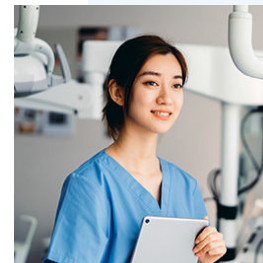




Tell me more about...

Data forensics for online testing



Contents

02

What is a data forensics program?

Page 10



04

Statistical indices

Page 18



01

Introduction

Page 4



03

Data forensics and your test security plan

Page 14

06

How does web crawling improve test security?

Page 26



08

A case study

Page 34



05

Data trends

Page 24



07

In conclusion...

Page 32



Introduction

Security measures have always been needed to protect the integrity of assessments, especially when they are used for high stakes purposes, such as educational qualifications, credentialing, and public safety employee selection. In these applications, test security is needed to guard against misconduct – whether testing takes place in-person or online.

Similar to changes in education and the workplace, testing has changed, with an increasing number of tests conducted online and taken remotely in a test taker's office or own home. In response to this trend, test security has become even more important. And many of the test security risks which may be amplified by online testing are now addressed by technological advances and innovative security measures.

One element of this enhanced security stems from the data that arises from computer-based testing, which presents an enormous opportunity for forensic analyzes. When people interact with tests and test questions in abnormal and fraudulent ways, this may be reflected in the data as irregularities in their answer choices, score patterns, and response times.

We can use this wealth of data to identify suspicious test taking behavior that may not be observed by test proctors – and take action – to detect misconduct and increase test security, for the benefit of testing organizations and test takers alike.





What is *data forensics*?



Data forensics, sometimes known as computer forensics, refers to the use of statistical methods to study or investigate digital data to detect anomalies. It is often used in incidents of financial fraud, data theft and other cyber security crimes.



How is data forensics used in *online testing*?



When it comes to online testing, the application of statistical detection methods helps us to recognize anomalies in testing data consistent with potential misconduct. Identifying these suspicious patterns helps us detect and investigate misconduct early.



**Proxy
testing**



**Item
harvesting**



**What are the
security risks?**

Collusion



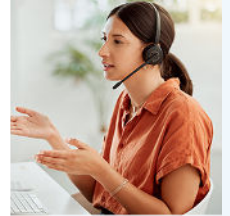
**Item pre-
knowledge**





"**Every** contact leaves a trace."

Locard's Principle of Forensic Science







"Locard was talking about physical evidence at a crime scene but we keep the same principle in mind when designing our statistical detection methods – we ask 'if a person was engaging in this specific type of fraudulent behavior, what traces might be left behind in their data patterns?' Then we design ways to detect those patterns."

Greg Hurtz, PhD

Professor of Psychology, California State University, Sacramento and Senior Research Scientist, PSI Services

What is a data forensics program?

Start off in the right way by making it easy for test takers to find what they need on their very first visit to your website. Are your testing program details obvious? Is essential information self-service with FAQs or videos? If they cannot find what they are looking for is it clear what to do?

A multi-faceted approach to data forensics that uses different tools is the most effective. For example, a program that incorporates different statistical indices and analytics, as well as web crawling to find test content on the internet.



