Usability of Antivirus Tools in a Threat Detection Scenario

Michael Körber[1], Anatoli Kalysch[1], Werner Massonne[1], and Zinaida Benenson[1]

Friedrich-Alexander-Universität Erlangen-Nürnberg
zinaida.benenson@fau.de

Abstract. Usability of antivirus (AV) tools has not received much attention yet. We conducted a laboratory study with 34 German students to investigate how they experience notifications and interventions of their AV when a threat is detected. During the study, a specifically designed harmless file triggered AV on participants' laptops. Out of 34 participants, 19 users noticed AV messages, and 8 of them understood that the message communicated threat detection concerning a specific file. Moreover, only 6 users understood that this file was removed by the AV tool. Additionally, most participants were distracted by Windows OS messages that were unintelligible to them. We investigate reasons for incomprehension in our sample, and give recommendations for improved user interaction design of AV tools.

This is the authors' version of the paper. The final publication is available at Springer via https://doi.org/10.1007/978-3-031-06975-8_18

1 Introduction

Security-related media and security experts have multiple times declared that antivirus (AV) is dead, or should not be used [16,25,27]. Nevertheless, strong majority of non-expert users have reported using antivirus and consider AV usage to be a good and actionable security advice [18,38,28,8,30]. At the same time, most of the surveyed security experts do not rate using AV as a "top" security advice [18,31,8]. These contradictions raise a question: What makes AV a good security tool from the non-experts' point of view?

Ion et al. [18, p. 333] note that high acceptance of AV "might be due to the good usability of the install-once type of solution that antivirus software offers". Indeed, AV tools are usually not actively *used*. After installation, an AV tool is supposed to run in the background. However, when threats are detected, the AV tool has to interact with users and their systems: show the threat detection notifications and delete or quarantine infected or suspicious files.

A usable AV tool should provide users with enough information, such that they can react adequately. Which actions are adequate will depend on the situation and on the technical sophistication of the user. Non-expert users considered in this work may want to ask for help, and therefore, should be able to explain

to other people what happened. Moreover, the users may want to warn other users about the threat, which at least requires them to understand which file is affected. Therefore, we consider the following main research questions in our study:

- RQ1: Which design and user interaction elements of AV notifications help users under real-life circumstances to understand that a particular file was removed by an AV tool, and that this file might be malicious?
- RQ2: What are users' attitudes to AV and their reasons for AV usage? Under which circumstances are users satisfied with their AV tools and feel well protected?

To answer the main research questions, we defined several subsequent questions that are presented in supplementary materials[1]. For example, in order for AV notifications to be helpful, at least the following steps are necessary: the users should be able to notice the notification, and understand why this notification was issued. Therefore, the corresponding subsequent questions are: Do the users notice the notifications of their AV tools? Can they explain that the notifications appeared because AV tool found a suspicious file?

We report the first (to our knowledge) laboratory study that investigates user experience when AV tools detect malware. The aim of this work is not to critique individual AV tools, but to show which notifications and user interaction elements support users well, and which need improvement.

Timeframe: The user study took place in 2017[2]. However, we tested in 2021 whether our results are still applicable: We used multiple Windows PCs with the AV tools considered in the study, and checked whether reactions of the AV tools changed over time. Whereas we registered changes in the notifications of some AV tools, we discuss in Section 5.2 that our results are still applicable and useful today, and represent a good starting point for further research.

2 Background and Related Work

2.1 Industrial AV Tests

To the best of our knowledge, no works on usability of AV tools have been published in academic research so far. However, some companies regularly publish AV tests, although these tests do not consider user interaction. For example, AV-TEST [5] measures usability as the number of false alarms using a system for automated deployment of programs from a large proprietary database of benign software samples [4]. However, an AV tool issues messages also when genuine threats are detected. Our study investigates whether these messages enable the users to understand what happened and to react adequately.

[1] https://www.cs1.tf.fau.de/research/human-factors-in-security-and-privacy-group/antivirus-usability

[2] Due to a collision of several unfortunate circumstances, the research team found resources for writing up the results only in 2021.

2.2 Attitudes and Behavior regarding Malware

Wash [37] investigated mental models of malware of US users through qualitative interviews. He found that various models (e.g., "viruses are buggy software") are connected to usage and non-usage of AV tools. Later, Wash and Rader [38] verified and extended these findings in a survey with a representative sample of the US population. Kauer et al. [20] replicated Wash's study in Germany, uncovering some additional mental models of malware.

Ion et al. [18] found in a survey that using AV is the most popular security measure for non-expert users in their sample. This result was later corroborated in a replication study by Busse et al. [8]. Whereas AV tools were much less popular among security experts in these studies, they still recommend AV usage to non-expert users [31]. Redmiles et al. investigated sources and acceptance of security advice qualitatively [29] as well as quantitatively [28]. AV is reportedly used by over 80% of US population, with family members' or media advice and negative experiences being important reasons for adopting AV. Overall, AV usage is considered actionable, effective and high priority security advice [30].

