

# THE TESTING INDUSTRY'S TOOLKIT TO COMBATING CHEATING ON TESTS

---

James A. Wollack, University of Wisconsin



# CHEATING 101

Cheating defined

Deterring cheating

Detecting cheating

Deciding how to handle cheating

# CHEATING DEFINED

An action taken by an individual to intentionally bias assessment results.



# CHEATING DEFINED

Cheating can occur by anyone with knowledge of or access to testing materials or the testing process

# CHEATING DEFINED

Cheating can occur by **anyone** with knowledge of or access to testing materials or the testing process

# CHEATING DEFINED

Cheating can occur by anyone with knowledge of or access to **testing materials** or the testing process

# CHEATING DEFINED

Cheating can occur by anyone with knowledge of or access to testing materials or the **testing process**



# WHY DO WE CARE?

Fundamental fairness for examinees

Public health, safety and well-being



# CHEATING IN CONTEXT



# STORIES FROM THE BATTLEFIELD

2010, American Board of Internal Medicine

139 individual physicians taking board certification exams found cheating. In total, more than one hundred items were compromised and thousands of physicians took the course and were exposed to stolen items.

2017, Civil Aviation Authority, UK

In spite of having a licensing exam and a language proficiency requirement, candidates with no proficiency in English were awarded licenses through a network of friends and colleagues without having to test.

# STORIES FROM THE BATTLEFIELD

August 9, 2018: Cleveland Police Department

15 Police cadets investigated for cheating on police licensing exam.

June 18, 2018: India: Police testing

23 individuals arrested for providing professional test-takers and high-tech listening devices to help applicants pass a two-day test to become police constables

Feb, 2018: England

Between 2012-2016, thousands of teachers were caught cheating to try to improve their pupils' test results, mostly by providing "improper assistance"

## WINDOW OF OPPORTUNITY



Nearly 14.35 lakh students are appearing at 80,000 centres across Bihar in the Class 10 examination conducted by the Bihar School Examination Board. The exams began on March 17. In Bhind district, exam centres are usually crowded with family and friends. Some even write answers for the examinees while being guarded by armed men. Books and 'chits' fly in and out of the windows of the building. District education officer Deepak Kumar Pandey spoke of incidents of violence at the centres. Bhind has become infamous for 'exam tourism' because of the widespread cheating, he added.

**SHAHROZ AFRIDI**

■ Family and friends climb walls to make answer chits available to those appearing for their Class 10 exams at a centre in Vaishali, Bihar, on Wednesday.

HT

“...exam centres are usually crowded with family and friends. Some even write answers for the examinees while being guarded by armed men. Books and “chits” fly in and out of the windows of the building.”

# STORIES FROM THE BATTLEFIELD

2017 Brazil

11 individuals arrested for using electronic devices during the Enem

Also, statistical analysis found large numbers of unusually similar answers in at least 1,125 Enem tests concentrated in dozens of different cities from 2011-2016.



# INFLUENCE OF TECHNOLOGY



# SCOPE OF CHEATING

Focus of cheating used to be on individual examinees seeking personal gain.

- Methods limited numbers involved and the extent of score bias

Shift to coordinated and entrepreneurial cheating efforts

- Today's methods involve many people and have opportunity to produce very large gain scores

# IMPACT OF CHEATING

Measurement

Societal

Financial





# CHEATING ACROSS THE TESTING LIFECYCLE



- Deter
- Detect
- Decide

# PREVENTING / DETERRING / STOPPING CHEATING: TEST DESIGN

- Single-form, linear test
  - CAT
  - Event based v. Windows based testing
  - Controlling exposure rates
    - Over time
    - Within person
    - Within window
- Deter
  - Detect
  - Decide
  - Impede

# PREVENTING / DETERRING / STOPPING CHEATING: **LIMITING ACCESS**

- Limiting paper during item development
- Restricting access to secure materials
  - Chain of custody
  - Security of electronic files



