Information Research - Vol. 31 No. iConf (2026)

Amulets, boxers, and cybersecurity: influencer metaphors in VPN advertising and the communication of cybersecurity risk

Yu-Wen Huang, Yu-Jie Lin, Kai-Hsiang Chou, An-Jie Li, Tsai-Hsuan Hsieh, Li-Fei Kung, and Wei Jeng

DOI: https://doi.org/10.47989/ir31iConf64145

Abstract

Introduction. VPNs primarily establish encrypted tunnels for network traffic and mask users' IP addresses; they do not inherently block malware, phishing, or denial-of-service attacks. Yet advertising content often portray VPNs as comprehensive security solutions, blurring public understandings of what these tools can and cannot do. This exploratory study examines how YouTube influencers in Taiwan frame VPNs in sponsored content and communicate cybersecurity risks.

Method. This study employs qualitative content analysis of sponsored VPN videos published by top Taiwanese YouTube influencers. Data include 20 videos and 99 close-watching observation notes documenting narrative strategies in promotional contexts.

Analysis. Using open coding and iterative qualitative analysis, we examine influencer narratives across four analytical dimensions informed by risk communication frameworks: technical accuracy, risk translation, metaphorical strategies, and audience engagement.

Results. This study identifies several recurring strategies in VPN-sponsored videos, including simplifying technical descriptions, reframing risks as consumer disadvantages, and using metaphors or everyday scenarios to make VPNs more relatable. While these tactics increased accessibility and entertainment value, they also risked obscuring the conditional and limited scope of VPNs.

Conclusion. The findings show how vernacular metaphors and promotional framings shape public understandings of digital security, extending cybersecurity communication research beyond news or policy discourse.

Introduction

Cybersecurity is fundamental to modern society, yet communicating its technical knowledge remains challenging. Rapid technological change has increased complexity, while regulatory and cultural developments lag behind, widening the gap between experts and the public. (Fortinet, 2022; Goldhammer, 2015). Informal learning has long been recognized as a primary way in which adults make sense of the world (Malcolm et al., 2003; Ollis, 2011). In the cybersecurity domain, empirical studies show that most users rely on friends, media, or online searches for advice and knowledge rather than official authorities, and these channels profoundly shape users' mental models and behavioral patterns (Redmiles et al., 2016; Wash, 2010). Since the 2000s, influencers have become pivotal in science communication, with YouTube as a leading streaming platform. Unlike traditional authorities, influencers cultivate trust through relatability and authenticity, translating technical concepts into everyday language and experiences. Recognizing this influence, companies increasingly collaborate with or sponsor influencers, making them central to brand promotion (Álvarez-Monzoncillo, 2023; Yuan et al., 2020).

Poorly designed cybersecurity messages can cause the public to underestimate risks or adopt unsafe behaviors, threatening the broader security ecosystem (Bada et al., 2015; Nurse et al., 2011). Ensuring that credible and impactful knowledge is communicated through digital channels has thus become a critical challenge for both research and practice.

Within this context, our study examines how influencers convey cybersecurity knowledge through digital narratives, focusing on Virtual Private Networks (VPNs) as a representative case. Although widely recognized by consumers, VPNs are technically abstract and often misunderstood. The industry's rapid growth has fuelled extensive collaborations with YouTube influencers, making VPN advertisements one of the most visible cybersecurity narratives for the public. VPN advertisements play a dual role: they implicitly educate the public about privacy and online threats while simultaneously employing commercial rhetoric that simplifies or exaggerates, highlighting the tension between marketing imperatives and educational value. This tension makes VPN advertisements an ideal site to study the framing of cybersecurity knowledge in public discourse.

Specifically, this study addresses the following research questions:

RQ1. What strategies do digital content creators employ when presenting VPNs as cybersecurity tools?

RQ2. How might these strategies frame cybersecurity risks and shape viewers' potential understandings of VPNs and their protective value?

