SE Labs

INTELLIGENCE-LED TESTING

SMALL BUSINESS ENDPOINT PROTECTION

APR - JUN 2019

SE Labs tested a variety of anti-malware (aka ‘anti-virus’; aka ‘endpoint security’) products from a range of well-known vendors in an effort to judge which were the most effective.

Each product was exposed to the same threats, which were a mixture of targeted attacks using well-established techniques and public email and web-based threats that were found to be live on the internet at the time of the test.

The results indicate how effectively the products were at detecting and/or protecting against those threats in real time.

MANAGEMENT
Director Simon Edwards
Operations Director Marc Briggs
Office Manager Magdalena Jurenko
Technical Director Stefan Dumitrascu

TESTING TEAM
Thomas Bean
Dimitar Dobrev
Liam Fisher
Gia Gorbold
Pooja Jain
Ivan Merazchiev
Jon Thompson
Dave Togneri
Jake Warren
Stephen Withey

IT SUPPORT
Danny King-Smith
Chris Short

PUBLICATION
Steve Haines
Colin Mackleworth

Website www.SELabs.uk
Twitter @SELabsUK
Email info@SELabs.uk
Facebook www.facebook.com/selabsuk
Blog blog.selabs.uk
Phone 0203 875 5000
Post ONE Croydon, London, CR0 0XT

SE Labs is BS EN ISO 9001 : 2015 certified for The Provision of IT Security Product Testing.

SE Labs is a member of the Microsoft Virus Information Alliance (VIA); the Anti-Malware Testing Standards Organization (AMTSO); and the Messaging, Malware and Mobile Anti-Abuse Working Group (M3AAWG).

AMTSO Standard reference:
https://tinyurl.com/sel2019q2

CONTENTS

Introduction 04
Executive Summary 05
1. Total Accuracy Ratings 06
Small Business Endpoint Protection Awards 07
2. Protection Ratings 08
3. Protection Scores 10
4. Protection Details 11
5. Legitimate Software Ratings 12
5.1 Interaction Ratings 13
5.2 Prevalence Ratings 14
5.3 Accuracy Ratings 14
5.4 Distribution of Impact Categories 15
6. Conclusions 15
Appendix A: Terms Used 16
Appendix B: FAQs 16
Appendix C: Product Versions 17
Appendix D: Attack Types 18

Document version 1.0 Written 12th July 2019

INTRODUCTION

The best security tests keep it real

Why it's important not to try to be too clever

Realism is important in testing, otherwise you end up with results that are theoretical and not a useful report that closely represents what is going on in the real world. One issue facing security testing that involves malware is whether or not you connect the test network to the internet.

The argument against this approach is that computer viruses can spread automatically and a test could potentially infect the real world, making life worse for computer users globally. One counter argument goes that if the tester is helping improve products then a few dozen extra infected systems on the internet is, on balance, worth it considering there are already millions out there. The benefits outweigh the downside.

Another counter argument is that viruses such as we understand them from the 90s are not the same as they are today. There are far fewer self-replicating worms and more targeted attacks that do not generally spread automatically, so the risk is lower.

Connecting to the internet brings more than a few advantages to a test, too. Firstly, the internet is where most threats reside. It would be hard to test realistically with a synthetic internet.

Secondly, for at least 10 years most endpoint security products have made connections back to management or update servers to get the latest information about current threats. So-called 'cloud protection' or 'cloud updates' would be disabled without an internet connection, effectively reducing the products' protection abilities significantly. This then makes the test results much less accurate when running assessments.

There are cases in which turning off the internet is useful, though. Last year we ran a test to check whether or not artificial intelligence could predict future threats. We ran our Predictive Malware Response Test without an internet connection to see if a Cylance AI brain, which had been built and trained three years previously, could detect well-known threats that had come into existence since then. You can see the **full report here**.

But that was a special case. When assessing any security product or service for real-world, practical purposes, a live and unfiltered internet connection is probably a useful and even necessary part of the setup. Naturally we have always used one in our testing, at one point even going as far as using consumer ADSL lines when testing home anti-malware products for extra realism. When reading security tests check that the tester has a live internet connection and allows the products to update themselves.

If you spot a detail in this report that you don't understand, or would like to discuss, please contact us via our Twitter or Facebook accounts.

SE Labs uses current threat intelligence to make our tests as realistic as possible. To learn more about how we test, how we define 'threat intelligence' and how we use it to improve our tests please visit our website and follow us on Twitter.

This test report was funded by post-test consultation services provided by SE Labs to security vendors. Vendors of all products included in this report were able to request early access to results and the ability to dispute details for free. SE Labs has submitted the testing process behind this report for compliance with the AMTSO Testing Protocol Standard v1.0. To verify its compliance please check the AMTSO reference link at the bottom of page three of this report or **here.**

Executive Summary

Product Names

It is good practice to stay up to date with the latest version of your chosen endpoint security product. We made best efforts to ensure that each product tested was the very latest version running with the most recent updates to give the best possible outcome.

