

AV-Comparatives e.V.



Testing Methodologies & Frequently Asked Questions

Note: this document is currently undergoing an update

Language: English

Last Revision: April 2009

www.av-comparatives.org

Table of Contents



Testing Methodology	3
Sorting procedure	4
Test Lab Security	6
Sources of samples	7
Conditions for participation	9
Contact point	13
About AV-Comparatives	14
Awards given by AV-Comparatives	15
Frequently Asked Questions (FAQ)	17
Copyright and Disclaimer	26

Please read this whole document carefully before asking anything or making incorrect assertions about the tests and the tested products. If you are still unsure about something after reading it, please ask us!

Testing Methodology

1. The operating system is installed on a PC and updated to the latest available service pack, including important updates.
2. An image and a snapshot utility will be installed. Some other utilities may be installed, based on what kind of test is going to be performed and which tools are needed (for example, tools to track changes made to the system, etc.).
3. An image of the operating system is created and cloned to other (identically specified) PC's.
4. The operating system is configured and the anti-virus product installed on the PC using default product settings.
5. An image for each PC is created and saved to an external hard disk.
6. When the test starts, all (anti-virus) products are updated at the same time, and images are actualized too. After that, PC's are disconnected from the internet and isolated. (Depending on the test method and product, an active connection or a simulated internet may be made available).
7. The set of samples (clean set, malicious samples test-set, etc.) needed for the actual test is introduced.
8. The products are tested according to the test scope (for example, first with default settings, and then with the most paranoid settings).
9. Missed samples and any faults discovered are sent to the anti-virus vendors (in accordance with the test conditions). This applies to the published main tests.
10. Vendors get some weeks to peer-review the results and – if needed – we correct the results before we publish them.

The images of the anti-virus products of the February and August tests will also be used for the retrospective test. The images are also needed in order to be able to check at any time the same scenarios as were used during the test-phase.

When the PC's are not used for testing, they are also used for sandboxing / sample analysis.

Sorting procedure

1. Samples from all sources are copied to the incoming server.
2. Encrypted and archived samples/collections are decrypted and extracted from archives.
3. Duplicate samples are weeded out.
4. File names are renamed to make sorting and maintenance more effective.
5. File extensions are renamed by a tool created in-house to its correct executable extension. Unrecognized file formats are given the extension “.VIR” and are moved to a separate location (for further inspection).
6. Samples are analyzed, using various tools (commercial tools, for example, but also tools used and maintained by the anti-virus community) in order to recognize known garbage or non-working samples. We also use several other static analyzers, PE parser, and so on, including our own in-house tools.
7. Most known adware, hacker and virus tools, components, hoaxes, jokes, virus simulators, commercial software, constructors, keygens (key generators), cracks, key loggers, engines, sniffers, unviable (bad, corrupted, inactive, damaged or intended) samples, virus source code, various garbage and disputed files, and so on. Basically, files and gray-area samples that should not be included in the main test-sets – are sorted out. Working adware, spyware, etc. is maintained separately for future tests based on such types of threat.
8. All PE malware is analyzed by a sandbox developed by people working at AV-Comparatives, and also by various commercial sandboxes, in order to exclude non-working samples and other garbage. Non-PE malware is also checked by some automated tools, but usually they need to be checked manually, as are some PE files that our sandbox was not able to categorize reliably. Viruses are verified by replication, but we do not always use the replicated samples for the tests – we use some of them to check whether viruses were added by the vendors with reliable accuracy, or whether some vendor only added some checksums in order to detect replicating viruses. The latter case may be considered as unacceptable by us and can lead to exclusion of the product concerned. If a file doesn't seem viral or malicious we don't include it. Instead, we move it to the “unwanted” database. (We also do this even if, for example, all anti-virus programs report the file as being infected – this means we don't rely on anti-virus programs to select which samples to include in the test-set, and we advise any other testers not to do that either). Our test-sets do not contain samples that do not work under Microsoft Windows NT/2000/2003/XP/Vista. Old macro samples (prior to Microsoft Office 97) are not included either. In addition, we no longer include compromised HTML files.
9. Verified samples are sorted into the various categories we use; as this task is often tricky, we also use (for example) VGrep to see how anti-virus vendors would classify a sample (e.g. as a backdoor or worm). Sorting is based on the majority verdict. For example, if most products classify a malicious program as a backdoor and one product classifies it as a worm, we classify it as a backdoor too. There are only a few exceptional cases where we do not agree with the way the majority of products classify some malware and in that case our own classification will be applied. In case of replicating or polymorphic malware, we take care not to include a disproportionate amount of the very same variant, in order to avoid flawed results. This is also a reason why our test-sets often are “smaller” than others.
10. All samples are at some point validated. As automated systems (not to mention humans, especially students...) are not fool-proof, it can nevertheless happen that grey-area or totally inappropriate files also slip in (but they do get removed later from the sets).

11. We freeze the February and August test-sets, usually a few days before the test starts, which means that many files which have not been fully analyzed by automated tools or by humans are also included in the test-set. While the tests are already running we continue to check the recently added samples too, and remove any bad samples from the test-set afterwards. As the vendors will also receive all samples they missed in the meantime, they may also get some bad samples, but they will be removed before the end of the test and not counted as misses in the published report (and vendors have some weeks to report faults and bad samples).
12. After the tests, we look again to see whether there are any samples that were not detected by any product. Usually we find 2-3 files that are indeed not detected by any product, and on examination those files always turned out to be bad samples. We therefore decided that samples determined to be undetected by all tested products will be removed from the test-set, and will not be counted as misses in the test actually performed (since they are garbage).
13. In the testing month, we focus our analysis on the samples that were missed by the tested products. We start from those samples that were missed by most products, as they have a higher probability of being non-working.
14. Files reported as bad by vendors will be removed, and the results will be corrected before they are published on the website. Due to the (approximately) two weeks (peer-) reviewing procedure, we are also able to include in our sets fresh malware, and to analyze the samples even when the tests are already started. This also gives vendors the opportunity to report back testing faults or inappropriate samples, though they are not obligated to do so. This all helps to ensure that in the end we publish correct results for our readers. Anyway, since we commenced this methodology in research published at the begin of this year, some bad samples may still be in the test-set, but considering the size of the test-set, they should be so few, that they have practically no significant effect on the results and no discernable impact on the rankings or awards given. Should we ever find out in our QA that the error margin was higher than anticipated, or high enough to have an impact on a ranking or award, we will publish that information.



