

INVESTIGATING EXAM MISCONDUCT: BEST PRACTICES BEFORE, DURING, AND AFTER THE EXAM

by Deborah J. Harris and Rachel R. Watkins Schoenig

The legal profession is one of many professions that require an applicant to demonstrate competence in some manner prior to being allowed to practice. If an examination is part of the application process—as is the case in the legal profession, which relies heavily on the bar exam to assess applicant competence—it is important that an applicant’s score on that exam be an accurate reflection of the applicant’s own achievement and abilities. Incidents of misconduct are a common threat to valid examination scores and, though the extent of misconduct is difficult to document, the consensus seems to be that incidents of misconduct are on the rise.

“Misconduct” is a broad term, ranging from copying answers to sharing memorized test questions with future examinees. As such, incidents of misconduct may occur before, during, and following an examination. Further, persons who engage in misconduct can include examinees, test providers, test administrators, end users, and others. The tools used to engage in misconduct have also increased in sophistication, as devices like cell phones, cameras, ear buds, and recording devices have become more readily accessible to examinees.

Interestingly, while the number of exams administered and the sophistication of misconduct methods have increased over recent years, the basic types of

misconduct—such as copying, using a surrogate (an individual who impersonates and takes the test in the stead of the individual actually registered to test), accessing prohibited materials during the exam or “live” test questions prior to the exam, and continuing to work after (or beginning to work before) time is called—have remained fairly constant. Likewise, the goal of the investigative process has remained constant; test sponsors still seek to gather the most relevant evidence available and then, considering the evidence as a whole, to render a decision as to whether misconduct occurred and, if so, to apply an appropriate sanction.

The purpose of this article is to provide guidance on best practices for assembling evidence related to an incident of alleged misconduct. We first discuss the practices and procedures that should be in place before the exam begins, followed by what may trigger a suspicion that misconduct has occurred. We then address gathering and evaluating the evidence of misconduct.

BEST PRACTICES BEFORE REGISTRATION: THE CONTRACTUAL AGREEMENT

Best practices for misconduct investigations start before the examination is actually given—in fact, even before the examinee registers for the exam. The first step to a solid misconduct investigation is

a clear, well-thought-out contract that sets forth the following:

- what constitutes misconduct;
- what process will be followed if misconduct is suspected; and
- the right of the test sponsor to impose certain consequences if it concludes that the test scores are invalid for any reason, including a misconduct determination.

Identifying What Constitutes Misconduct

The contract between the test sponsor and the examinee should clearly identify what constitutes misconduct. For example, accessing notes, having possession of electronic devices, bringing drinks or food into the exam room, and using pencils or calculators that were not provided by the test sponsor may be defined as misconduct. Additionally, if the test sponsor desires that the examinees not reveal the content of the exam, a clear confidentiality provision should be included in the contract. It may also be appropriate to include hypothetical examples of violations in the registration materials, to help ensure that examinees understand the practical implications of the contractual language.

Defining the Misconduct Review Process

The contract should also set forth the review process that will be followed in the event that the test sponsor suspects misconduct or concludes that, for

whatever reasons, the test scores are not valid. The process should be easy for both the test sponsor and the examinee to understand and follow. Often the process includes an opportunity to take a privately arranged retest or to provide evidence to a professional board or a panelist within the test sponsor's organization. Many processes also contain an arbitration option for the examinee to use to appeal the

conclusion of the professional board or panelist. If arbitration is offered as an option, the contract should specify the question for appeal.

Stating the Test Sponsor's Rights

The contract should reserve the test sponsor's right to withhold or cancel scores in the event that the test sponsor deems the scores invalid for any reason. The contract should also set forth any consequence or range of consequences that may be imposed if misconduct is found, such as prohibiting future testing or communicating a score cancellation or a misconduct

determination to third parties. In addition, the contract should reserve the test sponsor's right to confiscate prohibited items, including but not limited to electronic devices such as cell phones, cameras, and pen scanners.

BEST PRACTICES BEFORE THE EXAMINATION

Before the examination begins, several opportunities arise to gather relevant information that can be used during a subsequent misconduct investigation.

In addition to providing evidence, many of these practices also serve to deter misconduct in the first place.

