



## Anti-Virus Testing Websites

An overview on which testing sites  
can be trusted and which can not

Date: April 2007

Last revision: 20<sup>th</sup> April 2007

Authors: Andreas Clementi

Website: <http://www.av-comparatives.org>

## **Table of content**

1. The importance of independent antivirus testing	3
2. Trustworthy testing bodies	3
2.1 Testers with comprehensive test-sets	3
2.1.1 AV-Comparatives.org	4
2.1.2 AV-Test GmbH	4
2.2 Tests based mainly on the WildList	5
2.2.1 VirusBulletin	5
2.2.2 CheckVir	5
2.3 Certification bodies	5
2.3.1 ICSA Labs	6
2.3.2 West Coast Labs Checkmark	7
2.4 Tests in magazines	7
2.5 Academical tests	8
2.5.1 antiVirus Testing Center (aVTC) Hamburg	8
2.5.2 Virus Research Unit at the University of Tampere	8
2.5.3 Test Lab at Moscow State University	8
2.5.4 University of Magdeburg	8
2.5.5 University of Innsbruck	8
3. Non trustworthy testing bodies or flawed tests	9
3.1 Tests run by VXers	9
3.1.1 Virus.gr	9
3.2 Test results influenced by money / Association ID referrals	9
3.2.1 TopTenReviews, 6starreviews & No1reviews	9
3.3 Tests run by users and / or unexperienced peoples	10
3.3.1 Malware-Test	10
3.3.2 Consumer Reports	10
3.3.3 Tests based on Multi-engine scanner sites or samples from honeypots	10
4. Conclusions	11

## 1. The importance of independent antivirus testing

Without independent anti-virus testing sites, users would not know which solutions protects their computer better than others or which ones better suit their needs. Several institutions have made it their goal to deliver independent test results to the public, investing lot of time and money, as well as gaining over the years the necessary expertise to be capable of addressing the complex task of anti-virus testing.

## 2. Trustworthy testing bodies

Not many trustworthy and completely independent testing and/or certification institutions exist. The ones that are currently known to me, and which you can trust, are:

- AV-Comparatives.org ([www.av-comparatives.org](http://www.av-comparatives.org))
- AV-Test GmbH ([www.av-test.de](http://www.av-test.de))
- CheckVir ([www.checkvir.com](http://www.checkvir.com))
- ICSA Labs ([www.icsalabs.com](http://www.icsalabs.com))
- VirusBulletin ([www.virusbtn.com](http://www.virusbtn.com))
- West Coast Labs ([www.westcoastlabs.org](http://www.westcoastlabs.org))

They all differ a bit in what they test and which methods and test-sets they use, but in general the results are usually quite similar. Usually you can compare the results of the following tests together:

*Tests with large test-sets:* AV-Comparatives.org  $\leftrightarrow$  AV-Test GmbH

*Tests based on ITW-samples:* VirusBulletin  $\leftrightarrow$  CheckVir

*Certification bodies:* ICSA Labs  $\leftrightarrow$  West Coast Labs

It is anyway suggested to look not only at one test, but to compare various test results together and draw a conclusion based on them.

### 2.1 Independent Testers with comprehensive test-sets

AV-Test GmbH and AV-Comparatives.org both use large test-sets, containing in the range of 1 million samples. They do not limit their test-sets and add every functional piece of malware which can be and/or has been encountered in the real world. „To be effective, a detection test has to be comprehensive in its approach“ and not contain only samples listed on the WildList.

On the other hand, it is also important to measure when a sample gets detected, therefore it has to be combined e.g. with retrospective testing too.

### **2.1.1 AV-Comparatives.org**

Those tests are done by Andreas Clementi and his team. AV-Comparatives.org has been conducting anti-virus tests for many years and have made their findings publicly available on their website since 2004. They have very strict testing rules and test at regular intervals (every 3 months) using 16-18 of the best home user products with detection rates over 85% under the currently most used operating system. The test methodology and the detailed test reports/results are available free on their website. They deliver very comprehensive tests including (but not limited to): on-demand detection tests on large malware collections, retrospective/proactive tests to measure the capability of the anti-virus products to protect against new/unknown malware, polymorphic virus detection tests, false alarm tests, scanning speed test, etc. From time to time they also release other special tests. The results are published without big delay soon after the tests are completed. AV-Comparatives give out three awards: Standard, Advanced and Advanced+.