Lalonde Levesque et al. [23] investigated how devices get infected despite having an AV. The identified risk factors were high computer expertise, visiting many websites of particular categories, such as streaming, and installing a lot of software. Forget et al. [15] found that users with high engagement in security still get malware. Sharif et al. [33] investigated how customers of an AV company identify and describe virus incidents in customer support chats. They found that some users were surprised that they can get malware despite AV usage, and often could not precisely identify whether they got infected and how this happened. Overreliance on AV for protecting against a multitude of threats was reported by Krol et al. [21], and might be a likely factor of risky Internet behavior [9].

In contrast to the previous work, we consider how users interact with AV tools in a threat detection scenario.

2.3 Security Warnings

Security warnings inform users about possible dangers and offer a choice of at least two options on how to proceed, whereas notifications provide information about found and eliminated threats and usually offer either one option or none [6]. Design, wording and purpose of AV notifications is comparable to security warnings in many aspects. Reaction to security warnings and indicators, especially in web browsers, has been in focus of empirical studies for at least last 15 years [39,32,12,34,7]. These works considered existing warnings, made improvement suggestions and sometimes tested new warnings. Malware warnings in browsers and other tools have also been investigated [21,24,1].

Generally, a sizable amount of users consistently ignored warnings, did not trust them, and did not understand them across the studies. Ignoring warnings is tightly connected to the effects of habituation and generalization, especially if warnings have high false positive rates, or appear in non-critical situations [3,36]. Although some improvements in adherence could be reached through careful design and timing [14,19], warnings still remain an active research topic.

3 Method

3.1 Design and Usage of the 'Infected' File

We designed a harmless file that triggered AV tools of the participants. In the sequel, we refer to *'infected'* files using quotation marks to emphasize that these files were entirely harmless. This file was written in `C` and exhibited behavior typical for malware called *downloader*: It installs itself on the victim's PC and later downloads additional malicious files. Next, we uploaded this file from various IP addresses using the TOR network[3] to a malware detection service VirusTotal that automatically tests uploaded files against over 60 antivirus tools[4]. Uploads from different locations can indicate a malware campaign and prompt AV distributors to include the signature of the file into their databases. Whereas in the beginning only 10 out of 63 AV tools flagged our file as malicious, after two weeks of uploads 46 AV tools detected it.

To make the malicious file fit the cover story (usability study of Microsoft Word), we changed the file extension of the 'infected' file from `.exe` to `.doc`. This did not change its detection rate, because file extensions are not used for virus signature generation. The changed file could not be run anymore, but could be opened in Word as a binary file (i.e., it did not show any readable content).

Malware can be introduced via email attachments, file download (including drive-by downloads) and via USB drives or other portable media. We pretested all three scenarios using virtual machines and found that different email clients and browsers exhibit specific error messages that interact with AV notifications and actions, whereas USB drives only elicit error messages by Windows OS (as presented in Figure 3 in Section 4.1). Therefore, we decided to use USB drives in this exploratory study, and leave other scenarios to future work. Otherwise, to make sure that the 'infected' files work as expected, we would have needed to test beforehand a variety of email clients or browsers used by the participants. When using USB drives, we only needed to test Windows OS without any additional software. Additionally, if the participants opened their emails during the study, we could not avoid observing at least some details of their emails, which has privacy implications. Moreover, if any emails unrelated to the study arrived shortly before or during the study, they would have distracted the participants.

3.2 Study Design and Procedure

We conducted an observational exploratory study, because nothing was known about user interaction with AV tools beforehand. For example, we were unable to predict from the literature on security warnings whether users would notice and understand AV notifications, and whether they would notice and understand that a suspicious file was removed by their AV tool. Therefore, it was necessary to observe user behavior under conditions that should be as natural as possible,

[3] https://www.torproject.org
[4] https://support.virustotal.com/hc/en-us/articles/115002146809-Contributors

with the aim to create a foundation for further research in this area. In our study design, we used the principles from the guideline by Krol et al. [22, p. 23] on conducting user studies in security: (1) Give participants a primary task; (2) Ensure participants experience realistic risk; and (3) Avoid priming the participants.

In order not to prime the participants, we recruited them for a usability study of Microsoft Word. Accordingly, their primary tasks in the study were concerned with text processing, e.g., copying a part of one text into another text, inserting an image or changing formatting. The participants used their own laptops, such that the risk to their data and system was made as realistic and salient as possible. Participants' laptops were connected to the laboratory computer via a remote access program TeamViewer[5] in order to use screen capture.

The overview of the study is presented in Fig. 1. After a short introduction, participants signed the informed consent form and received two envelopes, each containing a USB drive with 8 files (various Word documents, PDFs and figures) and a list of tasks. Both task lists first asked the users to copy files from the USB drive to the laptop, and then to execute some text processing tasks, as described above. The topic of texts and pictures in all tasks was German monetary policy.

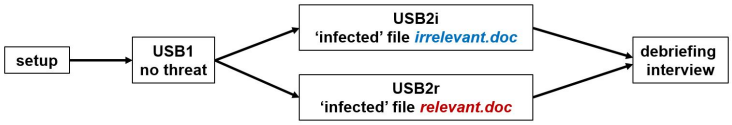


Fig. 1: Study design: participants first worked with USB1, which did not contain any 'infected' files. Their second task requierd them to copy files from USB2i or USB2r, where one file was 'infected'.