Test Lab Security

All the databases are encrypted by PGP and parts of them are at least hard encrypted by RAR3. The only person which can decrypt the files is the Chairman. One encrypted backup of the databases is kept in a highly secured building in Munich.

Only fully-trusted AV-Comparatives staff members get access to the samples for the purpose of analyzing them on protected, isolated systems. The room containing the workstations carrying unencrypted malware sets (as happens during the test period) is fully secured, under video control¹ with motion detection and alarm systems directly connected to the police and a private security service. Additionally, the area is checked several times at day and night by an external security service, in order to avoid unauthorized access. All media containing malicious software are clearly labelled as such.

AV-Comparatives sends (missed) samples only AFTER the main tests in February and August to trusted representatives of vendors whose products were publicly tested. We do not send any samples to unknown/untrusted vendors/individuals, no matter what they say or offer. We at AV-Comparatives consider malware as dangerous and take countermeasures to avoid any endangerment to the public (e.g. by any possibility that it gets into the wrong hands).

People wishing to submit malware collections to AV-Comparatives, should encrypt the archives by using the public PGP key available at <http://www.av-comparatives.org/clementi.key>



¹ <http://securitycam03.av-comparatives.org>

Sources of samples

AV-Comparatives have various sources from which it obtains samples. Like anti-virus vendors, we also use various **traps** and **honeypots** from all over the world, as well as samples downloaded from **malware downloaders** and **infected websites**. Furthermore, we get samples from the field which were collected by us or our partner companies (e.g. computer repair/cleaning services) on infected PC's belonging to home users and/or small/medium business companies. We also get samples from various **online scanning services** and (single and large) submissions from visitors² to our website, as well as **various organizations** that collect malware (internal and public security forums, honeypot projects, anti-malware initiatives, and so on). In order to have a test-set that is statistically valid and as large and representative as possible, AV-Comparatives also accepts samples from (security) **vendors**. Currently, samples submissions from about a dozen vendors are included in our tests and nearly dozen more vendors which are not included in our tests also contribute.

Any vendor is encouraged to send us samples they get from their customers, but no vendor is obliged to. While we are not going to disclose the names of the vendors which submit or do not submit their samples (partly because Non-Disclosure Agreements may apply), we can assure you that submitting samples to AV-Comparatives does not help a vendor to get a better score. As the test-set consists of samples from many various sources and vendors, a single vendor's contributions just make the test set more representative – in fact, there are some vendors who do not submit anything and score very highly, and some other vendors who submit a lot are at the bottom regarding detection rates. The reason for this may be that samples are usually shared between vendors anyway and most of the samples we get are usually already in some other collections, so it is impossible to tell how much is coming from which individual source and so on.

We also prefer not to disclose this information because of the possibility that some vendors may use it to mislead the public for PR reasons (this has happened several times in the past, for example when a vendor was unhappy with some test results or wanted to put pressure on a tester) or focus on specific sources. As we've said, any vendor is welcome to submit us their samples if they wish to. Last-minute submissions (especially "extraordinary" collections) from vendors are not accepted; this source of samples is usually frozen 2-3 weeks before the test starts, in order to avoid possible bias.

AV-Comparatives does not create, modify or repack any malware (for testing purposes or for any other purpose).

² In future AV-Comparatives will add a malware submission form on its website, so users can submit to us samples online instead by email

Sources of clean files:

CD's and DVD's from various **magazines** from various countries (mainly German, Italian and English computer magazines) and **well-known software** (incl. most downloaded software from some legal download sites). Main source for the clean sets are PC's owned by individual users and various (mainly European) SMB companies (maintained by our partner kompetenzzentrum.IT) which allowed us to use in our clean sets (without sensitive data). We also have access to the content of the application **servers** of an university (without personal data). Duplicates are weeded out and files keep their original file names.



Conditions for participation

Which products are to be tested is decided by the board of AV-Comparatives e.V. - AV-Comparatives prefers to include in its tests *only* anti-virus products with good detection rates. The product must use its own or licensed engines. The product must be able to finish a scan of the full database using the most secure possible detection settings within a reasonable time, without crashing or causing major problems. Products must be able to scan a subdirectory tree (depending from the type of test). The scanner should not move or change in any way the files or the system during the scan when running in report-only mode. The product should be a well-known anti-virus product used worldwide and should not produce too many false positives. The below TOS is an example and mainly applies to the main tests which results get published quarterly.

Additionally, the following Terms of Service agreement has to be accepted and signed:

Terms of Service for Anti-Malware Software Testing

This document contains Terms of Service (hereinafter referred to as "TOS") for Anti-Malware Software Testing by AV-Comparatives which are applied to tests performed by AV-Comparatives e.V. or its representatives (hereinafter referred to as "the Tester").

1) Test Methods.

The methods used by the Tester are described in a document published on the Test center website www.av-comparatives.org. The Tester reserves the right to improve and/or change the methods as necessary. Notice of such changes will be published on the www.av-comparatives.org website at least 30 (thirty) days before they take effect. Agreement with changes notified is implied by continuing to participate in testing, subject to terms in (2.)

2) Participation.