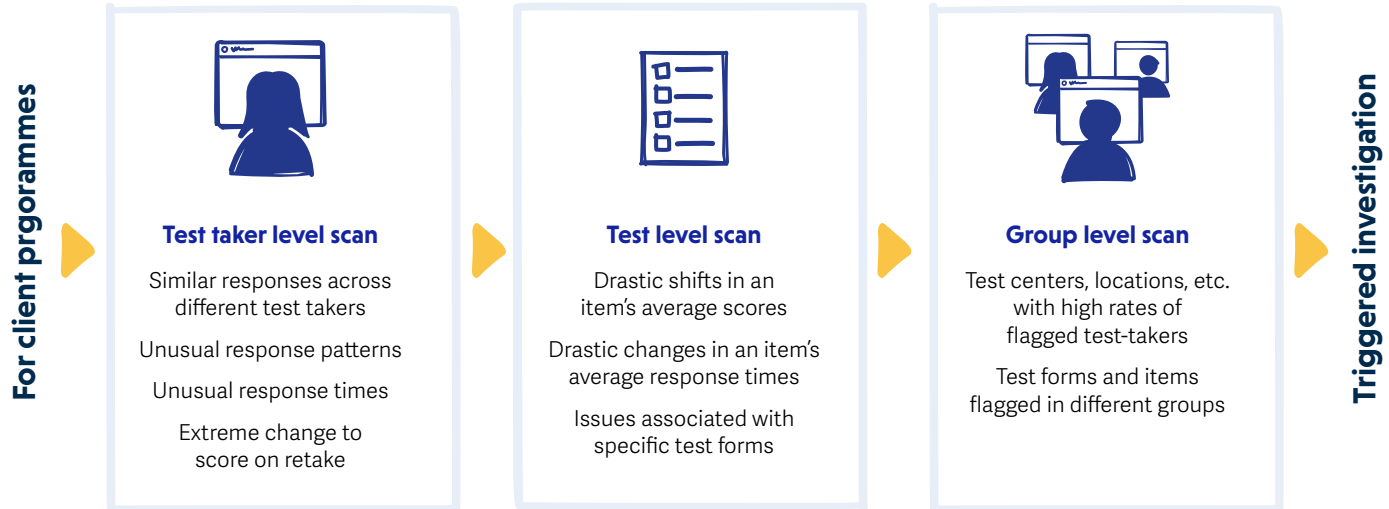
This involves analyzes across three different levels:

- ▶ **1. Test taker level** – to look for abnormal, irregular, or responses that are highly similar to other test takers.
- ▶ **2. Test level** – to analyze results and responses to a particular test, form, or item.
- ▶ **3. Group level** – to scrutinize a specific geographical area, test site, or other group defined by the testing organization.

Analysis takes place on a daily, monthly, or quarterly basis, depending on test taker volume and the security requirements of the testing program and organization.



What does a data forensics program *look for*?





"An effective data forensics program is a high-level partnership between the client and the provider. One that proactively looks for trends related to cheating, while applying expertise and insights relevant to the specific testing organization and industry."

Nicole Tucker

Director of Statistical Reporting and Analytics, PSI Services

Data forensics and *your test security plan*

An effective test security plan covers the whole test life cycle – from prevention during test development and delivery, to detection should any misconduct take place.



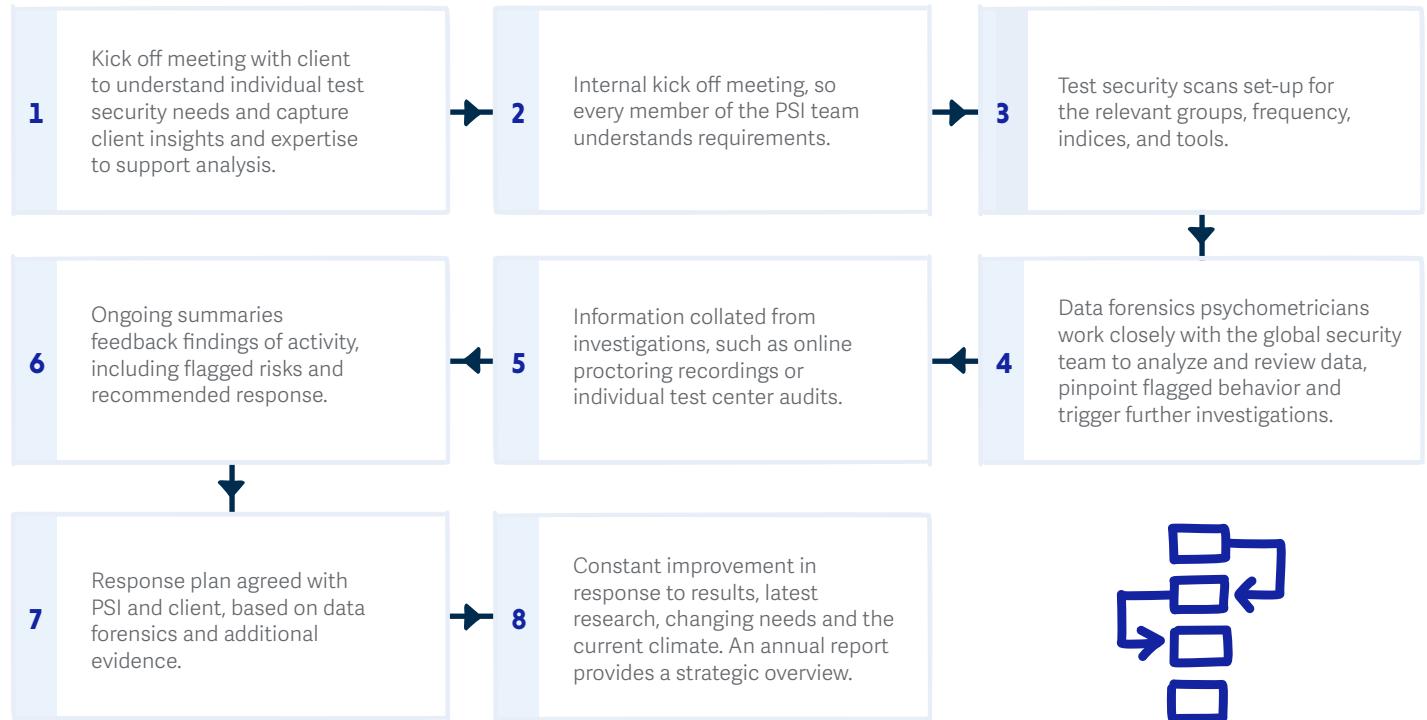
For example, a non-disclosure agreement with your subject matter experts, along with secure storage and transfer of test content, will increase security during test development. And tools such as online proctoring and the use of alternate or unique test forms will prevent misconduct during test delivery.