To explore these questions, we focus on Taiwan as a case study. Taiwan is among the most targeted regions for cyberattacks: government agencies reported 697 major incidents in 2023 (Ministry of Digital Affairs, 2024), and in 2024 daily attacks averaged 2.4 million (Reuters, 2025). Amid these persistent pressures, enhancing public cybersecurity awareness has become a national security priority. In this environment, YouTube has particular salience, reaching 88.5% of users in Taiwan—a level comparable to several Western countries (Kemp, 2024). Moreover, as Mandarin is the world's second most used online language, representing about one-fifth of global internet users (Internet World Stats, 2020), the Taiwanese context also speaks to the broader significance of the Mandarin-speaking market. Situated between official campaigns and the public, YouTube influencers warrant closer examination for their potential to act as either facilitators or barriers in the dissemination of cybersecurity knowledge.

This study employs qualitative content analysis and interdisciplinary theoretical perspectives to examine VPN advertisement videos on Taiwanese YouTube channels, focusing on their narrative frames, persuasive strategies, and how cybersecurity risks are rhetorically constructed. The findings aim to enrich the literature on cybersecurity communication by situating the analysis

within a Mandarin-speaking context and highlighting audience-specific dynamics, while also providing empirically grounded insights into communicative practices relevant to cybersecurity science communication, thereby advancing understanding of the broader communication ecosystem and contributing to discussions on public cybersecurity awareness.

Related works

In this section, we review empirical studies of VPN advertisements and theoretical models of risk communication, both of which inform our study.

VPN advertising studies

Akgul et al. (2022, 2024) have conducted the only systematic investigations of VPN advertising on YouTube to date. Their first large-scale study (2022) identified 243 promotional videos in a random sample and extrapolated to roughly 17,000 videos with over 4.4 billion views, underscoring the ubiquity of this practice. Content analysis revealed common strategies—security assurances, technical claims, threat-based appeals, and promises of cross-regional access—alongside frequent exaggerations and misleading statements that risk distorting users' mental models of cybersecurity. In a follow-up study, Akgul et al. (2024) examined the cognitive impact of such ads using the complete YouTube viewing histories of 217 participants. Combining machine learning–based exposure estimates with surveys of security and privacy mental models, they found that ad exposure was associated with brand familiarity and heightened threat perceptions, but not with understanding of technical functions. These findings suggest that VPN advertisements influence emotions and risk awareness more than technical comprehension.

Our study extends this line of research by shifting attention from large-scale quantitative analysis to the rhetorical construction of cybersecurity content through theory-driven qualitative close reading. By focusing on the Mandarin-speaking YouTube ecosystem, this study complements prior English-language research and contributes a cross-cultural perspective to understanding cybersecurity communication.

Frameworks for risk communication

While prior work has quantified the scale and cognitive impact of VPN advertisements, our study extends this line of inquiry by applying risk communication frameworks to examine the rhetorical construction of cybersecurity knowledge in these ads.

In the field of information security and privacy, Das et al. (2022) introduced the Security & Privacy Acceptance Framework (SPAF), which identifies three barriers to adopting protective behaviors: limited awareness of threats and responses, unwillingness to act due to perceived costs, and lack of ability stemming from insufficient knowledge and skills. The framework explains why expert recommendations often fail to translate into everyday practice. In this study, SPAF provides a key theoretical lens for evaluating whether influencers' narratives raise risk awareness, strengthen willingness to act, and address gaps in ability, thereby clarifying both the educational potential and the limitations of such content.

Beyond SPAF, this study also draws on Fischhoff's (1995) Seven Stages of Risk Communication, a framework from public health that outlines key tasks for effective risk communication, the seven stages and their definitions are summarized in Table 1.

Stage	Name	Definition
1	Get the numbers right	Communicators assume their job is done once they have calculated accurate quantitative risk estimates.
2	Tell them the numbers	Communicators believe that simply handing over the numbers in raw form suffices, even though they may not be understood.
3	Explain what the numbers mean	Communicators try to interpret and explain the numbers so that recipients can understand their significance.
4	Show them they've accepted similar risks	Communicators compare the unfamiliar risk with familiar ones, encouraging consistency in acceptance.
5	Show it's a good deal for them	Communicators emphasize the benefits and trade-offs, presenting the risk as fair or advantageous.
6	Treat them nicely	Communicators recognize the importance of respectful delivery and building trust beyond just accurate content.
7	Make them partners	Communicators involve the public as active participants in risk management and decision-making.

