"All of them claim to be the best": Multi-perspective study of VPN users and VPN providers

Reethika Ramesh
University of Michigan

Anjali Vyas
Cornell Tech

Roya Ensafi
University of Michigan

Abstract

As more users adopt VPNs for a variety of reasons, it is important to develop empirical knowledge of their needs and mental models of what a VPN offers. Moreover, studying VPN users alone is not enough because, by using a VPN, a user essentially transfers trust, say from their network provider, onto the VPN provider. To that end, we are the first to study the VPN ecosystem from both the users' and the providers' perspectives. In this paper, we conduct a quantitative survey of 1,252 VPN users in the U.S. and qualitative interviews of nine providers to answer several research questions regarding the motivations, needs, threat model, and mental model of users, and the key challenges and insights from VPN providers. We create novel insights by augmenting our multi-perspective results, and highlight cases where the user and provider perspectives are misaligned. Alarmingly, we find that users rely on and trust VPN review sites, but VPN providers shed light on how these sites are mostly motivated by money. Worryingly, we find that users have flawed mental models about the protection VPNs provide, and about data collected by VPNs. We present actionable recommendations for technologists and security and privacy advocates by identifying potential areas on which to focus efforts and improve the VPN ecosystem.

1 Introduction

Since their introduction over two decades ago, the use of Virtual Private Network (VPN) technologies has grown rapidly. With commercialization, VPN products have found their way into a regular Internet user's toolbox [16, 43]. Though the VPN ecosystem has expanded into a multi-billion dollar industry [33], questions regarding why VPNs have been adopted so widely are still unanswered. Is the popularity of VPNs grounded in an understanding of risks from the users' part? Is the rise of VPNs due to dwindling trust in Internet service providers? What benefits do users perceive to gain?

A majority of previous studies have found various issues in the technical implementations of VPNs [12, 15, 36, 54, 55]. Only limited prior work has delved into the human factors of VPN use: factors that contribute to retention of VPNs [29, 58], attitudes of university students and corporate users towards VPNs [3, 10, 11], and the widespread misconceptions of how privacy-enhancing tools work [48].

However, no study has combined both the users and VPN providers perspectives to answer fundamental questions about the VPN ecosystem. For instance, users using VPNs are essentially transferring trust from their network provider onto the VPN provider, but it is unclear as to what VPN features encourages them to make this shift? On the other hand, the VPN industry has been known to employ various marketing tactics [1] and dark patterns around discounts [21, 51], but it is yet unknown if these practices are bound to have any significant effect on VPN users. Moreover, the community has not yet understood VPN providers' incentives in sustaining such dark patterns, nor do we know what efforts they take to foster user confidence in an ecosystem plagued with mistrust. To gain a clearer picture of the inner workings of such a large consumer ecosystem, it is imperative to study both its users and its providers.

This is the first multi-perspective study that uses a *quantitative survey of **(n=1,252)** VPN users* in the U.S. along with *qualitative interviews of nine leading VPN providers*. We choose to survey 1,252 users, that have either used or currently use a VPN, to provide us with practical insights into our various lines of inquiry that we systematize into the following research questions:

RQ1: [Motivations] Why do users use VPNs?

RQ2: [Needs and Considerations] What factors around VPNs do users consider when choosing a provider?

RQ3: [Emotional Connection and Threat Model] How safe do users feel when browsing the internet with and without a VPN? (If and) From whom do users want to secure/conceal their online activity?

RQ4: [Mental Model] Do users have an accurate understanding of how VPNs work and what data they collect?

RQ5: [Perception and Trust] How do users perceive the VPN ecosystem?

RQ6: [Alignment between VPN users and providers] What are the key areas of (mis)alignment in priorities and incentives between the two?

We find that users rate speed, price, and easily understandable GUI, as the top requirements from VPNs rather than features such as the variety, number of available VPN servers, and their locations. We also find that in alignment with VPN providers' expectations, pricing plays a key role with users. Thus indicating that discounts, and marketing around pricing can have a significant effect on them. Prior research suggests that malicious marketing tactics [1] and dark patterns around discounts are common, which are often used to ensure customer lock-in [21, 51]; an example of such a dark pattern in the VPN ecosystem is shown in Figure 1.

Interestingly, we find that when it comes to choosing VPNs to use, more users seem to lean towards using search engines (61.1%), and recommendation websites (56.5%), rather than relying on more traditional methods such as word of mouth (5.7%). Furthermore, almost 94% of these users rate these websites trustworthy. On the other hand, our interviews with VPN providers highlights that the VPN recommendation ecosystem is mostly money motivated, with widespread malicious practices that include having paid review spots, and auctioning off the #1 spot. Some "review" sites have been reported to send emails to VPN providers asking for higher cost-per-action/click to get ranked on their list [56]. Users' reliance on such websites further amplifies our worries of an unregulated marketing ecosystem around VPNs.

Exploring reasons for why users use VPNs, we discover that users attach an emotional connection with using a VPN, namely a feeling of safety (86.7%), which was found by prior work to be a key factor in retention of VPN use [29]. Our intuition suggested that exploring users' threat models could explain why they attach such considerations; indeed, we find that 91.5% of users indicate they use VPNs for securing or protecting their online activity. When exploring who they aim to protect it from, we find that their top concerns are hackers/eavesdroppers on open WiFi networks (83.9%), advertising companies (65.4%), and internet service providers (46.9%). This marks a departure from known prior concerns such as government surveillance (30%), and indicates a shift of attitude towards surveillance capitalism and user privacy.

Given the emotional attachments and user dependency on VPNs for security and privacy concerns, we find that an alarmingly high proportion of users (39.9%) have a *flawed mental model* of what VPNs provide them and what data they collect. These users believe their ISP can still see the websites they visit over the VPN. More worryingly, we do not see significant difference between users of different expertise having flawed mental models (χ^2-test, p=0.0927, N=1252). From our VPN provider interviews, we find that providers also mention that they recognize the need for improving user knowledge, and consider effective education a key challenge. We also find that dark patterns in the industry may also be a key issue;

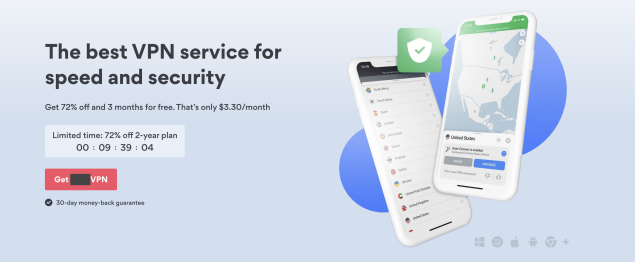

Figure 1: Example of dark pattern–using countdown timers.♣

multiple VPN providers mention "malicious marketing" is problematic, including preying upon users' lack of knowledge and overselling of service.

Continuing to explore the confusion surrounding the operations of VPN providers, we find that a significant portion of limited expertise users believe that the data is being collected for monetization, such as advertising (36.4%), user tracking (36.4%), and selling to third parties (33.6%). Although a majority of all users (79.2%) believe the main reason for data collection is internal analytics, confusion found amongst the limited expertise users may be even more widespread among the VPN users in the general public. Moreover, users also expressed high degrees of concern towards VPN providers selling their data (73.2%). This is yet another area of misalignment between users and providers, because multiple VPN providers believe that they clearly communicate their logging practices, and/or have released audits to prove this.

From our study of 1,252 users and 9 VPN providers, we present the following actionable recommendations for the VPN ecosystem: prioritizing user education, oversight on advertisements and marketing surrounding VPNs, coordinated efforts to bring attention to the flawed VPN recommendation ecosystem, and regulations to curb malicious marketing tactics that lead to false mental models and false expectations for users. We believe that our work will help security and privacy advocates such as EFF and CDT, technologists, and VPN providers alike, by calling attention to the key areas in the commercial VPN ecosystem.

2 Background & Related Work

2.1 Virtual Private Networks

Virtual Private Networks were initially created in 1996 [26] as a peer-to-peer tunnelling protocol developed in Microsoft to facilitate private communication in enterprise settings. Virtual private networks (VPNs) provided a way to create private connections between computers and transfer data between them securely over the public Internet. These are still the guarantees that VPNs provide for general users today. VPN products (VPNs hereon) create a secure connection, often called a "tunnel", to a secure server that then connects them to their intended destination. This tunnel typically provides

an extra layer of encryption that serves as protection from surveillance by the intermediate networks, bypasses access restrictions active in those networks, and hides the user's actual IP address from their destination service [20].

Commercial VPN providers make use of the available VPN protocols such as OpenVPN, L2TP, IPSec, IKEv2, and Wireguard [9, 14, 32, 50], or develop proprietary protocols which are typically extensions of existing ones, optimized to fit their particular needs and business model. VPNs offer different subscription models: paid/premium services, free to use services, and freemium models that offer limited free features and charge for premium features and services.

While some work has focused on analyzing technical aspects [12, 15, 36, 49], Weinberg et al. [54] focused on evaluating the claims of VPN server locations, and found at least one-third of the 2269 servers were definitely not in the country advertised, and another one-third probably were not in the location they claim. Investigating an often overlooked source of security advice, Akgul et al. [1] studied 243 YouTube videos containing VPN ads and find a number of concerning misleading claims, including over-promises and exaggerations which may lead to users forming inaccurate mental models of internet safety. There have also been news reports of data breaches, leaks and misuse by VPN providers, some of which were published on VPN recommendation websites [35, 42, 53, 57].