Participants first worked with USB1 that did not trigger their AV tool, such that they could familiarize themselves with the environment and the tasks. However, the USB drive in the second envelope (USB2r or USB2i) contained an 'infected' file. When the users inserted the second USB drive and started copying the files, their AV tool detected the threat and reacted by showing a threat detection message and moving the corresponding file into the quarantine.

Both second USB drives contained the same files, and one of these files was 'infected'. As all file names were in German, we call these files here *relevant.doc* and *irrelevant.doc* to improve readability. The tasks for both USB drives were the same. However, the 'infected' file on USB2r was *relevant.doc*, and it was relevant for the task: The participants were required to insert an image into it. In this case the participants would be unable to complete their tasks, because the file would be moved to quarantine. We were interested in how they would make sense of this situation. The 'infected' file on USB2i was *irrelevant.doc*, which was not needed for the tasks. In this case, we were interested whether the participants would notice the removal of the file by their AV tool, and understand

[5] https://www.teamviewer.com

what happened. This condition emulates situations where a file is removed by the AV tool, but some time later the user needs this file, or receives a message from some program that this file was not found.

After task execution, the participants were interviewed about their experience during the study and debriefed about its real goal. We also showed to them screen capture of their task execution. Finally, we asked about their experience with viruses and attitudes to AV tools. The study took 37 minutes on average. The participants were reimbursed with a 10 EUR Amazon voucher. The interview guide is presented in the supplementary materials[6].

3.3 Ethics

The study protocol was reviewed by two usability experts who did not participate in study design and execution, and approved by the data protection office at our university. We also run four pretests and adjusted the study protocol accordingly.

To minimize the experimenter effect, the participants were left alone in the lab after the study setup. We were concerned that they might experience negative emotions during the unexpected threat detection, blame us for giving them an unsafe USB drive, or feel helpless if a file they need for the task is missing. Therefore, we told the participants that they can contact us anytime in the neighboring room. We also observed the participants through a one-way window of the lab (they were informed about this observation), such that we could interfere in case a participant remained helpless for too long or appeared frightened[7].

Participants were fully debriefed about the real goal of the study in the post-interview. A specific ethical issue arose because the USB drives were used multiple times in experimental runs. Thus, if any participant had undetected malware on their laptop, it could spread to other laptops. Therefore, we "disinfected" USB drives after every usage by securely erasing them under Mac OS.

3.4 Data Collection and Analysis

Two fixed team members were present at each study run. One of them was the main contact for the participant, handled the study setup and conducted the interview. The other observed the study through the one-way window and on a monitor in the lab and made structured notes in an Excel sheet that contained core observation points. Later both researchers independently watched screen capture videos and listened to audio recording of the interviews. They noted down their observations on user interaction and the answers of the participants, and additionally transcribed especially interesting or important phrases verbatim. They then compared and discussed their notes to validate the observations and the interview data. Finally, core themes were extracted from the notes and

[6] https://www.cs1.tf.fau.de/research/human-factors-in-security-and-privacy-group/antivirus-usability

[7] None of the participants felt frightened or blamed us for the unsafe USB drive. All participants called us if they could not proceed with the task.

categorized. Additional team members watched the videos and listened to the interviews, and wrote short summaries of the cases. These notes were subsequently used in several team meetings to complement the analysis.

3.5 Participants and Their AV Tools

Participants were recruited via student mailing lists of economics, social sciences and engineering departments at two German universities. They were invited to take part in a pre-screening questionnaire for a lab study concerning usability of Microsoft Word. To avoid priming, AV tools were not mentioned in the recruitment email. The goal of the pre-screening questionnaire was to find people who use AV, and to pretest their AV tools. To this end, the questionnaire asked whether the participants use a Windows laptop, would use it in the lab study and let us install TeamViewer on it, and which tools in general they use on their laptop, including word processing, web browsers, email programs, and AV tools. After this block of questions, the participants were asked multiple questions about their Internet usage, and also demographic questions. This disguise of the study goal was successful, as none of the participants in the lab study reported that they suspected the study to be connected to AV tools.

We received 91 completed questionnaires, of which 44 participants qualified for the invitation to the study. Of these 44 participants, 34 took part in the study. The rest either did not react to the invitation, or canceled their slot. To reach the intended aim of 40 participants, six additional users were recruited in the building where the study took place. AV tools that were mentioned by the participants in the recruiting survey were tested beforehand with default settings. We copied the 'infected' file from a USB drive to a laptop, and the tools reacted as expected: they showed notifications and removed the file into quarantine. Unfortunately, during the user study, all three recruited Avast users and two Avira users did not receive any notifications, and their 'infected' files were not removed. We could not find out why this happened, especially as AV tools of these users were able to recognize and remove the corresponding files through manual scan in all cases. Furthermore, one participant switched off Windows Defender on his laptop. Data of these six users were excluded from the analysis.