Table1. Summary of Fischhoff' s (1995) seven stages of risk communication

We adapted Fischhoff's model from a sequential process into a set of qualitative coding dimensions that capture narrative strategies in VPN advertisements. Compared to SPAF's broader structure, the seven-stage model offers greater operational detail, enabling closer analysis of how information is constructed and communicated. Although originally developed in public health, it has been widely adapted across disciplines, and here we apply it to cybersecurity risk communication to assess the educational value and limitations of influencer narratives.

Method

This exploratory study employs qualitative content analysis to examine how YouTubers in Taiwan promote virtual private networks (VPNs) through sponsored content. The research process is divided into two stages: data collection and data analysis.

Data collection and selection

To systematically identify relevant content and reduce researcher bias, we developed a three-stage human–AI collaborative workflow that leverages generative AI tools (see Figure 1). This workflow allowed us to move beyond researcher familiarity, ensuring broader coverage and greater consistency in detecting VPN advertisements.

To ground our subsequent model training and evaluation in reliable empirical evidence, we first constructed a small manually annotated dataset. 10 channels were randomly selected from the top 100 Taiwanese YouTube influencers, covering a 180-day window from June to December 2024 and focusing on videos produced in Mandarin and based in Taiwanese audience. All 150 videos published during this period were annotated to determine (1) whether they contained any sponsorships and (2) whether the sponsorship was VPN-specific. This manually annotated dataset served as the basis for subsequent model training and evaluation.

Building on this dataset, we then applied the three-stage workflow: OpenAI Whisper was used for automatic speech-to-text transcription, the Gemini API for prompt-based retrieval to flag segments likely containing VPN sponsorships, and finally human evaluation to confirm ad presence and assess detection accuracy. This process reduced manual workload while maintaining high

precision. A detailed methodological account is provided in a companion study available via our OSF repository. (https://osf.io/9pe2z/overview?view_only=21ba995772b84a588c115919fa5e258d).

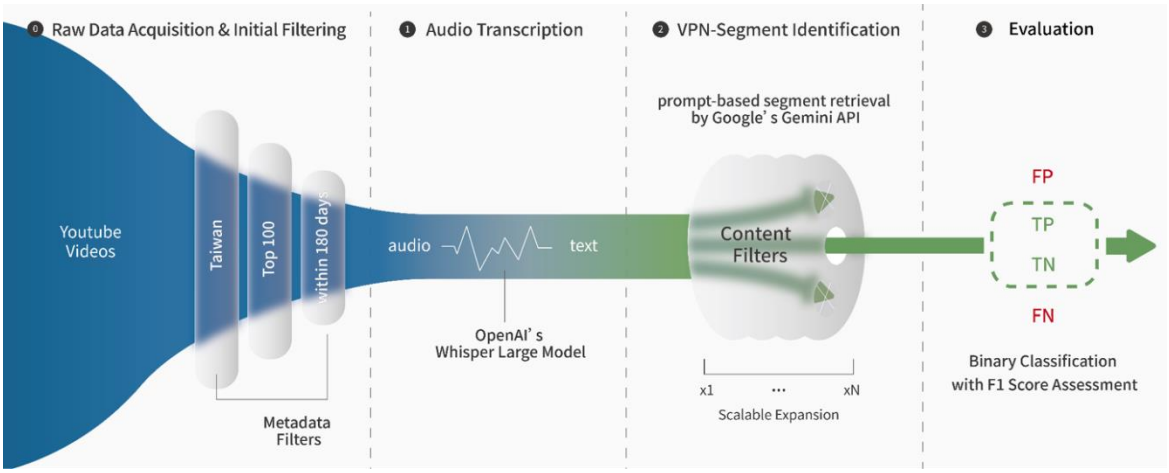

Figure 1. Human–AI collaborative workflow for data collection (adapted from Hsieh et al, 2025)

We compiled a dataset of Taiwan YouTube videos containing VPN sponsorship segments, which served as the basis for sampling. As this study emphasizes qualitative analysis and close reading, we ultimately selected 20 VPN sponsorship clips for in-depth examination. The sample size was determined by analytic considerations rather than representativeness the selected clips allowed for detailed examination of narrative and rhetorical strategies while reaching analytic saturation, as recurring patterns became evident across channels and video formats. The remaining videos are retained for future analysis.