2.2 User Adoption of VPNs and Other Tools

As users adopt more privacy-enhancing tools such as VPNs for a variety of reasons, their privacy needs become important to assimilate. The level of security and privacy a user needs may depend on myriad factors like the reasons for use, tolerance of failure, legality of these VPN services in the country of the user *etc*. Only few community efforts focus on providing threat model based VPN (and other tools) recommendations, such as the Security Planner [7]. We present the related work summarized in relevance to the topics studied:

Prior work studying VPN users: Namara et al. [29] conducted a study with 90 technologically savvy users and studied the adoption and usage of VPNs, and the barriers they encounter in adopting them. They find users with emotional reasons to use a VPN such as fear of surveillance or desire for privacy, are more likely to continue using them rather users who use it for practical reasons. Similarly exploring the factors that influence user decisions to adopt VPN apps, Sombatruang et al. [46] interviewed 32 users in UK and Japan and found that user review rating and price significantly influenced the choosing of a VPN to use.

Prior work exploring particular sub-populations: Binkhorst et al. [3] studied the mental models of 18 users in the context of corporate VPNs, and found that experts and non-expert users have similar mental models of VPNs, and experts also tend to have false beliefs on security aspects of VPNs. Dutkowska-Zuk et al. conducted a study focused on a specific sub-population of 349 university students to find how and why they use VPNs, and whether they understand the various privacy risks caused by VPNs [11]. They found that students are mostly concerned with access to content rather than privacy, and that most students did not use VPNs regularly. Extending this study, they looked at how these students compare to general VPN users in the awareness of risks of VPN use and how they adopt VPNs [10]. Specifically, they found that despite having different use cases, both groups had low understanding of the risks of data collection by VPNs, highlighting the need for better awareness campaigns.

Prior work studied user attitudes and use of privacy-enhancing tools: Various prior works have shown that users, particularly in the U.S. are aware of risks such as tracking, and are concerned about online tracking in different situations [5, 25, 34]. However, some prior work has shown that they are unclear on how to protect themselves [44]. Story et al. [48] highlighted this in their study of the use of and perceptions about web-browsing privacy related tools. In their survey of 500 U.S. users, they ascertain user perception of the protection provided by different tools across 12 different scenarios, and interestingly, they find that users having more experience using VPNs is associated with confusion about their protection. Further, studying the adoption and abandonment of 30 commonly recommended security and privacy practices, Zou et al. [58] surveyed 902 users and find that security practices were more widely adopted and privacy related practices were among the ones most commonly abandoned.

These prior work, though useful, are limited in the scale, topics studied, and have focused on particular sub-populations of VPN users. **In our work, we create a novel line of inquiry to study the motivations, needs, and considerations of VPN users in depth, and improve greatly upon the scale of users surveyed. We are also the first to conduct a study of VPN providers.** We augment insights from both users and providers to characterize any misalignments between them, which could be exploited by bad actors to further deepen problems in the VPN ecosystem. Given the wide reach of the VPN ecosystem, our study will help technologists and security and privacy advocates gain a deeper understanding of the key problem areas where they can focus their efforts.

3 Methods

We set out to study VPN users and providers to understand their unique perspectives on the VPN ecosystem and the issues surrounding it. We conduct a large-scale survey of VPN users as well as a qualitative interview of nine VPN providers.

3.1 User Survey

Small-Scale Interviews and Interactions. We believe that a successful large-scale quantitative study must be preceded by a smaller-scale qualitative study and community research

Themes	Definitions
Reasons for using VPN	Motivations for and reasons to use a VPN
VPN Use	General thoughts about commercial VPNs, what they look for
Threat model for using a VPN	Personal threat models for needing, using, and/or recommending a VPN
Mental Model of VPN	What is a VPN and what does it provide me? (Sketching exercise included)
Attitudes towards VPN services	What is lacking in current ecosystem, their perception of what the VPN ecosystem looks like
Improving ecosystem	Thoughts to improve ecosystem and boosting adoption and safe usage

Table 1: Six themes with their definitions. ♣

to extract key concerns. To that end, we conduct seven user interviews (4 men, 3 women, ages 18-45), and we participated in various VPN-focused community events with VPN providers and users in attendance in order to gather topics and research questions that interest the community.

For the small-scale user interview study, we design a questionnaire with open-ended questions to serve as the framework for each interview, and obtain approval from our Institutional Review Board (IRB). During the interview, we collect general demographic information, including gender, age range, occupation, country of residence, and level of education. Our introductory questions ask about the interviewee's awareness of their own threat model and of online risks such as trackers. Next, we ask about the perceived positives and negatives of VPN use. We then ask participants to sketch their understanding of how VPNs work while walking through the steps of setup and use, diagrammatically. The interview concludes with questions about how the VPN ecosystem can improve.

We recruit seven participants via a pre-interview survey at the Citizen Lab Summer Institute [28] that has global attendees who are passionate about technologies aiding Internet freedom, security, and user rights. Prior to the start of each interview, we obtain explicit consent for participation and permission to audio record it using an IRB-approved consent form. Participants are also given the chance to ask any questions before the interview begins and are allowed to stop at any point. After completion of the interviews, the first author transcribed all the recordings. Overall, the interviews lasted 15-20 minutes not including setup and conclusion.

Developing the Large-Scale Survey Instrument. After completing the interviews, we use an inductive open-coding method for analysis. Two members independently coded all the transcripts, and held a meeting to resolve any disagreements and create a codebook. The research team then met to collaboratively go through the codebook and identify emerging themes [4] and hence, we do not present inter-rater reliability for this case [23, 27]. We augment these with the knowledge extracted from attending several Internet freedom community gatherings organized around VPNs and VPN use including the IFF VPN Village [13]. Finally, we combine our work to arrive at six common themes, shown in Table 1.

Using these themes, we devise an initial survey instrument to study VPN users. The instrument contains questions aimed at understanding users' motivations, needs, and considerations when it comes to VPN services, and discerning their threat models, perceptions of VPNs, and understanding of how VPNs work. During the design phase, we also create a consent form and obtain IRB approval. Our survey questions only collect the information we need and do not involve the collection of any personally identifiable information.

Cognitive Pre-testing. In order to reduce the potential for biases that arise from the ordering and/or phrasing of questions, we conduct systematic pretesting in *three phases*, iteratively improving the survey between each phase.

First, we recruit test participants (from the target demographic, VPN users) at an Internet freedom, security and privacy focused event organized by the Open Tech Fund [31] to pretest the initial survey instrument and obtain unbiased opinions about the survey. The pretesting involves vocally stepping through the survey while a facilitator from our team takes notes. We use these notes to detect biases, signs of confusion regarding the intent of the question, as well as "leading" questions. In this round, 17 pretesters worked through the survey, and we learned that comparison-scale adjectives (None at all, Little, Somewhat) were unclear for participants. We amend the scales to avoid ambiguity and provided clearer distinctions e.g. we use Likert-type Scale when asking about concern or importance. The scale is provided on the Qualtrics software and is a psychometric scale developed for scaling responses in survey research [19]. We learned that participants had varying understandings of what "commercial VPNs" mean, and that participants were not sure if the questions pertained to personal or professional VPN use. To remove ambiguity, we define "commercial VPNs" on the survey landing page and present examples within the survey.

After refining our survey using the initial pretesting, we requested external user-study experts to go through our survey and provide feedback. They helped us refine our matrix style questions and simplify the organization of our survey.

After incorporating expert feedback, we run the last round with eight new pretesters. This round helped us refine some of the examples used in the survey, improve consistency of language, and disambiguate a handful of questions.

The Final Survey Instrument. The final survey instrument contains six parts, 28 questions (with sub-parts), and we incorporate one quality check (where they must confirm they use a VPN) and two attention checks. The survey starts with a demographic section, where we follow the community best practices for inclusive language, and also have a "prefer not to disclose" option for all demographic questions [39, 47]. Then, we ask users general questions about their VPN usage, reasons

for using a VPN, the resources used in discovering VPNs, importance of different criteria and features, their mental model of VPNs and the data it collects, their emotional connections tied to their use of a VPN (e.g. safety), and their expectations from a VPN provider. We specifically avoid using words such as privacy and security in the text, since these concepts are broad, subjective, and mean different things to different users. Instead we allow users to select from list of options, and we distill into certain buckets during analysis. The final survey instrument is available in our pre-print [37, Appendix C].

Analysis of Survey Data. For the quantitative data from the survey, we report the results summarizing the users' responses. We aim to understand how different subgroups of users answer the same question, i.e. users with different security and privacy expertise, and users who prefer to use certain subscription type (free or paid VPNs). We conduct χ^2-tests, where all the assumptions are satisfied in each case, to examine if users in different subgroups of the same type (e.g. expertise) answer questions differently. If there were significant differences between subgroups, we conduct pairwise comparison Z-tests (α=0.05), where we adjust the significance levels for multiple comparisons through the FDR-BH adjustment [2] and present how they compare to each other.

We analyze the survey participants' open-ended text box responses using inductive coding. A primary coder created an initial codebook and assigned codes to all responses. A second coder analyzed 20% of the responses for each coded question and ensure high inter-rater reliability [17]. Cohen's κ between the two raters is 0.81, 0.86, 0.75, 0.81 for each question in Appendix B, indicating moderate to strong agreement [6, 24]. The coders also coded responses for "Other" write-in options in questions, and present the responses in the results (§5).

3.2 Qualitative Interviews of VPN Providers

Interview Instrument. Using the same parent themes as mentioned in Table 1, we create a questionnaire to interview VPN providers. These questions aim to extract insights from the providers about their users, VPN users in general, their business decisions, and what they see as the main issues in the VPN ecosystem. We design the topics for the questions to be counterparts to the VPN user survey.

Interview Procedure. We design a semi-structured interview with eight broad open-ended questions, and five additional questions to ask in case we have time. We obtain IRB approval prior to conducting the interviews. Our questionnaire[1] serves as a framework for the interviews to ensure we maintain structure and consistency from one provider to another. However, we also explore statements made by the interviewees for clarity and insights.

