

How to Write Effective Multiple-Choice Test Problems Guidance for Founder Applicants

Preamble

As a requirement for becoming a Founder of the new Digital Forensics Certification process, you are required to submit fifteen (15) multiple-choice problems as potential test problems for future exams. These problems are part of the application process and will be evaluated against criteria that will be provided to you. Additionally, each problem must address appropriate domain(s) of knowledge and be adequately referenced. Points for your founder application will only be given for valid and appropriately referenced problems as determined by the certification committee.

This document is intended to help you to write high-quality multiple-choice test problems that measure the candidate's knowledge of digital forensics, while resisting success due primarily to testmanship.

A multiple-choice test problem consists of several components:

- A question (either ending with a question mark, or as a declarative statement with a blank for the answer)
- A correct answer
- Two or three plausible, but incorrect answers, known as "distractors"

A multiple-choice problem is best developed in stages. First, the question and correct answer are carefully crafted. Next, the distractors are created. Finally, the problem as a whole is scrutinized to ensure that it is fair, unambiguous, and resists being solved by testmanship alone. Testmanship employs the skill of eliminating distractors and the art of good guessing, allowing the exam-taker to score well without displaying significant knowledge of the subject matter. This, of course, is to be avoided.

A good multiple-choice problem has several characteristics:

1. The problem tests a specific objective, which should be clear from the question body
2. The question is unambiguous
3. The question is not arcane or over-specific
4. There is a single correct answer
5. No distractor is arguably correct
6. No distractor is obvious to the layperson
7. The collection of distractors do not self-eliminate
8. Other problems (previous and subsequent) do not provide clues to the current problem

Characteristic 1: The problem tests a specific objective, which should be clear from the question body.

Start by asking yourself, “What exactly am I testing with this problem?” Along with each problem you submit, you are required to provide the specific reference(s) to support the correct answer, so this is a good place to document those sources and define the specific knowledge that you are testing. You should also determine the level of cognition your problem addresses, summarized as follows (from Bloom’s Taxonomy¹):

1. Recall – rote memorization of facts.
Example: “Which items are ...”
2. Comprehension – understanding the meaning of material.
Example: “Which parameters would make tool x do y?”
3. Application – using learned material in new situations
Example: “In situation x, which y would be most appropriate?”
4. Analysis – breaking down knowledge into its component parts and understanding their relationships
Example: “What allows this set of commands to perform function z?”
5. Synthesis – putting knowledge together to form a new construct, exhibiting creativity. Since multiple-choice problems provide a fixed set of possible answers, no creativity is demonstrated; therefore, this level is generally not applicable.
6. Evaluation – making value judgments against specific criteria.
Example: “Which option would be best in this situation?”

An example of defining the initial question and answer follows. This classic example tests the candidate’s understanding of the labeling convention of drives, both in Linux and Windows systems:

Question: Given the Linux command “**dd if=/dev/hdb ...**”, which would be the corresponding command in Windows?

Answer: **dd if=\\.\PhysicalDrive1 ...**

Level: Comprehension

Reference(s): *(as appropriate)*

Characteristic 2: The question is unambiguous.

The question must be written simply and clearly, with no room for misinterpretation. The context must also be apparent, and all relevant information must be in the

¹ <http://www.nwlink.com/~donclark/hrd/bloom.html>

question. The candidate must also be tested on the subject matter, not on the ability to detect subtleties in the question's phrasing.

Poor example: "The four-step security model includes _____."

The flaw is that the candidate does not know whose "four-step security model" is being discussed. Even if a specific reference was given, the author's specific views may not be universally accepted among the digital forensics community, so such a question should be avoided.

Better example: "According to the Scientific Working Group on Digital Evidence, the term 'digital evidence' includes all of the following EXCEPT:"

Poor example: "When you send a job to the printer, a _____ file is created."

The flaw is that the candidate is not told what operating system is being used. Even if the question was modified to include "Windows", this may not be specific enough since the printer may dictate the format.

Better example: "Under Windows XP, when a job is spooled for a HP printer, which forensically-interesting file type is usually created?"

Characteristic 3: The question is not arcane or over-specific.

Except when dealing with specific KSAs (Knowledge, Skills, Abilities), exam problems should cover general knowledge of importance to the wide field of digital forensics and practitioners. Questions that have been historically used in examinations – in the early days when powerful tools were non-existent and drive sizes were more manageable – might no longer be applicable.

Poor example: "The deletion date of a file is found at what byte offset in the INFO2 file?"

