Microsoft

Transparency report

Examining AV-TEST results, January-February 2020

Prepared by

Microsoft Defender ATP Research Team

Table of Contents

1 Executive summary

This report provides a review of independent industry test results for Microsoft Defender Antivirus (formerly Windows Defender Antivirus), the next-generation protection component of Microsoft Defender Advanced Threat Protection (Microsoft Defender ATP), Microsoft's unified endpoint protection platform.

Over the last few years, Microsoft has been improving its performance in industry tests. Today, it consistently achieves high scores in these tests, demonstrating the strength of our protection capabilities and the innovations we continue to make in our security technologies.

While current antivirus tests don't necessarily reflect how attacks operate and how solutions are deployed in the real world, they can influence important business decisions. We are actively working with several leading industry testers to evolve security testing. Meanwhile, we're publishing this report to provide more details, insights, and context on test results. We'd like to be transparent to our customers and the industry about our wins as well as improvement plans because of these tests.

1.1 AV-TEST: Perfect Usability and Performance scores (January-February 2020)

Microsoft Defender Antivirus achieved perfect scores (6.0/6.0) in the Performance and Usability test module of AV-TEST's January-February 2020. The industry-leading antivirus solution has achieved this in all AV-TEST cycles.

In the Protection test module, Microsoft Defender Antivirus achieved a score of (5.5/6.0) in January-February 2020 Business User test cycle.

Learn More >>

2 Examining AV-TEST results

2.1 Summary of overall AV-TEST scores

The table below summarizes the overall test results for Microsoft Defender Antivirus in the January-February 2020 AV-TEST Business User test:

	January-February
Protection	5.5/6.0 (±0)
Usability	6.0/6.0 (±0)
Performance	6.0/6.0 (±0)

Table 1. Microsoft Defender Antivirus's overall antivirus test results in the January-February 2020 Business User test. AV-TEST uses Protection, and Usability, and Performance test modules.

2.2 Understanding Protection scores

Below are details on the Protection test scores.

	January-February
Real World testing	98.1% (394/402)
Prevalent Malware testing	100% (20,606/20,606)
Overall malware protection rate (all samples)	99.5% (20,798/21,008)
Overall Protection score >>>	5.5/6.0 (±0)
Overall Protection ranking >>>	2nd out of 15 (tied with 3 more)

Table 2. Summary of Protection scores for the January-February 2020 Business User test.

Microsoft Defender Antivirus detected 100% of malware files used in the Prevalent Malware in January-February 2020 test cycles from 20,606 files used.

The diagrams below show Microsoft Defender Antivirus detection rates in the Prevalent Malware and Real-World tests over one year. Microsoft Defender Antivirus achieved 100% in 12 out of the 12 monthly Prevalent Malware tests and 100% in 8 out of the 12 monthly Real-World tests.