We begin all the interviews by presenting the interviewees with an overview of our project. Using an IRB-approved consent form, we obtain explicit consent for participation and audio recording the interview. On average, the interviews were $\approx$44 minutes in length, not including set up and conclusion. We conclude the interviews by thanking the participants, and provide ways to contact us to learn more about our project.

Analysis using Qualitative Coding. The first author transcribed all the interviews and the analysis is done using inductive open-coding, and thematic analysis [4]. Although we have nine VPN providers, we have eight transcripts in total because two of the providers opted to interview together[2] and each of them answered each question independently. A primary coder coded all transcripts, and two additional coders independently coded five and three transcripts each. Then, the team went over each coded transcript together to reconcile any differences. We then collaboratively identify the emerging themes for each question, and common themes that appear across different questions. Since the team collaboratively analyzed the coded transcripts together to identify themes, we do not present inter-rater reliability [23, 27].

Since this interview is meant to shed light on the VPN provider's perspectives and form a clearer picture of the VPN ecosystem, we only report aggregate results after performing thematic analysis. We anonymize the comments and do not attribute statements to particular providers.

3.3 Recruitment

User Survey. In partnership with Consumer Reports, a leading consumer research and advocacy organization with over 6 million members, we launched our user survey on March 1, 2021. We ask VPN users to participate in our survey by distributing the recruitment message in Consumer Reports' tech-focused mailing list, subreddits such as `r/VPN`, `r/asknetsec`, `r/samplesize`, and on Twitter using the research team's own personal accounts. We also request participation from users on mailing lists belonging to Open Tech Fund and Internet Freedom Festival. We opt to recruit participants organically and to ensure anonymity, we did not offer any compensation for taking the survey.

VPN Provider Interviews. We reached out to 15 leading VPN providers; nine of whom agreed to our interview. We chose to contact commercial VPN providers based on their popularity in the U.S., and included non-commercial projects that develop VPNs for users, based on their involvement in the Internet freedom and anti-surveillance community. We did not compensate the interviewees for participation.

The VPN providers we interviewed are the following (in alphabetical order): CalyxVPN, Hide.me, IVPN, Jigsaw Out-

[1] The questionnaire is presented in our pre-print [37, Appendix D]

[2] They are (non-commercial) partner projects, with separate services.

line, Mullvad VPN, RiseupVPN, Surfshark, TunnelBear VPN, and Windscribe. We interviewed CEOs, CMOs, and/or researchers working in the company who were authorized to speak to us on behalf of the company.

3.4 Ethics

Our user study is approved as exempt from ongoing review under Exemption 2 as determined by our Institutional Review Board (IRB), and the VPN provider interview received a "Not Regulated" status. Furthermore, we draft a privacy policy document that was reviewed by experts from Consumer Reports, and add it to our website. We also provide information on our study's Qualtrics page and ensure that our participants, pretesters, and interviewees explicitly consent to the study.

We follow user survey best practices such as using mindful, inclusive language in collecting demographics data [47]. We also offer "prefer not to answer" as an option on our required demographics questions as per American Association for Public Opinion Research code of ethics [18, 39]. We did not collect any personally identifiable information from our participants, and our results from the VPN providers are anonymized as well.

We solicit participation as mentioned in §3.3 and to ensure anonymity, we offer no compensation for any of our studies. We deeply analyze the collected responses to ensure response quality, as we detail in §4. Audio-recordings of the interviews (both the small-scale user ones, and the VPN providers) were only accessed by the first author who did all the transcriptions.

3.5 Limitations

As with many user surveys, some of our comparisons rely on self-reported data, which is prone to biases. We take efforts to reduce these biases to our best extent, elaborated in §4, such as by explicitly explaining the different levels of privacy and security expertise in Q7.

Our participants are not fully representative of the global users of VPNs. Our respondents skewed older, male, and more educated than the general U.S. population; this reflects the main user population for VPNs, especially in the U.S. [52]. Our collaboration with Consumer Reports demonstrated to us that their user base, who formed a large part of our recruitment, are avid VPN users that express the need for recommendations and advice from experts. Though we study a more-educated and possibly more tech savvy user base, the issues that we identify in our results (e.g., inadequate understanding) lead us to believe that such problems may be *even more prevalent* among the general U.S. population. Therefore, we argue that our results serve as an upper bound, and our recommendations will benefit the larger, more-general user base as well.

We restricted our analysis to only people located in the U.S. While VPN users outside the U.S. have diverse and valuable perspectives, their use cases are also different. Future studies could specifically explore the perspectives of users from countries where VPNs are commonly used to circumvent censorship or other access restrictions.

We intentionally only include users of commercial VPNs, university VPNs (typically managed by the university or a third-party), and users of free, and non-commercial VPN services in this study. We do not include users of self-hosted VPN solutions or (managers and users of) workplace-specific VPNs. We leave it to future work to explore these specific subgroups of users, since they are typically more highly-skilled, and/or possess high levels of technical knowledge.

Demographic	**Respondents**	**%**
Man	1011	80.75%
Woman	202	16.13%
Prefer not to disclose	35	2.8%
Non-Binary	4	0.32%
Over 65	741	59.19%
56-65	260	20.77%
46-55	105	8.39%
36-45	56	4.47%
26-35	36	2.88%
Prefer not to disclose	34	2.72%
18-25	20	1.6%
Post-grad education	527	42.09%
College degree	508	40.58%
Some college, no degree	150	11.98%
High school or eqlt	41	1.20%
Other	15	1.2%
Prefer not to disclose	11	0.88%
High-expertise users (Knowledgeable/Expert)	511	40.81%
Moderately knowledgeable users	631	50.40%
Limited-expertise users (No or mildly knowledgeable)	110	8.79%
Total	1252	

Table 2: Demographics of the (n=1252) survey respondents.♣

4 Data Characterization and Validation

Survey Responses. In total, we collected the user survey responses for six months, from March 1 to September 1, 2021. We had a total of 1,514 valid, completed responses out of which 1,374 (90.8%) indicated they are in the U.S.. The second-highest country (China) had 23 participants, and 20 countries had only 1 participant each. We decided to focus on the U.S.-based participants (and VPN providers popular in the U.S.) as there is not enough sample to draw meaningful conclusions about other countries.

Quality Checks. We have three questions, one quality- and two attention-checks, to ensure high-quality responses. Among the 1,374 U.S.-based participants that finished the survey, 1,264 or 92% passed our generic quality check. Next, we filter out users that failed both of our attention checks (Q11 and Q21). Furthermore, we review open-ended responses from the 259 participants that failed at most one attention check, as done in [22], and find that over 95.8% of these users

had insightful responses. Hence, we consider **1,252 users** that passed at least one attention check for the rest of the analysis.

Participant Demographics. Ours is the largest survey of VPN users to date, and we report on the 1,252 valid, high-quality responses. Shortly after launching the survey, we served on the panel of a VPN workshop organized by Consumer Reports with over 1,500 enthusiastic users in attendance, and sent out our study recruitment message to them. We believe that our various recruitment methods ensure that we study users who are highly motivated about commercial VPNs and actively use them. Our participants skewed older, male, and highly educated. However, due to the high number of responses we obtained, we are still able to make significant conclusions from the data. Though our participants do not represent **all** VPN users, our results (§5) indicate concerning issues even amongst the more educated, more tech savvy users, implying that our recommendations likely will benefit the more general VPN user population. The demographics are described in Table 2.

Cross-validating Self-reported Expertise. We report our results for different sub-groups of users based on their self-reported expertise, and type of VPN subscription they generally use, shown in Figure 2. We bucket participants based on their reported expertise in security and privacy: high expertise users (knowledgeable, expert), moderate expertise users, and limited expertise users (no, mild). In order to mitigate self-reporting biases, we follow all the recommended survey design methodology best practices by including descriptive explanations for each expertise level. We craft these explanations using our expertise and incorporating feedback from user survey practitioners. We use the terms "security" and "privacy" in these descriptions to allow users to use their own judgements, and we use our threat- and mental model questions later to have the user expound on their definitions. Furthermore, we analyze the open-ended text box responses to cross-validate users' expertise, and find that high expertise users provided insightful details to add to their mental models (presented in Appendix B.4) and limited expertise users were more likely to admit they do not know what protection the VPN offers them (§5.4).

5 Results from the User Survey

Security and privacy advocates, and technologists need a deeper understanding of the VPN ecosystem, and the misalignment of understanding between the stakeholders (VPN users and providers) can be exploited by bad actors to further deepen problems in the VPN ecosystem. To investigate and illuminate such issues, in this study, we conduct quantitative and qualitative studies of VPN users and VPN providers.

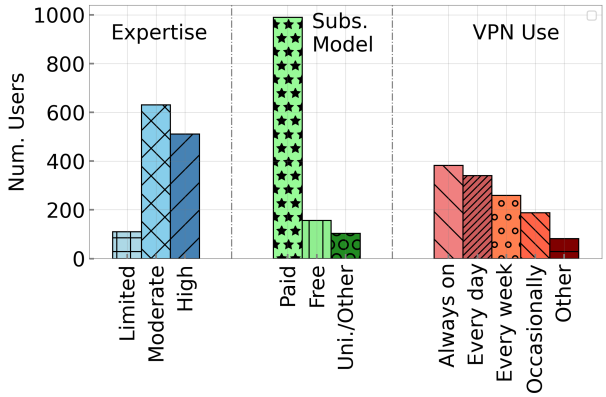

Figure 2: Overall statistics of the users' security and privacy expertise, their VPN subscription type, and VPN usage.♣

Based on the responses from our survey and interviews, we answer the following research questions:

5.1 RQ1: Motivations

First, we explore the reasons for which users use VPNs and allow them to choose multiple reasons. We provide them various options that we then distill into different categories.