While all these measures serve to **prevent** misconduct, a comprehensive test security plan will also include steps to **detect** misconduct. This is where data forensics, and operational tools such as web crawling, are invaluable.



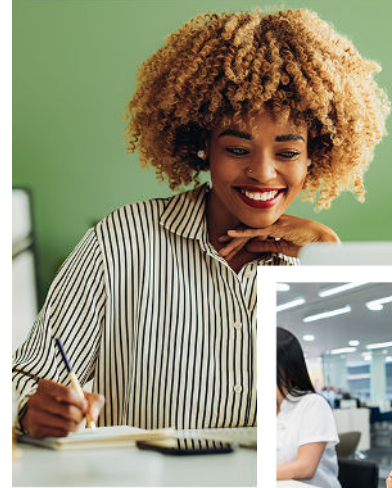
A customized approach *to data forensics*

The best test security plans are tailored to the unique requirements of the testing organization.





"Preventing test misconduct is crucial. But in the event of failure to prevent it, detecting its presence and impact on scores is equally important."



Statistical *indices*

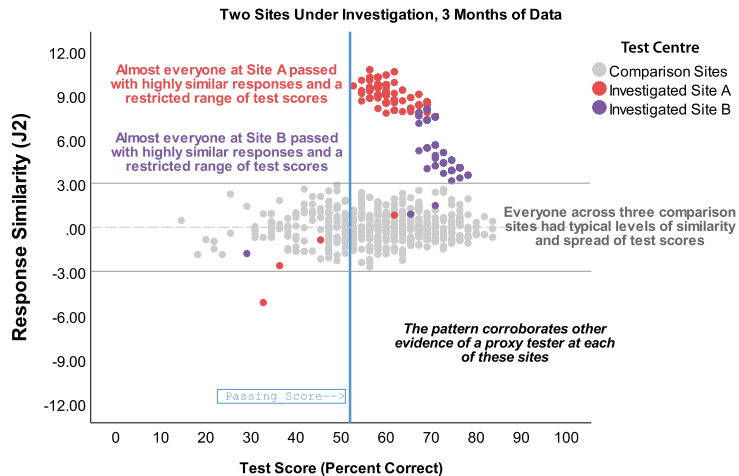
An index is a summary measure of a group of data points, reflecting a trend, effect, or feature. Familiar examples include the Gender Pay Gap or the Dow Jones stock market index. For online testing, patterns in the data can reveal evidence of potential misconduct. For example, a group of test takers who all give similar responses to test questions and pass a test, or who complete the test successfully in an improbably fast time.