# PREVENTING / DETERRING / STOPPING CHEATING: CHECK-IN

- Prohibited items
- Government-issued identification
- Biometrics



# PREVENTING / DETERRING / STOPPING CHEATING: TEST ADMINISTRATION AND PROCTORING

- Random seating
- Spacing
- Breaks



# PREVENTING / DETERRING / STOPPING CHEATING: **TEST ADMINISTRATION AND PROCTORING**

- Active monitoring
- Conflicts of Interest
- Training

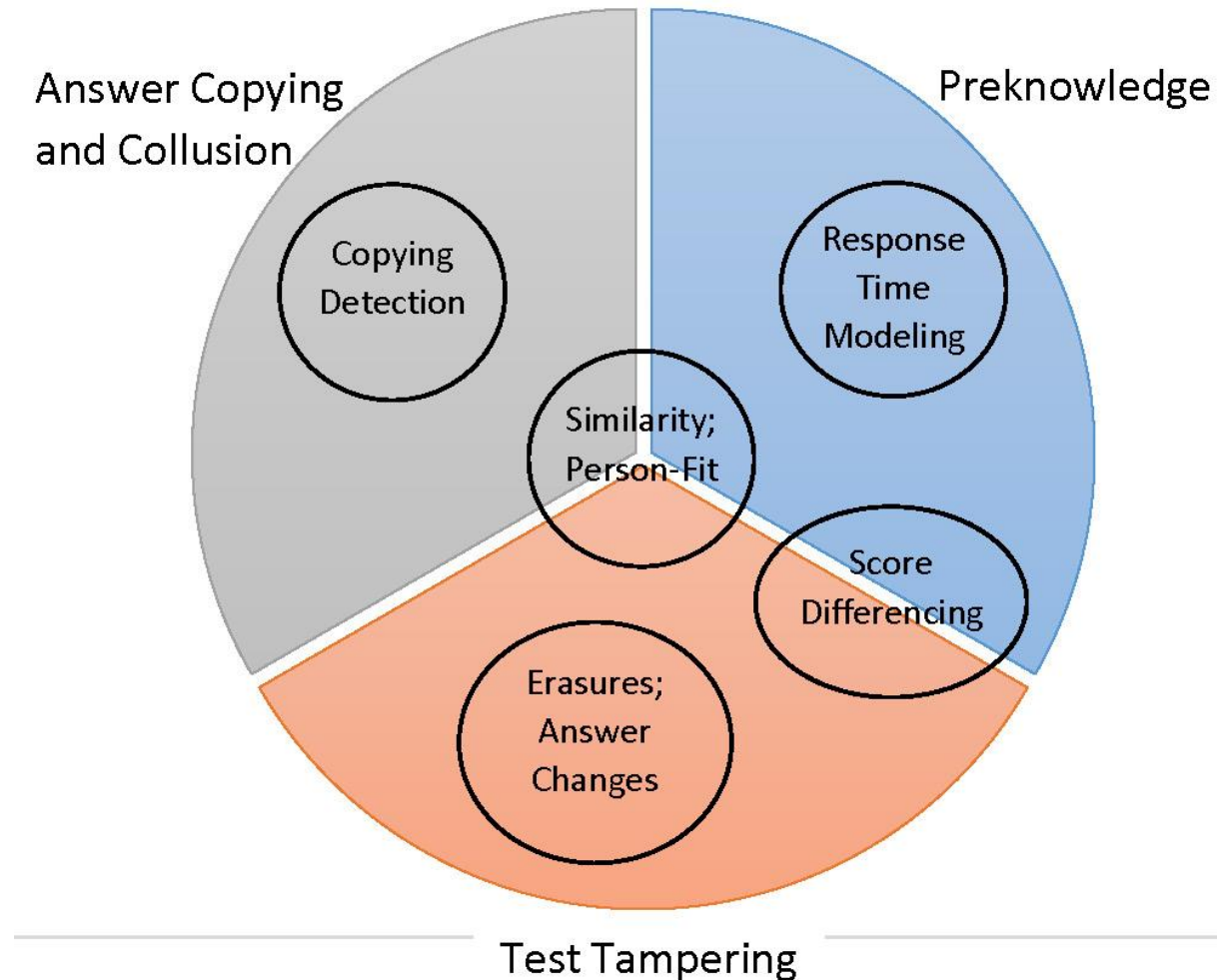


# DETECTING AND INVESTIGATING CHEATING

Statistical Methodologies



# DETECTING AND INVESTIGATING CHEATING: STATISTICAL DETECTION



# DETECTING AND INVESTIGATING CHEATING: STATISTICAL DETECTION

## Copying detection and similarity

- Comparison of the number of answer matches relative to the expected number of matches.
- Answer copying is directional—How often do we expect for this suspected copier, given his overall performance, to pick the same answers as the alleged source?
- Answer similarity is symmetric—a pair of examinees provides only one index value
  - Copying, collusion, preknowledge, test tampering
  - Similarity/Copying indexes have been studied more than other statistical methods and tend to be one of the most effective tools in our data forensics toolkit.

# DETECTING AND INVESTIGATING CHEATING: STATISTICAL DETECTION

## Score Differencing

- Identifies candidates whose performance varies significantly across two different sets of items: one believed to be largely secure and one which is possibly compromised.
  - Works best when compromise status is known
- Gain scores for repeat candidates are a special case of score differencing

# DETECTING AND INVESTIGATING CHEATING: STATISTICAL DETECTION

## Erasures and Answer Changes

- Answer changes are uncommon events
- About half the answer changes should be Wrong to Right (WTR)
  - Other half are Right to Wrong and Wrong to Wrong
  - Large numbers of WTR changes are suspicious
- Methods
  - Compare number of WTR changes at group level with the expected number
  - Because benign erasures are so uncommon, one doesn't need to tamper very much with data before it is detectable.