Any vendor of security software (hereinafter referred to as "the Vendor") has the right to decide whether to participate in tests performed by the Tester. If the Vendor decides to participate in tests performed by the Tester, the Vendor is obliged to send an application for inclusion in testing to the Tester by email or by fax. The application will contain notice that the Vendor accepts this TOS and the current methods published and used by the Tester. Furthermore the application shall be dated and signed by the authorized representative(s) of the Vendor and stamped by the Vendor's seal, or provided on official headed notepaper where a seal is not available. Applications will not be accepted without an authorized signature. The Vendor is obliged to deliver the original of the application to the Tester by first class business mail within 14 (fourteen) days after the delivery via email or fax. The Application shall remain in force until revoked by written notice to the Tester. Whether or not to test a product shall remain at the Tester's sole discretion.

3) Software, License Keys.

The Vendor is obliged to provide a full working product version and all necessary license keys to the Tester upon request. The Vendor is obliged to supply the Tester with the name of a person responsible for contact with the Test centre. The Tester shall not distribute the product or license

keys provided for testing purposes to any third party. Upon completion of testing, Tester shall return the software to the Vendor or certify in writing that all copies of the software have been destroyed. The Tester shall neither display the Vendor logos without specific written permission, nor use the Vendor's name or trademarks in a manner that implies endorsement by the Tester or the av-comparatives.org web site.

4) Fees.

The Vendor (or a third party) has to pay a fee for the various services provided (e.g. usage of logo in marketing material and time/work spent in providing the various services, etc.). The fee has to be paid quarterly after the tests are finished and already published.

5) Sample Submission.

The Tester will accept submissions of monthly collections from the Vendor. The Tester will not accept samples from the Vendor if the Vendor does not wish the Tester to send any missed samples to other participating vendors that are already getting all missed samples.

6) Restricted Distribution of Samples.

The Vendor may request that the Tester restricts distribution of samples to certain other vendor(s) where there is an issue of trust. The Vendor is obliged to identify clearly the other vendor(s) to which the Vendor wishes such a restriction to apply. The Tester will review such request individually, and after review by the Tester the Vendor will be informed as to whether the restricted distribution of samples will be applied. The Tester suggests that in such case the Vendor does not submit further samples for the duration of the review period. If the Vendor is not satisfied with the outcome, the Vendor may decide to do not send samples or to discontinue sending samples to the Tester. In cases where an issue of trust arises against the Vendor and a review by the Tester shows the concern to be valid, the Tester will provide a limited number of missed samples from any test, at the sole discretion of the Tester.

7) Missed Samples.

The Vendor must have an established virus lab in order to be entitled to receive missed samples after the on-demand tests. The Tester will provide the missed samples to the Vendor only if the Vendor's product is successfully able to identify a given minimum of the Tester's actual full test set during an on-demand scan with the most secure settings.

The Tester will provide missed samples to the Vendor so that the Vendor can verify the validity of the test results. The Tester will send samples missed by the Vendor's product, unless the Vendor is subject to a restricted distribution of samples as described in the section above (RESTRICTED DISTRIBUTION OF SAMPLES). If a distribution restriction has been applied, the Vendor will receive a limited number of samples selected by the Test center together with a list of missed samples in form of log with CRC32 checksums and - where possible - detailed reasons on why the restriction had to be applied, in order that the remaining missed samples can be requested from other vendors or located among samples in the Vendor's own lab.

8) Liability Limited.

The Tester will undertake to perform all tests with due care, according to the published methodology, and will make all reasonable efforts to ensure the correctness of the results. However, the Tester cannot be held liable for any inaccuracies which may occur. The Tester makes no warranty, express or implied, with regards to the test results, and disclaims all implied warranties of merchantability, title and fitness for a particular purpose. In no event shall the tester be liable, whether in contract, tort (including negligence) or otherwise, for any indirect, incidental or consequential damages (including, but not limited to, lost, savings or profit, lost data or business interruption even if test center is notified in advance of such possibility). This included damages incurred by the Vendor, the Vendor's customers or any third party. This limitation protects the Tester.

9) Product Exclusion reserved.

- I. The Tester reserves the right to exclude any product from testing. The Vendor will be informed about reasons leading to the product exclusion should such exclusion occur. The Tester reserves the right to publish the reason for product exclusion, but this will be done only if the Tester considers it absolutely necessary.
- II. Reasons for product exclusion may include, but are not limited to:
 - a. Knowingly providing samples to virus writers, or to any un-trusted party or to party without acceptable need, experience or discretion to handle samples in a safe way.
 - b. Engagement by the Vendor in illegal practices or practices generally considered harmful to the anti-virus industry or the general public.
 - c. Practices designed to deliberately bias, or lead to wrong, test results.

The Tester reserves the right to decide to allow the Vendor to continue to participate in the tests even if one of the reasons under (II) has occurred, but in this case the Vendor will receive only a limited number of missed samples or no samples, at the Tester's discretion.

10) TOS breach.

The Tester reserves the right to decide whether the Vendor is in breach of this TOS and how to proceed against the Vendor should such a breach occur.

11) Right to change TOS reserved.

The Tester reserves the right to change this TOS in future. Notice of changes will be published on www.av-comparatives.org website at least 30 (thirty) days before they take effect. Agreement with changes notified is implied by continuing to participate in testing, subject to terms in (2.)

12) Choice of Law.

THIS TOS SHALL BE GOVERNED BY AND CONSTRUED IN ACCORDANCE WITH THE SUBSTANTIVE LAWS OF THE REPUBLIC OF AUSTRIA, WITHOUT REFERENCE TO CONFLICT OF LAW PRINCIPLES. ANY ACTION SHALL BE INITIATED AND MAINTAINED IN A COURT OF COMPETENT JURISDICTION IN INNSBRUCK, AUSTRIA. The Tester and the Vendor irrevocably consent to the personal and venue of state courts within the selected jurisdiction.

13) Miscellaneous.