Pre-Exam Procedures

Test administration instructions and procedures provide important deterrents as well as important investigative evidence. For example, obtaining copies of government-issued photo identification and a photograph of the person actually presenting to test is helpful, as the photographs can be used later to determine if the registered examinee sent a surrogate to test. Similarly, obtaining fingerprints or palm vein scans of examinees during registration or prior to admission to the testing room can help determine if a surrogate has tested. Having an examinee sign an “honor contract” prior to beginning the exam reinforces the rules the examinee agreed to abide by, and also provides helpful handwriting samples that can be used to identify a surrogate.

Conducting Web Searches

Searching the Web for evidence of misconduct, by monitoring chat rooms or conducting web crawls, is also an advisable best practice prior to the examination. As chat rooms and social networking sites have increased in popularity, the incidents of examinees violating confidentiality rules and exposing copyrighted test content have also increased. Thus, web searches may reveal pretest item exposure, requests for items after testing, or even offers to act as a surrogate for the exam. This information can put the test sponsor on notice of the types of misconduct that may be attempted during the exam.

Testing Room Procedures

Prior to admitting examinees to the testing room, it is helpful to have a sign reminding examinees of prohibited items, such as cell phones, drinks, food,

or notes, and to provide examinees a secure place to store such items if necessary. After examinees are seated and before testing begins, it is also helpful to provide examinees with verbal reminders of prohibited items and conduct. This not only deters misconduct but also makes it easier to dismiss an examinee or to invalidate scores if these items are later found in the testing room.

The presence of video monitoring equipment may also help deter misconduct and provide useful evidence during a subsequent misconduct investigation. The video equipment can provide evidence concerning the test event, including visual evidence of examinees accessing prohibited materials or sharing information with one another.

WHAT INITIATES AN INVESTIGATION

Examinee conduct during testing may warrant immediate dismissal from testing or trigger a post-exam investigation. During the exam, vigilant proctors may witness examinees colluding (i.e., copying answers, texting, or leaving notes in the restroom) or accessing prohibited notes and cheat sheets. Additionally, other examinees may call suspicious behavior to the proctor’s attention or later report it to the test sponsor. Unusual behavior that does not rise to the level of immediate dismissal should be noted on an irregularity report that can be used later by the test sponsor to further evaluate scores. Having proctors document such unusual behavior is an advisable best practice during the examination.

Suspicious Behavior

Test proctors should observe examinees for signs of suspicious behavior. For example, a proctor should report an examinee who frequently asks to leave the room for breaks. Examinees who request to leave the exam room frequently may be accessing notes

planted in the hall or restroom prior to the exam, leaving notes for other examinees, or using cell phones to receive or provide exam assistance. In addition, an examinee could try to send in a surrogate following a break, so proctors should check identification upon reentry and make note of any examinee whose attitude or appearance seems to have changed. Similarly, a proctor should report an examinee who does not know his or her demographic information or who acts extremely nervous during check-in, as this may be indicative of a surrogate test taker.

The proctor should also report examinees who appear to finish extremely quickly or who appear to stay on the same page for an unusual amount of time. Examinees who appear to finish the test more quickly than would be expected may have had pretest exposure to items or may be using prohibited materials to help them complete the exam. On the other hand, examinees who appear to focus on only certain questions may be memorizing those questions with the intent to later sell the questions to future test takers.

Examinees who fidget, who feign stretching, or who appear to have roving eyes may be attempting to copy from other examinees. Assuming this activity is not so overt as to warrant immediate dismissal, the proctor should note it on an irregularity report. Examinees may use signals such as pen or foot taps or coughing to indicate to others the answer to a question. By watching and listening for patterns of noises or gestures, a proctor can identify potential evidence of coordinated misconduct activity during the exam.

EXAMINEES MAY USE SIGNALS SUCH AS PEN OR FOOT TAPS OR COUGHING TO INDICATE TO OTHERS THE ANSWER TO A QUESTION. BY WATCHING AND LISTENING FOR PATTERNS OF NOISES OR GESTURES, A PROCTOR CAN IDENTIFY POTENTIAL EVIDENCE OF COORDINATED MISCONDUCT ACTIVITY DURING THE EXAM.