Security and privacy are the main reasons why users use a VPN. We find that protection from threats, which we consider a *security motive* (82.1%, 1,027 of 1,252) and making public networks safer to use, which we term *privacy motive* (58.4%, 731) are the biggest reasons why users use VPNs. On the other hand, censorship circumvention (8.8%, 110) and file sharing such as torrenting (12.1%, 151) are among the least popular reasons. Our results are in contrast with [11] which finds university students prefer access to content (institutional, media streaming) over privacy, possibly due to the different priorities of the user populations. The overwhelming number of users that use VPNs for protection from perceived threats indicates the successful marketing of VPNs as a panacea for all security and privacy issues in the Internet.

Furthermore, 118 users also write-in additional reasons why they use VPNs (Appendix B.1). Users mention *privacy* (60.2%, 71 of 118; from ISP, tracking, surveillance, ad targeting) , *security* (12.71%, 15), being *offered the service* (10.1%, 12; by a company, with a purchase), *during travel* (7.6%, 9), and *anonymity* (2.5%, 3) as the main reasons for use.

Since finding a suitable VPN is not a trivial task, we ask users whether they had difficulty in selecting a VPN provider. Although the responses are almost evenly spread over the difficulty scale, we find differences between users with varying security and privacy expertise shown by a χ^2-test (p = 0.004206, with N=1251). As mentioned in 3.1, we perform pairwise z-tests (α=0.05) with FDR-BH correction to find how different user groups relate to each other.

High expertise users less likely to find VPN discovery very difficult, more likely to find it somewhat easy. We

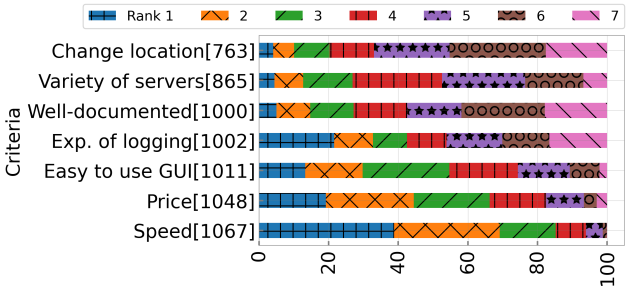

Figure 3: Importance levels users attach with criteria they look for in a VPN, presented along with number of users who chose it. Ranked from 1-most important to 7-least.♣

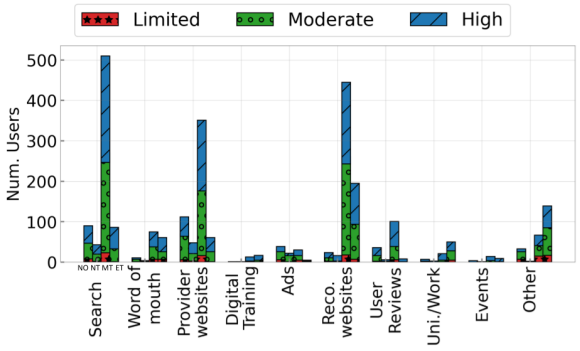

Figure 4: Trustworthiness of each resource as rated by users with different security and privacy expertise. Bars are No Opinion, Not-, Moderately- and Extremely-Trustworthy.♣

find that only 3.7% (19 of 511) of high expertise users find the discovery process very difficult which is significantly less than the 7% (44 of 631) of the moderate- and 11.9% (13 of 109) of the limited expertise users who find it so. High expertise users are significantly more likely to find the process somewhat easy (21.1%, 108 of 511, compared to 11% of the limited expertise users).

Furthermore, we find significant difference between users that use different subscription types (free, paid/premium, other) also shown by a χ^2-test (p = 0.000005, with N=1249). Understandably, university and "other" VPN users (most use a VPN provided as part of a software suite) are significantly more likely to say the process was somewhat or very easy (58.8%, 60 of 102) compared to 33.7% (334 of 990) of paid VPN users and 28% (44 of 157) of free VPN users. A portion of both the free VPN (40.8%, 64 of 157) and paid VPN users (34%, 337 of 990) find the process at least somewhat difficult. All of these findings are detailed in Table 5 in the Appendix.

5.2 RQ2: Needs and Considerations

To understand the needs that different users have, we ask them choose and rank criteria that they look for in a VPN. We ask the users to select the criteria they require in a VPN, and/or prefer to see in a VPN and then ask them to rank those criteria, from most important to least.

Speed, price, and an easy to use app are among the top three requirements in a VPN. We see that speed (72.6%, 909 of 1,252), price (55.4%, 694), and easy to understand app/GUI (44.1%, 553) are consistently among the top three requirements for VPN users, and over 216 users (17.3%) ranked clear explanation of logging and data practices as their number one, as shown in Figure 3. On the other hand, variety or number of servers (18.8%, 235 of 1,252), and using a VPN to change location for media sites such as Netflix (12.4%, 155) are among the lowest ranked requirements. We also find that logging data practices, which have received relatively little study, are ranked more highly than criteria like changing location for content or number of VPN servers, which have received more attention in the literature [54]. We highlight that understanding real-world user requirements can help shape future research focus.

Price is a big criteria for limited-to-moderate expertise users. Interestingly, users of all expertise rank speed equally highly as a top three criteria (no significant differences, p=0.348, N=1067). But limited-to-moderate expertise users are significantly more likely to rank *price* higher (χ^2-test, p=0.000150, N=1048); 71.1% (436 of 613) of these users rank it in their top three, compared to 59.3% (258 of 435) of high expertise users. This means that prices, discounts, and marketing around these factors is bound to have a vast effect on these users, similar to the study on UK and Japan users [46]. As we will demonstrate in §6, malicious marketing around pricing is common and dark patterns are often used to ensure customer lock-in.

On the other hand, **high expertise users rank clear explanation of logging significantly higher** (53.4%, 237 of 444 who chose it put it the top three) than all other users (33.8%, 164 of 485 moderate- and 34.2%, 25 of 73 limited expertise users) as shown by a χ^2-test ($p \ll 0.0001$, N=1002). Also, we find that significantly more high expertise users value an easy to understand GUI lower (only 38.6%, 158 of 409 high expertise users chose and rank it in their top three) compared to 64.4% (334 of 519) of the moderate- and 73.5% (61 of 83) of the limited expertise users, shown by a χ^2-test ($p \ll 0.0001$, N=1011). This indicates that high expertise users may be more confident in their ability to use a VPN application, and place higher value on the clarity of communication about the VPN service and the provider's data practices.

Users rely on search and recommendation sites rather than word of mouth to choose a VPN. Given these different needs and criteria, we explore what resources users use to discover and choose the right VPN for them. Users report that actively researching on the Internet (61.1%, 765 of 1,252), using recommendation websites (56.5%, 708), and reading the VPN providers' websites (48.1%, 602) are the top three ways they use to find a VPN for their needs. Users lean on these search engines and recommendation websites, rather

Population Subscription	Safety without VPN VS/SS/NO/**SU**/**VU**	**U%**	Safety with VPN **VS**/**SS**/NO/SU/VU	**S%**
Paid/premium	27/243/73/**(491/156)**	**65.4**$^{\uparrow}$	**(350/538)**/44/38/20	**89.7**$^{\uparrow}$
Free	9/37/20/(78/13)	58.0	(22/97)/**24**/11/3	75.8
Uni.&Write-in	10/36/12/(37/8)	43.7	(35/42)/18/7/0	75.5

Table 3: Number and % of users with different subscription types and their feeling of safety without and with a VPN (from VS-Very Safe to VU-Very Unsafe). $^{\uparrow}$indicates more likely than the other subgroups for that column.♣

than traditional methods like word of mouth from friends and family (5.7%, 167), or digital training workshops (1.19%, 35). This highlights the perils of an unregulated advertising and marketing ecosystem around VPNs, as we expound in §6.2.

Users rate recommendation websites as trustworthy sources. Interestingly, among the top three resources they use, more users rate recommendation websites as trustworthy compared to the other two; 93.9% (665 of 708) of them rate them moderately to extremely trustworthy. Figure 4 illustrates how users rate the trustworthiness of each of the resources. Notably, a high proportion of users whose work or school provides their VPN service rank it extremely trustworthy (61.1%, 55 of 90), highlighting that these users expect work/university VPNs to be of a high-quality.

Interestingly, 281 users use the "other" option to write-in other resources they may have used. From our qualitative coding of these responses, we notice that the VPN being offered as part of a software/security suite is the most common response (36.3%, 102 of 281). Other responses include: trusted service provider recommendations (9.6%, 27), and prior experience (5.3%, 15). Appendix B.2 contains all the codes.

5.3 RQ3: Emotional connection and Threat model

To understand if users attach emotional considerations such as a feeling of safety with using a VPN, we first ask them their perception of safety when browsing without a VPN and then, with a VPN. We find that there are significant differences between users that use different VPN subscription types (paid, free, and university and other) and their perception of safety without a VPN, (χ^2-test, $p = 0.0001$, N=1250). We explore differences between users with varying expertise levels in Appendix A. **Users indicate they feel unsafe without a VPN, especially those who use paid/premium VPNs.** Overall, users indicate that they feel unsafe (62.6%, 784 of 1,252) browsing the Internet without a VPN. Interestingly, we find that paid/premium VPN users are significantly more likely to feel at least somewhat unsafe when browsing without their VPN (65.4%, 647 of 990) as compared to users that use university and other VPNs (43.7%, 45 of 103).

