualificado o no para colaborar en la creación de un test.

Riesgo. Estimación de la probabilidad de que una amenaza se convierta de hecho en un fallo de seguridad y de la cantidad de daño que el fallo podría causar.

Robo del test. Cualquier comportamiento que intenta o consigue obtener el contenido del test de forma ilegal.

Robo organizado de ítems. Los intentos, con éxito o no, de conseguir el contenido del test de forma ilegal y en contra de las reglas de seguridad del programa.

Similitud de las respuestas. Comparación forense de datos entre los patrones de respuesta de dos o más personas con el fin de detectar colusión, suplantación o ayuda ilícita.

Supervisión de la web. Conjunto de métodos para buscar en Internet preguntas filtradas de exámenes y tests.

Supervisor. Persona responsable de la seguridad de un test durante su aplicación. Se le llama también vigilante.

Suplantador. Persona que hace el test que debiera hacer otra persona.

Tachaduras. Respuestas de una hoja de respuestas que han sido borradas.

Test de verificación. Test aplicado en un momento posterior para verificar el rendimiento de un evaluado en un test aplicado previamente.

Test "fallo de seguridad". Un test alternativo que se utiliza para reemplazar los ítems o un test filtrado.

Vigilante. La persona responsable de la seguridad de un test durante su aplicación. Se le llama también supervisor.

Vulnerabilidad. Un punto débil en la defensa de la seguridad de un programa de evaluación.

REFERENCIAS

Foster, D. F. & Miller, H. L., Jr. (2012). Global Test Security Issues and Ethical Challenges. In A. Ferrero, Y. Korkut, M. M. Leach, G. Lindsay, & M. J. Stevens (Eds.). *The Oxford Handbook of International Psychological Ethics* (pp. 216-232). Oxford: Oxford University Press.