For specific build numbers, see ***Appendix C: Product Versions*** on page *17*.

EXECUTIVE SUMMARY			
Products Tested	Protection Accuracy Rating (%)	Legitimate Accuracy Rating (%)	Total Accuracy Rating (%)
Kaspersky Small Office Security	98%	100%	100%
ESET Endpoint Security	98%	100%	100%
Symantec Endpoint Protection Cloud	98%	100%	100%
Microsoft Windows Defender ATP's Antivirus	96%	100%	99%
Sophos Intercept X	95%	100%	99%
Bitdefender Gravity Zone Endpoint Security	94%	100%	99%
McAfee Endpoint Security	96%	98%	98%
Trend Micro Worry Free Security Services	94%	97%	97%

Products highlighted in green were the most accurate, scoring 85 per cent or more for Total Accuracy. Those in yellow scored less than 85 but 75 or more. Products shown in red scored less than 75 per cent.

For exact percentages, see *1. Total Accuracy Ratings* on page *6*.

■ **The endpoints were generally effective at handling general threats from cyber criminals...**

Most products were largely capable of handling public web-based threats such as those used by criminals to attack Windows PCs, tricking users into running malicious files or running scripts that download and run malicious files.

■ **... and targeted attacks were prevented in many cases.**

All products were competent at blocking more targeted, exploit-based attacks. This was an uncommonly strong showing from the group.

■ **False positives were not an issue for the products**

All of the endpoint solutions were good at correctly classifying legitimate applications and websites. They allowed all of the legitimate websites and applications with the exception of **Trend Micro** and **McAfee's** products, which each blocked one.

■ **Which products were the most effective?**

Products from **Kaspersky Lab**, **ESET** and **Symantec** achieved extremely good results due to a combination of their ability to block malicious URLs, handle exploits and correctly classify legitimate applications and websites. Unusually in this rigorous test, all of the other products also performed well enough to achieve a AAA award.

1. Total Accuracy Ratings

Judging the effectiveness of an endpoint security product is a subtle art, and many factors are at play when assessing how well it performs. To make things easier we've combined all the different results from this report into one easy-to-understand graph.

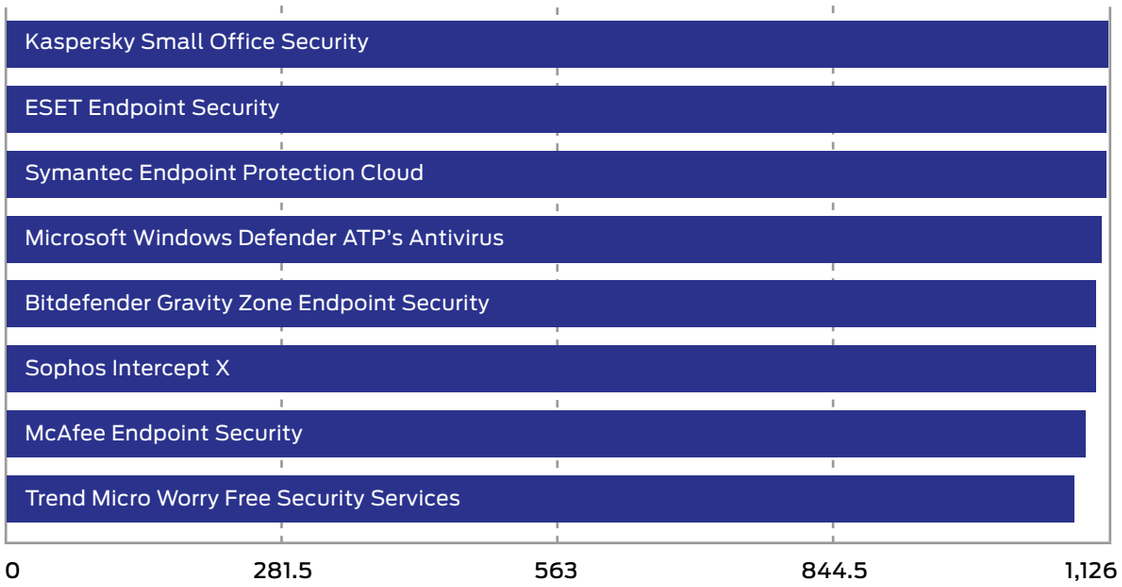
The graph below takes into account not only each product's ability to detect and protect against threats, but also its handling of non-malicious objects such as web addresses (URLs) and applications.

Not all protections, or detections for that matter, are equal. A product might completely block a URL, which stops the threat before it can even start its intended series of malicious events. Alternatively, the product might allow a web-based exploit to execute but prevent it from downloading any further code to the target. In another case malware might run on the target for a short while before its behaviour is detected and its code is deleted or moved to a safe 'quarantine' area for future analysis. We take these outcomes into account when attributing points that form final ratings.

