The Role of Professional Product Reviewers in Evaluating Security and Privacy

Wentao Guo, Jason Walter, and Michelle L. Mazurek, *University of Maryland*

https://www.usenix.org/conference/usenixsecurity23/presentation/guo-wentao

This paper is included in the Proceedings of the 32nd USENIX Security Symposium.

August 9–11, 2023 • Anaheim, CA, USA

978-1-939133-37-3

Open access to the Proceedings of the 32nd USENIX Security Symposium is sponsored by USENIX.

The Role of Professional Product Reviewers in Evaluating Security and Privacy

Wentao Guo, Jason Walter, and Michelle L. Mazurek
University of Maryland

Abstract

Consumers who use Internet-connected products are often exposed to security and privacy vulnerabilities that they lack time or expertise to evaluate themselves. Can professional product reviewers help by evaluating security and privacy on their behalf? We conducted 17 interviews with product reviewers about their procedures, incentives, and assumptions regarding security and privacy. We find that reviewers have some incentives to evaluate security and privacy, but they also face substantial disincentives and challenges, leading them to consider a limited set of relevant criteria and threat models. We recommend future work to help product reviewers provide useful advice to consumers in ways that align with reviewers' business models and incentives. These include developing usable resources and tools, as well as validating the heuristics they use to judge security and privacy expediently.

1 Introduction

Many Internet-connected devices and software have security and privacy vulnerabilities [6, 63, 75, 85]. This endangers consumers, who often lack the time, expertise, or motivation to evaluate security and privacy themselves across a staggering array of options [66, 100, 104].

Efforts, such as security and privacy labels [24, 39, 73], are underway to shift this burden toward professionals and institutions. Professional product reviewers, who publish information to help consumers decide what products to use [13], represent another potential path forward. Product reviews often distill the results of extended research and hands-on testing by independent experts. They influence consumers' perceptions and choices about products [64, 109]; like entertainment media [37] and VPN ads [1], they may also shape consumers' mental models of security and privacy regardless of the reviewer's expertise or intent. While not all product reviewers will conduct in-depth technical analyses of security and privacy, we hypothesize that with appropriate support, they are well positioned to help consumers choose Internet-connected products with better security and privacy.

Some advocacy groups seek to influence product reviewers' coverage of security and privacy. The Digital Standard is a framework for evaluating Internet-connected products intended to guide rigorous evaluation of security and privacy [86]. Civil rights groups have called on product reviewers to stop recommending Ring doorbell cameras in light of the company's partnerships with police departments for surveillance [28] and are tracking which organizations have done so [30]. However, there is currently no work systematically investigating how product reviewers evaluate security and privacy, and to what extent this role suits the business models and incentives involved in their work.

We fill this gap by conducting 17 interviews with professional product reviewers who evaluate Internet-connected devices and software, to understand whether and how they evaluate security and privacy, as well as what incentives, assumptions, and challenges they have. To prepare, we also analyzed security and privacy content in a small sample of 71 published reviews. We find that product reviewers consider a variety of security and privacy criteria and threat models, but we identify areas where consumers could benefit from more information. Reviewers use some techniques and tools to evaluate these criteria, but they are limited in time and expertise. While they have some incentives to evaluate security and privacy, reviewers face substantial disincentives and challenges that must be overcome if efforts to assist them are to be successful. Given limited resources, reviewers' assumptions—about products and about their audiences—inform what they prioritize. Based on our findings, we make recommendations for future research and for resource and tool development to support product reviewers in evaluating security and privacy.

2 Background

Here we provide background on professional product reviewers, defining the scope of this work. We then discuss existing research on the impact of professional product reviews, as well as existing security and privacy resources to help consumers choose safe products.

2.1 Professional product reviewers

Within the scope of professional product reviewers, we include reviewers at media companies (e.g., CNET[1]), nonprofits (e.g., Consumer Reports), and YouTube channels[2] (e.g., Marques Brownlee). As we are interested in shifting the burden of evaluating security and privacy to professionals, we exclude authors of user reviews, such as those aggregated on Amazon, and we focus on reviewers who interact with products beyond first impressions (as opposed to summarizing publicly available information or producing unboxing videos). Reviewers may primarily associate with domains outside of technology but still have expertise in certain products (e.g., a journalist who writes about parenting and also reviews baby monitors).

Product reviews are funded in many ways, but affiliate marketing has recently become especially impactful and nearly ubiquitous [13, 91, 98]. Affiliate reviews typically include a purchase link identifying the product reviewer, who earns a commission from each sale. Sponsorships are also a common source of compensation for independent reviewers in particular [4]; reviewers may be paid directly by companies for reviews, or they may be sent free or discounted products. By the policy of the U.S. Federal Trade Commission (FTC), reviewers must disclose certain compensation [20], although compliance on social media has historically been low [65].

While these business models raise legitimate concerns about trustworthiness, some reviewers take steps to mitigate bias, such as delegating business decisions to non-review staff. However, there is evidence of crooked business practices: e.g., some VPN review sites allegedly auction the top spot to the highest bidder [83]. Understanding how (dis)incentives affect evaluation of security and privacy across a wider range of Internet-connected products is an aim of this work.

2.2 Impact of professional product reviews

Website rankings indicate that many consumers consult product reviews about technology. A Tranco [57] list of the most popular websites between July 8, 2021, and July 7, 2022,[3] includes at least three websites focused on technology reviews in the top 1,000 globally, ranking CNET at 172, PCMag at 645, and TechRadar at 764. Similarweb [90] estimates that these three sites received 52, 23, and 33 million visits per month on average, respectively, between December 2021 and May 2022.[4]

Previous work suggests that professional product reviews do influence consumer behavior. Luo et al. found that the sentiment and volume of technically focused "expert blogs" (prominently featuring product reviews) are correlated with consumer perceptions of PC brands [64]. Analyses of download.com, which features CNET's professional review alongside user reviews, found that higher professional ratings lead to more user reviews and downloads of software [109], and that positive (but not neutral or negative) professional ratings lead to more downloads of software free trials [58]. Ramesh et al. found that 57% of surveyed VPN users had used recommendation websites to discover and choose among different VPNs, and 94% of these respondents considered these websites trustworthy [83].

Recent lab experiments, despite limited external validity, suggest that reviews attributed to an expert professional may have greater impact than user reviews in certain circumstances [49, 79, 80]. While some older work found that consumers considered reviews less useful when attributed to experts [60], we hypothesize that consumers' perceptions have evolved with increased awareness of fraudulent user reviews that are paid for secretly by product sellers [42, 47], written by online trolls for political reasons [10], or part of extortion schemes [70]. Indeed, two identical surveys conducted in 2011 and 2016 found that, over time, perceived source credibility and factual basis became more important in consumers' perceptions of review usefulness [31]. The impact of professional product reviews will likely continue to evolve, as existing technology review sites grow in traffic and revenue [78] and more major media organizations create their own product review operations [33].

2.3 Security and privacy consumer resources

Previous work suggests that consumers value security and privacy information if it is provided when choosing an Internet-connected product, but they do not often prioritize these factors in practice, especially as information is scarce. Emami-Naeini et al. found that some interviewees had considered security and privacy when purchasing an IoT device [25], while Zhang et al. found that few had considered data privacy when installing mobile apps [108]; participants in both studies said relevant information was difficult to find. Few participants surveyed by Ho-Sam-Sooi et al. mentioned security or privacy as a factor when deciding whether to buy a smart thermostat [43]. On the other hand, almost all participants interviewed by Emami-Naeini et al. also said they would pay more for a device if security and privacy information were provided [25]. Supporting this, various experiments have found that when people are given relevant information in accessible formats such as security labels and privacy checklists, they choose products with better security and privacy [51] and may be willing to pay more [43, 46, 99]. More specifically, Emami-Naeini et al. measured the differing impacts of individual security and privacy attributes on people's perception of risk and their willingness to purchase IoT devices [23].

[1] Examples of reviewer organizations in this paper should not be taken to imply anything about whether they were involved in interviews or not.

[2] In classifying some YouTubers and other creators as professional, we note that many produce video content requiring significant time and skill, and that they often share the same revenue sources as reviewers at more traditional organizations.

[3] This list is available at https://tranco-list.eu/list/Z2GLG

[4] We note that the Tranco and Similarweb estimates are not consistent with one another; we provide them simply to indicate rough orders of magnitude.

Some resources and tools are designed to help consumers learn more about the security and privacy of connected products. Mozilla's free online guide, *Privacy Not Included, covers publicly available information about connected consumer products [34]. Other media organizations consolidate discrete reviews, advice, and news into online security and privacy guides for consumers [26, 87]. Researchers have designed labels to communicate security and privacy information about privacy policies [50], Android apps [51], and Internet of Things (IoT) devices [24]. Similar privacy labels were recently adopted in the Apple [7] and Google Play [35] app stores. Various countries [21, 32, 39, 73] and private organizations [19, 44, 95] are developing programs to provide security and privacy information about IoT devices to consumers.

Some researchers have developed partially or fully automated frameworks for evaluating the security and privacy of Android apps [3, 85] and IoT devices [2]. They occasionally collaborate with journalists and professional product reviewers to inform consumers [18, 38].

Recent work suggests that Apple privacy labels do not yet inform consumers effectively. Li et al. reported that app developers make mistakes when creating Apple privacy labels and find the process time-consuming and overwhelming [61]. In large-scale analyses, Li et al. found that developers rarely create or update privacy labels unless forced [62], and Kollnig et al. found that most apps labeled as not collecting user data actually did so (perhaps unintentionally) through third-party tracking libraries [53]. Lay iPhone users interviewed by Zhang et al. found privacy labels useful but misunderstood them in many ways; most also had not heard of them [108]. While new resources like privacy labels can empower consumers to make better security and privacy decisions, ongoing barriers to effective implementation suggest that oversight by third-party experts, such as professional product reviewers, is complementarily important.

3 Analyzing a snapshot of product reviews

To inform our interviews, we analyzed a small sample of 71 published product reviews, focusing on what security- and privacy-relevant crieria they cover and what techniques and tools they use to evaluate them. As this was an exploratory activity, our findings are limited in scope and depth; however, we describe them here to give readers an impression of how security and privacy are covered in some reviews.

3.1 Review analysis method

For our dataset of product reviews, we focused on three common IoT devices: thermostats, locks (overtly security-related), and doorbell cameras (overtly security-related, with privacy controversies). For each, we devised one search string for list-style reviews and one for reviews of a specific, popular product; e.g., for thermostats, we used "best smart thermostats review" and "Nest thermostat review." Using private browsing mode, we downloaded the top five relevant results on Google, Bing, and YouTube for each search string, skipping reviews written by users (not professionals) or published before December 2017 (we collected reviews primarily in November and December 2021). Factoring in repeats across Google and Bing, our dataset contains 41 text and 30 video reviews. Table A1 in Appendix A counts the reviews in our dataset by source; we note that the text reviews are clustered in fewer sources, while the video reviews are more diffuse. Videos had 177,723 views on average, with a median of 88,816.