The Discovery of Prohibited Devices or Materials

In some instances, it is possible that observing suspicious behavior will warrant more than simply documenting the behavior; it may be important for the administrator to further investigate to discover whether an examinee has possession of a prohibited item. For example, if an examinee's behavior indicates an attempt to access a prohibited electronic device or use a pen scanner, the device should be confiscated and returned to the test sponsor for investigation. Locating and confiscating prohibited electronic devices and any other prohibited materials during the exam should be handled discreetly and according to the test sponsor's instructions to avoid unnecessary disruption of other examinees.

Suspicious Interaction with Objects or Clothing

While best practices suggest banning food and drink from the examination room, if they are allowed, examinees who have them should be observed to see if they are frequently looking at or rotating these items, as notes can easily be stored on labels or even written on items such as gum. Similarly, the proctor should be aware that other objects, such as rubber bands, erasers, and mechanical pencils, should be watched, as notes can be written on or hidden in them. If an object is found to contain notes, the object should be confiscated and returned to the test sponsor. If an examinee is found to have written notes on his or her body, the administrator should take a photograph of the writing or, alternatively, document what was written and where.

If a proctor notices an examinee wearing unusually bulky clothing or head coverings or playing with his or her clothing, the proctor should report this activity. The examinee may be attempting to record test items or to access a cheat sheet. Rolling a sleeve up and down or tugging on an earring may signal use of prohibited materials such as notes or use of an electronic device like an ear bud. Proctors should be alert to an examinee who drags a pen or pencil over the exam, as this could be an indication of a hidden scanner.

Conducting Routine Evaluations After the Exam

Routine audits or analyses of test scores and other data, such as erasures, after the exam may also raise suspicions that warrant further investigation. The test score audit may indicate unusual erasures, test score changes, test center pass rates, or answer pattern similarities between examinees. Identifying multiple examinees who registered and tested using the same address may trigger further investigation, as such information could indicate a ring of individuals attempting to steal (or harvest) test items. Further, information gleaned from third parties, by monitoring chat rooms, conducting web crawls, or receiving anonymous tips on a hotline established for reporting suspicious behavior, may also trigger post-exam score investigations.

GATHERING RELEVANT EVIDENCE

Once something has occurred to raise concerns regarding a test score or test event, the test sponsor must determine whether a formal investigation

and notice to the examinee are warranted. Because the test sponsor must invest resources in a formal investigation and perhaps in defending any sanctions applied at the conclusion of the investigation, both the nature of the misconduct and the ability to collect and review relevant evidence are important factors in determining whether to pursue a formal investigation. Therefore, before opening a formal investigation, the test sponsor often gathers

readily available relevant evidence concerning the test taker and the test event—both in support of and in contradiction to the misconduct allegation—with the goal of collecting as much information as practical to inform its decision whether to launch a formal investigation.

Evidence collected may include observational, physical, biometric, and statistical evidence. Information relevant for one type of misconduct may not be particularly informative for another. For example, a seating chart may be quite helpful in a suspected copying situation but less helpful in a situation where advance knowledge of test items is suspected. The type of information compiled will depend on the information gathered before, during, and after testing.

Observational Evidence

As mentioned in the discussion of possible investigation triggers, proctors can provide valuable observational evidence concerning potential misconduct. Firsthand accounts of suspicious behavior can be extremely helpful, especially if corroborated by a second observer, such as another proctor.

Physical Evidence

Physical evidence may be obtained from the test materials themselves, by comparing calculations, writing, or answers circled in the test booklet with responses or erasures on the answer sheet. Room rosters and seating diagrams can provide helpful information for a variety of purposes, such as establishing the proximity of suspected colluders to each other or determining if an obstruction may have limited a proctor's ability to observe the room. Limited visibility may explain how an examinee could have engaged in misconduct during testing, such as use of a cell phone, and avoided detection.

Biometric Evidence

Biometric evidence, such as examinees' fingerprints, palm vein scans, and retina scans, can provide evidence of surrogate testing. Handwriting samples and photographs can also provide helpful surrogate testing evidence. The reference sample may be collected before testing as part of the registration process or after testing as part of an investigation. An inability to match fingerprints, palm vein scans, or even handwriting samples provides strong evidence that a surrogate may have tested.