All videos analyzed in this study were published in 2024, with VPN sponsorship segments typically ranging from 50 to 150 seconds. The sample focused primarily on lifestyle and related categories, where commercial sponsorships are particularly prevalent. To ensure representativeness, we prioritized channels consistently ranked among Taiwan's top 50 YouTubers in 2024–2025, while also incorporating diversity in content themes, creator genders, and VPN providers. For within-channel comparison, multiple videos from the same influencers were included.

To ensure transparency, the dataset includes a full metadata list and corresponding video IDs, allowing readers to trace exactly which videos formed the analysis basis of our study. Meanwhile, to protect the identities of individual influencers, we deliberately withheld the mapping between IDs and exact videos. This approach ensures our study's reproducibility and replicability while preventing unnecessary identification or reputational concerns for specific influencers. The anonymized dataset is also available at our OSF repository mentioned above.

Data analysis: qualitative content analysis

This study employed open coding as the primary analytic method. Open coding, commonly used in qualitative research, emphasizes breaking down, labeling, and conceptualizing data without imposing predefined categories, thereby capturing emergent themes and patterns (Strauss & Corbin, 1990). In this study, seven researchers participated in the coding process, using a pre-designed open-coding worksheet to record observations of each video. For each video, we created a Word file that served as the worksheet, which contained two sections. In the first section, coders documented the video's metadata (e.g., video link, channel name, and start–end times of the coding session) and freely noted any observations while watching the clip. In the second section, coders responded to a set of guiding prompts that reflected our analytic framework.

This process resulted in a corpus of 99 coding documents across 20 analysed videos, with each annotation taking roughly 20–30 minutes on average. Each of the seven coders reviewed approximately 14–15 videos. Multiple coders' meetings were held throughout the process to iteratively refine the worksheet and align interpretations. During these meetings, discrepancies in coding were discussed and resolved through consensus, and the coding focus was adjusted accordingly.

In the initial coding phase, as noted in the literature review, we drew on the core principles of Das et al.'s (2022) security & privacy acceptance framework (SPAF) and operationalized Fischhoff's (1995) seven stages of risk communication by adapting its seven tasks into coding dimensions. However, after the first round of coding it became clear that most VPN advertisement videos, due to their brevity and commercial orientation, did not systematically address every element of risk communication. To better capture the rhetorical and communicative patterns that actually emerged in the material, we consolidated and refined the seven stages into four analytic dimensions:

Technical accuracy. Consolidating Fischhoff's '*get the numbers right*' and '*tell them the numbers*,' this dimension evaluates whether influencers explain VPN technologies and assesses the accuracy or potential errors in their claims.

Risk translation. Derived from '*show it's a good deal for them*,' this dimension examines how influencers frame or reframe cybersecurity risks in their promotional narratives, focusing on whether such risks are downplayed, shifted into other domains, or reconstructed as alternative concerns.

Narrative and metaphor strategies. Adapted primarily from '*explain what the numbers mean*,' and informed by elements of '*show them they've accepted similar risks*,' this dimension highlights the rhetorical techniques used to frame VPNs. It captures how influencers deploy metaphors, analogies, and narrative devices to make complex technologies more accessible, as well as the limits of such simplifications.

Audience engagement. Consolidating '*show it's a good deal for them*,' '*treat them nicely*,' and '*make them partners*,' this dimension considers how influencers' rhetorical strategies are designed to engage audiences—through persuasive appeals, interactive framings, and calls to action—and how these strategies may shape attitudes and behaviors toward cybersecurity threats and protective practices.

Results

The following section details the results of the open coding analysis, highlighting four dimensions that reveal the persuasive strategies and communicative tensions in VPN advertisements.

Technical accuracy: what VPNs can and cannot do?