Paid VPN users are more likely to feel safe with their VPN, while free VPN users likely to indicate no opinion. Subsequently, there are also significant differences between users with different subscription types and their perception of safety with a VPN, ($p \ll 0.001$, N=1250). While large sections of all populations feel somewhat or very safe using a VPN (86.7%, 1,086 of 1,252), we find that paid/premium users are significantly more likely to indicate they felt safe when using their VPN (89.7%, 888 of 990), compared to free VPN users (75.8%, 119 of 157) and university/other users (75.5%, 77 of 102), who are significantly less likely. Free VPN users are significantly more likely to indicate no opinion about security (15.3%, 24 of 157) as compared to the 4.4% of paid users alone (44 of 990), shown in Table 3. Overall, we find that a large number of users attach emotional considerations such as safety with VPN use, and hence are likely to continue using VPNs, according to prior work studying retention [29].

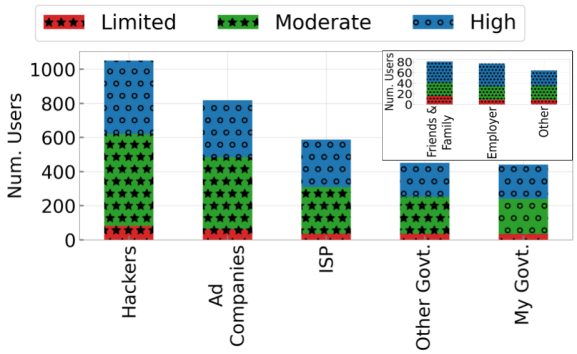

Figure 5: Entities from whom users with different security and privacy expertise want to protect their online activity.♣

A majority of users use VPNs to protect and secure their online activities. To understand users' threat models when it comes to using a VPN, we first ascertain whether users use a VPN to secure their online activities, and if yes, who they want to protect it from. Notably, 91.5% (1145 of 1,252) of users indicate they use VPNs for securing or protecting their online activity. When exploring who users aim to protect themselves from, we find that hackers/eavesdroppers on open WiFi networks (83.9%, 1,051 of 1,252), advertising companies (65.4%, 819), and internet service providers (ISP) (46.9%, 587) are the top three responses. Notably, only ≈30% of users are concerned about the U.S. government or other governments. This is intriguing because post Snowden's surveillance revelations in 2014, more users moved towards privacy tools such as VPNs and anonymity tools such as Tor [45]. Our results indicate a shift in user's attitudes, and show a growing concern towards corporate and advertisement surveillance. This shift could be due to the influence of the marketing around VPNs and the security advice to which users are exposed. Prior work also shows YouTubers often cite "the media" and "hackers" as common adversaries [1]. Figure 5 shows the number of users for each of these options.

High expertise users more likely to list their ISP in their threat model. We test each option independently to see if there are significant differences between users with

varying expertise. We find that significantly more high expertise users indicate their ISP as one of the reasons (54.4%, 278 of 511), as compared to other users (43.3%, 273 of 631 moderate-, and 32.7%, 36 of 110 limited expertise users) ($p \ll 0.0001$, N=1252). While no significant difference was found between users selecting advertising companies (χ^2, p=0.157, N=1252), significantly less proportion of limited expertise users indicate that hackers and eavesdroppers are a concern (73.6%, 81 of 110) as compared to 85.6% (540 of 631) of the moderate- and 84.1% (430 of 511) of the high expertise users, as confirmed by a χ^2-test (p=0.00695, N=1252).

5.4 RQ4: Mental Model

To evaluate users' mental model of VPNs, we ask them a scenario question which aims to elicit their understanding of the protections VPNs actually provide. In the given scenario in the question, the user concluding that their ISP learns what websites they visit while connected to a VPN indicates a flawed mental model.

Almost 40% of users have a flawed mental model. We find that a high portion of users (39.9%, 500) have a flawed mental model and believe their ISP can see the websites they visit over the VPN. Worryingly, we see no significant difference between users of different expertise based on the χ^2-test (p=0.0927, N=1252). Our results are concordant with previous work which find that users, even experts, have misconceptions about the protections certain tools offer [3, 48]. We initially also considered the 135 users who answered "Nobody [can see what website I visit]" as having a flawed mental model. But we instead opt for a conservative approach and did not include them because four users clarified their response using the textbox accompanying this question. They state that since their VPN says no logging, tracking, or sharing, *ideally* nobody should know what website they visited.

Limited expertise users are more likely to have an unclear mental model, while high expertise users more likely to add insightful details. We find significant difference between users that chose "I don't know" to this question, based on a χ^2-test ($p \ll 0.00001$, N=1252). We find that users with limited expertise are more likely to choose "I don't know" (30.9%, 34 of 110 users) compared to 16.3% (103 of 631) of the moderate- and 5.5% (28 of 511) of high expertise users. High expertise users are significantly more likely to use the "other" option and write-in their answer (14.9%, 76 of 511) as compared to 8.9% moderate- and 1.8% limited expertise users (p= 0.00003, N=1252). Analyzing these write-in responses, we find that high expertise users add insightful details such as *DNS providers* knowing what websites user visits, and *site owner* learning about the user using logins or cookies. They identify other threat actors such as the site's partners, search engine used to navigate to the site, government agencies, and browser fingerprinters; all of the codes are presented in Appendix B.4.

Expertise	NotSure	**NS%**	Typ/Dang/Misc/O.	**Typ.%**	**Dang.%**
High	132	25.83	326/35/217/58	86.02	9.23
Moderate	304	48.18	292/44/220/32	89.30	13.46
Limited	68	61.82	35/18/36/3	83.33	42.86

Table 4: Number and % of users who indicate the types of data they think VPN providers collect. Users can choose multiple options, and we exclude users who chose "not sure" (NS) from the other counts.♣

To understand if VPN users have a good idea about the data VPNs can collect about them, we present many options and ask users to indicate the various kinds of data they think a VPN provider collects about them. During the analysis, we bucket these options into: *typical*, *dangerous-unreasonable*, *miscellany*, *not sure*, and *custom input*. While the last two are self-explanatory, "typical" includes demographics and account holder information, VPN servers connected to, timestamps at when VPN is in use, and device type. We consider them typical since the data is readily available to a VPN provider. The "dangerous-unreasonable" bucket includes: private messages, audio/video recordings, and keystrokes from device, all of which are not usually collected by a VPN provider, unless they are operating a malicious service, while "miscellany" includes website visited, geolocation, and interests for ads. While a reasonable provider would not collect this type of data, it is possible that some VPN providers do collect them.

At least 40% users indicate they are unsure what data is collected, and ≈13% of the remaining users think unreasonable kinds of data are collected by VPNs. We find that 40.3% (504 of 1,252) of users indicate they are not sure what data is collected, limited-(61.8%, 68 of 110) and moderate expertise users (48.2%, 304 of 631) are significantly more likely to indicate uncertainty as compared to 25.8% (132 of 511) of the high expertise users (χ^2, $p \ll 0.0001$, N=1252). We exclude these users from the analysis and from the remaining 748 users, we see that in general users believe typical data (87.3%, 653) is collected by VPN providers. However, 13% (97 of 748) of users think VPNs collect dangerous-unreasonable data. The fact that users of all expertise levels have this belief, reiterates the need for better, more effective user education. Table 4 summarizes these results.

Finally, we explore the reasons why users think such data is being collected by VPN providers. A majority of respondents (79.2%, 992) believe the main reason is for internal analytics and quality of service reasons. Interestingly, significantly more limited expertise users believe that the data is being collected for advertising (36.4%, 40 of 110), as compared to 20.4% of moderate- and 16.4% high expertise users (χ^2,p=0.000014, N=1252). A significantly high portion of limited expertise users also believe data is used for user tracking (36.4%, 40, p=0.019), and selling to third parties (33.6%, 37, $p \ll 0.0001$), highlighting that limited expertise users believe VPNs use data collected about users for monetary benefit.

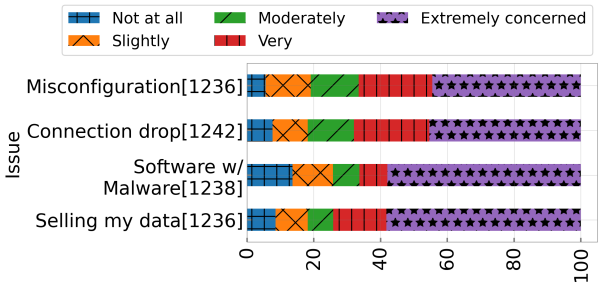

Figure 6: Users indicate their concern levels towards VPN-related issues, with the number of users who answered each.♣

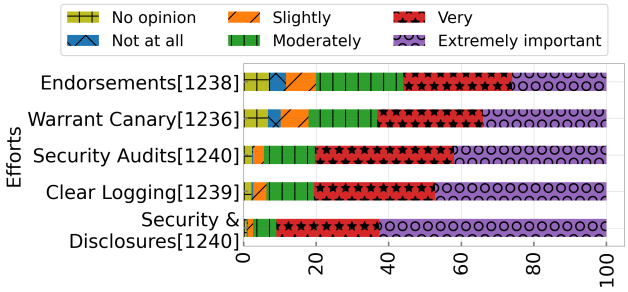

Figure 7: Users indicate the importance of trust-increasing efforts by VPNs, with number of users who answered each.♣

5.5 RQ5: Perception and Trust

In order to understand users' perception of the VPN ecosystem and its issues, we ask users to rate their concern levels towards VPN related issues. We find that users are very or extremely concerned about VPN providers selling their data (73.2%, 917 of 1,252), and the VPN software containing malware (65.6%, 821). Users also express higher degrees of concern towards more technical issues such as VPN software failing without warning (67.4%, 884), and misconfigured VPN services (65.7%, 823), illustrated in Figure 6. We find no statistically significant differences between users of different expertise or subscription types for these options (χ^2, p»0.05).

Finally, we ask users what level of importance they associate with efforts that VPN providers undertake to earn and increase trust from the user base. We find that users consistently rate security protocols and disclosure of breaches (62%, 776 of 1,252) as an extremely important effort, followed by having a clear logging policy (46.7%, 585), and independent security audits (41.6%, 521), as shown in Figure 7. While there may be other efforts that we do not list, we hope that VPN providers and researchers use these insights gleaned from the users' perspectives to inform their future efforts and campaigns to secure and foster user trust.