The remaining 34 participants were 23 years old on average, 25 identified as female, and 9 as male. Most of them studied economics (15) or social work (12), and the rest studied electrical engineering or design. They used the following AV tools: Windows Defender (9 participants), Avira (6), Sophos (6, with 4 using a fallback laptop[8]), AVG (3), Norton (3), Bitdefender (2), G Data (2), Microsoft Security Essentials (2), Kaspersky (1).

[8] If the participants could not use their own laptop (e.g., they forgot to bring it, or had technical issues), they used a "fallback" laptop with Windows 10 and the AV tool Sophos. Our university requires Sophos on university computers.

Table 1: User experience with AV tools; 34 users in total, 15 of them handled *irrelevant.doc* (*i* column), and 19 handled *relevant.doc* (*r* column)

	n	*i*	*r*	User experience: AV message
nn	15	6	9	did not notice that an AV message appeared on the screen
nu	5	2	3	did not understand that the message came from AV, but noticed it
pu	6	2	4	remembers some parts of the message (e.g., that a threat was detected, or no action is needed), but cannot fully explain why the message appeared (e.g., does not connect the message to the disappearance of the 'infected' file, or thinks that the USB drive or a file is defect, or protected)
u	8	6	2	explains that AV found a virus in a particular file and issued a message
	n	*i*	*r*	User experience: AV intervention
nu	21	8	13	either appears to be completely lost, or says that something on the PC or on the USB is broken, or otherwise cannot be accessed or copied
pu	7	3	4	explains that a virus was found, but does not fully understand the situation (e.g., does not know which file is affected, or says that "something" on the USB drive or on the PC is infected)
u	6	5	1	explains that AV found a virus in a particular file and removed it

nn = not noticed; nu = not understood; pu = partially understood; u = understood

4 Results

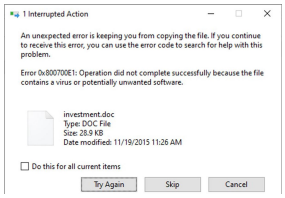
4.1 User Experience with Threat Detection

Descriptions of observed user experience and cumulative statistics are presented in Table 1. Out of 34 participants, 15 users did not notice AV messages. Out of the remaining 19 users, 8 understood the messages. Moreover, only 6 users understood that the the 'infected' file was removed by the AV tool. Although more users in the *irrelevant.doc* condition understood the situation, this occurred because the AV tools could not be balanced between conditions due to low number of participants for several AV tools. In the following, we first present three case studies to make clear how the combination of different factors shaped user experience. We then discuss each factor in more detail in Section 4.2.

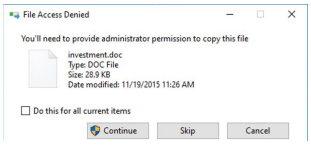
Case study 1: Windows Defender (9 users) Notifications remained almost the same between 2017 and 2021 (Fig. 2). They are small ($\sim$ 3% of the screen), black and do not present any details about the found threat. They appear in the lower right corner of the screen, and disappear without user interaction after 3 seconds. The Windows OS notification `Error` (Fig. 3a) appears at roughly the same time in the center of the screen. Although this notification contains the word "virus" in the end of its second paragraph, only P20 noticed this fact,

which led to his understanding the situation. All other participants either closed `Error` notification quickly, or clicked `Try Again` several times (the notification reappears in this case), and then closed it. During this time, the AV notification disappeared. P7 and P10 mentioned later that the `Error` message was too long.

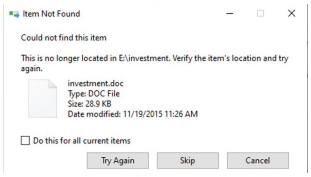
Fig. 2: Windows Defender notifications: 2017 (to the left) and 2021 (to the right).



(a) `Error`: *An unexpected error is keeping you from copying this file.*



(b) `Adm`: *You'll need to provide administrator permission to copy this file.*



(c) `notF`: *Could not find this item.*

Fig. 3: Windows OS notifications (replicated in English). All messages named the affected file (called here *investment.doc*).

Combination of a long Windows OS notification in the middle of the screen with quick disappearance of the AV notification in the corner caused six users to overlook the AV notification. Additionally, two users noticed it, but thought that it belongs to other programs. As P9 remarked: *[AV notification] neither differs in size nor in anything else from all other Windows 10 messages*. This resulted in the extremely low understanding of the situation (1 out of 9). Later, when shown the AV notification on the screen capture, several users were dissatisfied with the absence of details about the virus: *It tells me that something has happened, but not in which program or where [...] the message also does not tell me due to which file this happened* (P3). They also commented that they would expect the AV notification to be red and appear in the center of the screen.

Case study 2: Avira (6 users) Notifications remained almost the same between 2017 and 2021 (Fig. 4). They are of middle size ($\sim$ 8% of the screen), appear in the lower right corner, and users have to click on the red cross to close them. They show the name and path of the affected file, and the name of the malware. Windows OS notification `Adm` (Fig. 3b) appeared simultaneously with the

Avira notification, and if the users clicked `Continue`, the `notF` message (Fig. 3c) appeared. Four users did not notice the Avira notification at all, although it persisted till the end of their participation – they were too much distracted by the Windows OS notifications. When shown the screen capture video, they were surprised: *Oh, here it is! Oh my! I did not see it at all. Did it really appear?* (P21). Participants also reported habituation, as Avira seems to show a lot of information not connected to the threats: *I never look there [lower right corner], because Avira always shows ads there* (P21). Overall, only one user fully understood what happened.