For example, a product that completely blocks a threat is rated more highly than one that allows a threat to run for a while before eventually evicting it. Products that allow all malware infections, or that block popular legitimate applications, are penalised heavily.

Categorising how a product handles legitimate objects is complex, and you can find out how we do it in ***5. Legitimate Software Ratings*** on page ***12***.

TOTAL ACCURACY RATINGS			
Product	Total Accuracy Rating	Total Accuracy (%)	Award
Kaspersky Small Office Security	1,125	100%	AAA
ESET Endpoint Security	1,124	100%	AAA
Symantec Endpoint Protection Cloud	1,124	100%	AAA
Microsoft Windows Defender ATP's Antivirus	1,117	99%	AAA
Bitdefender Gravity Zone Endpoint Security	1,111	99%	AAA
Sophos Intercept X	1,110	99%	AAA
McAfee Endpoint Security	1,100	98%	AAA
Trend Micro Worry Free Security Services	1,090	97%	AAA



Total Accuracy Ratings combine protection and false positives.

Small Business Endpoint Protection Awards

The following products win SE Labs awards:

- **Kaspersky** Small Office Security
- **ESET** Endpoint Security
- **Symantec** Endpoint Protection Cloud
- **Microsoft** Windows Defender ATP's Antivirus
- **Bitdefender** Gravity Zone Endpoint Security
- **Sophos** Intercept X
- **McAfee** Endpoint Security
- **Trend Micro** Worry Free Security Services

2. Protection Ratings

The results below indicate how effectively the products dealt with threats. Points are earned for detecting the threat and for either blocking or neutralising it.

Detected (+1)
If the product detects the threat with any degree of useful information, we award it one point.

Blocked (+2)
Threats that are disallowed from even starting their malicious activities are blocked. Blocking products score two points.

Neutralised (+1)
Products that kill all running malicious processes 'neutralise' the threat and win one point.

Complete Remediation (+1)
If, in addition to neutralising a threat, the product removes all significant traces of the attack, it gains an additional one point.

Persistent Neutralisation (-2)
This result occurs when a product continually blocks a persistent threat from achieving its aim, while not removing it from the system.

Compromised (-5)
If the threat compromises the system, the product loses five points. This loss may be reduced to four points if it manages to detect the threat (see Detected, above), as this at least alerts the user, who may now take steps to secure the system.

Rating Calculations

We calculate the protection ratings using the following formula:

Protection Rating =
(1x number of Detected) +
(2x number of Blocked) +
(1x number of Neutralised) +
(1x number of Complete remediation) +
(-5x number of Compromised)

The 'Complete remediation' number relates to cases of neutralisation in which all significant traces of the attack were removed from the target. Such traces should not exist if the threat was 'Blocked' and so Blocked results imply Complete remediation.

These ratings are based on our opinion of how important these different outcomes are. You may have a different view on how seriously you treat a 'Compromise' or 'Neutralisation without complete remediation'. If you want to create your own rating system, you can use the raw data from ***4. Protection Details*** on page ***11*** to roll your own set of personalised ratings.

Targeted Attack Scoring

The following scores apply only to targeted attacks and are cumulative, ranging from -1 to -5.

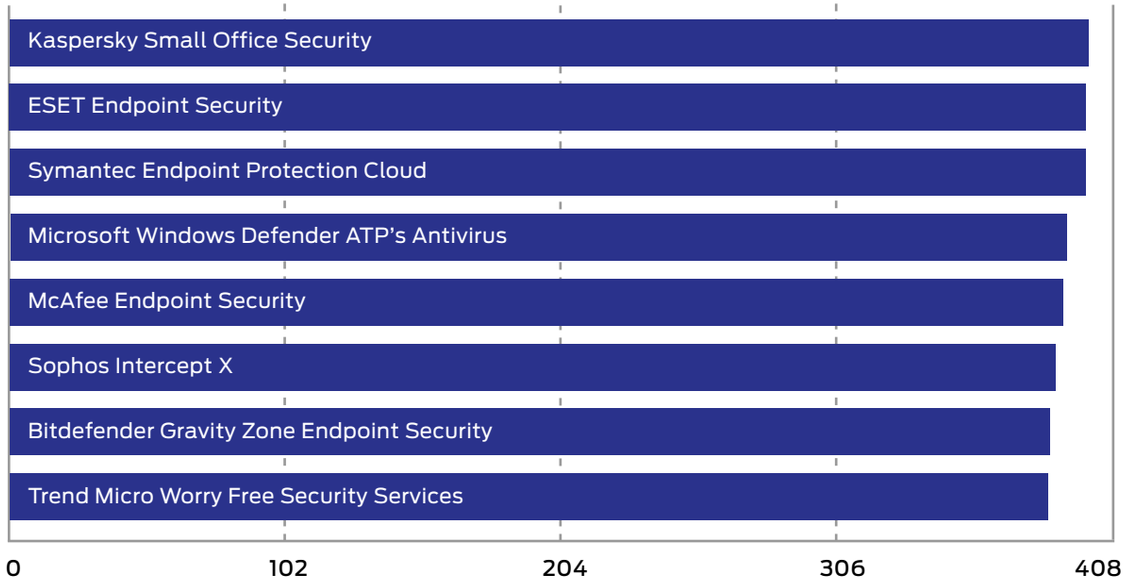
Access (-1)
If any command that yields information about the target system is successful this score is applied. Examples of successful commands include listing current running processes, exploring the file system and so on. If the first command is attempted and the session is terminated by the product without the command being successful the score of Neutralised (see above) will be applied.