Different indices are better suited to different scenarios – the best index depends upon what you are looking for. When used to examine test data, a combination of indices will be even more effective at identifying the presence of misconduct and its impact on test scores.

If you aren't a statistician, the results of data analyzes and indices can be difficult to interpret. Visualization techniques graphically and clearly present any anomalies consistent with misconduct, and they also assist with early detection.



Different types of statistical index and what they tell us

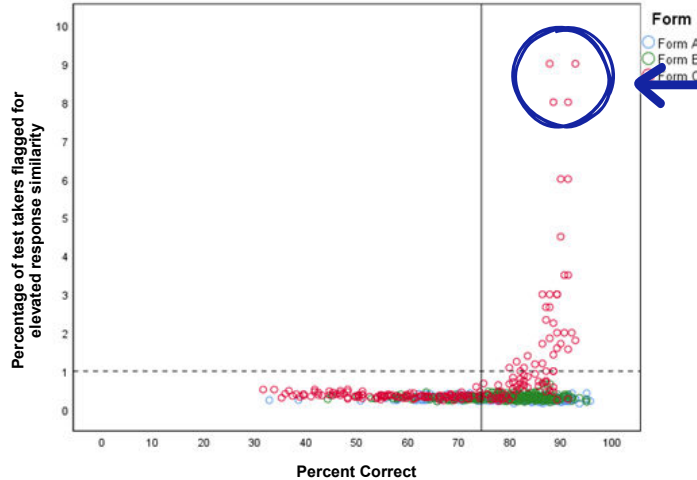


Response Similarity

Measures the degree of similarity between pairs of test takers in their responses, for example with proxy testing or access to the same pirated content.



"PSI has developed our own proprietary index of similarity that has proven to be successful in detecting certain types of cheating."



These pairs of test takers made almost identical errors. Many pairs of test-takers on one specific test form made a high number of identical errors, whereas this pattern did not occur on the other forms. This strongly suggests that one form was compromised.

Errors in common

Measures the degree of similarity between pairs of test takers for items they both answered incorrectly, for example with proxy testing or access to the same pirated content.

- ▶ **Item and test time** – flags unusually fast response times for a whole test or specific item.
- ▶ **Abnormal score patterns** – flags higher than average scores at a group level.
- ▶ **Pass rate elevations** – flags unusually high pass rates at a group level.





We also use *item level analysis*

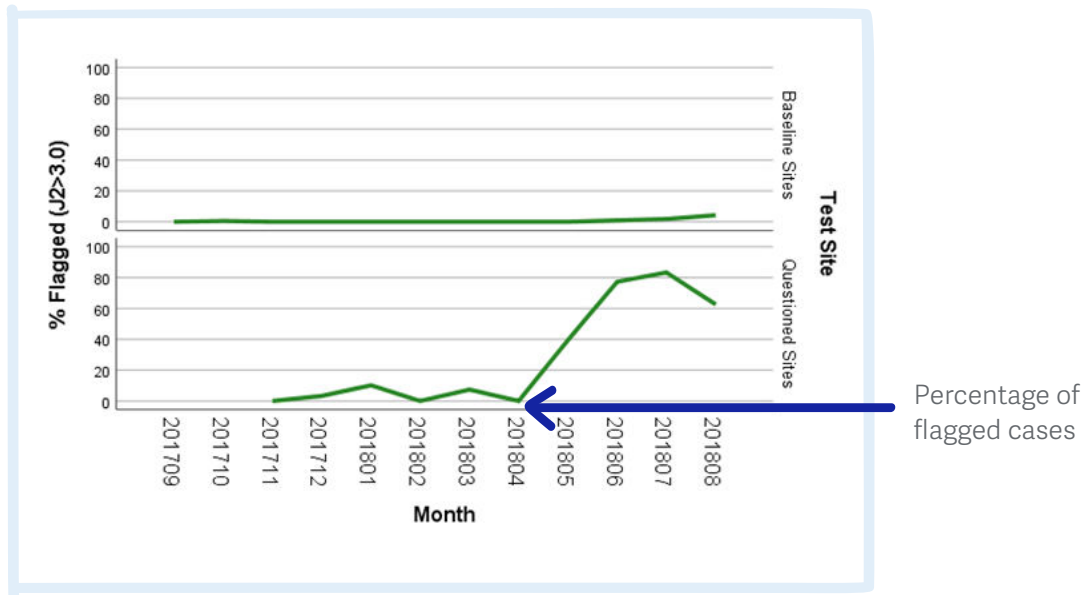
When groups of test takers have been flagged for potential misconduct, data forensics can show whether a certain subset of items is consistently involved. This indicates those items may have been exposed and need to be removed and replaced.