Critique: Tools now exist that extract this data in readily reportable forms, removing the need for the examiner to perform such low-level analysis. This kind of information should still be taught to students of digital forensics, so that they understand the internal workings of the tools and how systems manage the data, but such detail may not be particularly relevant for certification examinations.

Better example: "What information can be found in the INFO2 file?"

Poor example: “The command dd.exe is not compatible with which Windows OSes?”

The flaw with this question is that tools change, and additional capabilities do get added. Over time, the correct answer to the question may also change.

Better example: “When George Garner created the Windows version of command dd.exe, what additional capabilities did he include that were not in the Unix version?”

Characteristics 4 and 5:

4: There is a single correct answer.

5: No distractor is arguably correct.

You have probably taken a multiple-choice examination where it was emphasized that you were to choose the BEST answer. This is one of the most frustrating experiences for a test-taker, because it often means that care was not taken to properly formulate the problem, and your job is reduced to selecting the “Least Worst Answer.” Please avoid creating such problems.

Poor example: “An analysis of a full network packet capture can provide direct evidence of which of the following?”

- The programs used in the attack
- The vulnerability that was used to compromise the attacked system
- The private keys used to decrypt a secure session
- The attacker’s true identity and motive

One flaw in this example is in the interpretation of “direct evidence.” If taken at face value, there is no correct answer. Interpreting it to mean “strong evidence” or “evidence that can be leveraged for analysis” results in two arguably correct answers.

Better example: “An analysis of a full network packet capture can provide strong evidence of which of the following?”

- The vulnerability that was used to compromise the attacked system
- The private keys used to decrypt a secure session
- The attacker’s true identity and motive

Poor example: “What is the name of the U.S. Government’s statute that addresses the interception of electronic communications while traversing a network?”

- The Federal Wiretap Act
- The Electronic Communications Privacy Act
- The Electronic Privacy Act
- The Federal Data Protection Statute

This problem has multiple flaws. One is that the question asked for a statute, and only one was provided: the rest are “acts.” (Some people use the two terms interchangeably; others make distinctions.) The intended correct answer is “The Federal Wiretap Act”, which would fail to satisfy the literal interpretation of the question. Another flaw is that the Federal Wiretap Act is a subcomponent of the Electronic Communications Privacy Act, making both answers technically correct.

Better example: “What is the name of the U.S. Government’s statute or act that addresses the interception of electronic communications while traversing a network?”

- The Federal Wiretap Act
- The Realtime Electronic Collection Statute
- The Electronic Privacy Act
- The Federal Data Protection Statute

Characteristic 6: No distractor is obvious to the layperson.

In the preceding example about network packet capture content, the last two answers could be eliminated without much difficulty. Keys are not sent in the clear in secure sessions. Motive is never directly evident; it’s only inferred. Recognizing these simple facts leaves only two possible answers, and significant knowledge about information found in network packet captures was not required.

Characteristic 7: The collection of distractors do not self-eliminate.

In many poorly constructed problems, eliminating one distractor causes a cascade effect of eliminating others. Problems with “All of the above” or “None of the above” as answer choices are often subject to this undesirable phenomenon, as are answers with “all,” “always,” and “never.”

Some problems are of the type “choose all that apply”; in these cases, groups of answers are created so that only one group is the correct response. Great care must be taken so that if one of the choices is eliminated, the problem doesn’t become trivial.

Example: The tool suite FTK provides which of the following capabilities. [Select one radio button]

- A. Automatic indexing of all text fragments to speed-up searches
 - B. Automatic password cracking of all encrypted files
 - C. Network traffic analysis
 - D. Email message recovery
 - E. Carving of graphics files from unallocated disk areas
 - F. Automatic chat room dialog reconstruction
 - G. Report generation
-
- A, B, C, and D
 - A, B, D, E, F and G
 - A, C, D, E, F and G
 - A, D, E and G
 - B, D, E, and G
 - All of the above

In this example, if the candidate was to recognize the offensive word “all” in choice B, four of the six responses are immediately eliminated.

Also note that choice D appears in every response, so it contributes nothing to the problem but clutter.

Better Example: The tool suite FTK provides which of the following capabilities. [Select one radio button]

- A. Automatic indexing of all text fragments to speed-up searches
 - B. Automatic password cracking of all encrypted files
 - C. Network traffic analysis
 - D. Email message recovery
 - E. Carving of graphics files from unallocated disk areas
 - F. Automatic chat room dialog reconstruction
 - G. Report generation
-
- A, B, C, and D
 - A, B, E, F and G
 - A, C, D, E, F and G
 - C, D, E and G
 - B, E, and G

Characteristic 8: Other problems (previous and subsequent) do not provide clues to the current problem.