Statistical Evidence

Statistical evidence can help a test sponsor determine the likelihood that an observed result would have occurred by chance, the assumption being that unlikely events may provide more support for a finding of misconduct than likely events. As with the other types of evidence, test sponsors typically consider statistical evidence as part of the overall evidentiary picture and do not rely solely on statistical evidence to reach a conclusion that misconduct has occurred. Statistical evidence, however, often pro-

vides useful indicators of aberrant scores or scoring patterns that require further investigation.

Statistics gathered as evidence may include significant changes in scores from one test administration to another for examinees who retest. Additionally, statistics may show the relationship between the suspect test score and a predicted score based on other variables, such as grade point average. A comparison of the correct and incorrect responses of possible colluders is frequently performed in alleged copying cases.

Because it may be difficult to know what constitutes an unusual degree of similarity in the abstract, benchmark statistics are often used, such as the distribution of changes in test scores across all examinees. Similarly, the responses of a suspected collusion pair may be benchmarked across pairs of examinees unlikely to have colluded with each other, such as those who tested in different locations. The better the comparison data and the more discrepant the suspect data is with the comparison data, the more compelling the support for the alleged misconduct incident having taken place.

Unusual scoring patterns, such as a suspect answering hard items correctly and easy items incorrectly, may also be collected as evidence. The more extreme the statistical evidence for the suspect compared to other examinees who are not suspected of engaging in misconduct, the more compelling the evidence is in support of an irregularity. Of course, unlikely events do occur by chance, and many test sponsors require other evidence in addition to one type of statistical evidence before proceeding with a formal investigation. Multiple pieces of statistical evidence, however, such as an unusually large score gain, unusual answer patterns, and unusual

similarities with the responses of an examinee seated nearby, can be strongly persuasive.

EVALUATING THE EVIDENCE

Assuming the test sponsor determines that there is sufficient evidence to warrant a formal investigation, the agreed-upon procedures set forth in the contract between the test sponsor and the examinee must be followed. Typically, some step in the investigation process requires notification of the examinee suspected of engaging in misconduct and may provide the examinee an opportunity to offer evidence supporting the validity of the test score. For example, the examinee may be required to take a privately arranged retest and replicate the questioned score or may be allowed to submit other information supporting his or her ability to achieve the suspect score.

Examinee-Provided Evidence

The type of information submitted by examinees suspected of engaging in misconduct varies. Depending on the allegation, a suspect may submit school performance information (such as courses and grade point average), character evidence, or medical information that helps explain an improvement in scores or other discrepancies (for example, detailing the medication used by the examinee that helped him or her better focus on the exam, or detailing an injury that might explain differences in handwriting).

The manner in which the examinee submits this evidence is usually determined by the contractual agreement between the examinee and the test

sponsor, including whether the information is to be submitted in writing, what information is relevant, and what information the examinee may receive regarding the test sponsor's concerns.

Viewing the Evidence in Its Entirety

Once the examinee's evidence has been gathered, the test sponsor reviews all of the evidence compiled—both from the test sponsor's investigation and that submitted by the examinee—and evaluates the evidence as a whole. With the exception of extremely overt activity that typically results in immediate dismissal from the exam, no one piece of evidence is typically sufficient to conclusively establish misconduct.

For example, the test sponsor may evaluate an examinee's erasure patterns. An unusual number of erasures may mean that the examinee accidentally skipped an answer early in the test and the erasures were necessary to "get back on track" on the answer sheet. Alternatively, of course, those erasures may reflect an examinee's decision to change answers and copy from a fellow examinee. The erasure evidence taken in conjunction with extreme similarities in responses to another examinee's answer sheet, a lack of evidence of independent work, and proctor observations that the examinee seemed to be looking at the other examinee's answer sheet may warrant a conclusion that the test scores are not valid due to copying.

Evaluating the Quality of the Evidence

At this point, the quality of the evidence is also taken into account. For example, a letter from a

student telling the test sponsor that he received all A's in law school is not as compelling as seeing the grades on an official transcript. Further, whether information submitted by an examinee is persuasive will depend on the unique circumstances of the incident. For example, evidence that over the last six months the suspected examinee took additional coursework may be persuasive in explaining an unusual score increase, but not persuasive if similarity statistics, erasure analyses, and proximity to an alleged source seem to indicate copying.