This document constitutes the entire and exclusive agreement between the Tester and the Vendor with respect to the subject matter hereof and supersedes all other communications, whether written or oral. This document and entry (acceptance) in electronic form, or a hardcopy duplicate in good form, shall be considered an original document with authenticated signature admissible into evidence unless the document's authenticity is genuinely placed in question. Any provision found by a tribunal of competent jurisdiction to be illegal or unenforceable shall be automatically conformed to the minimum requirements of law and all other provisions shall remain in full force and effect. Waiver of any provision hereof in one instance shall not preclude enforcement of it on future occasions. Headings are for reference purposes only and have no substantive effect.

In some cases changes to the TOS are possible on vendors' request.

Contact point

Anyone can contact us through the contact form available on our website at <http://www.av-comparatives.org/seiten/contact.php>

Mails sent through the contact form usually get answered within one day (mean response time over the past four years was usually under one hour). Using the contact form is a secure way to avoid being marked as SPAM. We also get notified about the incoming message through the contact form by SMS, so please make sure to choose the right sender (*HomeUser, AV company, Magazine reviewer*) and to give a short but meaningful subject, as well as typing in your email address correctly. Please be aware that mails that need elaborated/long answers or mails that need (e.g.) board decisions may need more time to get an answer. Mails with inappropriate/unknown language or nonsense mails will not receive a reply. You can write us in the German, English or Italian languages. Please make sure to check your SPAM folder in case you miss a reply.

People working for a software vendor or journalists should use their company email address when contacting us.

Vendors which products are or were tested by AV-Comparatives, or well-known journalists/people get an email address (not available for the public) and a phone (and cell phone) number in order to reach us directly at any time and get a fast reply.

The address of the AV-Comparatives main office is:

AV-Comparatives e.V.
Erzherzog-Eugen-Strasse 3
6020 Innsbruck
AUSTRIA – Europe



The Golden Roof in Innsbruck

In order to avoid communication chaos, the usual main contact persons are (the management board): Andreas Clementi (Chairman) and Peter Stelzhammer (Vice-Chairman). Please write directly to the board instead of to employers, volunteers or students if you need something.

We usually have various contact points with security software vendors and usually one main contact point for each vendor (this way miscommunication or loss of information is avoided). If possible, please ask first within your company who is responsible for communicating with us, or ask your company if you are allowed to take decisions on its behalf before mailing us.



About AV-Comparatives

Q: Does AV-Comparatives get paid by the vendors of the tested products for the services that AV-Comparatives provides?

A: Like almost every other testing organization, also AV-Comparatives get paid for its work. While in the past we provided various services for free, since 2008 we ask for a fee for the services we provide, as we need to cover our (past and current) expenses. As we are aware that some peoples will try to discredit our work just because we openly admit that we do not longer provide our services for free, we have some rules that our readers should know: AV-Comparatives is a Non-Profit-Organization.

Neither the Chairman and CEO of AV-Comparatives nor any other peoples in the organization are profit sharing employers. Employers (including the management board) get a fixed salary. All money that AV-Comparatives receive is invested into AV-Comparatives (e.g. hardware, software, memberships in anti-malware related organizations, conferences/travel expenses, infrastructure and office expenses, website/traffic expenses, etc.) and not used to make profits. This ensures that we do not get influenced by money. There is no positive correlation between the fee of the various services and the results, so payment of a fee does not influence results. We also provide internal tests for vendors. A fee is charged, for example, for usage of our logo and reprints of our results in marketing materials, and various other internal services. Vendors who contribute the full yearly fee flat-rate get special discounted rates for other internal services.

People that contribute their time to AV-Comparatives also do other IT-Security related work or are students (AV-Comparatives collaborates with several IT consulting companies and some local academic institutions). Anyone who is involved or is going to be involved in AV-Comparatives has to sign an ethical contract before he/she is accepted to work on malware-related tasks, as well as agreeing to do not write in the name of AV-Comparatives in the public media. Important decisions as well as public statements are done by the whole board and not only by one individual. It must be signed by at least two members of the board.

In the year 2009 we plan some public initiatives, to raise the public awareness about the risks in the internet and the whole IT and countermeasures. We also work together with the chamber of commerce to educate companies to avoid risks and implement standard security practices.

AV-Comparatives uses only renewable energy from hydropower.

Registration:

Registered seat: Innsbruck, Austria

Court of registry: Federal Police Directorate Innsbruck

ZentraleVereinsRegister-Nr.: 017712505

AV-Comparatives e.V. is a Non-Profit Organisation (NPO).

Awards given by AV-Comparatives

AV-Comparatives gives to each product which was tested in the major yearly tests an award according to its scores in the test of February, May, August and November. The given awards range through ADVANCED+, ADVANCED, STANDARD and no award.

Products that were able to reach the STANDARD award in a test report can be understood to be good products with a good score, those with ADVANCED got a very good score and those with ADVANCED+ an excellent score. Products which did not get an award may still need some further improvement to reach the tested goal. An overview of given past awards can be found on our website³.



Currently (August 2008) the rules for the awards are as follow (as test-sets and methods change, also the award systems need to be updated from time to time):

Test report of February and August (overall detection rate tests):

To get ADVANCED+, over 97% of the whole test-set have to be detected during an on-demand scan with best possible settings.

over 97%	ADVANCED+
93-97%	ADVANCED
87-93%	STANDARD
under 87%	NO AWARD

An updated award system which will also consider the false alarm rate will be introduced and applied in the tests of 2009.

Test report of May and November (retrospective tests):

To get the Advanced+ award, a product must be able to detect at least 50% of new malware proactively and at the same time have only few false alarms.

	0-10%	10-25%	25-50%	50-100%
none - few	NO AWARD	STANDARD	ADVANCED	ADVANCED+
many	NO AWARD	NO AWARD	STANDARD	ADVANCED
very many	NO AWARD	NO AWARD	NO AWARD	NO AWARD

* proactive detection rate vs. amount of false alarms

³ <http://www.av-comparatives.org/seiten/overview.html>

Other test reports (e.g. performance tests, etc.) may also be awarded.

False alarms are an important issue and need to be taken into account when looking at detection rates. That's why e.g. in the retrospective tests false alarms lead to lower awards.