Action (-1)
If the attacker is able to exfiltrate a document from the target's Desktop of the currently logged in user then an 'action' has been successfully taken.

Escalation (-2)
The attacker attempts to escalate privileges to NT Authority/System. If successful, an additional two points are deducted.

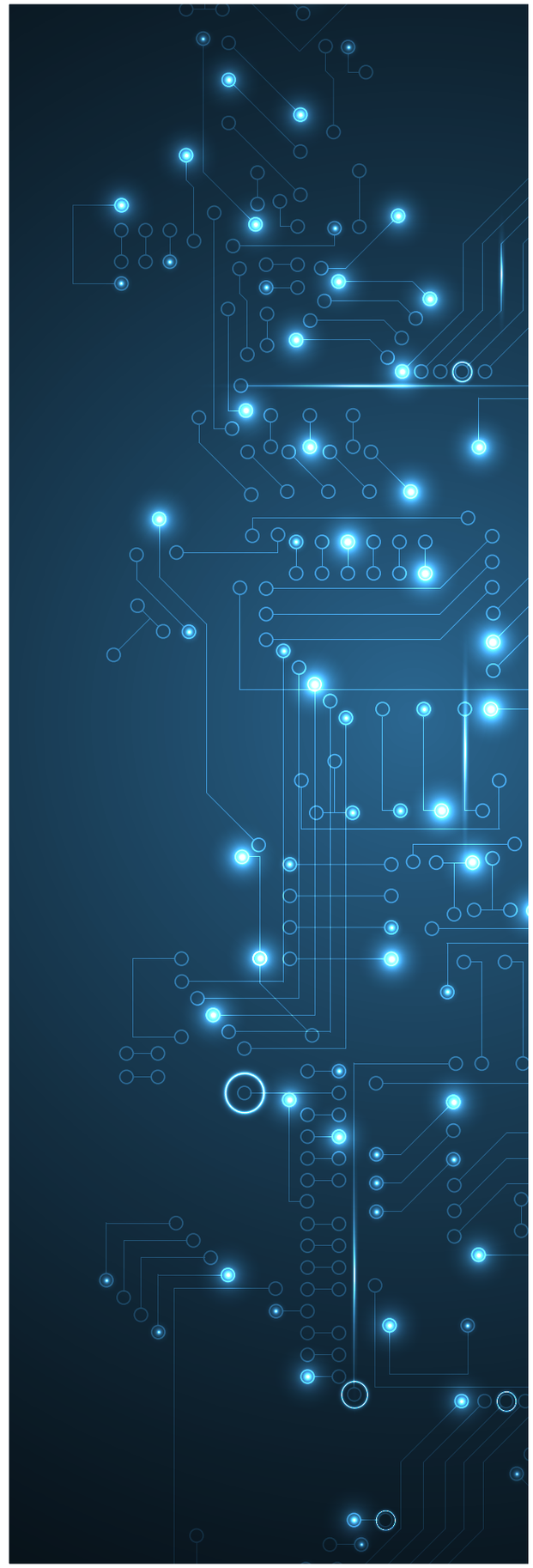
Post-Escalation Action (-1)
After escalation the attacker attempts actions that rely on escalated privileges. These include attempting to steal credentials, modifying the file system and recording keystrokes. If any of these actions are successful then a further penalty of one point deduction is applied.

PROTECTION RATINGS		
Product	Protection Rating	Protection Rating (%)
Kaspersky Small Office Security	399	98%
ESET Endpoint Security	398	98%
Symantec Endpoint Protection Cloud	398	98%
Microsoft Windows Defender ATP's Antivirus	391	96%
McAfee Endpoint Security	390	96%
Sophos Intercept X	387	95%
Bitdefender Gravity Zone Endpoint Security	385	94%
Trend Micro Worry Free Security Services	384	94%



Protection Ratings are weighted to show that how products handle threats can be subtler than just 'win' or 'lose'.