Data *trends*

By monitoring data over a period of time we can gauge when there is a change that might indicate misconduct. For example, analyzing the average time it takes to complete a test during a period of months can help us to identify if, and when, there is a change. This might be tests completed suspiciously quickly, indicating that test takers have pre-knowledge of the test content.



The graph below compares data trends between baseline and questioned test center sites. Time based analysis helps us to identify the point in time (shown by the arrow) when there was an increase in flagged behavior.

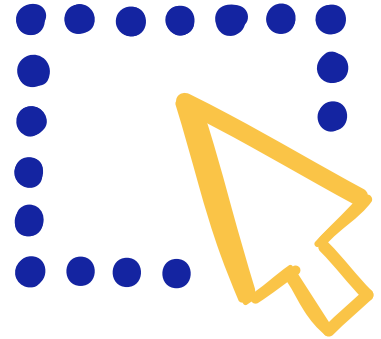


How does web crawling improve test security?

A web crawler, sometimes called a spider or a bot, is a computer program that systematically browses the internet searching for specific content. Web crawlers are used by search engines such as Google or Bing.

Web crawlers can be programd to automatically search the internet for references to proprietary test content. The crawler can find matching test content on websites, discussion boards or social media. This might include stolen items offered for sale, or past test takers who discuss the content on a social media site.

If copyrighted test content is found to be exposed, testing organizations can start the process of getting it removed from the internet, taking legal action if necessary. Organizations may also choose to remove or replace any test items a web crawler has shown to be exposed.





"Web crawlers take a proactive approach to test security. They are more accurate and efficient than the manual monitoring systems traditionally used by testing organizations. Web crawling is an important precautionary step for test content protection that complements other security steps."

John Weiner

Chief Science Officer, PSI Services



How does web crawling *increase* test security?

For test security, web crawling is the process of identifying potentially compromised test content posted on the internet. As a part of the web crawling service, PSI will acquire the suspicious materials and complete a matching procedure to confirm if materials are exact copies or highly similar to actual test content.

Web crawling service – *an overview*



- ▶ **1.** Identify key search terms from the active item bank.
- ▶ **2.** Schedule web crawl of suspicious websites and social media, including content hidden in PDF documents.
- ▶ **3.** If necessary, purchase suspected compromised content.
- ▶ **4.** Match suspected compromised content to actual test content using an index of item similarity.
- ▶ **5.** Notify client and internal teams and agree appropriate response.

Data forensics flags a potential issue.

What next?



Data forensics might detect evidence of suspicious behavior, but it doesn't stop here. There are a range of investigations that help to confirm whether misconduct has taken place. A test region, site, testing window or test taker associated with anomalous findings may trigger further in-depth analysis and investigation.

Investigative measures at test taker level, for example, include a review of session recordings, computer logs or registration data. A review of video or photos can also be used to investigate a test center, along with a site audit, drop-in inspections or secret shopper for tests administered at that site.

When data forensics and further investigations show misconduct has taken place, we work with a testing organization to formulate a response plan.

This might include deactivating a compromised test form or removing individual items from the bank, as well as legal measures to ensure test content is removed from websites or social media. In the most extreme cases of misconduct, test centers might be decommissioned, or test takers may be banned from future tests or have their credentials revoked.



"The more our clients work with us, the more confident they become in the results of our data forensics. With our combined expertise across test development, delivery, security, psychometrics and statistics, the PSI team brings a unique and highly customized approach to test security."

Nicole Tucker

Director of Statistical Reporting and Analytics, PSI Services

In conclusion...

Data forensics not only helps to confirm suspected issues with test security, it can also raise issues you aren't yet aware of. This helps you to make more informed, research-based decisions about your test security strategy and plan.