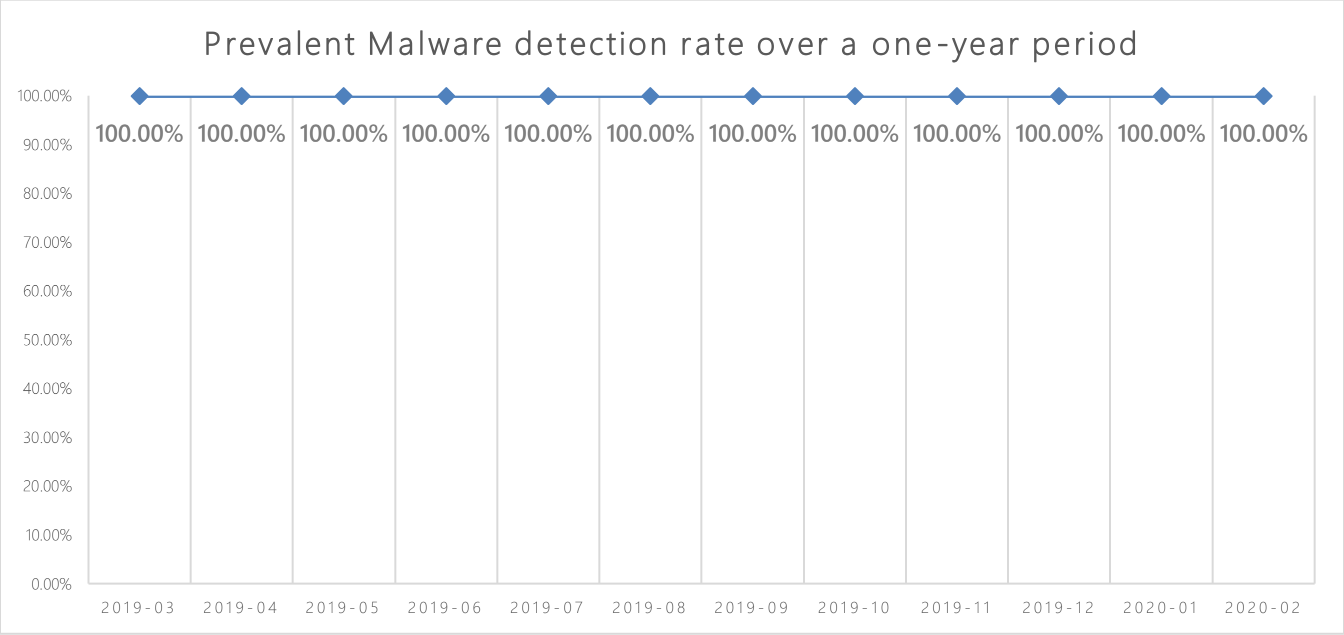

Figure 1.. Microsoft Defender Antivirus detection rates in AV-TEST "Prevalent malware" tests over one year

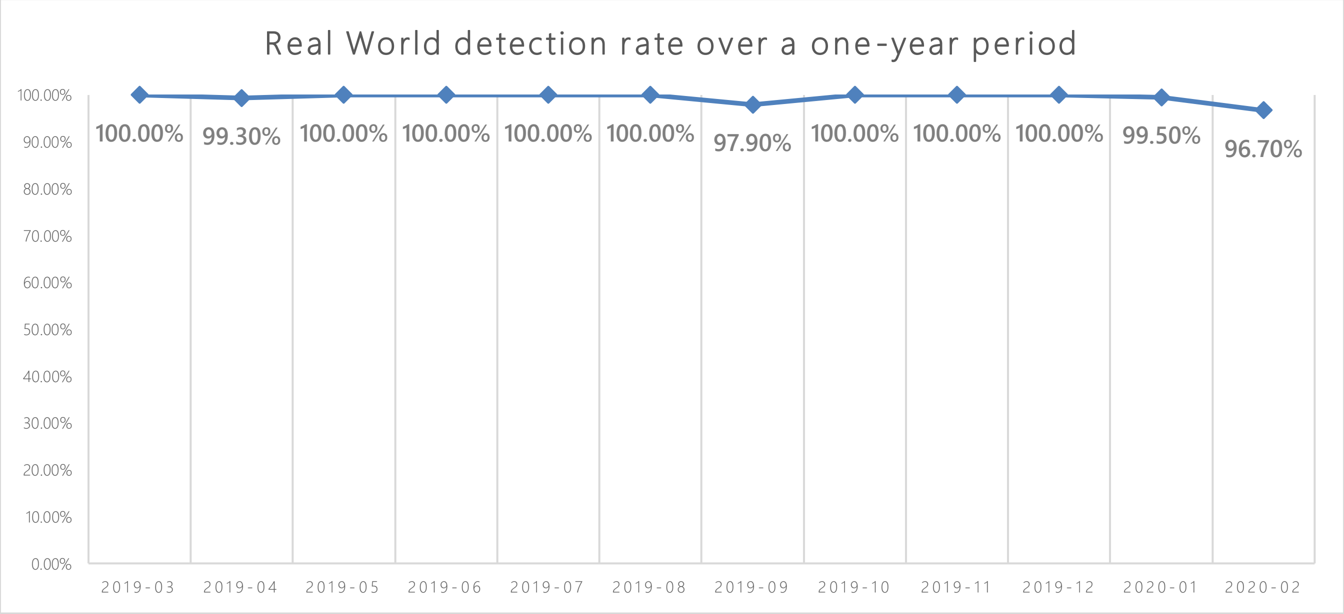

Figure 2.. Microsoft Defender Antivirus detection rates in AV-TEST "Real World" tests over one year

2.2.1 Missed samples are opportunities for improvement

The Microsoft Defender Research team takes missed samples as an opportunity to improve detection capabilities. For each missed sample, a team of researchers analyzes and assigns a proper label to the sample to make sure it is detected. Also, the team analyzes the root cause for the miss and drives long-term detection improvements and monitoring.