Fig. 4: Avira notifications: 2017 (to the left) and 2021 (to the right).

Opinions on the AV message differed. Thus, P21 wished the message to appear in the center, and P24 wished it was green, because she felt frightened by the red color. Generally, the information about file name and path was found very helpful. Still, some users found the message incomprehensible: *It does not tell me why this happened, and what should I do. And in "Details" it shows more things that I don't understand* (P11). Although P6 noticed the message and understood the situation well, she was unsure what happened to the file, and asked us to explain what is quarantine.

Case study 3: AVG (3 users) The notification is large and appears in the center of the screen (Fig. 5). It requires the user to choose an action: "Protect (recommended)" or "Ignore threat" (translated from German) and presents file name and path, as well as malware name. Two participants chose the "ignore" option, and one chose "protect". Even if the threat was "ignored", the file could not be accessed by users anymore. In all cases, participants were able to explain what happened, and understood the situation well. Two participants who ignored the threat explained that they would choose "protect" at home, but in the study, they trusted the lab environment. To summarize, the "ignore" option seemed to be unnecessary in 2017, as the file could not be accessed anyway, but the participants spent some time on the decision process. The change in 2021 (Fig. 5) seems

to be positive: The message still appears in the middle of the screen, informs the users about the affected file, does not disappear without user interaction, but does not require taking a decision anymore.

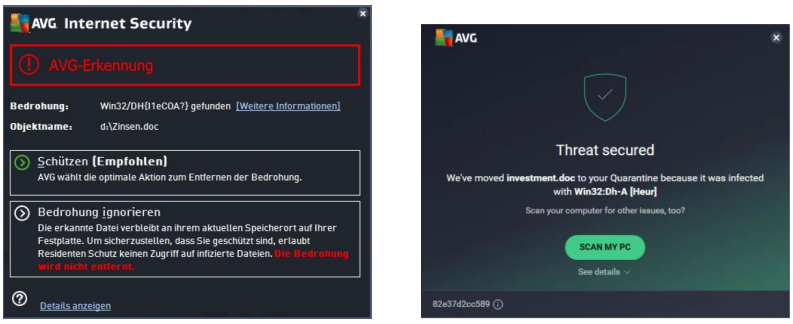


Fig. 5: AVG notifications: 2017 (to the left) and 2021 (to the right).

4.2 Important Elements of Notifications and User Interaction

Position, Size and Colors Most notifications appeared in the lower right corner of the screen. Several participants commented that the AV message should appear in the center. Especially Windows Defender and Sophos users, who saw small black AV notifications, wished that the notifications had a signal color. Central notifications were noticed and at least partially read by the users, and did not elicit any negative comments about their position, size or color.

Habituation and Generalization Habituation refers to frequent appearances of notifications, whereas generalization transfers habituation from various non-essential notifications to important ones [36]. Avira users reported habituation (Section 4.1), and messages appearing in the lower right corner prompted several participants to comment on generalization. P25 (Sophos) said: *One is used to clicking away error messages. I thought it is connected to the USB drive, because it appeared when I inserted it.* P17 (Norton) explained: *If something appears at the bottom right, I pretend that it is not there. [...] I always click away everything at the bottom right, because things often appear there.*

Interaction with Windows OS Notifications When AV notifications appeared at the lower right, some Windows OS notifications appeared around the same time in the center. This greatly distracted participants, such that they had difficulties to notice AV notifications, even if they did not disappear from the screen. Kaspersky, Security Essentials and Sophos exhibited a combination of a quickly disappearing AV message at the lower right with centrally placed Windows OS messages, similarly to Windows Defender, which resulted in confusion and low

understanding rates. Central AV notifications appeared over the Windows OS messages, and thus did not distract the participants. Norton was the only AV tool where no Windows OS messages appeared for all three participants, showing that AV tools can work without evoking Windows OS notifications.

User Interaction and Decisions As described in the AVG case, forcing users to decide what should happen to the 'infected' file was not helpful. However, AV notifications disappearing without any user interaction caused a lot of confusion. Most successful AV notifications were those that informed users about the threat detection and required user interaction to disappear.

Content Notifications containing name of the 'infected' file greatly contributed to the understanding of the situation. Also path information was helpful, as the users could determine that the affected file comes from the USB drive. Malware name, on the other hand, did not help. P29 (Bitdefender) said: *There was this message that something somewhere was detected as malware.* Bitdefender presented only malware name, but no information about the file. Norton users received a succession of several messages. Some of these messages informed the users about "security risk Downloader", which was incomprehensible to them. However, some other messages referred to the name of the 'infected' file (without path), which helped them to partially understand what happened.