Average 96%

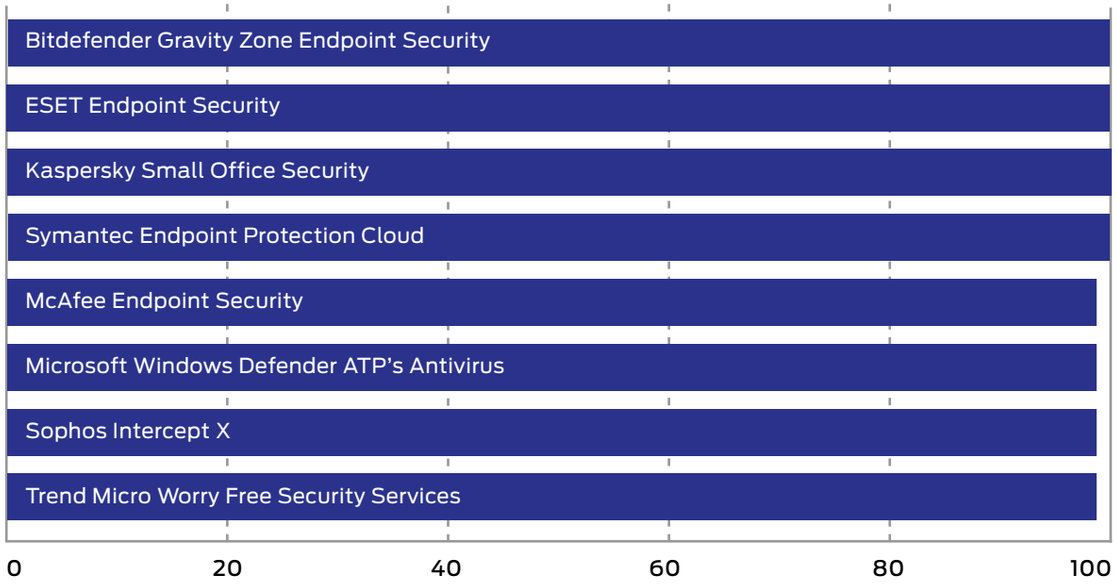


3. Protection Scores

This graph shows the overall level of protection, making no distinction between neutralised and blocked incidents.

For each product we add Blocked and Neutralised cases together to make one simple tally.

PROTECTION SCORES	
Product	**Protection Score**
Bitdefender Gravity Zone Endpoint Security	100
ESET Endpoint Security	100
Kaspersky Small Office Security	100
Symantec Endpoint Protection Cloud	100
McAfee Endpoint Security	99
Microsoft Windows Defender ATP's Antivirus	99
Sophos Intercept X	99
Trend Micro Worry Free Security Services	99



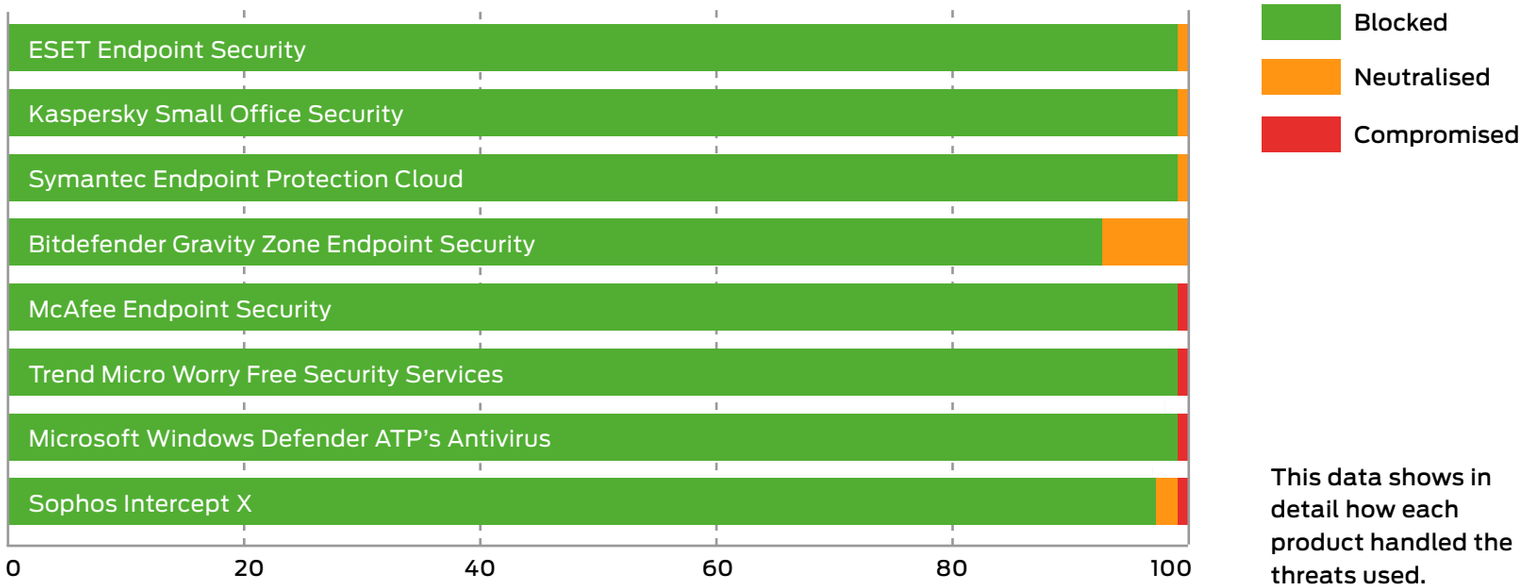
Protection Scores are a simple count of how many times a product protected the system.