Microsoft Defender Antivirus missed seven samples in this cycle. Below is the analysis of the misses and the improvements that were introduced as a result:

Missed Sample	Missed Reason	Improvement made
Sample 1-7	• Signal misclassification	• Improve cloud-based protection for this threat category

Table 3. Improvements made to Microsoft Defender Antivirus in response to this cycle's results

2.3 Understanding Usability scores

In Usability tests, AV-TEST includes clean file samples in the test population and checks whether antivirus products incorrectly classify them as malware (what is known as false positive, or FP). Below is a summary of the results for Microsoft Defender Antivirus in the Usability test.

	January-February
Number of misclassified files	1 (out of 9,65,259 samples)
Overall Usability score >>>	6.0/6.0 (±0)
Overall Usability ranking >>>	1st out of 15 (tied with 14 more)

Table 4. Summary of Usability test scores for the January-February 2020 Business User test

2.3.1 Analysis: What kinds of files were misclassified?

Our research team analyzed the sample that Microsoft Defender Antivirus misclassified and assigned proper determination. The team also analyzed the root cause of these misclassifications and worked with threat research teams to enhance detection accuracy.

Below is an example of a file that Microsoft Defender Antivirus misclassified in the test cycle. Based on our research and file prevalence data, the misclassified sample is not common in enterprise environments.

Sample	File prevalence (30 days)	Description	Digitally signed? (Y/N)
Sample 1	200	Full-fledged word processor application for the Tamil language	N
Sample 2	100	A word processor application for regional (Tamil) language	N

Table 5. Files that Microsoft Defender antivirus incorrectly classified as malware during January-February 2020 Business User test

Microsoft encourages software vendors to take steps to raise the level of trust both by security vendors and users alike. These steps include signing software with certificates issued by reputable Certification Authorities.

2.3.2 The synthetic nature of usability tests

Misclassifications in a synthetic test are not necessarily indicative of false positives in real-world scenarios. This is true when the test methodology discounts contextual elements that Microsoft Defender Antivirus uses for issuing a verdict. For example, when a file is tested, it is not downloaded from the vendor website. Both the original file name and the download site are contextual information that is removed in tests. We've seen many cases where a customer in the real world downloads a clean program from the vendor site without encountering any erroneous detection. However, when a tester gives the file a seemingly random name (e.g., its SHA-256 hash), removes the mark of the web, and doesn't download the file from the vendor website, some of our more aggressive machine learning models issue blocks that don't occur in the real world.

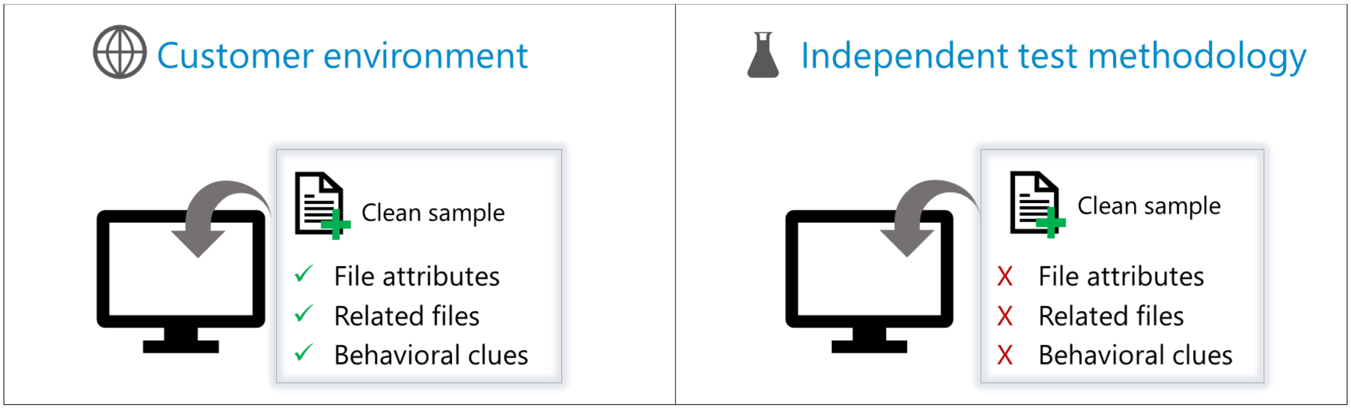

Figure 3.. In some cases, samples are incorrectly classified (false positive) in the synthetic test environment but not on customer machines.