Terminology Only G Data called the 'infected' file "virus". Other AV tools used terms "threat" (AVG, Bitdefender, Security Essentials, Sophos), "pattern" (Avira), "infected file" (Kaspersky), "security risk" (Norton), or "malware" (Windows Defender). Although previous work uncovered that "virus" is the most understandable term for non-experts [37,20], the terms used by participants' AV tools were understandable to them. However, some users commented that they don't know what is "quarantine".

4.3 AV Usage and Attitudes to AV Tools

When asked why they use AV, participants said that AV protects them from threats and provides a feeling of security. As many AV tools have extended functionality, some participants emphasized that their AV protects them from malicious websites: *My Norton also blocks some websites, then I don't go there* (P34). Users rely on the *expert function* of AV tools: *I find [using AV] sensible, because I cannot detect viruses by myself* (P33). However, several participants mentioned that AV cannot protect them from all threats, as one is never 100% secure. This effect is also reported by Wash and Rader [38]. 20 out of 34 users reported that they had a virus before (not necessarily on the same computer), and 11 of them lost data in consequence. 13 participants said that their present AV tool found threats previously, as it issued the corresponding notifications.

Participants expressed high satisfaction with their AV tools: they find them effective, trustworthy and usable, even if they were confused during the study. Many users think that paid AV tools offer better protection. Thus P15 (Avira)

said: *It's for free, so I think it does not provide the best protection [...] On the other hand, I'm content with what I have.* P16 (Norton) commented: *I'm paying for it. Then they are taking care that it is okay.* Windows Defender is a special case, as five out of nine users were surprised that they have it on their laptops. They thought to have other AV tools: Avast, Avira, McAffee and Norton. P9 would like to return to Avira, but he has no idea how to do this. P20 and P31 seemed not to grasp at first that their previously installed AV tools were not active: *Windows Defender is not an antivirus program for me, I have Avira. I would have reacted to an Avira message* (P31).

When asked why they use this particular AV, 20 users said that a third person (parents, siblings, partners, computer shop) installed it, and 13 users installed their AV themselves. The latter performed better in the study: 9 out of 13 users at least partially understood the situation. Out of 17 users whose AV was installed by other people, only five at least partially understood the situation. Thus, it seems that past experience with the AV tool was helpful[9].

Parents were most often mentioned as a source of AV advice, which agrees with previous research [29,28]. Thus, P14 uses Kaspersky, because her mother uses it as well. P28 reports: *My father looked it all up [different AV tools] and recommended G Data.* P17 said (jokingly): *My parents have a Norton packet with four licenses, and forced me to install it, too.* Using AV seems to be a rule that parents convey to their children, just like rules of good behavior: *I know this from an early age* (P18).

5 Discussion

5.1 Preliminary Recommendations for AV Tool Design

To summarize our analysis, a helpful AV message should appear in the center of the screen and contain some signal colors, such as green or red. It should not require a decision about what should happen to the infected file, but it should also not disappear without user interaction. Name and path of the infected file should be communicated. Malware name, on the other hand, should not be presented in the threat detection notification (it can be presented in a more detailed view, which all AV tools offer). Furthermore, it is paramount that AV tools do not evoke Windows OS messages.

These findings are in line with the recommendations for design of security warnings [6]. A warning should present all important contextual information – for AV notification, this means file name and path. The warning should be concise, present information from users' viewpoint and offload expert information into "Details" – for AV notifications, this expert information is malware name. It is recommended to require user interaction in case of important events – a central AV notification that requires user interaction to disappear seems to be suitable.

Our recommendations are preliminary and should be systematically tested. Especially placing security notifications in the middle of the screen and requiring

[9] Four participants that used the fallback computer are excluded from these analyses.

user action can be considered a controversial design decision, as it increases user effort [17]. However, confusion and incomprehension in case of threat detection also increases user effort and may be more disruptive to the users. This aspect needs further investigation.

5.2 Changes in AV Tools over Time

According to Microsoft, Windows Defender was installed on 50% of Windows devices worldwide in 2019 [26,35], and its threat notification remained almost unchanged between 2017 and 2021. Thus, its poor performance in our study is likely to remain important in 2021. Also Avira and Norton notifications remained almost the same. These tools account for 17 out of 34 users in our study.

User interface and interaction of AVG, Bitdefender, Kaspersky and Sophos changed considerably from 2017 to 2021. Whereas changes in AVG seem to be positive according to our analysis, changes in other AV tools seem to be ambiguous. These tools present file name and path to the users in 2021, which is a positive change. On the other hand, all notifications appear in the lower right corner and have a black or dark blue background, which makes them similar to various Windows 10 notifications, implying danger of generalization. Whereas Sophos notifications now require user interaction to disappear (positive change), Kaspersky and Bitdefender notifications disappear without user interaction. On the whole, these changes make especially Kaspersky and Bitdefender notifications similar to the Windows Defender, which performed extremely poorly in our user study. We conclude that our study offers useful and novel insights irrespective of changes in AV tools over time.