To analyze the security and privacy content of these reviews, we developed an initial qualitative codebook from other reviews of Internet-connected devices and software, also incorporating concepts from the Digital Standard [86]. Two researchers refined the codebook while collaboratively coding 21 reviews from our dataset; they then coded 9 reviews independently, achieving a Krippendorff's α of 0.83 averaged across each code that exhibited variation, which indicates good inter-rater reliability [55]. The two researchers then split all remaining reviews. Our codebook is in Appendix D of the supplementary materials.[5]

In general, we coded security- and privacy-relevant criteria even if they were not presented explicitly in that context (e.g., describing how multi-user access works for a smart lock). We reasoned that we were not equipped to judge whether the security and privacy implications would be apparent to consumers—this is a question for future work.

3.2 Review analysis results

Table 1 lists the most common security- and privacy-relevant criteria included in our dataset. More complete results are in Appendix E of the supplementary materials.

In our dataset, many criteria are included mainly in the context of a device's functionality: e.g., *human/AI processing* is usually present because we counted any mention of voice recognition, a way to control IoT devices, and *software updates* are often described as part of the setup process. Reviews for different product types cover different criteria corresponding to their core functionality: e.g., *audit log/notifications* is in most reviews of locks and doorbell cameras because notifying owners of people at the door is a popular feature, but it is rarely in reviews of thermostats. When mentioned at all, threat models are also typically tied to functionality (e.g., an intruder bypassing a smart lock to break into the house), although one review mentions botnets, describing them as a "larger societal problem" causing harm to banks and other institutions [81].

Some criteria unrelated to core functionality arise occasionally, with *encryption* and *multi-factor authentication* mentioned in 20–30% of lock and camera reviews: e.g., one re-

[5] Supplementary materials are located at https://osf.io/m2pe7/?view_only=e6a8443956704fe2b380cfce1def1204.

Table 1: All criteria included in more than 10% of the reviews for any product type. The table lists the percent of reviews with each criterion, per product type.

	Therm. (N = 23)	Lock (N = 24)	Camera (N = 24)
Human/AI processing	100	83	88
Audit log/notifications	9	79	79
Limiting/controlling data handling	22	4	92
Multi-user access control	9	83	12
Functional bugs	35	38	29
Locale of data storage/processing	4	12	83
Software updates	26	29	29
Can withhold action/capabilities	57	17	12
Data retention	0	0	67
Encryption	4	21	29
Multi-factor authentication	4	21	21
Password sec/priv	0	42	0
Sec/priv reputation of company	0	0	38
Data minimization/justification	4	4	25
Other data sharing	4	4	25
Sec/priv for special classes of data	9	8	17
Usability/accessibility for sec/priv	0	8	12
Recovery	0	17	0

view lists "no end-to-end encryption" as a reason to avoid a camera [82]. *Security/privacy reputation of company*, *data minimization/justification*, and *other data sharing* are mentioned in 25–40% of camera reviews. This mainly relates to criticism of Ring cameras for sharing data with police [71]: one review questions "whether a company with both financial and operational ties to law enforcement" should be trusted with sensitive personal data [92].

Few reviews mention techniques and tools used to evaluate devices. 56% give some indication of *"living with"* a device for an extended period of time or in a realistic environment. Fewer than 10% each report checking *customer feedback*, sharing a device with *multiple simultaneous users*, *communicating with the company*, or reading *policies/documents*. Rarely, reviews mention challenges: one reviewer explains that they do not cover privacy policies or user agreements in detail because "it's impossible for us to read and analyze every single one of these agreements" [89].

Throughout this analysis, we remained unsure whether the limited discussion of security and privacy was because these aspects were not evaluated at all or because they were a part of the review process that was simply not prioritized for communicating to the audience. Thus, in our interviews we included a section focusing on how product reviewers decide what to communicate in a review and what to leave out.

4 Product reviewer interviews

To dig deeper into how and why product reviewers evaluate security and privacy for Internet-connected products, we conducted 17 semi-structured interviews with 18 product reviewers over video call between February and May 2022, with these research questions:

1. What security and privacy criteria do product reviewers evaluate?
2. How do incentives and assumptions influence their approach?
3. What techniques and tools do they use?
4. What challenges do they face? What resources and tools do they need to be more effective?

4.1 Interview method

Recruitment. We recruited participants who were 18 years or older, spoke English, and had published at least ten reviews of Internet-connected devices and software.[6] We used our best judgment to only include reviews where the reviewer demonstrated they had interacted with products themselves beyond first impressions. In two instances we had two participants from the same organization: P1 and P2, and P9a and P9b. P9a and P9b review as a pair and were interviewed together.

To carry out recruitment, we compiled a list of eligible product reviewers from media companies, nonprofits, and YouTube channels that we were familiar with. We also used search engines extensively to find new reviewers and organizations, focusing on three types of Internet-connected products: smart home devices (e.g., thermostats and security cameras), wearables (e.g., watches and sleep trackers), and software (e.g., tax filing programs and photo storage services). We reached out directly to 144 individuals and organizations in this dataset using publicly provided contact information, yielding 15 participants. We also recruited 2 participants via an industry contact and 1 through snowball sampling. We continued recruiting until we reached saturation [40].

During recruitment, we avoided mentioning security and privacy, in order to reduce sample bias and avoid priming; we framed the interview as a study of how reviewers evaluate connected devices and software in general. During the interview, we did not bring up security and privacy until after asking participants generally what kinds of criteria they consider, to see whether they would be mentioned unprompted. While participants could have learned that we were security and privacy researchers from information online, there was no indication that any knew this other than P1 and P2, who are in security- and privacy-focused roles and were recruited via our industry contact.

Interview design. Our interviews were semi-structured, meaning that we broadly followed a protocol but adapted it and asked follow-up questions as appropriate for each participant. Interviews included four main parts. First, we asked

[6] P2 did not satisfy the publication requirement but had significant reviewing experience as a program manager.

background questions: e.g., about the participant's business model and experience. Second, we asked what criteria (first general, then security- and privacy-specific) they evaluate, as well as security and privacy criteria they consider important but do not evaluate. Third, we asked about techniques and tools they use to evaluate security and privacy, as well as challenges they had encountered. We also prompted them to imagine any hypothetical tools, resources, regulations, or industry norms they would want to help them evaluate security and privacy. Fourth, we asked participants how they communicate security and privacy information and how they approach negative reviews. Our interview protocol is in Appendix B. The recorded portion lasted 48 minutes on average.

At the end of the interview, participants completed a four-minute survey on their security and privacy knowledge, organization, and demographics (Appendix C). Participants received a $50 Amazon gift certificate as compensation; three refused compensation due to their organizations' policies.

Analysis. We recorded audio of our interviews, which was transcribed automatically and corrected manually. The first two authors coded four transcripts collaboratively, developing a qualitative codebook from scratch guided by the research questions. After four interviews, the high-level structure of the codebook was largely stable, so the first two authors coded the rest of the interviews separately, meeting after every two or three to resolve differences and update the codebook. Our goal was to identify and discuss qualitative themes pertaining to our research questions [12], not make quantitative claims, so we did not calculate inter-rater reliability [67]. Our codebook is in Appendix F of the supplementary materials.

Ethics. This study was reviewed and approved by the University of Maryland Institutional Review Board. We obtained informed consent, including for automatic transcription, and we told participants that they could skip any question they were uncomfortable with. We have taken care to not release identifying data about participants or their organizations.

Limitations. As this is a qualitative study with 17 interviews, our findings may not generalize to all product reviewers. To our knowledge, all participants have a primarily English-speaking, U.S. or European audience. There may be self-selection bias, as monetary compensation may provide limited incentive for busy professionals; compounding this, some organizations prohibit compensation and even participation. As a result, participants may disproportionately be enthusiastic about product reviewing, feel comfortable talking about their work because they believe it meets a high standard for quality or ethics, or want to help improve the state of the field; each of these sentiments was expressed by multiple participants.

Some participants may have over-emphasized their evaluation of security and privacy out of social desirability bias, especially if they believed this reflected the quality of their work. To mitigate this, we stressed at the beginning of interviews that there were no right or wrong answers. Throughout the interview, we repeatedly reassured participants that it was fine not to have substantive answers to questions about how they evaluate security and privacy, encouraging them to speak about barriers preventing them from doing so.

Table 2: Information on the number of years participants have been reviewing, the number of views for a typical review, and the number of reviewers at their primary organization.

	Years	Views	Team size
P1	3–5		30+
P2	1–2	1,000–9,999	30+
P3	10–14	10,000–99,999	4–9
P4	3–5		4–9
P5	15–19	10,000–99,999	20–29
P6	3–5		20–29
P7	3–5	1,000–9,999	1
P8	1–2	10,000–99,999	1
P9a	6–9	1,000–9,999	2–3
P9b	6–9	1,000–9,999	2–3
P10	40+		1
P11	15–19	10,000–99,999	2–3
P12	1–2	1,000–9,999	1
P13	1–2	1,000–9,999	1
P14	3–5	1,000–9,999	1
P15	10–14		4–9
P16	3–5	10,000–99,999	4–9
P17	15–19		2–3

4.2 Participant information

Table 2 contains self-reported information about participants' product reviewing experience and professional circumstances. For privacy, we report only aggregate statistics on demographics in Table A2 of Appendix A.

In the post-interview survey, for some additional context, participants self-reported their knowledge about online security and privacy (via a question developed by Faklaris et al. [27]). They consider themselves generally knowledgeable; 6 strongly agreed that "I am extremely knowledgeable about all the steps needed to keep my online data and accounts safe," 11 agreed, and 1 answered neutrally. We then measured participants' online security and privacy knowledge using six questions from a 2019 Pew Research Center poll [104]; responses are summarized in Table 3. Participants are generally knowledgeable, answering each question with greater accuracy than a representative sample of Americans in 2019. The only question answered correctly by fewer than 16 out of 18

Table 3: Summary of participants' responses to security and privacy knowledge questions, compared to responses given by a representative sample of Americans in 2019. No participant answered "Not sure" to any question.

	Participants	U.S. public [104]	
Question answer (rephrased from multiple choice answer)	% correct *(N of 18)*	% correct	% unsure
Cookies allow websites to track user visits and site activity	94 *(17)*	63	27
Advertising is the largest source of revenue for most major social media platforms	94 *(17)*	59	32
Privacy policies are contracts between websites and users about how a site will use user data	94 *(17)*	48	27
"https://" in a URL means that information entered into the site is encrypted	89 *(16)*	30	53
Phishing scams can be encountered through social media, websites, emails, and text messages	94 *(17)*	67	15
Private browsing mode stops someone on the same computer from seeing one's online activity	56 *(10)*	24	49

was about the purpose of private browsing mode; 10 answered correctly, while 6 held the common misconception [106] that online activities would be hidden from visited websites, and 2 chose other incorrect answers.

4.3 General product reviewing practices

In the next several sections, we describe our interview findings. For context, we report the number of interviews in which a theme or point appeared, out of a total of N = 17 (i.e., we do not double-count P9a and P9b, who review and interviewed together). As participants answered questions focusing on different aspects of their work that they personally found most relevant, counts should not be taken as measuring the true prevalence of these themes among our participants.

All participants had reviewed smart home or wearable devices; some also cover associated apps, standalone software, and more traditional devices including computers, phones, tablets, and routers. Five focus specifically on Apple or HomeKit products. Our participants publish video reviews (N = 4), written reviews (N = 3), or both (N = 10); some also mentioned other formats, such as social media or podcasts. They generally described their audiences as consumers or tech enthusiasts, while P10's audience also includes corporate buyers.