Currently (as of August 2008) the labels for the amount of false alarms are given as follows:

none or very few	0 - 3
few	4 - 15
many	16 - 100
very many	101 - 500
crazy many	over 500

At the end of each year, products are allocated an award in a summary test report, where products are nominated in various tested aspects (overall detection rate, proactive detection rate, false alarm rate, scanning speed, etc.⁴). To be designated product of the year, a product needs to get better scores than other products in most of the various tests done during the year. The label "Best product of the year" indicates only that the product was better than other products in most tests provided during the year⁵. More details about the summary awards will be given in the December report.

Since this year (2008), vendors of products receiving awards in the summary reports will get a certification plaque to display, for example, in corporate offices.

In 2009 the award logos and website appearance will change.



⁴ We plan to add performance tests, dynamic tests and some other tests in future.

⁵ To know which product is best for you, please try out the software on your system. The "best product" for any user and any need and situation does not exist: we just tell you which products scored better than others in regard to some aspects of the software.

Frequently Asked Questions (FAQ)

1) I am a publisher/journalist and would like to use the test results that AV-Comparatives provide on its website. Do I have to pay something and what are the rules I have to follow?

You are allowed to use the published test results free of charge, but you should conform with the following rules:

- give the source (www.av-comparatives.org) and the date of when the test was performed (e.g. February 2008). You should always use the most recent test results
- it is suggested that you let us proof-read your article before you publish it, in order to be sure that the results are interpreted correctly and not misused
- we would like to know in which magazine etc. our results are going to be published

2) I am a publisher/journalist and would like AV-Comparatives to test some products for us to be published in a magazine or similar. Is that possible?

Yes, it is possible, but we do it very rarely. If you cover the additional testing expenses we can do the tests. But we suggest asking other independent labs first. A list of some well-known testing labs can be found in the links section of our website.

3) I am a website/forum owner and would like to post the full detailed results, screenshots of the results or host the test reports on our server. Am I allowed to do this?

No, you are not allowed to do this without written agreement from the management board of AV-Comparatives e.V. – please send us a short message to discuss ways in which you can do it (free of charge).

4) Where do AV-Comparatives get the tested product solutions from and how do AV-Comparatives ensure that the latest available updates are being used in the tests?

The products are usually either downloaded from the vendors' website or submitted by the vendor, by sending us a mail attachment or FTP/HTTP download address, along with the necessary license keys.

Products/Signatures get updated according to the user manual (usually by online update or, very rarely, by manually downloaded latest official updates). The various products get updated at the same time. As we are in contact with the vendors and let them know which update has been frozen and used for the test, you can be sure that we use the latest available updates. Please note that we use the latest available updates in the tests of February and August; in the retrospective tests, we use those updates again (in retrospective tests products are not updated).

5) On which dates were the tests conducted and when were the products updated?

All this information is included in every report. In the big yearly tests, the products get usually updated in the first week of February and in the first week of August. The tests are performed immediately after that and are concluded when the report is published on the website.

6) Are the products tested on virtual machines like (for example) VMware, or on real machines?

AV-Comparatives test the products on real machines.

7) Which versions and what settings of the products were used?

The product versions and used settings are mentioned in the reports. Most tests provided by AV-Comparatives are done with the highest, most secure settings. Tests with default detection settings are also provided and noted as such in the reports. AV-Comparatives usually include the paid versions of the home user standalone product, by agreement with the vendors. When a vendor prefers that another version is included or (e.g.) lower (less secure) settings are used, we do this and note it in the test reports.

8) Would the free product versions score different than the paid versions included in the tests?

The paid product versions often include more features, options and support than the free versions of the same vendor, but the engine and signatures are the same and would get nearly the same scores in most of our tests. Some products don't include adware/spyware detection in their free version, but as we do not include these types of badware in the sets, their free version would score as highly as the paid version in the on-demand tests against the malware test-sets.

9) On what workstations are the tests being performed?

The detection tests are currently performed on Intel Core 2 Duo E8300/2.83 GHz, 2 GB RAM workstations, completely identical for each tested product.

10) Can AV-Comparatives conduct tests for products that require a real, live internet connection?

Yes, products that require a live internet connection (e.g. for in-the-cloud technologies) are tested with a live internet connection. Currently our tests include only one product that uses in-the-cloud technology also during on-demand scans. A testing standard for in-the-cloud products is right now under development by AMTSO. When we test products with in-the-cloud technologies, we first test the products without an Internet connection (in order to get a baseline scenario) and then, separately, we execute a scan with highest settings over the missed malware when the ITC has an Internet connection. To avoid any time advantage for in-the-cloud products, we test those products before or at the same time when the other products get their last updates. To give a full picture, for such products we publish in our reports both results (with and without in-the-cloud).

11) Can you name me six other established testing institutions (apart AV-Comparatives) that you consider noteworthy?

Below are (in no specific order) six established and recognized companies that are noteworthy. Please note that they all provide different kind of tests and are independent from each other. We suggest that you do not rely on just one test lab only and to look instead to as much different testing organizations as possible to get an overview of products' capabilities and consistencies.

- Virus Bulletin (www.virusbtn.com)
- NSS Labs (www.nsslabs.com)
- ICSA Labs (www.icsalabs.com)
- West Coast Labs (www.check-mark.com)
- AV-Test (www.av-test.de)
- CheckVir (www.checkvir.com)

12) What file extensions are present in the test-set?

The majority (over 95%) of the files present in the current test-sets are PE files with EXE or DLL extension. For example, all files in the categories backdoors, trojans and windows viruses are PE files.

13) Who ensures that the tests performed at AV-Comparatives are scientifically and statistically valid?

AV-Comparatives e.V. collaborates with local academic institutions, which provide us scientific consulting and supports us by holding courses for interested students who, for example, try to develop tools for us to automate some types of testing.