5.3 Usability and User Acceptance of Security Tools

Our study shows that AV tools are not quite as usable as can be assumed from their popularity [18]. Nevertheless, participants considered AV to be usable and trustworthy, although not necessarily protecting them from all dangers. AV tools play the role of security experts that have knowledge and skills to recognize malcious files – which users would not be able to do on their own. Therefore, using AV seems to be a sensible protection measure. Returning to Microsoft's statement that more than 50% of Windows devices use Windows Defender [26], it is not quite clear whether this high usage rate is conscious, as some participants were not aware that they have Windows Defender. This raises a question of how *invisible* should a security tool be? Invisible, automated security has serious usability limitations [11], whereas visible security might offer better user experience, as is known for message encryption [13] and e-voting [10].

Recommendations for using particular AV tools usually refer to their malware detection rates, false positive rates and performance [2,4], but not to their usability. The key takeaway from this user study extends to all security tools with a user interface: Attention should be paid to the user experience and usability in cases where the tool detects an attack. These cases are most important for the users, and should be carefully tested.

5.4 Limitations

Conducting a lab study implies that participants might behave differently compared to their usual environment. To increase ecological validity, we gave the participants a task not related to security. Furthermore, 30 out of 34 participants used their own laptops, which accounted for realistic risk. We verified our realistic risk assumption by asking users which consequences a virus infection would have for their laptops, and 23 participants said that they would face serious problems. When we asked if participants would react differently at home, seven of them were unsure, and eight said "yes". They explained that they trusted the lab environment or wanted to complete the task, and therefore were less cautious. We used deception in recruitment to elicit most natural reactions of participants. In debriefing interviews, all participants said that they did not suspect that the study is concerned with AV tools.

We considered the scenario of inserting USB drives, but did not test drive-by downloads or email attachments. Therefore, these malware distribution scenarios need further investigation. Furthermore, as the user study was conducted several years ago, it is possible that users' reactions might have changed over time, as they learned more about their AV tools, and about the digital world generally.

Our sample was very young, well educated (university students) and skewed towards female participants. Therefore, our results may not generalize to older or less educated populations, and especially to those less knowledgeable about AV tools than our participants. Moreover, we could test only a limited number of AV tools, and most tools were tested only with a very small number of participants. Therefore, derived design guidelines should be further tested.

6 Conclusion and Future Work

We conducted a user study that observed user interactions with nine AV tools in a threat detection scenario, and uncovered serious user experience deficits. Our results, obtained in 2017, remain valid in 2021, as design of user interfaces and interactions of AV tools seems to lack evidence-based recommendations. Our study serves as a necessary starting point for further investigations. The next step is to validate our findings in a controlled experiment that systematically compares design elements from Section 4.2 with the goal of providing evidence-based recommendations for design of AV tools.

Acknowledgments: We thank Thilo Voigt for essential support in conducting the user study, Katrin Proschek for support in the usability lab and for the idea of the cover story, Martin Ortlieb and Stefan Brandenburg for help with study design, Stella Wohnig for assistance with data analysis, the anonymous reviewers for their valuable comments, and Simone Fischer-Hübner for shepherding.

References

1. Almuhimedi, H., Felt, A.P., Reeder, R.W., Consolvo, S.: Your reputation precedes you: History, reputation, and the chrome malware warning. In: Symposium On Usable Privacy and Security (2014)
2. Anti-Malware Testing Standards Organization. https://www.amtso.org (2021)
3. Anderson, B.B., Kirwan, C.B., Jenkins, J.L., Eargle, D., Howard, S., Vance, A.: How polymorphic warnings reduce habituation in the brain: Insights from an fMRI study. In: ACM Conference on Human Factors in Computing Systems (2015)
4. AV Test Modules. https://www.av-test.org/en/about-the-institute/test-procedures/test-modules-under-windows-usability/ (2021)
5. The best antivirus software for Windows Home User. https://www.av-test.org/en/antivirus/home-windows (2021)
6. Bauer, L., Bravo-Lillo, C., Cranor, L., Fragkaki, E.: Warning design guidelines. Tech. rep., CMU-CyLab-13-002 (2013)
7. Bravo-Lillo, C., Cranor, L.F., Downs, J., Komanduri, S.: Bridging the gap in computer security warnings: A mental model approach. IEEE Security & Privacy **9**(2), 18–26 (2010)
8. Busse, K., Schäfer, J., Smith, M.: Replication: No one can hack my mind revisiting a study on expert and non-expert security practices and advice. In: Fifteenth Symposium on Usable Privacy and Security (2019)
9. Christin, N., Egelman, S., Vidas, T., Grossklags, J.: It's all about the benjamins: An empirical study on incentivizing users to ignore security advice. In: International Conference on Financial Cryptography and Data Security (2011)
10. Distler, V., Zollinger, M.L., Lallemand, C., Roenne, P.B., Ryan, P.Y., Koenig, V.: Security-visible, yet unseen? In: ACM Conference on Human Factors in Computing Systems (2019)
11. Edwards, W.K., Poole, E.S., Stoll, J.: Security automation considered harmful? In: Proceedings of the 2007 Workshop on New Security Paradigms (2008)
12. Egelman, S., Cranor, L.F., Hong, J.: You've been warned: an empirical study of the effectiveness of web browser phishing warnings. In: ACM Conference on Human Factors in Computing Systems (2008)
13. Fahl, S., Harbach, M., Muders, T., Smith, M., Sander, U.: Helping Johnny 2.0 to encrypt his Facebook conversations. In: Symposium on Usable Privacy and Security (2012)
14. Felt, A.P., Ainslie, A., Reeder, R.W., Consolvo, S., Thyagaraja, S., Bettes, A., Harris, H., Grimes, J.: Improving SSL warnings: Comprehension and adherence. In: ACM Conference on Human Factors in Computing Systems (2015)
15. Forget, A., Pearman, S., Thomas, J., Acquisti, A., Christin, N., Cranor, L.F., Egelman, S., Harbach, M., Telang, R.: Do or do not, there is no try: user engagement may not improve security outcomes. In: Symposium on Usable Privacy and Security (2016)
16. Goodin, D.: Antivirus pioneer Symantec declares AV "dead" and "doomed to failure". Ars Technica (2014)
17. Herley, C.: More is not the answer. IEEE Security & Privacy **12**(1), 14–19 (2013)
18. Ion, I., Reeder, R., Consolvo, S.: ... no one can hack my mind: Comparing expert and non-expert security practices. In: Symposium On Usable Privacy and Security (2015)
19. Jenkins, J.L., Anderson, B.B., Vance, A., Kirwan, C.B., Eargle, D.: More harm than good? how messages that interrupt can make us vulnerable. Information Systems Research **27**(4), 880–896 (2016)