Participants evaluate products through day-to-day use in realistic settings (N = 11) and through testing in a lab or home studio (N = 8). Two do not evaluate products directly: P17 has experience testing products but currently conducts online research and interviews users, while P2 is a product manager overseeing review processes. Seven participants mentioned following a standard evaluation framework. The length of testing before publishing varies greatly by reviewer and by product, ranging from hours to a month or more. Many have updated reviews, generally by revising them directly (N = 10), although on YouTube reviews must be updated in other ways, such as adding a pinned comment or updating the description (N = 6). When asked how they determine what product to review next, participants gave a variety of factors, including picking products that are popular (N = 9), are interesting (N = 7), or have brand presence (N = 6).

Business model. Affiliate marketing (N = 13) was the most common source of compensation reported, followed by ads (N = 9), sponsorships (N = 6), free products (N = 4), and subscriptions (N = 3). While we did not prompt participants for measures they take to reduce potential resultant bias, nine mentioned these organically, including maintaining separation between product reviewers and people who make business decisions (N = 5), avoiding sponsored reviews (N = 4), avoiding ads (N = 2), buying all products themselves (N = 2), and not using products outside of reviewing work (N = 2).

4.4 Security and privacy criteria

Only four participants mentioned security or privacy when asked what criteria they evaluate. Of these, P1 and P2, who focus on reviewing security and privacy, listed many relevant criteria; P9a mentioned checking for local and cloud storage because that was a privacy concern their audience cares about; and P8 mentioned security only to say that they do not cover it much unless there is a glaring issue.

After prompting, all participants reported at least sometimes evaluating security and privacy. Some always do: for smart home devices, P1's organization follows a testing framework covering authentication, encryption, security over time, and more, and P6's organization always considers two-factor authentication and whether data storage and processing are on the device or in the cloud. However, others rarely do: P8 said, "If I see something [security- or privacy-related] that seems glaring, I'll call it out, but my audience really isn't that type." Some organizations have separate security and privacy experts who help with reviews (N = 6). Relatedly, five participants saw security as someone else's job: P5 said, "Those kinds of vulnerabilities ['gaping security flaws'] are things that security experts find later on or get publicized as an exploit, which isn't something I can necessarily test." All criteria mentioned by at least two participants are listed in Table 4.

Reviewers prioritize criteria differently. With many possible security and privacy criteria, reviewers must prioritize. Eight participants mentioned tailoring priorities to the type of product. Sometimes, this means prioritizing different criteria:

Table 4: All security and privacy criteria evaluated by at least two participants.

Count	Criterion
11	What data is collected, and how is it shared/used?
7	Reputation of company/product (e.g., breaches)
6	Locale of data storage/processing (local or cloud)
5	Encryption
5	Known vulnerabilities
5	Multi-factor authentication
5	Physical shutters and shut-off switches
4	Data controls
4	Measures to secure data against hacking
4	Transparency about data handling
3	Authentication
3	Data deletion
3	Geographic locale of data storage/processing
3	Justification of data handling
3	Length of data retention
2	Full platform compatibility
2	Security over time
2	Software updates
2	Usable security and privacy features

P1 said, "The most important thing for IP cameras is the audio data and video data. But for smart TVs, people are more caring about if they get monitored while they're watching TV." Other times, this means elevating security and privacy for more concerning products: P6 said, "Cameras are scrutinized a lot harder than other devices, because that's actually a gateway into your home. And like how terrifying was it two years ago when Ring cameras were getting hacked and actually like communicating with children in their homes?" And sometimes, this means deprioritizing security and privacy if they were perceived as forgone: P10 said, "If I'm reviewing security cameras, it's a big deal. If I'm reviewing a social media app—you've thrown away your privacy."

Participants differed in the overall weight given to security and privacy. P6 said their organization would not recommend products that violate a "baseline" of security, such as transmitting maps of a user's home insecurely. Similarly, P17 said that at a previous organization, they would not recommend a product with missing security and privacy information, even if it scored well on all other criteria. In contrast, other participants viewed security and privacy as just two of many factors: P10 mentioned a formula with various weighted criteria that includes security for certain products. Still others generally did not consider security and privacy: P15 said that given a limited word count, privacy is one of the first aspects their organization will cut, because functionality is more important.

Reviewers sometimes recommend how to configure or use products. P1 guides audiences through data controls such as opting out of data sharing, and P3 teaches audiences to prevent devices from communicating with the Internet using a HomeKit router. P11 provides general security advice in their reviews, emphasizing the importance of installing updates, using password managers, and enabling multi-factor authentication. Given limited space, however, some create or link to separate security and privacy content: P6's organization provides how-to guides on deleting recordings, resetting devices, and more. P6 views providing this advice as part of responsible reviewing.

4.5 Incentives and responsibilities for covering security and privacy

The extent to which product reviewers cover security and privacy, as well as what they prioritize, is influenced by perceived incentives, disincentives, and responsibilities.

Protecting their reputation is an incentive for reviewers to evaluate security and privacy. Ten participants said their reputation is tied to their reviews, meaning that recommending insecure or invasive products could hurt their credibility. P8 said, "I would hate to have a glowing review about a product, and then two weeks later, they have a data breach because they did something stupid—they didn't encrypt something … [it] makes you look bad." Some described this as a responsibility: P10 said not evaluating important security and privacy criteria would be cheating their readers. However, one participant came to the opposite conclusion: P12 said concern for their reputation makes them hesitant to comment on security and privacy without more expertise, because making a wrong assertion could discredit them and endanger their audience.

Some see audiences as uninterested in security and privacy. Six participants cited lack of audience interest as a disincentive for including security and privacy in their reviews. P8 said, based on the questions their audience asks, "There are definitely people out there that want to see [security-related content], but there's a lot of people that don't care." P11 said, "If I do a video about [security vulnerabilities] … it doesn't get searched for; it doesn't get watched; nobody cares about it," and P6 recalled creating security- and privacy-focused content that received "horrible traffic." Even when manufacturers highlight security and privacy features in reviewer guides, P4 said their organization often skips them because they're not "sexy," adding that "No one's really complained."

Even when reviewers do evaluate security and privacy, they may avoid reporting their findings to their audiences due to lack of interest. As mentioned in Section 4.4, P6's organization examines all smart home devices for a security baseline. However, they may not write about this, because their audience may not understand or care. If a product does not meet the baseline, "We wouldn't even be recommending the device in the first place." When asked, P6 guessed that some but not all audience members are aware of this rule.

Some audiences are interested in security and privacy. On the other hand, audience interest is sometimes an *incentive*

to evaluate security and privacy. P14 said, "Sometimes the audience will be wondering, what if I don't want this camera on, or how can I protect myself?" More narrowly, P17 said, "A lot of commenters and people emailing you will be like, 'This thing's dialing into China.'" And particular audiences may care: P2 said their organization increased their coverage of security and privacy in order to appeal to younger audiences and people with children.

Incentives against negative reviews impede publication of security and privacy concerns. Though reviewers have some incentives to consider security and privacy, these do not necessarily translate into incentives to *publish* reviews of products with major concerns. Many participants said they avoid publishing negative reviews (N = 10) or picking bad products to review in the first place (N = 8): P11 said, "I'm not comfortable right now with the security of [a particular IoT device], so I just won't even [review] it." Thus, consumers may not receive information about security and privacy risks.

Participants gave different justifications for avoiding negative reviews. P16 said reviewing bad products doesn't prevent people from buying them and may even be counterproductive by giving them more attention: "there's still going to be people that buy a bad product because it's cheap. ... I don't want to give it air time." P7 suggested consumers will avoid bad products even without the help of negative reviews, citing a poorly received product where "nobody reviewed it, and it's not selling." And multiple participants said consumers are less interested in negative content: P11 recalled posting a video highlighting bad products that weren't reviewed, but they said those types of videos did not attract much traffic. Similarly, P4 said their organization avoids publishing negative reviews because people will get the message from the headline without reading the review or buying the product, depriving them of revenue from ads and affiliate marketing.

Negative reviews are important to some. Four participants mentioned that avoiding negative reviews would decrease their credibility, and three felt it was their duty to publish negative reviews that were in consumers' best interest. For example, P13 said when audiences question their integrity based on the fact that they receive affiliate revenue, they "often point to other videos where I would make more money if I'd pushed that product, and I don't."

Reviewers who do not publish negative reviews may still mention security and privacy concerns for products that receive a neutral or positive review. In a review of a device that shared sensitive data with servers by default, P3 warned their audience, "If you aren't comfortable with that, don't get this product, or here's a way that you can limit that." However, overall, we find that incentives against publishing negative reviews more likely dissuade reviewers from informing consumers of security and privacy concerns.

Thorough evaluation is limited by finances. Across different kinds of organizations, financial constraints limit security and privacy evaluation. P7 said creating content on platforms like YouTube takes time and often doesn't pay well, which means that "putting something through its paces, like legitimately testing something ... is just not possible." And P1 pointed out that searching for security and privacy issues often leads to "finding nothing in the end," meaning that "the constraints of the budget really make a huge difference" to whether they can investigate a potential issue.

Some aim to protect consumers by helping companies detect issues. Given incentives against negative reviews, some participants expressed a duty to protect consumers in other ways. Eight mentioned disclosing issues or vulnerabilities they discover to companies. P8, who tries to avoid publishing negative reviews, identifies as a "de facto beta tester" and has delayed publishing reviews until bugs they reported are fixed. They said, "If I do have a closer relationship with a company, I can maybe help drive them in a more pro-consumer direction." P7, who avoids publishing negative reviews altogether, wanted to be more like a beta tester, with earlier access to products so their feedback could be taken into account before release. In some cases, companies may not be receptive to this feedback: P1 said, "Sometimes, when we find out some issue on the device—say, your device security is not up to date ... the company will say, 'No, it's good!'" To address this, they suggested a law to intervene when companies "ignore security researchers who are just being kind and trying to help."

4.6 Assumptions guiding prioritization

Along with incentives, product reviewers' assumptions about security and privacy play an important role in their choice of priorities. (We defer comment on the reliability of these beliefs to Section 5.)

Apple, HomeKit, and other intranet-based products are seen by some as secure and private. Eight participants believe Apple and HomeKit-certified products to be inherently secure and private, including all five who focus on those brands. For example, P16 said they did not think much about security, because they "just implicitly trust" HomeKit-certified products, and P7 trusted HomeKit-certified products despite the concerns of their audience: "You got a lot of people who don't trust, like, new, unknown Chinese brands. And I say, well, Apple has obviously verified, because they have access to HomeKit. So, if Apple trusts them, I trust them."

Participants trust HomeKit devices in part because they are configurable to allow only local storage and processing, which significantly limits potential attacks. P3 said, "With HomeKit, everything is like inherently secure. Stuff runs local; there's no external server calls. If someone was trying to hack into your smart home devices, they're gonna have to basically

be in your house, on your home Wi-Fi network." Relatedly, especially for locks, P8 said they encourage their audience to use local mesh networking protocols, such as Zigbee, Z-Wave, and Thread, explaining that they are "effectively 100% secure" because a compromised device "can't hop over to the Internet and go out and get instructions."