2.3.3 Criteria for evaluating files may vary across vendors and testers

The criteria for classification can vary between antivirus vendors and testers depending on their policies. Some files identified as clean by some vendors could be files that Microsoft Defender Antivirus identifies as a potentially unwanted application (PUA) and thus would be blocked. Microsoft's policy aims to protect customers against malicious software while minimizing the restrictions on developers. The diagram below demonstrates the high-level evaluation criteria Microsoft uses for classifying samples:

- Malicious software: Performs malicious actions on a computer.
- Unwanted software: Exhibits the behavior of adware, browser modifier, misleading, monitoring tool, or software bundler
- Potentially unwanted application (PUA): Exhibits behaviors that degrade the Windows experience
- Clean: We trust that the file is not malicious, is not inappropriate for an enterprise environment, and does not degrade the Windows experience

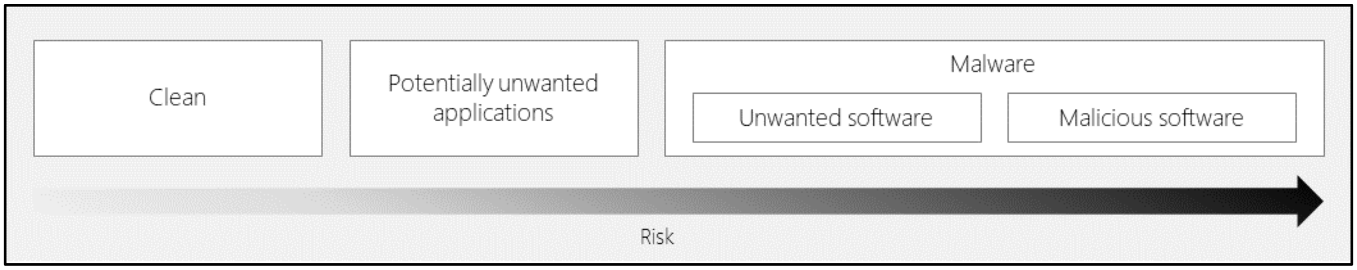

Figure 4.. Microsoft's high-level sample classification criteria

2.4 Understanding Performance scores

The table below summarizes Performance test results.

	January-February
Overall Performance test score >>>	6.0/6.0 (±0)
Performance ranking >>>	1st out of 15 (tied with 5 more)

Table 6. Summary of Performance test scores for the January-February 2020 Business User test

The table(s) below presents Microsoft Defender Antivirus's performance test results compared to industry averages during the January-February 2020 test cycle. Performance is measured by the average impact of the product on computer speed; therefore, a smaller number is favorable. Green boxes indicate areas where Microsoft Defender Antivirus performed better than or the same as the industry average; red boxes indicate performance lower than the industry average.

Operation*	**Standard PC**	*Industry average*	**High-End PC**	*Industry average*
Launching popular websites	8%	*19%*	7%	*15%*
Downloading frequently used applications*	2%	*2%*	0%	*1%*
Launching standard software applications	13%	*16%*	8%	*11%*
Installation of frequently used applications	24%	*26%*	22%	*23%*
Copying of files (locally and in a network)	0%	*6%*	0%	*11%*

Table 7. The average impact of the product on computer speed in daily usage January-February 2020

*The description for these operations is given by AV-TEST and might not be aligned with what Microsoft's data indicates as realistic.

2.4.1 Areas that matter the most to customers

Microsoft Defender Antivirus performed better than the industry average in most areas and had a limitation in the area that AV-TEST labels as "*Installation of frequently-used applications.*" There are several factors to consider for driving the right conclusion out of these test results:

- **Consider the frequency of the action**
  Most users in enterprise environments are information workers whose common user activities include:
  - Browsing the web
  - Using email clients
  - Processing documents
  - Accessing network resources

  Users spend substantially less time installing new applications compared to the activities listed above. This is true for all user segments, but especially for enterprises, where software installation is usually governed by usage policies. Microsoft Defender Antivirus is optimized for delivering high levels of performance during high-frequency actions. Performance is a priority area for the Microsoft Defender Antivirus team, and we're working to improve it even further.

- **Consider the level of risk**
  Microsoft Defender Antivirus is designed to perform thorough scanning during the software installation process. This could have a performance cost. One reason for this is that software installation is a relatively complex operation that touches different areas of the operating

system. A thorough inspection is necessary to reduce the risk of introducing malicious software on the system.

- **What impactful areas are not being tested?**
  Several areas are not being tested for performance by AV-TEST that are critical to user experience. Examples include:
  - Shutdown and startup
  - Universal Windows app launch
  - Battery consumption