4. Protection Details

These results break down how each product handled threats into some detail. You can see how many detected a threat and the levels of protection provided.

Products sometimes detect more threats than they protect against. This can happen when they recognise an element of the threat but aren't equipped to stop it. Products can also provide protection even if they don't detect certain threats. Some threats abort on detecting specific endpoint protection software.

PROTECTION DETAILS					
Product	Detected	Blocked	Neutralised	Compromised	Protected
ESET Endpoint Security	100	99	1	0	100
Kaspersky Small Office Security	100	99	1	0	100
Symantec Endpoint Protection Cloud	100	99	1	0	100
Bitdefender Gravity Zone Endpoint Security	100	92	8	0	100
McAfee Endpoint Security	100	99	0	1	99
Trend Micro Worry Free Security Services	100	99	0	1	99
Microsoft Windows Defender ATP's Antivirus	99	99	0	1	99
Sophos Intercept X	99	97	2	1	99



This data shows in detail how each product handled the threats used.

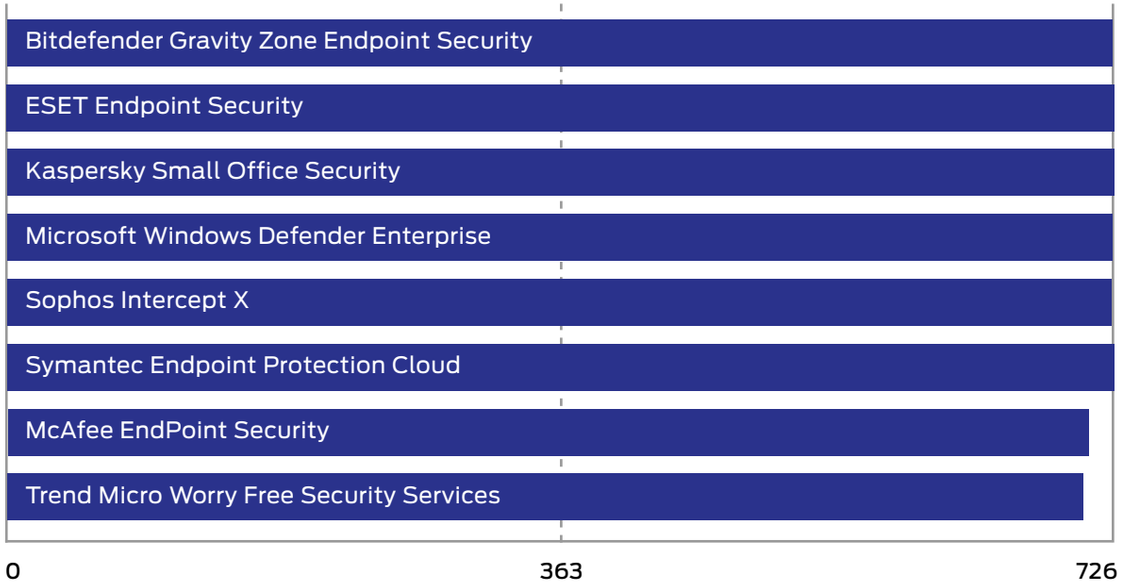
5. Legitimate Software Ratings

These ratings indicate how accurately the products classify legitimate applications and URLs, while also taking into account the interactions that each product has with the user. Ideally a product will either not classify a legitimate object or will classify it as safe. In neither case should it bother the user.

We also take into account the prevalence (popularity) of the applications and websites used in this part of the test, applying stricter penalties for when products misclassify very popular software and sites.

To understand how we calculate these ratings, see ***5.3 Accuracy Ratings*** on page ***14***.

LEGITIMATE SOFTWARE RATINGS		
Product	**Legitimate Accuracy Rating**	**Legitimate Accuracy (%)**
Bitdefender Gravity Zone Endpoint Security	726	100%
ESET Endpoint Security	726	100%
Kaspersky Small Office Security	726	100%
Microsoft Windows Defender Enterprise	726	100%
Sophos Intercept X	726	100%
Symantec Endpoint Protection Cloud	726	100%
McAfee EndPoint Security	710	98%
Trend Micro Worry Free Security Services	706	97%



Legitimate Software Ratings can indicate how well a vendor has tuned its detection engine.

5.1 Interaction Ratings

It's crucial that anti-malware endpoint products not only stop – or at least detect – threats, but that they allow legitimate applications to install and run without misclassifying them as malware. Such an error is known as a 'false positive' (FP).

In reality, genuine FPs are quite rare in testing. In our experience it is unusual for a legitimate application to be classified as 'malware'. More often it will be classified as 'unknown', 'suspicious' or 'unwanted' (or terms that mean much the same thing).