<http://www.av-comparatives.org>

### **2.1.2 AV-Test GmbH**

Those tests are done by Andreas Marx and his team. AV-Test GmbH has been conducting anti-virus tests (and tests on other security products like firewalls, etc.) for many years at regular intervals and publishing the results in various online and printed magazines, on behalf of the anti-virus vendors or magazines. AV-Test GmbH is the biggest anti-virus test centre in the world and delivers very comprehensive tests under various platforms, including (but not limited to): outbreak tests, packer/archive tests, on-demand and on-access tests on large collections, ITW-tests, scanning speed, impact on system performance, etc. Their test methodology is published, but is not available for free. Results can be found in most popular computer magazines, but are currently not available on their website.

<http://www.av-test.de>

*AV-Comparatives.org and AV-Test GmbH do both deliver transparent test results and use accepted test methods, allowing to the anti-virus vendors to cross-check and verify the results, e.g. by sending them the samples they missed.*

## **2.2 Independent Tests based mainly on the WildList**

The WildList (<http://www.wildlist.org>) contains viruses that are widely spreading worldwide. The WildList is released monthly, but usually with a delay of some months. This means that anti-virus vendors usually have some months to detect those viruses that are reported on the WildList. The WildList consists mainly of viruses and worms, although it does not include some of today's significant threats like Trojans and similar malware, which is nowadays often found in the field. It is also important to note that some vendors have access to the WildCore collection while other vendors (and AV-Comparatives) don't, which makes tests based on replicated samples of the contained samples somehow biased. Any respectable product should be able to pass ITW tests.

### **2.2.1 VirusBulletin**

The VB100 tests are run by John Hawes, and take place every other month on a different platform. Anti-virus products are tested against the VB test-sets, the main part of which is a set of samples of viruses listed on the WildList. Other test-sets, including a polymorphic set, are used for the test, but for a product to pass the test and get the famous VB100 award, products are only required to detect all ITW-viruses, both on-demand and on-access, and to not raise any false alarms against the clean set of files used by Virus Bulletin. Detailed test results, including speed measurements, are published in the monthly Virus Bulletin magazine, which has an annual subscription fee of \$175. The magazine also contains other news, analysis and interesting articles from the anti-virus field. The raw results of the VB100 tests (pass or fail) are available for free on the website (a short, free registration is needed). Past issues with full review details are also available on the website.

<http://www.virusbulletin.com>

### **2.2.2 CheckVir**

These tests are done by Ferenc Leitold and his team. CheckVir performs on-demand and on-access tests against infectious threats contained on the WildList (about 80% consists of viruses listed on the last three WildLists and a maximum 20% of viruses from any WildList). To get the STANDARD qualification a product must detect all viruses from the test-set used. To get ADVANCED qualification a product must additionally be able to repair the infected files. <http://www.checkvir.com>

### **2.3 Certification bodies**

Certifications are very important – they set the standard which has to be met in order that products can be used in some areas. Having a certification is important for the anti-virus vendors, but is not so meaningful for home users. An anti-virus vendor explained in a forum how certifications done by some certification bodies works:

*THE CONTENT IN THIS REPORT HAS BEEN REMOVED (IMMEDIATELY AFTER PUBLICATION) ON REQUEST OF A CERTIFICATION BODY (WE DO NOT TELL WHICH). ACCORDING TO THEM, THE INFORMATION WAS INACCURATE. TO AVOID PROBLEMS, WE AGREED TO REMOVE THE CONTENT.*

Due to the reasons you do not see here anymore, in my opinion, home users who want to know if their anti-virus product protects their PC's well against samples listed on the Wildlist, should check out the results of VirusBulletin and CheckVir, as they test only once and show to the public also failed outcomes.

There are many specific certification awards (like Spyware, Trojan, Cleaning, Firewall, etc.) given by the following two certification bodies, but we focus here on the main anti-virus certifications. Both ICSA Labs and West Coast Labs provide Wildlist testings also.

#### **2.3.1 ICSA Labs**

„The ICSA Labs Anti-Virus Product Certification program and the criteria modules do not contain specific requirements for products or product components that eliminate other non-replicating, malicious software. Therefore the Anti-Virus Certification Criteria does not include requirements for handling malicious and non-malicious spyware, adware, backdoors, trojan horses, and other such non-replicating software.“ In order to pass the ICSA Labs Anti-Virus Product Certification program, the tested product has to detect (on-demand and on-access) 100% of the malware listed on the WildList and 90% of other „Zoo“ samples. However, it is unknown how many samples are included in „Zoo“ test-set. ICSA Labs shows only who passed the test. The website does not show who failed the test or how many submissions were needed until the product successfully passed the test criteria.