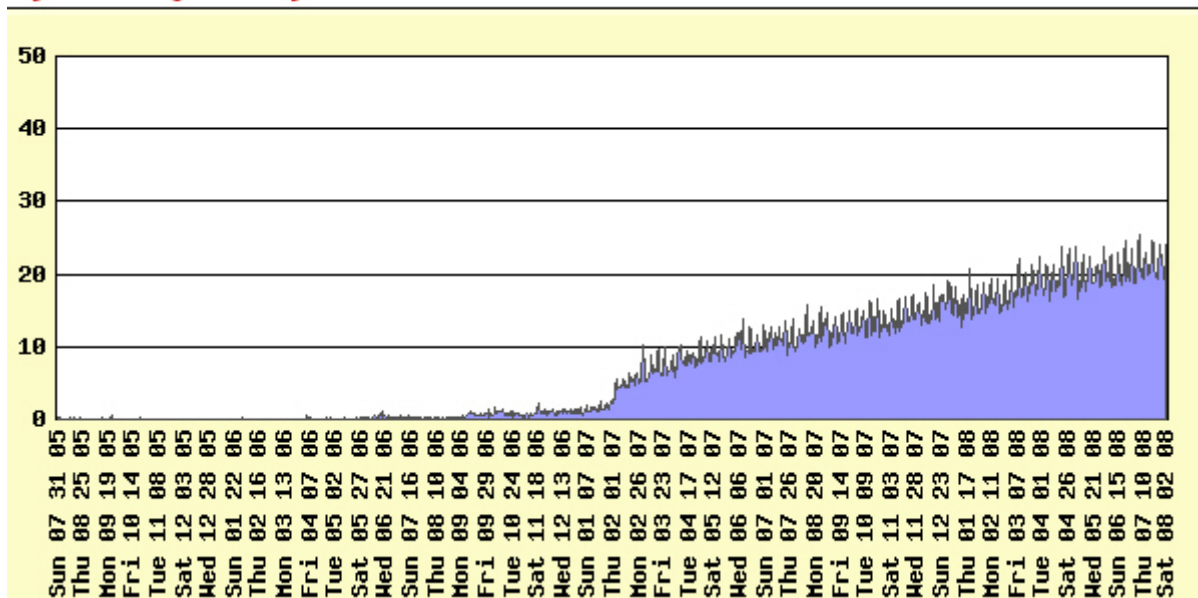
14) What is a retrospective test?

The retrospective test (which is performed on-demand) is used to test the proactive detection capabilities of scanners. It does give an idea of how much new malware a scanner (compared to other scanners) can detect (for example by heuristic/generic detection), before a signature is provided for the malware.

The on-demand detection test results are usually applicable also for on-access protection, but not for on-execution protection. In order to measure on-execution protection, dynamic tests are needed.

15) Under which operating system are the tests currently usually performed, and why?

About 75% of our website visitors still use Windows XP and only about 20% are currently using Vista (see graph below; updated begin of August 2008). As soon as the majority of the users use Vista, we will test under Vista (or whatever the most used operating system after XP will be).

By Percentage of Daily Total

16) Will AV-comparatives provide in future only dynamic tests, and dismiss all other tests?

No. We will start to provide dynamic tests in 2009, in collaboration with local academic institutions. Several products already provide technologies whose full capabilities can be evaluated only by providing dynamic tests (where live threats are introduced to the system by its normal infection vectors and with a live simulated internet connection).

As AV-Comparatives provide comparisons of Anti-Virus products, products which offer already those technologies and pass the dynamic tests will be recompensed, but the other products will be not penalized. Due the complexity of dynamic tests, the test-sets used will probably be small. Also, the important false alarm rate evaluations will be limited.

Such (dynamic) protection tests are very important and necessary, but they do not replace the extensive testing of detection tests. They are just another important aspect that needs to be evaluated. We will include the dynamic tests as complementary to the other tests, not as a replacement. This way, any user, regardless of how he uses, prefers or configures his anti-virus products, will benefit from the evaluation data.

17) Are clean files reported by products as suspicious due to the use of packers also counted as false alarms in the false positive tests?

Clean files reported as suspicious (without any further explanation or misleading labeling) due to the use of packers are also counted as false positives, because in the malware test-set files reported as suspicious are also counted as detections and we have to be consistent over the various tests to be fair and not to mislead anyone. Some products report information about packers (not real detections), which are not counted as detections or as false alarms.

18) What does SET A and SET B mean in the on-demand tests?

We think that for users it may be more interesting to know how well the various products compare to each other in regard to detection of recent malware.

Due to the fact that SET B contains malware that has been around for the last few (currently nine) months and SET A contains malware that was mainly around before, SET B is usually covered very well by any anti-virus product. To get a "PASSED" classification against SET A, a product must detect over 97% of that test-set. If 97% is not reached, it gets a FAIL and the reached percentage in parenthesis.

The below picture shows what SET A and SET B in the August 2008 test looks like:



In future the periods will probably be shortened further. The most interesting SET on which readers should look is SET B, which contains actual malware. The awards are mainly given/based on scores in SET B (but to get ADVANCED+ a product also needs to score PASSED against SET A).

By reducing/splitting the test-set, it should be clear that percentages may seem lower than they really are. Example:

- 1 SET A + SET B = 1000000 samples
Product X detects in total 900000 samples (90%)
- 2 SET A = 500000 samples, SET B = 500000 samples
Product X detects ~100% of SET A, and 80% of SET B.

80% is lower than 90%, but in fact the detection rate remained the same. Please keep this in mind before jumping to any wrong conclusions. This is also one of the reasons why we always remind readers that we provide COMPARATIVES where product performance is compared to other products. Percentages/data alone may confuse some readers, so it is better to rely on the awards we give, as they already combine various factors and are easier to compare. Basically we suggest considering the results as ordinal scaled and not metric.

19) Why don't you give a false alarm percentage rate?

Giving a percentage for false alarms is in our opinion senseless and highly misleading. We give instead details about which products gave more false alarms than other products and on which files the false alarms occurred.