Participants cited a variety of other reasons for trusting Apple and HomeKit. P3 trusts Apple's certification process for HomeKit devices, describing it as "the gold standard in testing." They added, "While we've heard a multitude of stories of security and privacy issues with Amazon assistant and Google Assistant products, that has never been the case for a HomeKit one." P7 trusted HomeKit devices because of the end-to-end encryption used to transmit and store data, as well as the "kind of ridiculous" certification process. They specifically called out the comparatively low number of HomeKit-certified devices, unlike other smart home platforms, where a multitude of devices implies "super low standards." Participants also mentioned trusting Apple, as a company with a reputation for security and privacy to uphold: P12 said, "Amazon and Google are data companies, and fundamentally that makes me not trust them as much. ... You are the product. Whereas with Apple, they have a long-standing reputation that actually privacy is at the center of what they do."

Due to their trust in HomeKit, some participants do not cover security and privacy unless issues arise: P7 said, "People who are watching my content understand the level of security and privacy that comes along with HomeKit in general. So, I used to talk about it a lot, but I stopped, because there's no need to—because everybody understands that it's just as secure and as private as can be. ... Unless there's an issue. If there is an issue, I'll talk about it." Two participants perceive unusual HomeKit integration as a security and privacy issue unto itself: P16 said products that require an app other than Apple Home raise red flags.

Participants sometimes explain to their audience or remind them that HomeKit devices are safe (N = 2), but they also try not to repeat this too much across reviews (N = 2): P12 described a "tacit assumption or ... maybe false hope" that audiences would already know this from watching their other videos. In trying not to bore their dedicated audiences, though, reviewers risk omitting important security and privacy context for others who happen upon their reviews and may not share the same expectations.

Reviewers make assumptions about price and prominence. Six participants said free or cheap products typically have privacy trade-offs: P3 said, "If something is really affordable or has some sort of monthly free cloud option ... you're the product." Relatedly, P1 said their organization associates investment in security and privacy features, such as encryption and vulnerability disclosure programs, with competence: "If companies are able to put money in those areas, then we trust that the companies are able to make the device well."

Three participants believe that prominent brands are less likely to have security and privacy issues; conversely, P13 tells their audience that bigger companies are bigger targets. These assumptions may influence the level of scrutiny reviewers give to different products.

Threat models inform how reviewers evaluate and communicate security and privacy. Participants expressed many beliefs about what threats are realistic. Three asserted that simple attacks, such as dictionary attacks on passwords or simply picking a lock, are more common; this was given as a rationale for prioritizing simple threats over complex ones when evaluating security and privacy. P17 said the app for a smart device is often more concerning than the device itself, adding that they wished they were better at packet sniffing in order to investigate these concerns. P11 argued that by making products easy to use, companies can leave them vulnerable; in reviews, they recommend against connecting network-attached storage devices to the Internet, in order to prevent ransomware and other problems. P11 also argued that attacks by third parties (e.g., hackers distributing ransomware) are more concerning than what companies do with user data; as such, they may focus on different aspects of security and privacy than reviewers such as P4, who only mentioned companies that "try to figure out ways to sneakily do things" with their own products as a threat.

Some believe security and privacy are impossible or impractical. Five participants said all Internet-connected products are fundamentally insecure. Most simultaneously acknowledged that protective measures are still valuable: P10 added, "We can put layers of trying to protect ourselves, ... and I expect that from the products." However, P5 was more fatalistic, saying anything that connects to the Internet is a security hole: their organization doesn't prioritize security for most products, because vulnerabilities are "a fundamental of the [IoT] category and just of all network technology."

Similarly, seven participants expressed hopelessness about privacy. Some argued privacy erosion is inescapable in modern life generally: P16 said, "We've got Amazon spying on us, Google spying on us as well—it's a global thing," which contributes to their belief that it's "not worth the hassle" to focus on privacy in reviews. Others argued that Internet-connected products inherently sacrifice privacy: P9a said, "If you're someone that's starting to put smart home stuff in your house, you've already given up on the privacy thing."

Three participants argued that protecting security and privacy is technically possible but often inaccessible to most consumers. P8 personally believes it is important for a home network to isolate IoT devices, but they said doing so is beyond most people's capability: "You have to go buy an enterprise-level firewall; you have to segment out your Wi-Fi; you've gotta figure out which ports you have to open; you gotta test it ... it's not easy." Overall, these beliefs have clear potential to

Table 5: Techniques and tools used by participants to evaluate security and privacy.

Count	Technique or tool
5	Ask company questions
4	Examine data in transit
3	Reading privacy policies and other docs
2	Static and dynamic analysis
1	Automated privacy policy analyzer
1	Check customer feedback
1	Document manager for privacy policies
1	Evaluate code and libraries
1	Intuition
1	Monitor network security using firewall
1	Security and privacy label
1	Tool to visualize network traffic
1	Track privacy policy updates via hashes
1	Verify effectiveness of security and privacy features

shape reviewers' judgment about the extent to which security and privacy are worth evaluating.

Since total security and privacy is impossible, it makes sense that six participants emphasized tradeoffs as personal decisions for consumers to make for themselves. They often saw their role as informing these decisions without making strong recommendations: P13 said, "I just try to lay all that out there and then try to give my opinion, which generally errs on the side of, like, 'Hey, if you want to survive in this world, this day and age, you're going to probably use some products that are connected.'" Similarly, P14 said, "You have to figure out, how much privacy do you want to give up for the features that you want to get? So, yeah, that's what I would say. Not so much as like, 'Don't buy this just because of that.'"

4.7 Techniques and tools

Participants reported using 14 different types of techniques and tools to evaluate security and privacy, listed in Table 5. Seven participants reported techniques and tools that involved interacting directly with products: e.g., using Wireshark or routers with network monitoring capabilities to examine data in transit, or using static analysis tools to evaluate code. In addition to technical tools, some participants apply intuition from years of experience: P6 said, "After using a lot of smart home apps and services, you start to develop a little bit of a spidey sense of what looks like something you want to connect to your home network or not." Five reported asking manufacturers questions about products. Four said they review privacy policies and other written documentation, with two using custom tools to automatically distill important components of privacy policies or track their changes over time. On the other hand, four participants explicitly mentioned not conducting penetration testing out of practicality.

Limited expertise and time impede reviewers' use of techniques and tools. Participants reported numerous challenges, both to evaluating security and privacy and to reviewing Internet-connected products in general. Limited security and privacy expertise was a common barrier (N = 11), as was limited time and concentration (N = 6). These were often described as reasons for not evaluating security and privacy in the first place: P17 said, "If a company tells you that something is two-way encrypted, I don't have the skillset to prove it one way or the other." P8 does not use tools such as Wireshark and syslog, which can be used to monitor relevant device and program behavior, because "it takes knowledge and time." However, some participants did mention experiencing these challenges while actually attempting to evaluate security and privacy: P6 recalled their "eyes glossing over a little bit during privacy policy reading" early in their career, due to unfamiliarity with "jargon."

Lack of transparency is a common barrier. Nine participants cited lack of transparency or honesty from companies as a challenge, especially for techniques that rely on asking questions or reading documents. P3 was reluctant to rely on Apple privacy labels, noting instances where labels did not match behavior. P8 said many small companies are particularly opaque about backend practices; conversely, P17 pointed out that large companies have the resources to fight transparency. Citing Apple's history of litigation to prevent jailbreaking, they said, "You have to take them at their word and their reputation ... they don't have any open-source code; they don't disclose a lot more than is necessary governmentally about their products and how the security works on them." P17 also gave an example from their own experience of a device made by "one of those like shell companies inside of a shell company inside of a shell company," which would not answer questions about its data model.

P1 said vague or contradictory privacy policies are difficult to evaluate: "We do see a lot of companies using really vague language and kinda talking good things in the first paragraph; then in the second paragraph they basically just contradict themselves and then leave an open-ended loophole at the end for themselves to do whatever they want. ... I feel powerless when we evaluate it."

4.8 Desired tools and resources

Participants suggested tools that would aid in security and privacy evaluation. While we prompted participants to imagine tools they would personally use, we note their responses are nonetheless hypothetical.

Better tools to inspect network traffic (N = 10). P17 described a tool to connect with a device and report what servers and IP addresses it contacts, what data it sends and receives, and more, comparing this to a hacker's blog post breaking down flaws in an IoT device. They expounded, "Here's everything about this device that this little magic mouse can figure

out. . . . What do you think about the fact that it's sending unencrypted traffic to Turkey—or to Indiana? What do you think about the fact that it's running a Linux kernel that's like five years out of date?" Other participants described more narrowly scoped tools: P16 wanted a tool to determine if devices were "dialing home" to an external server, especially in another country, and P1 wanted to decrypt data between a device and the Internet to see what data is sent to whom.

While most did not reference existing tools, P13 wanted "a tool like Wireshark that I actually understood better." As a "wannabe network person" and not a security expert, they consider Wireshark "cool to know but not worth my time to learn." Ideally, they would like information to be collected automatically and presented in a dashboard. Similarly, P3 said they currently use routers to monitor which IP addresses a device connects to and how frequently, but they also want to know what data is being transmitted.

Resources and tools to verify encryption (N = 3). P10 described a third-party organization to evaluate encryption: "I don't want to have to just trust the website saying this is an encrypted site. I want a certification, from my trusted overseeing organization." P7 wanted a tool that would not only verify encryption but also facilitate understanding; they called out end-to-end encryption as "marketing fluff" that left them asking, "What does that mean?"

Automated monitoring of network traffic for suspicious behavior (N = 4). P9a described a "firewall on steroids" that would use AI to detect unexpected behavior. P8 described using machine learning to detect concerning traffic, such as "weird traffic going to Russia" or a light switch transmitting "three gigs of data in the last 24 hours."

Assistance evaluating the full life cycle of data (N = 7). Participants wanted to know what happens to personal data over the course of its existence, beyond network traffic. P4 wanted a tool to reveal not only what data is sent where, but also where it is stored and how it is used. Similarly, P8 wanted an "x-ray" into "everything that the device is doing, every stage of its operation," including how data is stored and consumed on the backend.

Rather than a tool, four participants wanted companies to disclose this information. This drew comparison to existing disclosure and transparency requirements: P13 suggested the FTC might implement such a program, P8 compared it to SEC filings designed to prevent insider trading, and P9a pointed to the GDPR as an example.

Some participants mentioned using privacy policies to understand the data life cycle. P15 wanted a tool to analyze privacy policies and other documents and produce an accessible summary of data collection, use, and controls, similar to Apple's privacy labels. P9b and P15 also wanted a requirement that privacy policies be written in plain, accessible language.

Labels (N = 7). Participants expressed interest in security and privacy labels or ratings, citing as analogues the French government's repairability index and UL standards for electronics products. Some envisioned a single numeric score, sometimes with a link to a website with more information: P7 and P17 both described a color and a number from 1 to 10 indicating general security and privacy risk, and P8 wanted a number from 1 to 4 indicating how data is stored and shared. On the other hand, one participant whose organization already creates security and privacy ratings said they look forward to providing information for labels.

Automated detection of vulnerabilities (N = 3). P3 described a hypothetical tool that would produce a rating from 1 to 10 on how "hackable" a product is, while P12 described a tool that would show them how an attacker might try to break into a system and what the attacker might be able to do. P1 wanted extensions to existing static analysis tools to better analyze firmware and iOS specifically.

5 Discussion

In this section, we characterize the kind of security and privacy advice that product reviewers provide, including suggestions for future research and outreach. We also recommend technical tools, resources, and other changes to support product reviewers in evaluating security and privacy without requiring them to become experts or devote significantly more time.