<http://www.icsalabs.com>

### **2.3.2 West Coast Labs Checkmark**

The certification process is done by Chris Thomas and his team. West Coast Labs – like ICSA Labs – is an independent organization which tests anti-virus products. „For a product to be certified Anti-Virus Checkmark Level 1, the product must detect all viruses on the WildList.“ „All products registered for the Anti-Virus Checkmark Level 1 will be tested against the "In the Wild" list which was published not less than two months prior to the product release date.“ „On the first occasion on which a product is tested, West Coast Labs conducts a preliminary test free of charge. All tests from that time on are charged at the agreed rate. Retesting of a product which has failed to pass the test is charged at the agreed rate.“ „For a product to be certified to Anti-Virus Checkmark, Level Two, the product must be able to comply with Checkmark Level One and, in addition, disinfect all viruses in the wild which are capable of being disinfected.“ There are only about one dozen viruses in the wild which are capable of being disinfected. West Coast Labs shows only who passed the test. The website does not show who failed the test or how many submissions were needed until the product successfully passed the test criteria. <http://www.westcoastlabs.org>

### **2.4 Tests in magazines**

I was not sure if I should list tests of magazines under „trustworthy tests“, because what can be trusted, in my opinion, is only the test results. It depends if the data has been delivered by independent testing bodies (usually AV-Test GmbH or AV-Comparatives.org) or based on the magazine's own in-house tests. Magazines often review also other features like price, colors of the GUI, usability, perceived system impact, etc. and then give a final score by mixing anti-virus detection results with the subjective data. Due to that, you may see some magazines with identical anti-virus test results but with different winners at the end. All in all this is logical, as the magazines have to sell you something new. One bad thing about magazine tests is that when the printed magazine is released, the results are already about three months old. The worst case in magazine reviews is when the tests are performed on lousy test-sets by the reviewers themselves, as the test-sets are usually too small, wrongly selected and not well-maintained (they often contain a lot of garbage files).

## **2.5 Academic tests**

Academic tests are going to be rare nowadays. Here are listed the most known ones, although none are active currently.

### **2.5.1 antiVirus Testing Center (aVTC) University of Hamburg**

Those tests were done by Klaus Brunnstein and his students (including, at one time, Vesselin Bontchev). It had a ten year of activity (from 1994 to 2004), but the tests stopped in 2004 as Prof. Dr. Klaus Brunnstein is now too busy with other projects. The tests were very detailed and had good methodologies. However, the main goal of the tests were to show to the students how to perform tests and use valid test methods, and not to deliver real-time test results to use to compare anti-virus solutions. The scanners were not updated the same day and test-sets contained also some garbage (which is normal in large test-sets). Therefore the results should not be used to evaluate anti-virus software. aVTC will remain the first reference of comprehensive anti-virus testing with very good test methods.

<http://agn-www.informatik.uni-hamburg.de/vtc>

### **2.5.2 Virus Research Unit at the University of Tampere**

Those tests were done by Marko Helenius. The Virus Research Unit has not recently conducted anti-virus product evaluations. However, it is their stated intention to continue the analyses in the future if resources allow.

<http://www.uta.fi/laitokset/virus>

### **2.5.3 Test Lab at Moscow State University**

The Test Lab at Moscow State University aimed to do tests under a wide range of typical real-life situations. Their results were never published on the Web. Unfortunately, the project now seems to have been terminated.

### **2.5.4 University of Magdeburg**

This was taken over by the AV-Test.de GmbH company some years ago.

<http://www.av-test.de>

### **2.5.5 University of Innsbruck**

This was taken over by the AV-Comparatives.org project some years ago.

<http://www.av-comparatives.org>

### **3. Non-trustworthy tests and / or flawed tests**

Unfortunately, there are many flawed tests floating around on the Net. Some are flawed on purpose, some others due to negligence. Here are the most known ones:

#### **3.1 Tests run by VXers**

A VXer is usually a virus collector who exchanges virus samples with other unknown people with the goal of increasing his own collection. Most VXers are just collectors and have no experience in analyzing malware.

##### **3.1.1 Virus.gr**

Is maintained by a VXer known as VirusP (Antony Petrakis). VirusP releases tests every half year. The samples are not checked for functionality, the products are not updated the same day/moment and vendors have no chance to verify the validity of the results. The selection of the samples is done by using anti-virus scanners. Additionally, those products which are used for virus trading or that have many unique virus names in their databases are favoured and influence the results. For several years it does not appear as though VirusP has made attempts to improve his tests adequately.