It is not uncommon to find professional examinations that contain a clue or the answer to one problem within the question or possible answers of another problem on the same exam. Of course, this situation must be avoided. Please carefully review the collection of problems you have created, giving extra attention to overlapped content amongst your problems in an effort to spot such clues. Remove or revise the problem set as needed to avoid such hints.

Guidance for Writing Good Answers and Distractors

There exists a great online resource² that provides guidance for writing answers and distractors. This section summarizes that resource and provides additional insight.

- **Distractors must be plausible, but incorrect, answers to the question. Include common errors that exam-takers tend to make.**
- **Answers and distractors must be parallel in grammatical structure**
- **Answers and distractors should be approximately the same length**

Poor example: “When establishing the forensics management processes, what three items should be included?”

- Define guidelines, policies, and procedures for the Forensic Incident Response team to follow before starting the investigation
- Scope, authority, and limitations on the examiner
- Expectations, procedures for the Forensics Incident Response team, list of vendors
- Who, what, where, and when of the investigation

Critique: The first choice is the correct answer, but it isn't parallel to the others: the others contain sets of nouns, while the first choice is a directive (“do this”). It is also noticeably more elaborate than the rest. The third choice is very vague regarding the vendors, so could be dismissed as implausible. Finally, the question asks for three items, but the final choice has four, making this distractor implausible.

Better example: “When establishing the forensics management processes, what three items should be included?”

² “Writing Multiple-Choice Questions that Demand Critical Thinking,” University of Oregon, Teaching Effectiveness Program [http://tep.uoregon.edu/resources/assessment/multiplechoicequestions/mc4critthink.html]

- Guidelines, policies, and procedures for the Forensic Incident Response team
- Scope, authority, and limitations on the Forensic Examiner
- Expectations and procedures for the Forensics Incident Response team, and a list of vendors of the organization's equipment and software
- What, where, and when of the investigation

Critique: Parallelism is achieved. The third choice is longer than the rest, but is necessarily so to convey the intended content.

- **Limit the number of distractors to 2 or 3. Additional distractors generally only serve to clutter the exam, waste time, provide unnecessary opportunities for challenge, and seldom improve the effectiveness of the examination process.**

In the previous example, the fourth choice could be easily eliminated without impacting the examination. It exists to ensure that the student recognizes the difference between pre-incidence planning (the root of the question), and incidence response. Since the second choice also addresses responding to a specific incident, the fourth choice is not needed.

- **If there are “magic words” in the question and answer, include them in the distractors so that the answer isn't obvious by comparison.**

Poor example: “When using the command dd to copy information from the main physical memory, the input would be referenced as:”

- \\.\VirtualDisk
- \\.\PhysicalMemory
- \\.\PhysicalDisk
- \\.\C:\Windows\pagefile.sys

Better example, providing distractors that are similar to the correct answer: “When using the command dd to copy information from the main physical memory, the input would be referenced as:”

- \\.\VirtualMemory
- \\.\PhysicalMemory
- \\.\MainMemory
- \\.\C:\Windows\physical.mem

Pulling it All Together

Here is the original example problem, carried through to completion:

First, develop the question and answer:

Question: Given the Linux command “**dd if=/dev/hdb ...**”, which would be the corresponding command in Windows?

Answer: **dd if=\\.\PhysicalDrive1 ...**

Level: Comprehension

Reference(s): *(as appropriate)*

Second, develop a set of plausible distractors:

Distractors:

dd if=\\.\PhysicalDrive2 ...
dd if=\\.\PhysicalDriveB ...
dd if=\\.\PhysicalDisk1 ...
dd if=\\.\PhysicalDisk2 ...
dd if=\\.\C: ...
dd if=\\.\D: ...

Next, pare down the number of distractors and randomize the presentation order:

Given the Linux command “**dd if=/dev/hdb ...**”, which would be the corresponding command in Windows?

- dd if=\\.\PhysicalDrive2 ...**
- dd if=\\.\PhysicalDrive1 ...**
- dd if=\\.\PhysicalDisk2 ...**
- dd if=\\.\D: ...**

Finally, critique the problem as a whole against the eight characteristics. In particular, the answer should not be obvious to the lay-person, no distractor is subject to challenge as being correct, and the distractors do not lend themselves to “self elimination.”

Suggestions for Your Fifteen Problems

Make it easy on yourself: focus on your forte. Attempt to span the spectrum of your specialty (e.g., management, practitioner, corporate, law enforcement, something else). Please aim for a mix of high-level and detailed problems. The goal is to create 15 high-quality problems that someone like yourself should be able to answer in order to be considered knowledgeable in this field.