Reviewers have mixed incentives to evaluate security and privacy, but they are uniquely positioned to guide consumers to safe choices. Product reviewers are incentivized to evaluate security and privacy in order to protect their reputation. However, they are disincentivized in other ways: many do not review bad products or publish negative reviews, they may leave out information because audiences are not interested, and evaluating security and privacy is plain difficult. While our participants are generally well positioned to understand their current audiences, future research could explore consumers' knowledge and interests more rigorously and at scale, in an effort to find more effective ways to communicate security and privacy information. Some reviewers assume that negative reviews would have little impact or even do harm by providing "air time" to bad products; prior work suggests this may not be true, finding that negative user reviews can have large dissuasive effects [74, 105, 110]. Further data specifically in the contexts of security and privacy and of professional product reviews could motivate reviewers to change their approach, although financial incentives to avoid negative reviews would still remain.

Nevertheless, product reviewers are very diverse: for every disincentive that several participants mentioned, someone else reported an opposite *incentive*. Despite the disincentives,

reviewers are technical professionals who currently provide important security and privacy advice to consumers, even if it is infrequent or skewed. They fill a crucial gap: given the sheer quantity of Internet-connected products, they are arguably the only people evaluating the security and privacy of popular products at all with regularity and timeliness. We see great potential in efforts by the research community to help product reviewers guide consumers, in ways that mesh with their existing processes, business incentives, and resources.

Reviewers cover meaningful criteria, but there are gaps. Participants reported evaluating a wide variety of security and privacy criteria. Overall, they tend to prioritize protection of user data, which can be difficult for reviewers to evaluate, and to deprioritize indicators that a product can maintain security and privacy over the long term, such as audits, update support, and presence of bug bounties; these were neglected by the published reviews we analyzed as well. These long-term indicators, however, can often be evaluated without new techniques or tools: e.g., by looking online or asking a company directly. We recommend that more product reviews include these longer-term criteria, so that consumers know which products are likely to *remain* functional and safe. Various groups could play a role in encouraging product reviewers to incorporate this information, including the FTC, which regulates monetized content; the Digital Standard, which could provide information on suggested priority and ease of evaluation in its list of criteria; and companies themselves, who could highlight security and privacy features that distinguish their products from others. Importantly, reviewers may need to explain to consumers *why* these features are protective; Emami-Naeini et al. found that some people believe security audits and updates indicate poor security [23]. Product reviewers themselves may also benefit from this messaging, as we found that some may share these misconceptions: P9a described firmware updates as a sign of "crappy engineering" indicating that products were released before they were ready.

Threat models seldom include botnets or misuse by legitimate users. Our participants discussed potential adversaries including manufacturers, hackers, and governments, overwhelmingly with the goal of accessing user data. No participants, and only one review we analyzed, mentioned botnets or any other threat primarily targeting someone other than the product owner. However, botnets are one of the most likely attacks IoT device users will experience, as they are employed at scale to conduct distributed denial of service attacks [6], spread spam and ransomware [68], and mine cryptocurrency [54].

It is perhaps unsurprising that reviewers rarely consider botnets, as their audiences may not care; these attacks often cause little direct, observable harm to the end user. However, they may have important secondary effects, such as disabling crucial protections in Windows Defender [68], and can be detrimental to others and to society at large. Future research could measure how botnet infections affect end users, from collateral security and privacy harms to increased energy costs and device wear, as well as consider the effect of altruistic messaging, in order to motivate product reviewers and end users to consider this threat.

Our participants also did not discuss preventing misuse by physically co-located users, such as guests using smart devices to access private information without permission. Accordingly, no interviewees and few analyzed reviews mentioned multi-user access control or parental controls as a security or privacy criterion. Again, this is unsurprising, as people often deprioritize these concerns and put their trust in social norms [107]. However, work by Moh et al. indicates this kind of everyday misuse is widespread for smart home devices [69]. Product reviewers could help inform potential new users about these considerations, and they could elevate products with more usable and secure access controls.

Defeatist attitudes toward security and privacy do not benefit consumers. While understandable, reviewers' beliefs that security and privacy are impossible or impractical obscure the more complicated reality: there are meaningful, if imperfect, steps consumers can take to reduce risk. Systematic evaluations of IoT devices have found, for example, that while many devices rate poorly on security and privacy, some do significantly better [2, 63]. Product reviewers could help empower consumers—who themselves frequently also believe that unwanted data collection [8] and hacking [37] are inevitable and uncontrollable—by providing security and privacy advice while acknowledging the undue burden that falls on consumers to protect themselves. Real-world success stories should be highlighted as evidence that security and privacy are worth caring about; in addition to outreach from advocates and companies, trade shows could feature these stories, as many product reviewers regularly attend events such as CES, which brings together people interested in consumer electronics.

Security heuristics can be useful but may not be valid. Our participants used heuristics, such as a manufacturer's reputation or whether a product connects to the Internet, to quickly evaluate security and privacy concerns. These assumptions often have some basis in fact and allow reviewers to provide some kind of security and privacy guidance without complex tools and analysis. However, they are not universally reliable and could lead to misinformed recommendations.

Ideally, security and privacy labels generated by a trustworthy third party might supplant these heuristics and provide accessible, reliable information for reviewers who face barriers doing their own evaluation. App privacy labels are a step in the right direction, but our own study and related work show that more work is needed to earn the trust of product reviewers and consumers. Many efforts to design IoT labels are

underway [24, 39, 73]; we recommend that product reviewers be considered important stakeholders who could digest information from an eventual label on behalf of consumers.

Specifically, several participants expressed strong trust in HomeKit-certified devices. This seems partially justified: certified devices are required to follow certain security practices, including in encryption [5]. Nevertheless, vulnerabilities have been reported in HomeKit itself [72, 97], in its integration with other IoT management frameworks [45], and in numerous HomeKit-certified devices [14–17, 88]. Some problems could be avoided if end users follow reviewers' recommendations to use these devices only through the Apple Home app or disable remote processing, but it is unclear to what extent they do so. Similarly, while other local mesh networking protocols in properly configured devices should prevent many Internet-based attacks, vulnerabilities still exist [48].

Overall, more work is needed to evaluate HomeKit certification and other frameworks, such as Matter, rigorously, at scale, and as actually used in practice; this would provide important insight into when reviewers' heuristics make sense and what other security and privacy criteria to prioritize. As more information emerges, ensuring it is well publicized is crucial; anecdotally, we observed that participants were closely tuned in to news about IoT security, with many mentioning the Wyze camera hack, Apple's certification processes for HomeKit devices, and Matter's planned security features.

Better tools could help evaluate security and privacy. Given the strong time and audience-interest limits many reviewers face, only some will be willing or able to independently evaluate security and privacy; for these few, it is crucial to provide tools that are extremely usable and accessible.

Most commonly, participants asked for tools to help understand network traffic quickly and easily, despite limited expertise. Many existing tools, such as packet sniffers, can be confusing even for experts [76, 103]. Instead, reviewers could use tools that provide summaries and visualizations of network traffic [101, 102], or even better, automatically flag concerning behaviors [22, 36]. Because many reviewers live with devices for a period of time, signals of misbehavior in real time [96] could also be useful. Further work is needed to make tools like these usable (and maintainable) in practice, and to make reviewers aware of them.

Relatedly, participants wanted to understand how data is stored, shared, and used after it leaves the home. While some existing work attempts to evaluate these issues in a black-box manner [2, 3], in many cases this is infeasible without companies making significant infrastructural changes to improve data handling and transparency [11]. The best path forward here may be regulations and norms that improve transparency, responsiveness, and truthfulness from manufacturers and enable audits to validate provided information.

Some participants imagined black-box tools that would automatically determine security properties, such as "hackability." While achieving this in the general case is intractable, program analysis tools can provide relevant insights [29, 52, 56, 59, 77]. Like packet sniffers, however, many program analysis tools are difficult even for experts to use [93, 94], and most are designed for use during software development. Reviewers would need new tools, designed for easy setup without source code, that provide interpretable snapshots of security and privacy characteristics.

While no participants requested tools to help understand privacy policies, we note that such tools could be useful to reviewers who already review policies manually. Several such tools have been developed [9, 41, 84], but they may not adequately meet the needs of reviewers, and reviewers may not be aware of them.

6 Conclusion

For this paper, we conducted 17 interviews with product reviewers and analyzed 71 published reviews in order to understand the role of professional product reviewers in evaluating the security and privacy of Internet-connected devices and software. We characterize the criteria and threat models reviewers consider, as well as how they fit into their review process. We find that while reviewers have incentives to consider security and privacy, they face significant disincentives and challenges, including audience disinterest, lack of expertise, potentially unreliable assumptions, and a dearth of effective and usable tools for evaluation. We make recommendations for further research and tool development to support product reviewers, within the constraints of their practices and incentives, in shifting the burden of security and privacy evaluation away from end users and toward professionals.

Acknowledgments

We thank our participants for their time and insights, our reviewers for their constructive feedback, and all who gave us advice or helped with recruitment, including Nora McDonald, who helped contextualize our qualitative analysis. This research was supported by the SPLICE research program under NSF SaTC award #19555805.

References

[1] Omer Akgul, Richard Roberts, Moses Namara, Dave Levin, and Michelle L. Mazurek. Investigating influencer VPN ads on YouTube. In *Proceedings of the 2022 IEEE Symposium on Security and Privacy*, pages 876–892, May 2022. https://doi.org/10.1109/SP46214.2022.9833633.

[2] Omar Alrawi, Chaz Lever, Manos Antonakakis, and Fabian Monrose. SoK: Security evaluation of home-based IoT deployments. In *2019 IEEE Symposium on Security and Privacy*, pages 1362–1380, May 2019. https://www.doi.org/10.1109/SP.2019.00013.

[3] Omar Alrawi, Chaoshun Zuo, Ruian Duan, Ranjita Pai Kasturi, Zhiqiang Lin, and Brendan Saltaformaggio. The betrayal at Cloud City: An empirical analysis of cloud-based mobile backends. In *Proceedings of the 28th USENIX Security Symposium*, pages 551–566, August 2019. https://www.usenix.org/conference/usenixsecurity19/presentation/alrawi.

[4] Edgar Alvarez. YouTube stars are blurring the lines between content and ads. *Engadget*, July 2017. https://www.engadget.com/2017-07-25-youtube-influencers-sponsored-videos.html.

[5] Mahmoud Ammar, Giovanni Russello, and Bruno Crispo. Internet of Things: A survey on the security of IoT frameworks. *Journal of Information Security and Applications*, 38:8–27, February 2018. https://www.sciencedirect.com/science/article/pii/S2214212617302934.

[6] Manos Antonakakis, Tim April, Michael Bailey, Matt Bernhard, Elie Bursztein, Jaime Cochran, Zakir Durumeric, J. Alex Halderman, Luca Invernizzi, Michalis Kallitsis, Deepak Kumar, Chaz Lever, Zane Ma, Joshua Mason, Damian Menscher, Chad Seaman, Nick Sullivan, Kurt Thomas, and Yi Zhou. Understanding the Mirai botnet. In *Proceedings of the 26th USENIX Security Symposium*, pages 1093–1110, August 2017. https://www.usenix.org/conference/usenixsecurity17/technical-sessions/presentation/antonakakis.

[7] Apple. Privacy: Labels. https://www.apple.com/privacy/labels/.

[8] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. Americans and privacy: Concerned, confused and feeling lack of control over their personal information. Technical report, Pew Research Center, November 2019. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/.