6 Perspectives of the VPN Providers

In this section, we present exploratory results from our VPN provider interviews and summarize the key issues and themes, with number of providers per theme in brackets. We compare these insights with results from the user survey, and highlight the key areas where the two are misaligned.

6.1 Key Themes

Key Efforts. We learn from providers that they focus on cross-platform security development (6/9), product simplicity (4/9), and usability (5/9) of their product. They also mention that they try to be reliable, gain trust over time (5/9), and practice transparency (5/9). We also noticed many VPN providers mentioned offering additional features, such as filtering, ad- and tracker-blocking similar to anti-virus software, indicating that VPNs are evolving beyond their normal functionality to retain users. From a mental model perspective, this could potentially be harmful as it sets an over-expectation of security and privacy, while users are already unclear about protections that standard VPNs offer them.

High-level Challenges. When asked about the biggest challenges in the industry, providers explain that *building trust* (6/9) is hard because there is a large number of providers and little transparency. We find that providers agree that problems generally stem from lack of trust, focusing on features and not privacy, and overestimation and overselling of service. Providers also mention that users do not understand risks (7/9), and that it is their responsibility to do better in user education and ensure honesty in their disclosures to users.

User Base. When asked about their user base and whether they conduct studies to understand them, almost all providers explain that having a *privacy-focused service deters user studies*, and that they try not to learn about their users (7/9). Instead, they typically depend on inbound user feedback such as in-app surveys or support tickets. We notice that commercial providers mention that they prefer privacy-centric users, and that western users are more likely to be paying customers.

Pricing & Marketing. Providers mention that development, labor and marketing are the main factors affecting pricing (5/9). Other factors include deals with server and cloud providers, organization build, technical means, and infrastructure. They mention that growing the user base is imperative as it creates economies of scale. Providers also note the existence of malicious practices around discounts that are not user-friendly (5/9), like marketing gimmicks to lock users. Multiple providers remark that it is the norm of the industry (3/9), one of whom says:

> "I think it's not good for consumers but why everyone does it, because everyone else does that."

A majority of providers agree that marketing plays a big role; noting that the marketing costs are high, and the competition

is harsh. Regarding marketing methods, many providers mention that they do ethical marketing by being involved with the user community, relying on user reviews and word of mouth.

VPN Review Ecosystem. We discover that a main theme from the interviews is the issue of the VPN review ecosystem. One provider calls it a "parasitic industry" and a majority of providers (6/9) remark that the review ecosystem mostly runs on money, e.g. paid reviews, and cost-per-action (CPA). They also explain that VPNs or their parent companies may own different review sites [38], many review sites even *auction the #1 spot*, and do reviews for money. Multiple providers also mention that Google search results are unreliable, and that there are few good reviewers left; one provider says:

> "You honestly cannot find even one ranking site that is honest, if you just tell people that...so that people know"

Dark Patterns in the Industry. Another recurring theme was about dark patterns in the industry. Since most of these patterns are usually not readily apparent to users and researchers, we also explicitly ask a question about them. We divide the issues mentioned by various providers into:

Operational Issues (7/9): These include VPN providers having anonymous or unknown owners, having deceptive subscription models, and tracking users on their own sites, which was also highlighted in a recent report [30]. Providers also remarked on aggressive and unethical marketing such as retargeting users with VPN ads, and relying on users forgetting to cancel subscriptions. On the other hand, providers mention that VPNs get attacked as well (by other providers, bad users, and by those who abuse free VPN services).

Malicious Marketing (6/9): Providers mention several issues, that we term as *malicious marketing*, including the use of affiliate marketing, preying upon users' lack of knowledge, and overselling of service including selling anonymity even though that is not a VPN guarantee. They also foster a false sense of security around VPNs through misinformation, fearmongering, dishonest non-expert reviews, and lying to users in disclosures. One provider, on fearmongering:

> "The best ways to get people to pay for something is to scare them and to tell them that they need security"

Factors Enabling Dark Patterns (4/9): Providers bring up several challenges that exacerbate these practices, such as the fact that the VPN ecosystem has no accountability, lacks transparency, and has few marketing and advertisement standards. Since the VPN industry is spread over multiple jurisdictions, it is hard to regulate. One provider calls it *the wild west*:

> "You know we could just say literally anything...there's absolutely no oversight. There's no one to tell you, *"Ah, you can't say that because that's not true."* There's no regulation, there's no kind of governing body"

6.2 RQ6: Alignment between VPN users and providers

We highlight several key areas where VPN users and providers are misaligned in their understandings and incentives, in addition to issues that both parties agree on. By highlighting these issues, we hope that technologists, and security advocates prioritize users' challenges, and focus on key problem areas. We arrange these issues from most aligned to least.

Privacy-centric Users. We note that providers explicitly mention that they prefer and cater to privacy-centric users, which aligns with the findings from our survey where over 91% of users mention that they use VPNs for security and/or privacy. Since providers mention they respect privacy and are unable to conduct user studies of their own, it is imperative for researchers to develop an understanding of VPN users.

Users' Mental Model of VPNs. Providers say that users have flawed mental models of VPNs (6/9) and our survey concurs that ≈40% of users do indeed have a flawed mental model. Providers and the security advocacy community should hence place high priority on user education. Providers mention that challenges in improving users' mental models include the lack of positive reinforcements (visual signs that a VPN is working), constant exposure to negative experiences (increased encounters of CAPTCHAs, media sites blocking VPN use), and striking a balance in technical communication. We emphasize that user-onboarding, clear communication, and responsible advertising are key drivers for change.

Importance of Pricing. From our user survey, we see that pricing is among one of the highest priorities for users, especially for limited-to-moderate expertise users. However, providers on the other hand mention that certain malicious marketing gimmicks are often used—such as fearmongering, fake countdown timers, and being always on sale—to lock-in users. We fear that since pricing is key for users, malicious tactics used by certain providers may chain users to a service that may not necessarily meet security standards. We strongly urge that advocates focus on regulations to protect consumers.

Users' Reliance on Review Sites. Despite most providers agreeing that the review ecosystem is not objective about the services and is instead largely motivated by money, our survey shows that users strongly rely on them and believe they are trustworthy. Though our survey studies only U.S. users, the VPN providers believe that the western population (including U.S.) are more likely to pay for their subscriptions. It is important to deter the exploitation of these users by informing them of the nature of the review ecosystem and how the reviews and rankings are made. As we highlight from the providers' interviews, a lot of the malicious marketing

preys on users' misunderstandings. Hence, shedding light on these behaviors in the review ecosystem is crucial to ensure that they do not continue profiting off users via paid reviews and CPA. One provider says:

> "[Running costs have reduced] in the last 10 years, yet [VPN] prices are all the same. Why is that? Well it's because the VPN review sites are getting all the money."

Users' View on Data Collection. We find that over 40% of users are not sure exactly what data is being collected about them by VPN providers. Of the remaining users, we find that 13% think that VPNs collect dangerous or unreasonable kinds of data. On the other hand, multiple VPN providers say that they clearly communicate their logging practices, or that they do no logging and have audits to prove it. From our survey, we also find that having a clear logging policy is among the top important indicators for increasing trust with users. Alongside improving users' mental models of how VPNs work, this is another key issue that VPN providers can address by better informing users about their operation.

7 Actionable Recommendations

Traditional approaches to regulating including standardization by government bodies may not be the best solution for VPNs because the providers and VPN servers span multiple jurisdictions. Another approach can be self-regulation within the industry. However, though coalitions look good on paper, it is necessary to bring enough providers together, and ensure oversight in order to hold these coalitions accountable. One provider, on why having such an alliance is hard:

> "[VPN providers] don't want to be held accountable for the [mistakes] of other providers...there's not a lot of trust."

Even if providers do form coalitions, we find that they do not really hold to their own self-regulated principles. In our prior work, we also find that the lack of regulation and standardization leads to VPN providers offering varying levels of security and privacy [36].

We strongly recommend that FTC and other government organizations exert oversight on VPN advertising and curb malicious tactics used by VPNs, because such aggressive and misleading ad campaigns could degrade users' mental models about VPNs. An example of successful oversight is NordVPN's ad being banned in the UK for misleading users [8]. In addition, we advocate for coordinated efforts from the industry, academia, and consumer protection organizations to bring attention to the flawed VPN recommendation ecosystem.

Finally, our study also shows that user education campaigns regarding VPNs and the VPN ecosystem must be prioritized. We find key areas that need the most improvement: users' mental model of what a VPN provides, what data it can collect, and the threat models for which VPNs can be most useful. Since the user population surveyed in our study is on average older and more educated, our results suggests that incomplete and flawed mental models may be even more prevalent among the general U.S. population. We urge security and privacy advocates such as the EFF and CDT, consumer protection agencies such as the FTC, and community initiatives such as IFF to devote their efforts towards VPN user education, raise awareness, and advocate for VPN industry oversight.

8 Discussion & Conclusion

VPNs have quickly gained popularity as a security and privacy tool for regular Internet users. Commercial VPNs are now a multi-billion global industry with numerous VPN providers, and apps on almost every platform. In our interviews with them, multiple providers mention that setting up a VPN and offering a service is not technically difficult, especially with the existing open source solutions [9, 32], and highlight that many VPN companies have unknown or anonymous ownership. One provider says there is a low bar to entry:

> "Technically it's not that hard to run a VPN...two people in a basement with a half decent power....can run a VPN."