# DETECTING AND INVESTIGATING CHEATING: STATISTICAL DETECTION

## Response Time (RT) Methods

- RT varies a lot across items as a function of reading load, cognitive load, computational load, and natural between-person differences
  - Variability in baseline RT patterns poses a serious detection problem
- Item response models for RT
  - Can detect by finding irregular RT patterns, especially across sets of items believed to be compromised or secure
  - RT data paired with response accuracy data appears to be a promising area for detection of preknowledge

# DETECTING AND INVESTIGATING CHEATING: STATISTICAL DETECTION

## Item Harvesting

- Perhaps the most serious type of cheating is item harvesting
- Harvester's score isn't necessarily expected to be high, but purpose of harvesting is to share items with subsequent examinees to improve their performance
  - Has potential to significantly influence many candidates
- Published literature does not include any methods to detect harvesting
  - What does it look like?

# DETECTING AND INVESTIGATING CHEATING: **OTHER METHODS**

Statistics

Proctor Irregularity Reports

Hotlines

Webcrawls

Data analysis

Logging and monitoring systems

# DECIDING HOW TO ADDRESS CHEATING: **PREPARING FOR THE INVESTIGATION**

Must first make a decision about whether to investigate

Test Security plan



# DECIDING HOW TO ADDRESS CHEATING: INVESTIGATING

Statistics



# DECIDING HOW TO ADDRESS CHEATING: INVESTIGATING

Data, Document, and Digital forensics

Access Logging and Monitoring Reports



# DECIDING HOW TO ADDRESS CHEATING: INVESTIGATING

Irregularity reports



# DECIDING HOW TO ADDRESS CHEATING: INVESTIGATING

Interviews



# DECIDING HOW TO ADDRESS CHEATING: INVESTIGATING

Mystery shopping / audits



# DECIDING HOW TO ADDRESS CHEATING: INVESTIGATING

Webcrawls / Social Media



# DECIDING HOW TO ADDRESS CHEATING: **RESOLVING**

Evaluating the evidence

Who?

What?

How much evidence is needed?

# DECIDING HOW TO ADDRESS CHEATING: RESOLVING

Determining consequences





# THANK YOU

---

James A. Wollack  
University of Wisconsin  
jwollack@wisc.edu