[9] Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Cranor, Shomir Wilson, Florian Schaub, and Norman Sadeh. Finding a choice in a haystack: Automatic extraction of opt-out statements from privacy policy text. In *Proceedings of The Web Conference 2020*, WWW '20, pages 1943–1954, April 2020. https://doi.org/10.1145/3366423.3380262.

[10] Tanya Basu. Anti-vaxxers are weaponizing Yelp to punish bars that require vaccine proof. *MIT Technology Review*, June 2021. https://www.technologyreview.com/2021/06/12/1026213/anti-vaxxers-negative-yelp-google-reviews-restaurants-bars/.

[11] Eleanor Birrell, Anders Gjerdrum, Robbert van Renesse, Håvard Johansen, Dag Johansen, and Fred B. Schneider. SGX enforcement of use-based privacy. In *Proceedings of the 2018 Workshop on Privacy in the Electronic Society*, WPES'18, pages 155–167, October 2018. https://doi.org/10.1145/3267323.3268954.

[12] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006. https://doi.org/10.1191/1478088706qp063oa.

[13] Eliza Brooke. Why so many recommendation sites promise to help you find the best stuff. *Vox*, December 2018. https://www.vox.com/the-goods/2018/12/11/18131224/recommendations-best-strategist-wirecutter-buzzfeed-reviews.

[14] Robert Byers, Chris Turner, and Tanya Brewer. CVE-2020-6007. https://nvd.nist.gov/vuln/detail/CVE-2020-6007, August 2020.

[15] Robert Byers, Chris Turner, and Tanya Brewer. CVE-2021-27954. https://nvd.nist.gov/vuln/detail/CVE-2021-27954, August 2021.

[16] Robert Byers, Chris Turner, and Tanya Brewer. CVE-2021-35067. https://nvd.nist.gov/vuln/detail/CVE-2021-35067, October 2021.

[17] Robert Byers, Chris Turner, and Tanya Brewer. CVE-2022-27152. https://nvd.nist.gov/vuln/detail/CVE-2022-27152, April 2022.

[18] Rachel Cericola. How Wirecutter vets the security and privacy of smart home devices. https://www.nytimes.com/wirecutter/blog/smart-home-security-privacy/, September 2020.

[19] CTIA Certification. IoT cybersecurity certification. https://ctiacertification.org/program/iot-cybersecurity-certification/.

[20] Federal Trade Commission. The FTC's endorsement guides: What people are asking. https://www.ftc.gov/business-guidance/resources/ftcs-endorsement-guides-what-people-are-asking, August 2020.

[21] Cyber Security Agency of Singapore. Cybersecurity Labelling Scheme. https://www.csa.gov.sg/Programmes/certification-and-labelling-schemes/cybersecurity-labelling-scheme/about-cls.

[22] Sohaila Eltanbouly, May Bashendy, Noora AlNaimi, Zina Chkirbene, and Aiman Erbad. Machine learning techniques for network anomaly detection: A survey. In *Proceedings of the 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies*, pages 156–162, February 2020. https://doi.org/10.1109/ICIoT48696.2020.9089465.

[23] Pardis Emami-Naeini, Janarth Dheenadhayalan, Yuvraj Agarwal, and Lorrie Faith Cranor. Which privacy and security attributes most impact consumers' risk perception and willingness to purchase IoT devices? In *Proceedings of the 2021 IEEE Symposium on Security and Privacy*, pages 519–536, May 2021. https://ieeexplore.ieee.org/document/9519463/.

[24] Pardis Emami-Naeini, Janarth Dheenadhayalan, Yuvraj Agarwal, and Lorrie Faith Cranor. An informative security and privacy "nutrition" label for Internet of Things devices. *IEEE Security & Privacy*, 20(2):31–39, 2022. https://doi.org/10.1109/MSEC.2021.3132398.

[25] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. Exploring how privacy and security factor into IoT device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, May 2019. https://dl.acm.org/doi/10.1145/3290605.3300764.

[26] Engadget. Engadget's guide to privacy. https://www.engadget.com/buyers-guide/personal-security/.

[27] Cori Faklaris, Laura A. Dabbish, and Jason I. Hong. A self-report measure of end-user security attitudes (SA-6). In *Proceedings of the Fifteenth Symposium on Usable Privacy and Security*, pages 61–77, August 2019. https://www.usenix.org/conference/soups2019/presentation/faklaris.

[28] Todd Feathers. Civil rights groups want tech sites to stop reviewing Amazon's Ring cameras. *Vice*, March 2021. https://www.vice.com/en/article/z3vpw3/civil-rights-groups-want-tech-sites-to-stop-reviewing-amazons-ring-cameras.

[29] Bo Feng, Alejandro Mera, and Long Lu. P^{2}IM: Scalable and hardware-independent firmware testing via automatic peripheral interface modeling. In *Proceedings of the 29th USENIX Security Symposium*, pages 1237–1254, 2020. https://www.usenix.org/conference/usenixsecurity20/presentation/feng.

[30] Fight for the Future. Rescind Ring. https://rescindring.com, 2022.

[31] Raffaele Filieri, Charles F. Hofacker, and Salma Alguezaui. What makes information in online consumer reviews diagnostic over time? The role of review relevancy, factuality, currency, source credibility and ranking score. *Computers in Human Behavior*, 80:122–131, March 2018. https://doi.org/10.1016/j.chb.2017.10.039.

[32] Finnish Transport and Communications Agency. Finland becomes the first European country to certify safe smart devices: New Cybersecurity label helps consumers buy safer products. https://www.traficom.fi/en/news/finland-becomes-first-european-country-certify-safe-smart-devices-new-cybersecurity-label, November 2019.

[33] Sara Fischer. Exclusive: WSJ debuts new commerce site "Buy Side". *Axios*, June 2022. https://www.axios.com/2022/06/11/wsj-new-commerce-site-buy-side.

[34] Mozilla Foundation. *Privacy Not Included: A buyer's guide for connected products. https://foundation.mozilla.org/en/privacynotincluded/.

[35] Suzanne Frey. Get more information about your apps in Google Play. https://blog.google/products/google-play/data-safety/, April 2022.

[36] Chenglong Fu, Qiang Zeng, and Xiaojiang Du. HAWatcher: Semantics-aware anomaly detection for appified smart homes. In *Proceedings of the 30th USENIX Security Symposium*, pages 4223–4240, 2021. https://www.usenix.org/conference/usenixsecurity21/presentation/fu-chenglong.

[37] Kelsey R. Fulton, Rebecca Gelles, Alexandra McKay, Yasmin Abdi, Richard Roberts, and Michelle L. Mazurek. The effect of entertainment media on mental models of computer security. In *Proceedings of the Fifteenth Symposium on Usable Privacy and Security*, pages 79–95, August 2019. https://www.usenix.org/conference/soups2019/presentation/fulton.

[38] Thomas Germain. Mental health apps aren't all as private as you may think. https://www.consumerreports.org/health-privacy/mental-health-apps-and-user-privacy-a7415198244/, March 2021.

[39] GOV.UK. New smart devices cyber security laws one step closer. https://www.gov.uk/government/news/new-smart-devices-cyber-security-laws-one-step-closer, January 2022.

[40] Greg Guest, Arwen Bunce, and Laura Johnson. How many interviews are enough?: An experiment with data saturation and variability. *Field Methods*, 18(1):59–82, February 2006. https://doi.org/10.1177/1525822X05279903.

[41] Wentao Guo, Jay Rodolitz, and Eleanor Birrell. Poli-see: An interactive tool for visualizing privacy policies. In *Proceedings of the 19th Workshop on Privacy in the Electronic Society*, pages 57–71, November 2020. https://doi.org/10.1145/3411497.3420221.

[42] Taylor Hatmaker. Amazon sues Facebook group admins over fake reviews. *TechCrunch*, July 2022. https://social.techcrunch.com/2022/07/18/amazon-lawsuit-fake-reviews-facebook/.

[43] Nick Ho-Sam-Sooi, Wolter Pieters, and Maarten Kroesen. Investigating the effect of security and privacy on IoT device purchase behaviour. *Computers & Security*, 102:1–12, March 2021. https://doi.org/10.1016/j.cose.2020.102132.

[44] ioXt. ioXt Certification for IoT products. https://www.ioxtalliance.org/get-ioxt-certified.

[45] Yan Jia, Bin Yuan, Luyi Xing, Dongfang Zhao, Yifan Zhang, XiaoFeng Wang, Yijing Liu, Kaimin Zheng, Peyton Crnjak, Yuqing Zhang, Deqing Zou, and Hai Jin. Who's in control? On security risks of disjointed IoT device management channels. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, CCS '21, pages 1289–1305, November 2021. http://doi.org/10.1145/3460120.3484592.

[46] Shane D. Johnson, John M. Blythe, Matthew Manning, and Gabriel T. W. Wong. The impact of IoT security labelling on consumer product choice and willingness to pay. *PLoS ONE*, 15(1):1–21, January 2020. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0227800.

[47] Ryan Kailath. Some Amazon reviews are too good to be believed. They're paid for. *NPR*, July 2018. https://www.npr.org/2018/07/30/629800775/some-amazon-reviews-are-too-good-to-be-believed-theyre-paid-for.

[48] Georgios Kambourakis, Constantinos Kolias, Dimitrios Geneiatakis, Georgios Karopoulos, Georgios Michail Makrakis, and Ioannis Kounelis. A state-of-the-art review on the security of mainstream IoT wireless PAN protocol stacks. *Symmetry*, 12(4):1–29, April 2020. https://doi.org/10.3390/sym12040579.

[49] Hean Tat Keh and Jin Sun. The differential effects of online peer review and expert review on service evaluations: The roles of confidence and information convergence. *Journal of Service Research*, 21(4):474–489, 2018. https://doi.org/10.1177/1094670518779456.

[50] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, pages 1–12, July 2009. https://doi.org/10.1145/1572532.1572538.

[51] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. Privacy as part of the app decision-making process. In *Proceedings of the 31st Annual CHI Conference on Human Factors in Computing Systems*, pages 3393–3402, April 2013. https://doi.org/10.1145/2470654.2466466.

[52] Mingeun Kim, Dongkwan Kim, Eunsoo Kim, Suryeon Kim, Yeongjin Jang, and Yongdae Kim. FirmAE: Towards large-scale emulation of IoT firmware for dynamic analysis. In *Proceedings of the Annual Computer Security Applications Conference*, ACSAC '20, pages 733–745, December 2020. https://doi.org/10.1145/3427228.3427294.

[53] Konrad Kollnig, Anastasia Shuba, Max Van Kleek, Reuben Binns, and Nigel Shadbolt. Goodbye tracking? Impact of iOS app tracking transparency and privacy labels. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, pages 508–520, June 2022. http://doi.org/10.1145/3531146.3533116.

[54] Radhesh Krishnan Konoth, Emanuele Vineti, Veelasha Moonsamy, Martina Lindorfer, Christopher Kruegel, Herbert Bos, and Giovanni Vigna. Minesweeper: An in-depth look into drive-by cryptocurrency mining and its defense. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, CCS '18, pages 1714–1730, October 2018. https://doi.org/10.1145/3243734.3243858.

[55] Klaus Krippendorff. Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3):411–433, 2004. https://doi.org/10.1111/j.1468-2958.2004.tb00738.x.