For users however, exposure to risk of surveillance, reports of ISPs selling data, and increasing access restrictions have all led to an increased awareness of online risks. VPNs are marketed as technological solutions to many of these issues, though not all users will be able to verify these claims. In simple terms, a user using a VPN is simply transferring trust, say from their Internet provider, onto the VPN provider. Internet service providers (ISPs) have been around for longer and have many regulations globally. However, such regulations and advocacy has not yet caught up to the VPN industry.

In this paper, we conduct studies on VPN users and providers and present actionable recommendations on important problem areas in the VPN ecosystem. Our interviews with VPN providers helps open up communication between academia and companies developing privacy-enhancing tools, which can lead to transfer of knowledge, foster collaboration, and help develop solutions for issues in the ecosystem that ultimately impacts users. We highlight that understanding real-world user needs and requirements can help shape future research focus. We hope that by shedding light on issues such as the ones rampant in the VPN review ecosystem, we raise awareness and encourage investigation, advocacy, and regulation to improve the entire VPN ecosystem for the better.

9 Acknowledgment

The authors are grateful to Michelle Mazurek for her valuable input, and Karen Jaffe, Philipp Winter and Jane Im for their feedback on the survey. We thank Ben Moskowitz, Leah Fischman, Yael Grauer, and the reviewers for their feedback for their feedback. This work was made possible by the Open Technology Fund, Consumer Reports Digital Lab Fellowship, and National Science Foundation grant CNS-2141512.

References

[1] O. Akgul, R. Roberts, M. Namara, D. Levin, and M. L. Mazurek. Investigating Influencer VPN Ads on YouTube. In *IEEE Symposium on Security and Privacy*, 2022.

[2] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 1995.

[3] V. Binkhorst, T. Fiebig, K. Krombholz, W. Pieters, et al. Security at the end of the tunnel: The anatomy of VPN mental models among experts and non-experts in a corporate context. In *31st USENIX Security Symposium (USENIX Security 22)*, 2022.

[4] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 2006.

[5] F. Chanchary and S. Chiasson. User perceptions of sharing, advertising, and tracking. In *Eleventh Symposium On Usable Privacy and Security (SOUPS)*, 2015.

[6] J. Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 1960.

[7] Consumer Reports. Security Planner, 2021. https://securityplanner.consumerreports.org/.

[8] G. Corfield. NordVPN rapped by ad watchdog over insecure public Wi-Fi claims. The Register, 2019. https://www.theregister.com/2019/05/01/nordvpn_tv_ad_rapped_advertising_standards_authority/.

[9] J. A. Donenfeld. Wireguard: Next generation kernel network tunnel. In *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2017.

[10] A. Dutkowska-Żuk, A. Hounsel, A. Morrill, A. Xiong, M. Chetty, and N. Feamster. How and why people use virtual private networks. In *31st USENIX Security Symposium (USENIX Security 22)*, 2022.

[11] A. Dutkowska-Zuk, A. Hounsel, A. Xiong, M. Roberts, B. Stewart, M. Chetty, and N. Feamster. Practicing Safe Browsing: Understanding How and Why University Students Use Virtual Private Networks. *arXiv e-prints*, 2020.

[12] M. Ikram, N. Vallina-Rodriguez, S. Seneviratne, M. A. Kaafar, and V. Paxson. An analysis of the privacy and security risks of android VPN permission-enabled apps. In *Proceedings of the 2016 Internet Measurement Conference*, 2016.

[13] Internet Freedom Festival. Announcing the VPN Village!, 2020. https://internetfreedomfestival.org/vpn-village/.

[14] C. Kaufman et al. Internet key exchange (IKEv2) protocol. Technical report, RFC 4306, 2005.

[15] M. T. Khan, J. DeBlasio, G. M. Voelker, A. C. Snoeren, C. Kanich, and N. Vallina-Rodriguez. An empirical analysis of the commercial VPN ecosystem. In *Proceedings of the Internet Measurement Conference (IMC)*, 2018.

[16] A. Kochovski. The Top 25 VPN Statistics, Facts & Trends for 2021. Cloudwards, 2021. https://www.cloudwards.net/vpn-statistics/.

[17] J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 1977.

[18] J. Lazar, J. H. Feng, and H. Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.

[19] R. Likert. A technique for the measurement of attitudes. *Archives of psychology*, 1932.

[20] Mark Smirniotis. What Is a VPN and What Can (and Can't) It Do? New York Times Wirecutter, 2021. https://www.nytimes.com/wirecutter/guides/what-is-a-vpn/.

[21] A. Mathur, G. Acar, M. J. Friedman, E. Lucherini, J. Mayer, M. Chetty, and A. Narayanan. Dark patterns at scale: Findings from a crawl of 11k shopping websites. *Proceedings of the ACM on Human-Computer Interaction*, (CSCW), 2019.

[22] P. Mayer, Y. Zou, F. Schaub, and A. J. Aviv. "Now I'm a bit angry:" Individuals' Awareness, Perception, and Responses to Data Breaches that Affected Them. In *Proceedings of the USENIX Security Symposium (USENIX Security)*, 2021.

[23] N. McDonald, S. Schoenebeck, and A. 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*, (CSCW), 2019.

[24] M. L. McHugh. Interrater reliability: the kappa statistic. *Biochemia medica*, 2012.

[25] W. Melicher, M. Sharif, J. Tan, L. Bauer, M. Christodorescu, and P. G. Leon. Preferences for web tracking. *Proceedings on Privacy Enhancing Technologies (PoPETS)*, 2016.

[26] Microsoft. Microsoft Leads Initiative for Virtual Private Networks Across the Internet, 1996. https://news.microsoft.com/1996/03/04/microsoft-leads-initiative-for-virtual-private-networks-across-the-internet/.

[27] M. B. Miles, A. M. Huberman, J. Saldana, et al. Qualitative data analysis: A methods sourcebook. *Thousand Oaks, CA: Sage*, 2014.

[28] Miles Kenyon. Citizen Lab Summer Institute 2019. Citizen Lab, 2019. https://citizenlab.ca/2019/02/citizen-lab-summer-institute-2019/.

[29] M. Namara, D. Wilkinson, K. Caine, and B. P. Knijnenburg. Emotional and Practical Considerations Towards the Adoption and Abandonment of VPNs as a Privacy-Enhancing Technology. *Proceedings on Privacy Enhancing Technologies (PoPETS)*, 2020.

[30] A. Ng. How Private Is My VPN? The Markup, 2021. https://themarkup.org/ask-the-markup/2021/08/12/how-private-is-my-vpn.

[31] Open Technology Fund. Open Technology Fund, 2021. https://www.opentech.fund/.

[32] OpenVPN. A Modern Private Network, Built for the Cloud, 2021. https://openvpn.net/.

[33] PR Newswire. New Analysis from Global Industry Analysts Reveals Steady Growth for Virtual Private Network (VPN), with the Market to Reach $77.1 Billion Worldwide by 2026, 2021. https://www.prnewswire.com/news-releases/new-analysis-from-global-industry-analysts-reveals-steady-growth-for-virtual-private-network-vpn-with-the-market-to-reach-77-1-billion-worldwide-by-2026--301368222.html.

[34] E. Rader. Awareness of behavioral tracking and information privacy concern in Facebook and Google. In *10th Symposium On Usable Privacy and Security (SOUPS)*, 2014.

[35] Rae Hodge. Why you should be skeptical about a VPN's no-logs claims. CNET, 2020. https://www.cnet.com/tech/services-and-software/why-you-should-be-skeptical-about-a-vpns-no-logs-claims/.

[36] R. Ramesh, L. Evdokimov, D. Xue, and R. Ensafi. VPNalyzer: Systematic Investigation of the VPN Ecosystem. In *Network and Distributed System Security*, 2022.

[37] R. Ramesh, A. Vyas, and R. Ensafi. "All of them claim to be the best": Multi-perspective study of VPN users and VPN providers. *arXiv preprint arXiv:2208.03505*, 2022. https://arxiv.org/abs/2208.03505.

[38] D. Rankovic. ExpressVPN critics are concerned about ties to former intelligence agents and adware. Reclaim the Net, 2021. https://reclaimthenet.org/expressvpn-sale-new-owners/.

[39] E. M. Redmiles, Y. Acar, S. Fahl, and M. L. Mazurek. A summary of survey methodology best practices for security and privacy researchers. Technical report, University of Maryland, 2017.

[40] Reethika Ramesh, Anjali Vyas, Roya Ensafi. VPN Provider Interview Questionnaire. https://web.archive.org/web/20220928200259/https://vpnalyzer.org/vpn-interview-questionnaire.pdf.

[41] Reethika Ramesh, Anjali Vyas, Roya Ensafi. VPN User Survey Instrument. https://web.archive.org/web/20220928200224/https://vpnalyzer.org/user-survey-instrument.pdf.

[42] D. Ruiz. 21 million free VPN users' data exposed. Malwarebytes Labs, 2021. https://blog.malwarebytes.com/cybercrime/privacy/2021/03/21-million-free-vpn-users-data-exposed/.

[43] Sarah Coble. VPN Usage in US Quadruples. Infosecurity Magazine, 2020. https://www.infosecurity-magazine.com/news/vpn-usage-in-us-quadruples/.

[44] F. Shirazi and M. Volkamer. What deters Jane from preventing identification and tracking on the Web? In *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, 2014.

[45] C. Smith. The Snowden effect: Three years after Edward Snowden's mass-surveillance leaks, does the public care how they are watched? *Index on Censorship*, 2016.

[46] N. Sombatruang, T. Omiya, D. Miyamoto, M. A. Sasse, Y. Kadobayashi, and M. Baddeley. Attributes affecting user decision to adopt a Virtual Private Network (VPN) app. In *International Conference on Information and Communications Security*. Springer, 2020.

[47] K. Spiel, O. L. Haimson, and D. Lottridge. How to do better with gender on surveys: a guide for HCI researchers. *Interactions*, 2019.