20) Do AV-Comparatives provide tests about the performance of anti-virus products (e.g. system resource impact, etc.)?

Yes.

21) Is the product with the lowest score in your test a worse product?

No. The products included in the main tests are already a selection of very good security products. Even the products at the bottom of the e.g. 16 tested products are still good products, but were surpassed by other products in the specific test due to better results.

22) I heard AV-Comparatives also tests products with in-the-cloud technologies. How does that work?

In-the-cloud tests require a live internet connection to the vendor's server, where the data gets analyzed. We provide also tests for such technologies, but we are aware that black sheep exist in the AV industry too, and it would be quite easy to cheat in such tests.

Therefore, (also due to possible privacy risks to users) we look at what data is transmitted to the servers: if the data is encrypted, we want to know the decryption key as part of the NDA. We will not test or comment on products which we know to pose a privacy risk or where the vendors were caught cheating.

23) What kind of malware is included in the "other malware" category?

It contains mainly Rootkits, but also some Exploits, DDoS, Flooders, Sniffers and Nuker tools, as well as some IRC malware and macro Trojan droppers.

24) Can I conclude that if an Anti-Virus product is not tested by AV-Comparatives it is a bad product that was not tested because it is bad?

You cannot draw any conclusions about a product just because it is not represented in the tests, as there are several various possible reasons for that, of which low scores constitute only one. For example, a vendor could, in general, wish not to be tested by independent testers, might not have a product ready (or in accordance with some of the requirements included in the TOS), or maybe the number of participants already reached the limit and we had to postpone its inclusion.

25) Do AV-Comparatives use the GUI or the command line versions of the products in its tests?

We use the GUI versions to generate the test result data. Should we need to use the command line version (and base findings on that alone) it will be noted in the report (for example, as results could vary considerably due to the use of command-line scanners).

26) Why is product X not included in the test?

In order to make a selection of which scanners to include in the tests, some conditions are stipulated. Scanner X probably did not fulfill some of these conditions. Another possibility is that the developers don't want to be tested by us. The tested products are just a selection, and we don't complain that it is all-inclusive. So please do not continue to ask why product xyz was not tested. We maintain the right to choose which products to include or exclude from the tests.

27) Are adware, spyware, potentially unwanted/dangerous applications, etc. also included in the test-sets of AV-Comparatives?

No, such forms of badware are currently intentionally not included in the AV-Comparatives test-sets. Based on the large quantity of such badware and due to users' interest in knowing detection rates, AV-Comparatives may in future provide a separate test using such potentially unwanted software.

28) When is the next test to be released on the website?

The major test reports are usually published online during March, June, September and December. Other tests or single-product tests are delivered from time to time, so you need to check for their release in the Comparatives section of our website. The yearly Summary Report is usually released in Mid-December.

29) How many bad samples are in the test-set and do they have an impact on the given awards?

Before the results are published, the test-set used is reviewed again and lot (several thousands) of inappropriate samples are removed. Furthermore, vendors are able to report bad samples. The test-set on which the published results are based on is, at the end, in a quite good 'clean' state and any remaining inappropriate files have a very low impact on the results. We did in the past run a QA⁶ over the August 2007 test-set (on which the published results were based) and as expected, there was no impact on the given awards and ratings. Should there be in future any significant impact on given awards, we will publish and note mistakes we discover retrospectively.

⁶ <http://www.av-comparatives.org/seiten/ergebnisse/QA2007.pdf>

30) Who is the AV-Comparatives delegate who attends conferences and meetings outside of Austria?

The founder of AV-Comparatives is the usual delegate, as he already knows most people working for security software vendors and he is one of the main contact points. Meetings in Austria are usually attended by at least the members of the management board of AV-Comparatives. Vendors are also welcome to visit us in our main office for a meeting (several representatives of various vendors already visited us in our new office this summer).

31) What happens if a vendor participates in the major tests but then decides to drop some tests (which are part of the major tests)?

Vendors can not drop from the big tests during the year. Should it happen anyway that a vendor “prohibits” us from publishing the results, we will not publish the results, but the non-participation will be considered as a “no-award” (as it could be that a vendor does not wish to get results published because its product would have scored miserably).

32) Why do AV-Comparatives not provide tests based on the Wildlist?

We do not choose to criticize the current state of the Wildlist and yet at the same time continue to provide tests based on it. As soon as the Wildlist gets reorganized and more meaningful, we may consider providing tests based on it (if we are allowed to). In the meantime you will find enough other testers who provide test results based on the Wildlist.

33) Are the samples used by AV-Comparatives samples that exist only in labs and a no time posed a risk to users?

No, the era where “zoo” samples were submitted by the malware authors only to labs, and which therefore existed only in labs and collections, ended years ago. Nowadays malware is created by criminals for financial reasons and they do their best to avoid their samples ending up in labs. Almost all malicious samples can nowadays be considered to be, at some point in time, in the field and posing risks to users. In past the argument that most samples exist only in labs was used by some vendors if their product scored low and they told their users that what they do not detect is not important.

34) Do AV-Comparatives offer re-testing to vendors or are the published results based on products’ first run?