#### **3.2 Test results influenced by money / AssociationID referrals**

##### **3.2.1 TopTenReviews, 6starreviews & No1reviews**

An anti-virus vendor explained quite well how such „reviews“ work: „[...] this is the way some people earn money. If you carefully scrutinize the "Buy Now" links for the top 5 rated products, you'll see that they all have an Affiliate ID. Which means that the author is getting like 20% commission for each transaction realized via this link [...] The author sends an email to 15 AV companies, asking them to become an affiliate partner. Some of them respond, some don't. Then he stitches up a web site with more or less junk info [...], comparing the various products on the market, but paying special attention to placing those that he has affiliation for on top places.“ So each year you see the same results, the person behind it just changes the year part of date and takes care that the products for which he earns more money are placed on top. Additionally, the tables contain wrong and outdated information about the products. No one should rely on or even visit such sites.

### **3.3 Tests run by users and / or unexperienced peoples**

On various forums and sites you can read about anti-virus tests done by various users. Unfortunately, you can not rely on such tests, due the usual small sample size, the non-analyzed samples (test-sets contain much garbage) and because you do not know who is really behind the test (it could be someone which works for an anti-virus company). The same applies to tests done by journalists who only conduct sporadic tests from time to time.

#### **3.3.1 Malware-Test**

Malware-test uses samples collected from a honeypot, without analyzing the samples for functionality. The products are not updated the same day and products are run with different settings and detections are counted wrongly, delivering different results (up to 12% different) for products which should score equally if the tests are done properly. All in all, a completely flawed test.

#### **3.3.2 Consumer Reports**

In 2006, ConsumerReports published a completely flawed test based on 5500 self-created modifications of malware and tried to sell their flawed test method as the only way to properly test the ability of anti-virus products to find new malware. In reality, they ignored established and well-documented test methods that are used e.g. by AV-Test GmbH and AV-Comparatives.org to test exactly this capability (retrospective testing). The moral of this story: do not trust test results just because they come from a well-known magazine. Magazine reviewers do not have the needed skills and equipment to be able to make unbiased tests. They should stick to comparing prices etc. and leave the work of testing detection to more experienced, established and independent testers.

#### **3.3.3 Tests based on multi-engine scanner sites or samples from honeypots**

Some problems with such tests are the following:

- samples gathered from honeypots or submitted to such sites are very often corrupted and are not verified for functionality
- The majority of the samples are spyware samples or other tools
- The sample size is too small and not really randomly selected
- The settings used by the multi-engine scanner sites differ from the ones used in the home user products

## **4. Conclusions**

The best thing to do is not rely only on one test result, but to check various independent testing sites over a period, draw conclusions about detection rates and then combine them with your own experience from using trial versions, taking into account resource usage, GUI, compatibility, etc. Additionally, when you read tests/reviews, you should ask yourself the following questions:

1. How big is the sample size? What does the sample group represent?
2. Who is behind the anti-virus test? Has/Have she/he/they enough experience/knowledge/resources to perform such tests?
3. Is the test-set free of garbage and other things like tools, etc.? Did the tester verify the samples in any way?
4. Are the results reproducible, can they be verified? Do the anti-virus vendors get the missed samples after the test in order to be able to verify the results?
5. Which settings were used? Were all product tested using the same settings?
6. When was the test performed? Are the results already outdated?
7. Were the products tested under the same conditions and were they all updated at the same time/day?
8. How were the products tested? Is the test methodology known/published and generally accepted by other anti-virus researchers and vendors?
9. Has the tester any financial interest in the ranking outcome of the test?
10. What is the goal of the test? What was tested? What do the results tell?

### **Copyright and Disclaimer**

This publication is Copyright (c) 2007 by AV-Comparatives. Any use of the results, etc. in whole or in part, is ONLY permitted after the explicit written agreement of AV-Comparatives, prior to any publication. AV-Comparatives and its testers cannot be held liable for any damage or loss which might occur as result of, or in connection with, the use of the information provided in this paper. We take every possible care to ensure the correctness of the basic data, but a liability for the correctness of the test results cannot be taken by any representative of AV-Comparatives. We do not give any guarantee of the correctness, completeness, or suitability for a specific purpose of any of the information/content provided at any given time. No one else involved in creating, producing or delivering test results shall be liable for any indirect, special or consequential damage, or loss of profits, arising out of, or related to, the use or inability to use, the services provided by the website, test documents or any related data.

Andreas Clementi, AV-Comparatives (April 2007)