[56] Melina Kulenovic and Dzenana Donko. A survey of static code analysis methods for security vulnerabilities detection. In *Proceedings of the 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics*, pages 1381–1386, May 2014. https://doi.org/10.1109/MIPRO.2014.6859783.

[57] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczynski, and Wouter Joosen. Tranco: A research-oriented top sites ranking hardened against manipulation. In *Proceedings of the 2019 Network and Distributed System Security Symposium*, February 2019. https://www.doi.org/10.14722/ndss.2019.23386.

[58] Young-Jin Lee and Yong Tan. Effects of different types of free trials and ratings in sampling of consumer software: An empirical study. *Journal of Management Information Systems*, 30(3):213–246, 2013. https://doi.org/10.2753/MIS0742-1222300308.

[59] Li Li, Tegawendé F. Bissyandé, Mike Papadakis, Siegfried Rasthofer, Alexandre Bartel, Damien Octeau, Jacques Klein, and Le Traon. Static analysis of Android apps: A systematic literature review. *Information and Software Technology*, 88:67–95, August 2017. https://doi.org/10.1016/j.infsof.2017.04.001.

[60] Mengxiang Li, Liqiang Huang, Chuan-Hoo Tan, and Kwok-Kee Wei. Helpfulness of online product reviews as seen by consumers: Source and content features. *International Journal of Electronic Commerce*, 17(4):101–136, 2013. https://doi.org/10.2753/JEC1086-4415170404.

[61] Tianshi Li, Kayla Reiman, Yuvraj Agarwal, Lorrie Faith Cranor, and Jason I. Hong. Understanding challenges for developers to create accurate privacy nutrition labels. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–24, April 2022. https://doi.org/10.1145/3491102.3502012.

[62] Yucheng Li, Deyuan Chen, Tianshi Li, Yuvraj Agarwal, Lorrie Faith Cranor, and Jason I. Hong. Understanding iOS privacy nutrition labels: An exploratory large-scale analysis of App Store data. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI EA '22, pages 1–7, April 2022. https://doi.org/10.1145/3491101.3519739.

[63] Franco Loi, Arunan Sivanathan, Hassan Habibi Gharakheili, Adam Radford, and Vijay Sivaraman. Systematically evaluating security and privacy for consumer IoT devices. In *Proceedings of the 2017 Workshop on Internet of Things Security and Privacy*, pages 1–6, November 2017. https://doi.org/10.1145/3139937.3139938.

[64] Xueming Luo, Bin Gu, Jie Zhang, and Chee Wei Phang. Expert blogs and consumer perceptions of competing brands. *MIS Quarterly*, 41(2):371–396, June 2017. https://ssrn.com/abstract=2268209.

[65] Arunesh Mathur, Arvind Narayanan, and Marshini Chetty. Endorsements on social media: An empirical study of affiliate marketing disclosures on YouTube and Pinterest. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–26, November 2018. https://dl.acm.org/doi/10.1145/3274388.

[66] Aleecia M. McDonald and Lorrie Faith Cranor. The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society*, 4(3):543–568, 2008. https://heinonline.org/HOL/P?h=hein.journals/isjlpsoc4&i=563.

[67] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):72:1–72:23, November 2019. https://doi.org/10.1145/3359174.

[68] Microsoft 365 Defender Threat Intelligence Team. Phorpiex morphs: How a longstanding botnet persists and thrives in the current threat environment. https://www.microsoft.com/security/blog/2021/05/20/phorpiex-morphs-how-a-longstanding-botnet-persists-and-thrives-in-the-current-threat-environment/, May 2021.

[69] Phoebe Moh, Pubali Datta, Noel Warford, Adam Bates, Nathan Malkin, and Michelle L Mazurek. Characterizing everyday misuse of smart home devices. In *Proceedings of the 2023 IEEE Symposium on Security and Privacy*, May 2023. https://doi.org/10.1109/SP46215.2023.00089.

[70] Christina Morales. Restaurants face an extortion threat: A bad rating on Google. *The New York Times*, July 2022. https://www.nytimes.com/2022/07/11/dining/google-one-star-review-scam-restaurants.html.

[71] Sara Morrison. Amazon's Ring privacy problem is back. *Vox*, July 2022. https://www.vox.com/recode/23207072/amazon-ring-privacy-police-footage.

[72] Nathaniel Mott. 'DoorLock' vulnerability can force iOS devices to endlessly reboot. *PCMag*, January 2022. https://www.pcmag.com/news/doorlock-vulnerability-can-force-ios-devices-to-endlessly-reboot.

[73] National Institute of Standards and Technology. Recommended criteria for cybersecurity labeling for consumer Internet of Things (IoT) products. NIST CSWP 24, National Institute of Standards and Technology, February 2022. https://csrc.nist.gov/publications/detail/white-paper/2022/02/04/criteria-for-cybersecurity-labeling-for-consumer-iot-products/final.

[74] Anna Naujoks and Martin Benkenstein. Who is behind the message? The power of expert reviews on eWOM platforms. *Electronic Commerce Research and Applications*, 44:1–10, 2020. https://doi.org/10.1016/j.elerap.2020.101015.

[75] Nataliia Neshenko, Elias Bou-Harb, Jorge Crichigno, Georges Kaddoum, and Nasir Ghani. Demystifying IoT security: An exhaustive survey on IoT vulnerabilities and a first empirical look on Internet-scale IoT exploitations. *IEEE Communications Surveys & Tutorials*, 21(3):2702–2733, 2019. https://doi.org/10.1109/COMST.2019.2910750.

[76] Erik E. Northrop and Heather R. Lipford. Exploring the usability of open

source network forensic tools. In *Proceedings of the 2014 ACM Workshop on Security Information Workers*, SIW '14, pages 1–8, November 2014. https://doi.org/10.1145/2663887.2663903.

[77] Paulo Nunes, Ibéria Medeiros, José Fonseca, Nuno Neves, Miguel Correia, and Marco Vieira. An empirical study on combining diverse static analysis tools for web security vulnerabilities based on development scenarios. *Computing*, 101:161–185, 2019. https://doi.org/10.1007/s00607-018-0664-z.

[78] Laura Hazard Owen. Wirecutter, which makes money when you shop, is going behind The New York Times' paywall. https://www.niemanlab.org/2021/08/wirecutter-which-makes-money-when-you-shop-is-going-behind-the-new-york-times-paywall/, August 2021.

[79] Do-Hyung Park. Consumer adoption of consumer-created vs. expert-created information: Moderating role of prior product attitude. *Sustainability*, 13(4):2024, January 2021. https://doi.org/10.3390/su13042024.

[80] Daria Plotkina and Andreas Munzel. Delight the experts, but never dissatisfy your customers! A multi-category study on the effects of online review source on intention to buy a new product. *Journal of Retailing and Consumer Services*, 29:1–11, 2016. https://dx.doi.org/10.1016/j.jretconser.2015.11.002.

[81] Molly Price and David Priest. Best smart locks of 2021: August, Yale, Schlage and more. *CNET*, November 2021. https://www.cnet.com/home/security/best-smart-locks/.

[82] Mike Prospero. Best video doorbells in 2022: Top smart doorbell cameras rated. *Tom's Guide*, January 2022. https://www.tomsguide.com/us/best-video-doorbells,review-4468.html.

[83] Reethika Ramesh, Anjali Vyas, and Roya Ensafi. "All of them claim to be the best": Multi-perspective study of VPN users and VPN providers. In *Proceedings of the 32nd USENIX Security Symposium*, May 2023. https://www.usenix.org/conference/usenixsecurity23/presentation/ramesh.

[84] Abhilasha Ravichander, Alan W Black, Thomas Norton, Shomir Wilson, and Norman Sadeh. Breaking down walls of text: How can NLP benefit consumer privacy? In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, volume 1, pages 4125–4140, August 2021. https://doi.org/10.18653/v1/2021.acl-long.319.

[85] Joel Reardon, Álvaro Feal, Primal Wijesekera, Amit Elazari Bar On, Narseo Vallina-Rodriguez, and Serge Egelman. 50 ways to leak your data: An exploration of apps' circumvention of the Android permissions system. In *Proceedings of the 28th USENIX Security Symposium*, pages 603–620, August 2019. https://www.usenix.org/conference/usenixsecurity19/presentation/reardon.

[86] Consumer Reports. The Digital Standard. https://thedigitalstandard.org/standard/, 2020.

[87] Consumer Reports. Guide to digital security & privacy. https://www.consumerreports.org/digital-security/online-security-and-privacy-guide/, July 2022.

[88] Daniel Romero. Technical advisory: Multiple vulnerabilities in Nuki smart locks (CVE-2022-32509, CVE-2022-32504, CVE-2022-32502, CVE-2022-32507, CVE-2022-32503, CVE-2022-32510, CVE-2022-32506, CVE-2022-32508, CVE-2022-32505). https://research.nccgroup.com/2022/07/25/technical-advisory-multiple-vulnerabilities-in-nuki-smart-locks-cve-2022-32509-cve-2022-32504-cve-2022-32502-cve-2022-32507-cve-2022-32503-cve-2022-32510-cve-2022-32506-cve-2022-32508-cve-2/, July 2022.

[89] Dan Seifert. Nest Thermostat review: More simple than smart. *The Verge*, December 2020. https://www.theverge.com/21725036/google-nest-thermostat-2020-review.

[90] Similarweb. https://www.similarweb.com/.

[91] Ben Smith. You've never heard of the biggest digital media company in America. *The New York Times*, August 2021. https://www.nytimes.com/2021/08/15/business/media/red-ventures-digital-media.html.

[92] Dale Smith. Ring Video Doorbell 4 review: A competent gadget from a company with a shaky reputation. *CNET*, July 2021. https://www.cnet.com/home/security/ring-video-doorbell-4-review-a-competent-gadget-from-a-company-with-a-shaky-reputation/.

[93] Justin Smith, Lisa Nguyen Quang Do, and Emerson Murphy-Hill. Why can't Johnny fix vulnerabilities: A usability evaluation of static analysis tools for security. In *Proceedings of the Sixteenth Symposium on Usable Privacy and Security*, pages 221–238, August 2020. https://www.usenix.org/conference/soups2020/presentation/smith.

[94] Justin Smith, Brittany Johnson, Emerson Murphy-Hill, Bill Chu, and Heather Richter Lipford. How developers diagnose potential security vulnerabilities with a static analysis tool. *IEEE Transactions on Software Engineering*, 45(9):877–897, September 2019. https://doi.org/10.1109/TSE.2018.2810116.

[95] UL Solutions. UL verified IoT device security rating. https://www.ul.com/services/ul-verified-iot-device-security-rating.

[96] Parth Kirankumar Thakkar, Shijing He, Shiyu Xu, Danny Yuxing Huang, and Yaxing Yao. "It would probably turn into a social faux-pas": Users' and bystanders' preferences of privacy awareness mechanisms in smart homes. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, pages 1–13, April 2022. https://doi.org/10.1145/3491102.3502137.

[97] Khaos Tian. Your home was not so secure after all. https://medium.com/hackernoon/your-home-was-not-so-secure-after-all-af52fbd6777c, December 2017.

[98] Jeffrey A. Trachtenberg. Gannett invests to boost product-review site, hoping to rival New York Times's Wirecutter. *The Wall Street Journal*, July 2021. http://www.wsj.com/articles/gannett-invests-to-boost-product-review-site-hoping-to-rival-new-york-timess-wirecutter-11625407202.