Presenting the Evidence

If a suspect admits to misconduct, the consequences set forth in the agreement between the test sponsor and the examinee should be followed. If the suspect does not admit to misconduct, then the test sponsor must decide whether the evidence is strong enough to pursue sanctions and, if so, how best to present the evidence. Depending on the agreement between the examinee and the test sponsor, the evidence collected in the investigation may be presented to an internal reviewer, a review board, or an arbitrator for a decision. Regardless of the reviewer, the evidence should be considered as a whole before a decision is rendered.

To facilitate interpretation and decision making, evidence should be presented clearly and concisely. For example, statistical evidence should not rely on specialized statistical language, and care should be taken to avoid common misinterpretations. The chance of two examinees having the same response to a multiple-choice item with four possible responses should be considered in the context of how popular the particular response was for examinees in general; if 98% of the examinees chose the same response, most examinee pairs, even those not in a collusion relationship, would show answer agreement for that item.

In presenting the evidence—whether to the examinee or to the reviewer—pictorial depictions are sometimes more readily interpreted than text. For example, a graph with an “X” for a particular examinee pair and dots for pairs of examinees at large may make it easier to interpret how extreme the differences were between a suspect pair and non-suspect pairs.

Reaching and Documenting the Decision

There is no threshold for how much evidence, or what type of evidence, will be persuasive. As mentioned earlier, some evidence is more relevant for particular incidents of alleged misconduct than others. Regardless of the type of evidence, the preponderance and consistency of evidence is often strongly persuasive.

Even if it is determined that misconduct has occurred, other factors may weigh into the test sponsor's decision to apply sanctions. Many factors become a part of the test sponsor's decision-making process, such as


- the overall persuasiveness of the evidence;
- the ability to clearly present the evidence;
- the cost of applying sanctions and defending the decision (including time and negative publicity); and
- the cost of not applying sanctions (including reputational damage and tacit encouragement for others to engage in misconduct).

Finally, the test sponsor must strive to show that the agreed-upon review process was followed in good faith. Thus, the test sponsor should maintain a file that reflects the evidence in the investigation and communications to and from the examinee. Further, the file should reflect the steps taken by the test

sponsor to adhere to the review process. This information will be important should the test sponsor's conclusion of misconduct be challenged in arbitration or litigation.

SUMMARY

Best practices in compiling evidence of misconduct start prior to the exam with a well-crafted agreement between the examinee and the test sponsor. This agreement forms an important framework for investigations into alleged incidents of misconduct and testing irregularities, both by defining what constitutes misconduct and by setting forth the investigation process and possible consequences if misconduct is found. Following the agreement ensures that incidents are handled consistently and that examinees are treated equitably. Well-devised pre-exam procedures, vigilant observation of examinees during the exam, and post-exam reviews of data and proctor reports all contribute to a sound system of collecting evidence in the event of exam misconduct.

In situations in which taking an examination is necessary to enter into a profession, it is important that an examinee's score have integrity and accurately represent the examinee's actual abilities. Because incidents of misconduct are a threat to valid exam scores, the implementation of best practices to identify and investigate instances of exam misconduct is critical to ensure legitimate access to the profession. 

REFERENCES

- Cizek, G.J. (1999). *Cheating on Tests: How to Do It, Detect It, and Prevent It*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Dows, M. S. (2005). One Administrator's Thoughts on—and Experiences with—Security of Test Materials. 74 *THE BAR EXAMINER* 3:6–13.
- Wollack, J. A., & Fremer, J. J. (in press). *Handbook of Test Security*. London: Taylor & Francis Group.



DEBORAH J. HARRIS is Assistant Vice President for Measurement and Reporting Services for ACT, Inc., a not-for-profit organization that provides a broad array of assessment and other solutions geared toward helping people achieve education and workplace success. She is responsible for providing psychometric consultation and ensuring the effective delivery of measurement and reporting services. Harris holds a doctorate from the University of Wisconsin–Madison and is an adjunct professor at the University of Iowa.



RACHEL R. WATKINS SCHOENIG is Director of Test Security for ACT, Inc. Before joining ACT, she led the in-house legal department for a publicly traded company involved in international manufacturing and franchising, where she handled numerous compliance, litigation, and intellectual property matters. Prior to that, as a partner in private practice, she represented clients in complex commercial litigation matters. Schoenig holds a J.D. from Drake University Law School.