



**Anti-Malware Testing Standards Organization
The Fundamental Principles of Testing**

The following represent a summary of the principles applicable to anti-malware testing that should be followed by testers, publications and vendors. These principles are based on our belief that everybody involved in such testing must behave ethically, test properly and communicate in a fair and accurate way. For additional information, please review guidelines for each item, beginning on page 2, below.

- 1. Testing must not endanger the public.**
- 2. Testing must be unbiased.**
- 3. Testing should be reasonably open and transparent.**
- 4. The effectiveness and performance of anti-malware products must be measured in a balanced way.**
- 5. Testers must take reasonable care to validate whether test samples or test cases have been accurately classified as malicious, innocent or invalid.**
- 6. Testing methodology must be consistent with the testing purpose.**
- 7. The conclusions of a test must be based on the test results.**
- 8. Test results should be statistically valid.**
- 9. Vendors, testers and publishers must have an active contact point for testing-related correspondence.**



**Anti-Malware Testing Standards Organization
Guidelines to the Fundamental Principles of Testing**

Principle 1: Testing must not endanger the public.

This principle is fundamental to the charter and purpose of AMTSO and each of its members. The public has the right to expect that the development and sale of anti-malware products, the review of such products and publication of those reviews are all done, fundamentally, to protect them. Thus, the foremost principle of testing anti-malware products is that neither the products nor the related testing should endanger the public. In furtherance of this principle, testers must follow appropriate procedures to avoid accidental release of samples at all times. In addition, new malware must not be created for testing purposes.

Q. What are considered to be “appropriate procedures”?

A. It is expected that any testing environment will utilize industry-standard best practices to ensure that malware samples are not accidentally released and that risks to the public are avoided.

Q. What is meant by “creation of new malware”?

A. This reference has historically referred to the creation of new viruses or strains of malware, one objection being based on the principle that there are more than enough samples available in the wild for everyone. This mandate has been complicated by introduction of packers and virtual machines, inviting the question as to whether utilizing these vehicles could be deemed to change the characteristics of pre-existing malware to the point that it could be deemed “new.” There are legitimate reasons to change existing malware characteristics for testing purposes – this principle is not included in order to preclude such testing. To be clear, however, this principle is included to demonstrate unanimous disapproval by AMTSO of the idea of the creation of new viruses or other malware and the related risk to the public. If you wish to contact AMTSO about these matters please send an inquiry to principles@amtso.org for more information.

Principle 2: Testing must be unbiased.

We believe that anti-malware testing, by its nature, should be unbiased – each product must be treated equally. Whether the test is commissioned by a vendor to support a marketing message or by a major magazine to run a story on product efficacy, it is the obligation of the tester to conduct the test in an ethical manner, and to present truthful and unbiased results.

There are many circumstances where vendors may provide financial incentives to a publication or tester. These incentives are neither unusual nor by definition unethical, and may be obtained through testing commissions or advertising revenue, for instance. Although generally innocuous, to avoid the appearance of impropriety, we believe that these relationships, when significant, should be disclosed. Thus, to meet this principle 2, we

encourage testers and publishers to publicly disclose the existence of any such significant financial relationships with a reviewed party or affiliate.

Q. What would constitute a significant financial incentive?

- A. The intent of this principle is to avoid bias and conflicts of interest in product testing and reporting. Thus, this disclosure should include any relationship that could potentially influence the tester, including: (i) whether the publication or tester has received revenue from a vendor or affiliate with regard to any particular test, and (ii) whether the publication or tester receives a significant portion of its overall revenue from a particular vendor. While testers are asked to disclose the source of their samples in the testing details, provision of samples in general is not considered a financial incentive.

Q. How should this disclosure be made?

- A. Ideally, each tester and publication will provide this disclosure as a footnote to each published report, or will provide a link or other reference to where such information can be found.

Principle 3: Testing should be reasonably open and transparent.

AMTSO recognizes that some publications may not be always comfortable with the disclosure of the methodology of published tests. However, AMTSO feels strongly that having open and transparent testing is critical to enforcement of these fundamental principles, and to ensuring reliability and consistency in anti-malware testing. As a result, we believe that any test released to the public must be accompanied by, or reference the location of, details regarding the test and testing methodology.