20. Kauer, M., Günther, S., Storck, D., Volkamer, M.: A comparison of american and german folk models of home computer security. In: International conference on human aspects of information security, privacy, and trust (2013)
21. Krol, K., Moroz, M., Sasse, M.A.: Don't work. can't work? why it's time to rethink security warnings. In: International conference on risk and security of internet and systems (CRiSIS) (2012)
22. Krol, K., Spring, J.M., Parkin, S., Sasse, M.A.: Towards robust experimental design for user studies in security and privacy. In: Learning from Authoritative Security Experiment Results (LASER) (2016)
23. Lalonde Levesque, F., Nsiempba, J., Fernandez, J.M., Chiasson, S., Somayaji, A.: A clinical study of risk factors related to malware infections. In: ACM SIGSAC Conference on Computer and Communications Cecurity (2013)
24. Modic, D., Anderson, R.: Reading this may harm your computer: The psychology of malware warnings. Computers in Human Behavior **41**, 71–79 (2014)
25. O'Callahan, R.: Disable Your Antivirus Software (Except Microsoft's). http://robert.ocallahan.org/2017/01/disable-your-antivirus-software-except.html (2017)
26. Popa, B.: Microsoft's Antivirus Defending More than Half of Windows PCs. Softpedia (2019)
27. Purdy, K., Klosowski, T.: You Don't Need to Buy Antivirus Software. Wirecutter https://www.nytimes.com/wirecutter/blog/best-antivirus/ (2020)
28. Redmiles, E.M., Kross, S., Mazurek, M.L.: How I learned to be secure: a census-representative survey of security advice sources and behavior. In: ACM SIGSAC Conference on Computer and Communications Security (2016)
29. Redmiles, E.M., Malone, A., Mazurek, M.L.: I think they're trying to tell me something: Advice sources and selection for digital security. In: IEEE Symposium on Security and Privacy (2016)
30. Redmiles, E.M., Warford, N., Jayanti, A., Koneru, A., Kross, S., Morales, M., Stevens, R., Mazurek, M.L.: A comprehensive quality evaluation of security and privacy advice on the web. In: USENIX Security (2020)
31. Reeder, R.W., Ion, I., Consolvo, S.: 152 simple steps to stay safe online: Security advice for non-tech-savvy users. IEEE Security & Privacy **15**(5), 55–64 (2017)
32. Schechter, S.E., Dhamija, R., Ozment, A., Fischer, I.: The emperor's new security indicators. In: IEEE Symposium on Security and Privacy (2007)
33. Sharif, M., Roundy, K.A., Dell'Amico, M., Gates, C., Kats, D., Bauer, L., Christin, N.: A field study of computer-security perceptions using anti-virus customer-support chats. In: ACM Conference on Human Factors in Computing Systems (2019)
34. Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N., Cranor, L.F.: Crying wolf: An empirical study of SSL warning effectiveness. In: USENIX Security (2009)
35. Tung, L.: Top Windows Defender expert: These are the threats security hasn't yet solved. ZDNet (2019)
36. Vance, A., Eargle, D., Jenkins, J.L., Kirwan, C.B., Anderson, B.B.: The fog of warnings: how non-essential notifications blur with security warnings. In: Symposium on Usable Privacy and Security (2019)
37. Wash, R.: Folk models of home computer security. In: Symposium on Usable Privacy and Security (2010)
38. Wash, R., Rader, E.: Too much knowledge? Security beliefs and protective behaviors among united states internet users. In: Symposium On Usable Privacy and Security (2015)
39. Wu, M., Miller, R.C., Garfinkel, S.L.: Do security toolbars actually prevent phishing attacks? In: ACM Conference on Human Factors in Computing Systems (2006)