Data forensics is a convenient and cost-effective addition to your test security toolkit. It's also flexible and can be adapted to suit the needs of your testing program and organization.





At PSI we are the data forensics experts. Our psychometricians know testing, our statisticians know data, and our security team know the most effective ways to act.

We don't make it complicated. Our team helps you understand your data and make the right choices about what to do with it. We always support you to ensure your testing program is fair, and no test taker has an unfair advantage because of misconduct.



A case study

Data forensics not only helps to confirm suspected issues with test security, it can also raise issues you aren't yet aware of. This helps you to make more informed, research-based decisions about your test security strategy and plan.

Data forensics is a convenient and cost-effective addition to your test security toolkit. It's also flexible and can be adapted to suit the needs of your testing program and organization.





The challenge

A testing organization had suspicions that some educators might be involved in sharing stolen test questions with test takers. They approached PSI for assistance with data forensics analysis to support further investigations.

The approach

PSI's highly experienced Data Forensics Psychometricians completed an analysis of test taker responses to look for:

-  **1.** Excessively similar responses between different test-takers. This might indicate direct collusion/copying, prior access to test content or proxy testing by the same person.
-  **2.** Unusual patterns of correct item responses. For example, when entering memorized correct answers to difficult questions, entering memorized incorrect answers to easy questions, or entering answers randomly, as if not taking the test seriously.

In addition to test response data, PSI had access to information about which educator each test taker was associated with. This allowed us to pay special attention to high concentrations of test takers flagged by our analysis, by educator. Two educators were highlighted as having high percentages of flagged test takers and these results were shared with the client.

The results

The testing organization used this information to launch an investigation and test content was found in the possession of one of the flagged educators. It was found that the educator had supplied compromised materials to students as part of the study materials packet.





"Our client was extremely appreciative of the analysis we delivered, which helped to guide their investigative efforts and lead to a positive outcome."

Nicole Tucker

Director of Statistical Reporting and Analytics, PSI Services

References

- i Hurtz, G. M., & Weiner, J. A. (2019). Analysis of test-taker profiles across a suite of statistical indices for detecting the presence and impact of cheating. *Journal of Applied Testing Technology*, 20, 1-15.
- ii Weiner, J. A., & Hurtz, G. M. (2017). A comparative study of online remote proctored versus onsite proctored high-stakes exams. *Journal of Applied Testing Technology*, 18, 13-20.

References

- i Hurtz, G. M., & Weiner, J. A. (2019). Analysis of test-taker profiles across a suite of statistical indices for detecting the presence and impact of cheating. *Journal of Applied Testing Technology*, 20, 1-15.
- ii Weiner, J. A., & Hurtz, G. M. (2017). A comparative study of online remote proctored versus onsite proctored high-stakes exams. *Journal of Applied Testing Technology*, 18, 13-20.

Your trusted **testing partner**

Every day our clients support millions of people to realize their dreams, reach their potential, and improve their life chances. They care about their test takers – and we share that responsibility.

Our unwavering focus is on delivering frictionless and fair test taker experiences, without compromising test integrity, through...

Secure test delivery

- ▶ Authorized global **test center** network.
- ▶ Secure and scalable **remote testing** Live and Record & Review **online proctoring**.
- ▶ Flexible **multi-modal** test delivery.
- ▶ Testing **windows or continuous testing** on-demand.

Rigorous test development

- ▶ Legally defensible and **valid test content**.
- ▶ **Job analysis** and **test** content specification.
- ▶ **Subject Matter Expert** (SME) recruitment, training and management.
- ▶ Secure **item authoring, banking and test generation** software.

Expertise in testing science

- ▶ Experienced **psychometricians**.
- ▶ Specialist **test developers**.
- ▶ **Data forensics** and web crawling.

Our willingness to listen and adapt means clients can either benefit from a full testing service, or access solutions at any stage of their testing journey.

Dreams deserve

We understand every test is about more than the result. It's about a dream. A dream the test taker believes is worth striving for. And we believe that too. Their dreams deserve trusted science, technology and operational expertise. They deserve PSI.



Connect with an expert today.

psiexams.com 