Details regarding the specific test should include the following information:

1. Which solutions were tested?
2. How were the solutions obtained and updated?
3. How were the samples or test cases obtained and validated?
(See also principle 5.)
4. What versions of the products were used?
5. What product settings/configurations were used?
6. When and under what conditions was the test conducted?
7. What environment was the test conducted in? (for example, the operating system/environment version, service packs applied, and other programs that were running at the time)

Details regarding the specific testing methodology should include the following information:

1. How were the test samples or test cases selected?
2. What were the sources of malicious and innocent samples or test cases?
3. How were the malicious and innocent samples or test cases applied?
4. How was the response of the solutions measured?
5. Was the test “apples to apples” (comparing products of similar type and functionality), or “apples to oranges” (comparing products of significantly different type and/or functionality)?
6. If “apples to oranges”, how were the various solutions compared?
7. How were the results calculated and interpreted?

Q. Where should the test and testing methodology be disclosed?

A. Ideally, this information will be included in the published report, either in the body of the report or by a link to the relevant information. If publications are unable or unwilling to include this information, testers can themselves make this information available on their website with a reference to a specific or general test.

Q. Must testers provide feedback and/or samples to vendors?

A. No. However, AMTSO encourages testers to provide vendors with constructive and adequate feedback in a timely fashion about specific faults and deficiencies (e.g. crashes, false positives, false negatives, etc.) This feedback can be in the form of technical details, reproduction steps, log files, memory dumps, samples, etc.

Principle 4: The effectiveness and performance of anti-malware products must be measured in a balanced way.

It is difficult – and can be misleading – to summarize product efficacy with a single measurement. Testers are encouraged to present multiple measurements of product performance in different areas in order to allow users to make an informed decision.

For instance, testers should appropriately balance false negative and false positive test cases. A product that is successful at detecting a high percentage of malware but suffers from a high false positive rate, may not be “better” than a solution which catches less malware but which generates less false positives.

Principle 5: Testers must take reasonable care to validate whether test samples or test cases have been accurately classified as malicious, innocent or invalid.

It has often been the case that seemingly reliable testing results are, in fact, not valid, because the samples used in the tests were misclassified. For example, if a tester determines that a product has a high rate of false positives, that result could be wrong if some samples were wrongly classified as innocent. Thus, it is our position that reasonable care must be taken to properly categorize test samples or test cases, and we especially encourage testers to revalidate test samples or test cases that appear to have caused false negative or false positive results.

Similarly, care should be taken to identify samples that are corrupted, non-viable or that may only be malicious in certain environments and conditions.

Principle 6: Testing methodology must be consistent with the testing purpose.

Tests must address the intended or stated purpose of the publisher’s related review or article. We believe that publishers should state the objective of their tests clearly, and that test methodology should be consistent with the stated test objective. (For example, publishing test results in a consumer-targeted magazine without making it clear that the test was conducted on corporate products because this does not simulate target user experience.)

For additional reference, David Harley has published a paper that analyses in detail the problems with one test that displayed a certain inconsistency between the test objective described and the methodology used.

See http://www.smallblue-greenworld.co.uk/AV_comparative_guide.pdf.

Principle 7: The conclusions of a test must be based on the test results.

This principle addresses a common high level problem with publishing conclusions alongside testing data that are not supported by those data. (For example, drawing broad and/or inaccurate conclusions from narrow test data.)

Principle 8: Test results should be statistically valid.

Testers should use a sufficient quantity of test samples, test cases or scenarios for results to be statistically sound. In addition, the tester's analysis of measurement errors is important and should be published. In general, AMTISO recommends using as many test scenarios as possible.

For additional reference, Igor Muttik has published a paper that analyses in detail how insufficient quantity of samples or test cases can produce random test results.

See http://www.mcafee.com/common/media/vil/pdf/imuttik_VB_conf_2001.pdf

Principle 9: Vendors, testers and publishers must have an active contact point for testing-related correspondence.

An "active contact point" is a current and monitored point of accessibility (via phone, fax or email) provided by vendors, PR departments, testers and publishers. Relevant correspondence regarding the subject product, test or testing methodology should be answered by the vendor, tester or publisher within a reasonable timeframe.