[99] Janice Y. Tsai, Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research*, 22(2):254–268, June 2011. https://pubsonline.informs.org/doi/abs/10.1287/isre.1090.0260.

[100] Joseph Turow, Yphtach Lelkes, Nora A. Draper, and Ari Ezra Waldman. Americans can't consent to companies' use of their data. Technical report, Annenberg School for Communication, University of Pennsylvania, February 2023. https://www.asc.upenn.edu/sites/default/files/2023-02/Americans_Can%27t_Consent.pdf.

[101] Juraj Uhlár, Martin Holkovič, and Vít Rusňák. PCAPFunnel: A tool for rapid exploration of packet capture files. In *Proceedings of the 2021 25th International Conference Information Visualisation*, pages 69–76, July 2021. https://doi.org/10.1109/IV53921.2021.00021.

[102] Alex Ulmer, David Sessler, and Jörn Kohlhammer. NetCapVis: Web-based progressive visual analytics for network packet captures. In *Proceedings of the 2019 IEEE Symposium on Visualization for Cyber Security*, pages 1–10, October 2019. https://doi.org/10.1109/VizSec48167.2019.9161633.

[103] Fábio Luciano Verdi, Hélio Tibagí de Oliveira, Leobino N. Sampaio, and Luciana A. M. Zaina. Usability matters: A human–computer interaction study on network management tools. *IEEE Transactions on Network and Service Management*, 17(3):1865–1878, April 2020. https://doi.org/10.1109/TNSM.2020.2987036.

[104] Emily A. Vogels and Monica Anderson. Americans and digital knowledge. Technical report, Pew Research Center, October 2019. https://www.pewresearch.org/internet/2019/10/09/americans-and-digital-knowledge/.

[105] Bettina von Helversen, Katarzyna Abramczuk, Wiesław Kopeć, and Radoslaw Nielek. Influence of consumer reviews on online purchasing decisions in older and younger adults. *Decision Support Systems*, 113:1–10, September 2018. https://www.sciencedirect.com/science/article/pii/S0167923618300861.

[106] Yuxi Wu, Panya Gupta, Miranda Wei, Yasemin Acar, Sascha Fahl, and Blase Ur. Your secrets are safe: How browsers' explanations impact misconceptions about private browsing mode. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, pages 217–226, April 2018. http://doi.org/10.1145/3178876.3186088.

[107] Eric Zeng and Franziska Roesner. Understanding and improving security and privacy in multi-user smart homes: A design exploration and in-home user study. In *Proceedings of the 28th USENIX Security Symposium*, pages 159–176, August 2019. https://www.usenix.org/conference/usenixsecurity19/presentation/zeng.

[108] Shikun Zhang, Yuanyuan Feng, Yaxing Yao, Lorrie Faith Cranor, and Norman Sadeh. How usable are iOS app privacy labels? *Proceedings on Privacy Enhancing Technologies*, 2022(4):204–228, 2022. https://doi.org/10.56553/popets-2022-0106.

[109] Wenqi Zhou and Wenjing Duan. Do professional reviews affect online user choices through user reviews? An empirical study. *Journal of Management Information Systems*, 33(1):202–228, 2016. https://doi.org/10.1080/07421222.2016.1172460.

[110] Marc Ziegele and Mathias Weber. Example, please! Comparing the effects of single customer reviews and aggregate review scores on online shoppers' product evaluations. *Journal of Consumer Behaviour*, 14(2):103–114, 2015. http://onlinelibrary.wiley.com/doi/abs/10.1002/cb.1503.

A Tables

Table A1: Sources of the reviews we analyzed.

Text Reviews		Video Reviews	
CNET	7	CNET	3
PCMag	6	Home Tech Decisions	3
Tom's Guide	4	Smart Home Solver	3
BestReviews	3	Life Hackster	2
Wirecutter	3	The 5 Best	2
Consumer Reports	2	Top 5 Picks	2
Reviews.org	2	Apple Insider	1
SafeHome.org	2	Automate Your Life Shorts	1
TechRadar	2	Detroit Tech	1
The Verge	2	Everyday Chris	1
Android Authority	1	James	1
Digital Trends	1	Locksmith Recommended	1
Reviewed	1	Modern Dad	1
Reviews.com	1	One Hour Smart Home	1
Safewise	1	Rizknows	1
The Guardian	1	Security.org	1
The Smart Cave	1	Serg Tech	1
Trusted Reviews	1	Steve Does	1
		Tech With Brett	1
		Techs You Can't Live Without	1
		Terry White	1

Table A2: Demographics of our interview participants, reported in aggregate.

		Count
Age	18–24	1
	25–34	4
	35–44	7
	45–54	2
	55–64	2
	65+	2
Gender	Female	3
	Male	15
Race	Asian	1
	White	16
	Other	1
Education (complete or pursuing)	High school	1
	Bachelor's degree	12
	Graduate or professional degree	4
	Prefer not to state	1
Country	United States	16
	Other	2

B Interview protocol

Big picture
I'll start with some background questions. If we can try to do this rapid-fire, that'll let us get to the main part of our interview faster.

1. What is your current role, professionally, when it comes to reviewing products?
2. Would you please describe the organization or publication for which you review products? For example, who is the audience?
3. What is the business model with respect to reviews? For example, affiliate marketing, sponsorships, ads, and subscriptions.
4. Do you publish written reviews, video reviews, something else, or a mix of formats?
5. How long have you been doing similar work reviewing products?
6. Would you please describe briefly what kinds of products you currently review?
7. Do you have a process for determining which products to review next?
8. In what setting do you review products? For example, in a lab, in your own home, or somewhere else?
9. How long do you typically test or use a product before publishing a review?
10. Do you ever revise or update a review after it is published?

Criteria
Now I'd like to talk about criteria that you may use for evaluating a product: in other words, information about a product that you use to judge its quality.

1. When reviewing a product, what kinds of criteria do you write about or take into account?
2. When reviewing a product, do you consider criteria related to security and privacy, and how do you prioritize them compared to other criteria such as a product's features and cost?
3. What are some of the most important security- and privacy-related criteria that you consider when reviewing a product?
4. Are there other security- and privacy-related criteria that you consider important but don't evaluate?

Techniques and tools
Now, let's pivot to the techniques and tools you use to evaluate products: in other words, what you do when testing a product, and how you do it. We're thinking about this broadly; it could be anything from reading a document to using a software tool to analyze a device.

1. How do you learn about techniques and tools for evaluating products in general?
2. What techniques and tools, if any, do you use to evaluate security and privacy?
3. For a minute, let's pretend you could have any tools you wanted to help you review the security and privacy of Internet-connected devices, apps, or other software. Don't worry about how they might technically work; think of a tool as a black box where you know how to set it up and what information about a product it will tell you. Can you think of any tools that you would use? What would they tell you about a product? How would you use them?
4. Now, let's pretend you could institute any regulations or industry-wide practices you wanted to help you review the security and privacy of Internet-connected devices, apps, or other software. Can you think of any that you would make use of? How would they change the way you review products?
5. Can you remember any other times that you ran into difficulties when trying to use a technique or tool to evaluate security and privacy?
6. Are there other techniques or tools that you're aware of for evaluating security and privacy that you would like to use but don't?

Communication, impact, and incentives
Next, I'd like to talk about your process for writing reviews and communicating your findings to consumers.

1. How frequently do you discuss security and privacy in your reviews, if at all?

- [if not at all]:
  - Do you ever consider discussing security and privacy?
  - Why do you not include this?
- [else]:
  - How do you decide whether or not to discuss security and privacy in a review?
  - Do you ever make security- or privacy-related recommendations about how to use or configure a product?

2. How often do you publish a negative review or recommend against a product, for any reason?
3. Could you describe how you decide between publishing a negative review and not publishing a review at all?
4. Do security and privacy findings affect whether you recommend a product or not? Has security or privacy ever been the deciding factor in your recommendation?
5. Are there other ways in which you aim to have an impact on consumers, aside from people directly reading your reviews?

Conclusion

1. Finally, is there anything else you'd like us to know about your work reviewing Internet-connected devices, apps, and other software, and about reviewing security and privacy in particular?

C Post-interview survey questions

Self-identified level of expertise

1. Please rate your agreement or disagreement with the following statement.
   I am extremely knowledgeable about all the steps needed to keep my online data and accounts safe.
   - Strongly agree | Agree | Neutral | Disagree | Strongly disagree

Security and privacy knowledge questions

1. If a website uses cookies, it means that the site...
   - Can see the content of all the files on the device you are using
   - Is not a risk to infect your device with a computer virus
   - Will automatically prompt you to update your web browser software if it is out of date
   - Can track your visits and activity on the site
   - Not sure
2. Which of the following is the largest source of revenue for most major social media platforms?
   - Exclusive licensing deals with internet service providers and cellphone manufacturers
   - Allowing companies to purchase advertisements on their platforms
   - Hosting conferences for social media influencers
   - Providing consulting services to corporate clients
   - Not sure
3. When a website has a privacy policy, it means that the site...
   - Has created a contract between itself and its users about how it will use their data
   - Will not share its users' personal information with third parties
   - Adheres to federal guidelines about deceptive advertising practices
   - Does not retain any personally identifying information about its users
   - Not sure
4. What does it mean when a website has "https://" at the beginning of its URL, as opposed to "http://" without the "s"?
   - Information entered into the site is encrypted
   - The content on the site is safe for children
   - The site is only accessible to people in certain countries
   - The site has been verified as trustworthy
   - Not sure
5. Where might someone encounter a phishing scam?
   - In an email
   - On social media
   - In a text message
   - On a website
   - All of the above
   - None of the above
   - Not sure
6. Many web browsers offer a feature known as "private browsing" or "incognito mode." If someone opens a webpage on their computer at work using incognito mode, which of the following groups will NOT be able to see their online activities?
   - The group that runs their company's internal computer network
   - Their company's internet service provider
   - A coworker who uses the same computer
   - The websites they visit while in private browsing mode
   - Not sure

Organization details

1. How many people work on reviewing products at your primary organization, to the best of your knowledge?
   - 1 | 2–3 | 4–9 | 10–19 | 20–29 | 30+ | I don't know
2. Please provide any details we should know. _____
3. How many views does a typical product review of yours receive, to the best of your knowledge?
   - 0–99 | 100-999 | 1,000-9,999 | 10,000–99,999 | 100,000–999,999 | 1,000,000+ | I don't know
4. Please provide any details we should know. _____

Participant demographics

1. What is your age?
   - 18–24 | 25–34 | 35–44 | 45–54 | 55–64 | 65+ | Prefer not to state
2. What is your gender?
   - Male | Female | Other _____ | Prefer not to state
3. What is your race?
   - White | Hispanic or Latino | Black or African American | American Indian or Alaska Native | Asian | Native Hawaiian or Pacific Islander | Other _____ | Prefer not to state
4. What is the highest level of formal education that you have completed or are currently pursuing?
   - No high school degree | High school graduate, diploma or equivalent (for example, GED) | Trade, technical, or vocational training | Associate's degree | Bachelor's degree | Graduate or professional degree | Other _____ | Prefer not to state
5. What country (or countries) do you work in?
   - ☐ United States
   - ☐ Other _____

Supplementary materials containing Appendices D–F are located at https://osf.io/m2pe7/?view_only=e6a8443956704fe2b380cfce1def1204.