[48] P. Story, D. Smullen, Y. Yao, A. Acquisti, L. F. Cranor, N. Sadeh, and F. Schaub. Awareness, adoption, and misconceptions of web privacy tools. *Proceedings on Privacy Enhancing Technologies (PoPETS)*, 2021.

[49] W. J. Tolley, B. Kujath, M. T. Khan, N. Vallina-Rodriguez, and J. R. Crandall. Blind In/On-Path Attacks and Applications to VPNs. In *Proceedings of the USENIX Security Symposium (USENIX Security)*, 2021.

[50] W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zorn, and B. Palter. Layer two tunneling protocol (L2TP). Technical report, RFC 2661, August, 1999.

[51] University of Chicago. Dark Patterns: UChicago/Princeton Research Reveals the Dirty Tricks of Online Shopping, 2019. https://cs.uchicago.edu/news/dark-patterns/.

[52] A. Vigderman and G. Turner. 2021 VPN Usage Statistics. Security.org, 2021. https://web.archive.org/web/20211011192217/https://www.security.org/vpn/statistics/.

[53] vpnMentor. Report: No-Log VPNs Reveal Users' Personal Data and Logs, 2021. https://www.vpnmentor.com/blog/report-free-vpns-leak/#Timeline-of-Discovery-and-Owner-Reaction.

[54] Z. Weinberg, S. Cho, N. Christin, V. Sekar, and P. Gill. How to Catch when Proxies Lie: Verifying the Physical Locations of Network Proxies with Active Geolocation. In *Proceedings of the Internet Measurement Conference (IMC)*, 2018.

[55] D. Xue, R. Ramesh, A. Jain, M. Kallitsis, J. A. Halderman, J. R. Crandall, and R. Ensafi. OpenVPN is Open to VPN Fingerprinting. In *31st USENIX Security Symposium (USENIX Security 22)*, 2022.

[56] Yegor Sak. We're not paying for #1. Windscribe Blog, 2021. https://blog.windscribe.com/were-not-paying-for-1-25b4e55ca10/.

[57] Zack Whittaker. NordVPN confirms it was hacked. TechCruch, 2019. https://techcrunch.com/2019/10/21/nordvpn-confirms-it-was-hacked/.

[58] Y. Zou, K. Roundy, A. Tamersoy, S. Shintre, J. Roturier, and F. Schaub. Examining the adoption and abandonment of security, privacy, and identity theft protection practices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

Population	VD/SD/NE/SE/VE	**V. Difficult %**	**S. Easy %**
High expertise	19/139/158/ 108 /87	3.7$^{\downarrow}$	21.1$^{\uparrow}$
Moderate expertise	44/180/198/116/93	7.0	18.4
Limited expertise	13/22/39/12/23	11.9	11
Population	Diff/Neither/Easy	**Difficult %**	**Easy %**
Paid/Premium	337/319/334	34	33.7
Free	64/49/44	40.8	28.0
Other (Uni./other)	16/26/60	15.7	58.8$^{\uparrow}$

Table 5: Number and % of users from different user groups indicate how difficult it was decide on a VPN to use (from VD-Very Difficult to VE-Very Easy). Symbols indicate $^{\uparrow}$more, and $^{\downarrow}$less likely than the other rows in the column.♣

Population Expertise	Safety without VPN VS/SS/NO/SU/VU	**S%**	Safety with VPN VS/SS/NO/SU/VU	**S%**
High	(22/133)/48/231/77	30.3	(202/244)/30/21/13	87.3
Moderate	(19/167)/44/324/77	29.5	(179/378)/38/27/9	88.3
Limited	**(5/17)**/13/52/23	**20.0**$^{\downarrow}$	**(27/56)**/18/8/1	**75.5**$^{\downarrow}$

Table 6: Number and % of users with different security and privacy expertise and their feeling of safety when browsing without and with a VPN (from VS-Very Safe to VU-Very Unsafe). Symbols indicate $^{\uparrow}$more, and $^{\downarrow}$less likely than the other rows in the column. Highlighted values indicate that they contribute to the relevant percentage.♣

A Appendix: Emotional connection with VPN for different user expertise (RQ3)

As shown in 5.3, in general, users indicate they feel unsafe without a VPN. We find that there are no significant differences between users with varying expertise levels and their perception of safety without VPNs, as shown by a χ^2-test (p = 0.085, N=1252). We notice that less limited expertise users indicate that they feel at least somewhat safe without a VPN (only 20%, 22 of 110 as compared to 30.3% of the high- and 29.5% of the moderate expertise users).

While this is not a statistically significant difference, we explored the reason they do not feel safe without a VPN by analyzing their textual response immediately after this question. Limited expertise users who responded (98 of 110) mainly express worry (about hacking, tracking, and more), and confusion about what VPN offers, and explain scenarios where they feel unsafe. S99 says:

> "One never knows when either the so-called good guys or the bad guys are lurking about, just waiting to pounce. In my book, I want to be safe rather than sorry[...]"

In general, users indicate they feel safer browsing the Internet with a VPN. In a different section of the survey, we ask them about the perception of safety while using a VPN. We find significant differences between users with varying expertise levels and their perception of safety with VPNs as well, shown by a χ^2-test (p=0.003, N=1252). While large sections of all populations feel somewhat or very safe (86.7%, 1,086) using a VPN, limited expertise users are *significantly less* likely to indicate they felt safe using a VPN (75.5%, 83 of 110) compared to 88.3% (557 of 631) of moderate- and 87.3% (446 of 510) of high expertise users), summarized in Table 6. We also find that instead limited expertise users were significantly more likely to indicate they had no opinion on safety while using a VPN (16.4%, 18 of 110), possibly due to confusion on what a VPN provides. S1153 says, who indicated no opinion says:

> "I feel both somewhat unsafe and somewhat safe"

B Codes from qualitative survey responses

B.1 Reasons for use

Privacy from ISP (22), **Privacy:** Privacy (17), from tracking (10), from tracking and ads targeting (5), surveillance (3), securing browsing history (3), hiding location (2), from ads (2), selling my data (2), hacking (2), from attribution (1), banking (1), during searching (1), ISP and large companies (1), **Security:** during banking (4), hackers (4), as a principle (2), paranoia (1), confidential/sensitive data (2), OpSec (1), hackers/surveillance and bad actors (1), protection (1), **Offered the service:** by Norton (4), free with other service (3), with router (1), by ISP (1), for low price (1), with device (1), from employer (1), **While travelling:** surveillance countries (2), protection from local actors (2), censoring countries (2), in general (2), don't trust hotels (1),**Anonymity** (3), **Access geo-restricted content** (2), **Work with tech** (1), **Safeguard device** (1), **No-log VPN** (1), **for IPTV** (1), **For work/uni** (1), **Browsing from different locations** (1)

B.2 Other Resources Used

Part of Software/Security Suite (102), **Trusted service provider** (27), **Prior Experience** (15), **Reviews:** Consumer Reports (13), Offered the service for free (13), Introduced as part of my job (8), thatoneprivacyguy (7), **Own testing** (7), Trying the trial option (7), **Word of Mouth:** from technical staff (6), Recommended by service (2), Computer Clubs (2), Colleague (2), University (1), Indiegogo (1), Meetings (1), Geek Squad (1), Friend/Family (2), Computer services company (1) **Trust:** the Mac App Store (5), the Google play store (1), Tech YouTubers (1), Leo Laporte (1), privacytools.io (1), Bloggers, Apple News+ (1), Expert reviewer (1), **No choice in VPN provided** (4), **Company Announcements** (5), **Reviews:** Reviewer Kim Komando (3), Specific to MacOS (3), PCMag (3), News Articles (3), Recommendation Sites (2), Trusted sources (2), NYT (2), Local tech advisor (2), testmy.net (1), ZDNet (1), Recommendation on YouTube (1), Print Magazines (1) **Advertising:** Ads on trusted podcast (3), Promos (3), in specific site (1) **No Research** (3), **Provider's website** (1)

B.3 Feeling of Safety without VPN: Limited expertise users

Worry: hacking (13), tracking (6), exposed personal details including IP and location (10), unsafe in today's world (4), bad actors (3), dark web/net (2), ISP access data (1), open to exploitation (1), malware (1), less protected (1), happened to others (1), breaches (1), ID theft (1), prevention (1), fear threats and financial data (1), **Confusion:** don't understand (10), what does ISP do with data (1), service stopped working (2), **Safety:** no prior issues (5), I'm careful (3), I have antivirus (2), using VPN makes me safer (2), use trusted provider (3), my device is safe (1), I am trusting (1), added protection (1), **Scenario:** only unsafe in public networks (4), I use it if I have it (1), no reason (1), HTTPS isn't always available (1), **No worry:** I feel okay (2), **Understanding:** with research (1), its supposed to hide me (1), anyone can see my traffic (1), **Specific needs** (1), **Needs:** trade-offs (1), harder for hackers (1), make me safer (1).

B.4 High-expertise users response to mental model

DNS Provider (7), **Site:** if entered personal info (6), has access to cookies (4), might know (5), if insecure protocol (1), **VPN Provider:** logging (5), depending on service (4), surely knows (3), alone cannot hide you (2), audit trail (1), **ISP knows if DNS leaked** (4), **Other actors:** example site's partners (3), large companies like Google/Facebook (3), Browser (2), search engine (2), ad networks (2), IDSs (1), badly implemented tech (1), **Threat actors:** tracking (3), government agencies (2), third party cookies (2), browser fingerprinters (2), hackers (1), **Idk:** nobody if no logging (2), any hop in between VPN and site (1).

C Survey Instrument and Interview Questionnaire

We release the user survey instrument and VPN provider interview questionnaire used in this study in our arXiv preprint version [37, Appendix C, D] and on our website [40,41].