We use a subtle system of rating an endpoint's approach to legitimate objects, which takes into account how it classifies the application and how it presents that information to the user. Sometimes the endpoint software will pass the buck and demand that the user decide if the application is safe or not. In such cases the product may make a recommendation to allow or block. In other cases, the product will make no recommendation, which is possibly even less helpful.

If a product allows an application to install and run with no user interaction, or with simply a brief notification that the application is likely to be safe, it has achieved an optimum result. Anything else is a Non-Optimal Classification/Action (NOCA). We think that measuring NOCAs is more useful than counting the rarer FPs.

	None (Allowed)	Click to Allow (Default Allow)	Click to Allow/Block (No Recommendation)	Click to Block (Default Block)	None (Blocked)	
Object is Safe	2	1.5	1			A
Object is Unknown	2	1	0.5	0	-0.5	B
Object is not Classified	2	0.5	0	-0.5	-1	C
Object is Suspicious	0.5	0	-0.5	-1	-1.5	D
Object is Unwanted	0	-0.5	-1	-1.5	-2	E
Object is Malicious				-2	-2	F
	1	2	3	4	5	

INTERACTION RATINGS		
Product	None (Allowed)	None (Blocked)
Bitdefender Gravity Zone Endpoint Security	100	0
ESET Endpoint Security	100	0
Kaspersky Small Office Security	100	0
Microsoft Windows Defender ATP's Antivirus	100	0
Sophos Intercept X	100	0
Symantec Endpoint Protection Cloud	100	0
McAfee EndPoint Security	99	1
Trend Micro Worry Free Security Services	99	1

Products that do not bother users and classify most applications correctly earn more points than those that ask questions and condemn legitimate applications.

5.2 Prevalence Ratings

There is a significant difference between an endpoint product blocking a popular application such as the latest version of Microsoft Word and condemning a rare Iranian dating toolbar for Internet Explorer 6. One is very popular all over the world and its detection as malware (or something less serious but still suspicious) is a big deal. Conversely, the outdated toolbar won't have had a comparably large user base even when it was new. Detecting this application as malware may be wrong, but it is less impactful in the overall scheme of things.

With this in mind, we collected applications of varying popularity and sorted them into five separate categories, as follows:

1. Very High Impact
2. High Impact
3. Medium Impact
4. Low Impact
5. Very Low Impact

Incorrectly handling any legitimate application will invoke penalties, but classifying Microsoft Word as malware and blocking it without any way for the user to override this will bring far greater penalties than doing the same for an ancient niche toolbar. In order to calculate these relative penalties, we assigned each impact category with a rating modifier, as shown in the table above.

LEGITIMATE SOFTWARE PREVALENCE RATING MODIFIERS	
Impact Category	**Rating Modifier**
Very High Impact	5
High Impact	4
Medium Impact	3
Low Impact	2
Very Low Impact	1

Applications were downloaded and installed during the test, but third-party download sites were avoided and original developers' URLs were used where possible. Download sites will sometimes bundle additional components into applications' install files, which may correctly cause anti-malware products to flag adware. We remove adware from the test set because it is often unclear how desirable this type of code is.

The prevalence for each application and URL is estimated using metrics such as third-party download sites and the data from Alexa.com's global traffic ranking system.

5.3 Accuracy Ratings

We calculate legitimate software accuracy ratings by multiplying together the interaction and prevalence ratings for each download and installation:

Accuracy rating = Interaction rating x Prevalence rating

If a product allowed one legitimate, Medium impact application to install with zero interaction with the user, then its Accuracy rating would be calculated like this:

Accuracy rating = 2 x 3 = 6

This same calculation is made for each legitimate application/site in the test and the results are summed and used to populate the graph and table shown under ***5. Legitimate Software Ratings*** on page ***12***.

5.4 Distribution of Impact Categories

Endpoint products that were most accurate in handling legitimate objects achieved the highest ratings. If all objects were of the highest prevalence, the maximum possible rating would be 1,000 (100 incidents x (2 interaction rating x 5 prevalence rating)).

In this test there was a range of applications with different levels of prevalence. The table below shows the frequency:

LEGITIMATE SOFTWARE CATEGORY FREQUENCY	
Prevalence Rating	**Frequency**
Very High Impact	26
High Impact	39
Medium Impact	16
Low Impact	10
Very Low Impact	9
TOTAL	**100**

6. Conclusions

Attacks in this test included threats that affect the wider public and more closely-targeted individuals and organisations. You could say that we tested the products with 'public' malware and full-on hacking attacks. We introduced the threats in a realistic way such that threats seen in the wild on websites were downloaded from those same websites, while threats caught spreading through email were delivered to our target systems as emails.

All of the products tested are well-known and should do well in this test. While we do 'create' threats by using publicly available free hacking tools, we don't write unique malware so there is no technical reason why any vendor being tested should do poorly.

Consequently, it's not a shock to see all products handle the public threats very effectively. Even the weaker products protected the target systems in the vast majority of cases. Targeted attacks were also handled well by all.