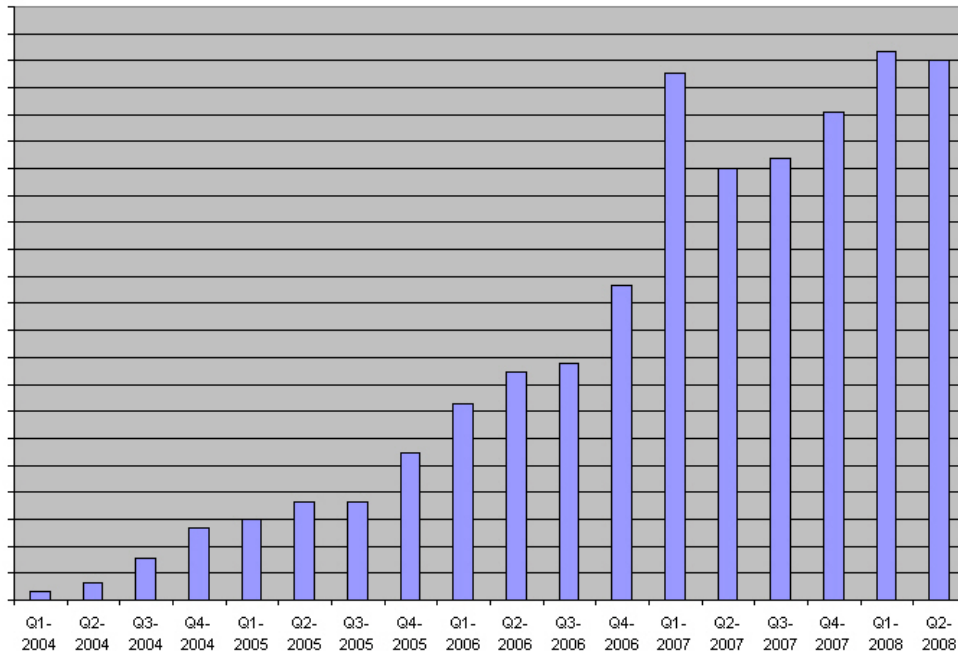
No, AV-Comparatives does not allow re-testing in case of bad results – published results are equally applicable to all vendors based on a product’s score in its first try (irrespective of its results). AV-Comparatives also does not provide so called “sponsored reviews”, where only the good features of a product are praised. Vendors that want to get tested by us should keep in mind that we publish negative aspects of products, too, and publish results even if a product scores low.

35) Does AV-Comparatives use any third party results (e.g. from other testers) in its reports?

No, we do not use third party results, mainly because in the case of third party results, we cannot guarantee the accuracy and independence of the results. We can reproduce and allow to be reproduced (e.g. by the University of Innsbruck or even vendors) only tests that we did by ourselves.

36) How popular is the AV-Comparatives.org website?

AV-Comparatives started as an insider tip in late 2003 and got more and more popular over the past years. Nowadays, AV-Comparatives is considered to be one of the most popular and frequently referenced anti-virus testing websites. Below we show a graph of our growth in popularity, divided into three monthly periods:



37) We are an Anti-Virus vendor and would like to put a survey on the AV-Comparatives website. Is this possible?

Yes, it is possible, but only under some conditions:

- We do not want to get paid for linking to your survey, but we want our visitors to have a chance to win something if they participate in the survey. So, you should randomly pick out some (up to 3, depending on the value of the price) participants and give them a gift. Licenses for your product are not considered as a gift.
- The survey should be of general interest and not specifically for your product (i.e. vendor independent questions), as well as approved by us. It is not necessary to disclose the vendor during the survey, but it must be disclosed when the survey is finished.
- The survey should not be too long (that is, a maximum of 20 questions).
- The survey can stay on the website for not longer than eight weeks. Surveys will not be linked during November, December and January. The link will usually be placed on the main site and in the weblog.
- There can be a maximum of three surveys on the website per year and only one survey per year per company.
- AV-Comparatives would like to have the permission to (re)publish the outcome of the survey.
- The results of the survey should contain a disclaimer, as well as the source of the visitors (e.g. that the survey was filled out by visitors of the AV-Comparatives website).

38) Why do AV-Comparatives limit the number of participants in the main tests?

We prefer to include in our tests only well-known products which fulfil some minimum requirements. If we were a normal company, testing more products would mean more money for us, but as AV-Comparatives is in any case a Non-Profit-Organization, we are more interested in and focused on providing good tests on a smaller number of products rather than (weak) tests on a large number of products.

39) Why do AV-Comparatives only use some few millions of samples in its tests?

This is for several reasons: mainly because we do not want to include in our test-sets adware, spyware, tools, dialers, components, garbage, old samples, or samples that do not work on current operating systems. We also do not want to skew the results by including (for example) 300000 samples of e.g. Netsky.q just because they have different MD5's. AV-Comparatives has over ten million samples (June 2008), but we do not focus on quantity, and that's why only a few million malware samples are used in our tests.

40) Can you please include some beta products in the major tests, too?

No, the products should be at least Release Candidates. Release Candidates are included only at the request of the vendor and with the agreement of AV-Comparatives. The tested engine/product must either already have been released as final in another of the vendor's products (e.g. the enterprise version) or be released as final before the test report gets published (usually within 5 weeks). It's only possible to include a release candidate for the tests done in the second half of the year. Beta products can be tested separately (and noted as such in the report) but not included in the major test reports. For instance it has happened in the past that a vendor wanted us to include their beta product in the test because it was going to be released as final soon, and then, when their product scored unfavorably they made the excuse that we tested a beta version and not the final version.

41) Does AV-Comparatives send out newsletters or similar?

Yes. To register for our newsletter, please go to www.av-comparatives.info

42) I would like to get some information not included in this document. I also found an error or outdated information in this document. What should I do?

Please send us your questions and we will decide whether we will include them in the next update of this document.

Copyright and Disclaimer

All information in this document and on the website is copyright protected © 2004-2008 by AV-Comparatives e.V. Any use of the content of this document, etc. in whole or in parts, is ONLY permitted with the explicit written agreement of AV-Comparatives e.V. representatives, prior to any use and publication. The authors cannot be held liable for the correctness of the content etc. given on this site or in this document. We don't give any guarantee of any kind. We are under no circumstances liable for any consequential damage including, but not limited to, capital/profit loss or other direct or indirect damage that could arise. We reserve the right to modify or deny access to this publication at any time and assume no responsibility for anything. No part of this publication may be reproduced in any form or stored in a retrieval system without the prior written permission of AV-Comparatives e.V. representatives. Mentioned products are trademarks by their respective holders.

AV-Comparatives e.V. (August 2008)