Products from **Bitdefender**, **ESET**, **Kaspersky Lab** and **Symantec** protected against all of the public and targeted attacks. **Bitdefender** scored lower because it neutralised some threats rather than blocking them outright. They also handled the legitimate applications correctly.

Products from **Trend Micro** and **McAfee** stopped all the public threats and all but one of the targeted attacks. **Microsoft Defender** and **Sophos Intercept X** stopped all of targeted attacks and allowed one public threat through each.

The leading products from **Kaspersky Lab**, **ESET**, **Symantec**, **Microsoft**, **Sophos**, **Bitdefender**, **McAfee** and **Trend Micro** all win AAA awards.

Appendices

APPENDIX A: Terms Used

TERM	MEANING
Compromised	The attack succeeded, resulting in malware running unhindered on the target. In the case of a targeted attack, the attacker was able to take remote control of the system and carry out a variety of tasks without hindrance.
Blocked	The attack was prevented from making any changes to the target.
False positive	When a security product misclassifies a legitimate application or website as being malicious, it generates a 'false positive'.
Neutralised	The exploit or malware payload ran on the target but was subsequently removed.
Complete Remediation	If a security product removes all significant traces of an attack, it has achieved complete remediation.
Target	The test system that is protected by a security product.
Threat	A program or sequence of interactions with the target that is designed to take some level of unauthorised control of that target.
Update	Security vendors provide information to their products in an effort to keep abreast of the latest threats. These updates may be downloaded in bulk as one or more files, or requested individually and live over the internet.

APPENDIX B: FAQs

A **full methodology** for this test is available from our website.

- The products chosen for this test were selected by SE Labs.
- The test was unsponsored.
- The test was conducted between 25th March to 3rd June 2019.
- All products were configured according to each vendor's recommendations, when such recommendations were provided.
- Malicious URLs and legitimate applications and URLs were independently located and verified by SE Labs.
- Targeted attacks were selected and verified by SE Labs.
- Malicious and legitimate data was provided to partner organisations once the test was complete.
- SE Labs conducted this endpoint security testing on physical PCs, not virtual machines.
- The web browser used in this test was Google Chrome. When testing Microsoft products Chrome was equipped with the Windows Defender Browser Protection browser extension (**https://browserprotection.microsoft.com**).

Q What is a partner organisation? Can I become one to gain access to the threat data used in your tests?

A Partner organisations benefit from our consultancy services after a test has been run. Partners may gain access to low-level data that can be useful in product improvement initiatives and have permission to use award logos, where appropriate, for marketing purposes. We do not share data on one partner with other partners. We do not partner with organisations that do not engage in our testing.

Q I am a security vendor and you tested my product without permission. May I access the threat data to verify that your results are accurate?

A We are willing to share a certain level of test data with non-partner participants for free. The intention is to provide sufficient data to demonstrate that the results are accurate. For more in-depth data suitable for product improvement purposes we recommend becoming a partner.

APPENDIX C: Product Versions

The table below shows the service's name as it was being marketed at the time of the test.

PRODUCT VERSIONS			
Provider	**Product Name**	**Build Version (start)**	**Build Version (end)**
Bitdefender	Gravity Zone Endpoint Security	6.6.8.119	6.6.10.148
ESET	Endpoint Security	7.0.2100.4	7.0.2100.4
Kaspersky Lab	Small Office Security	19.0.0.1088(d)	19.0.0.1088 (e)
McAfee	Endpoint Security	Agent Version: 5.6.0.878, Endpoint Security: 10.6	Agent Version: 5.6.0.878, Endpoint Security: 10.6
Microsoft	Windows Defender ATP's Antivirus	Threat definition version: 1.291.301.0	Antimalware Client Version (4.18.1904.1) Engine Version (1.1.15900.4) Antivirus Version (1.293.2614.0) Anti-spyware Version (1.293.2614.0)
Sophos	Intercept X	Core Agent (2.2.2), Endpoint Advanced (10.8.3), Sophos Intercept X, Device Encryption (1.4.103)	Core Agent (2.3.0) Endpoint Advanced (10.8.3.441) Sophos Intercept X (2.0.14) Device Encryption (1.4.103)
Symantec	Endpoint Protection Cloud	22.16.4.15	22.17.1.50
Trend Micro	Worry Free Security Services	20.0.1049	20.0.1049

APPENDIX D: Attack Types

The table below shows how each product protected against the different types of attacks used in the test.

ATTACK TYPES			
Product	**Web-Download**	**Targeted Attack**	**Protected**
Bitdefender Gravity Zone Endpoint Security	75	25	100
ESET Endpoint Security	75	25	100
Kaspersky Small Office Security	75	25	100
Symantec Endpoint Protection Cloud	75	25	100
Microsoft Windows Defender ATP's Antivirus	74	25	99
Sophos Intercept X	74	25	99
Trend Micro Worry Free Security Services	75	24	99
McAfee Endpoint Security	75	24	99