Our analysis indicates that VPN promotional videos generally lacked clear explanation of what VPNs actually can and cannot do. VPNs primarily protect network traffic by establishing encrypted tunnels and concealing the original source and destination of packets from intermediate observers through encapsulation and encryption (NIST, 2020); they do not inherently block malware, phishing, or denial-of-service attacks. Yet most videos relied on simplified functional claims that encouraged the impression that '*using a VPN equals safety*,' while offering little guidance for understanding the actual scope and limits of VPN protection. Although not necessarily factually incorrect, the absence of contextualization created a latent risk of misleading audiences.

A common strategy was to highlight VPNs through embellished and high-level functional claims, such as '*establish an encrypted channel, protect privacy, block hackers*' (Y03) or '*when I use public Wi-Fi, it helps me avoid the risk of personal data leakage*' (Y09). These narratives emphasized the

connection between VPNs, privacy, and security but they remained at the outcome level rather than explaining why encryption enhances protection or how Wi-Fi may expose users to risks. Without a clear understanding of VPNs' boundaries, influencers may present their capabilities as broader than they are. This does not imply deliberate misrepresentation but reflects uncertainty about technical limits.

Other descriptions were vague or potentially misleading. Examples included claims that VPNs can '*change your IP anytime, anywhere*' (Y14), '*overcome regional restrictions and government censorship*' (Y16), or '*encrypt your identity online, protect your digital footprint… nobody will know what you have seen*' (Y06). Some videos also have stacked jargon or conflated functions, as in '*VPN helps secure your online activities, such as blocking malicious ads, man-in-the-middle attacks, and password attacks*' (Y10). Such statements overstate VPN capabilities and blur the line between core function and add-on services; ad blocking, for instance, is closer to firewall or proxy functions than to VPN itself.

Risk translation: how are threats reframed in VPN ads?

Our analysis examined how cybersecurity risks were presented in VPN advertisements and found that depictions of risk scenarios were generally non-specific. When considering how risks were described, we drew on Akgul et al. (2024), who classify threats into attacks, adversaries, and assets, which guides us to identify that most references in the videos concerned attacks. For example, one video claimed that VPNs can '*block various online threats, such as phishing, password attacks, malicious ads, and malware*' (Y09), typically presented as stacked terms without further detail or contextual elaboration. By contrast, adversaries were described far less frequently and usually in arbitrary catch-all labels, most often as 'hackers' or 'bad actors.' In terms of assets, the most common references were 'personal data', 'IP address' or 'digital footprints.'

Some videos often invoked risk scenarios with strong or alarming language but provided little technical related descriptions or contextual clarification. For instance, '*using a VPN can hide your IP address; if bad actors have your IP, they can track you down*' (Y19) framed IP exposure as a direct risk, though in practice its implications are largely tied to subsequent privacy leakage rather than physical traceability. Similarly, '*protect your data when connecting to an unencrypted network*' (Y15) described in vague rather than precise language yet could appear alarming to non-expert audiences. Other examples include '*no need to fear when using Wi-Fi*' (Y10), which suggested a risk without identifying its source, and '*hiding your real IP reduces the risk of DDoS attacks*' (Y07), which is technically partially correct but also overstates VPNs' protective scope, especially given that DDoS attacks are not a common risk for typical users.

Interestingly, we observed that some redefined what counts as a '*risk*' by shifting the language from security breaches to consumer losses. A central device enabling this rhetorical shift is the VPN's IP substitution function: because VPNs allow users to appear as if they are browsing from another location, shopping platforms and streaming services that segment their offerings by national or regional markets can be tricked into offering different prices or content. In this framing, the danger was not stolen passwords or leaked data, but the possibility of paying more than others or missing out on better deals. Rather than emphasizing what might happen if a VPN is not used in terms of security breaches, these narratives positioned the absence of VPNs as a source of economic loss or missed benefits in everyday consumption. Examples included *encouragement to use VPNs for cheaper airline tickets* (Y09), *cross-border shopping discounts* (Y06), or to '*find the best price, save your wallet, and stop being taken advantage of internationally*' (Y08). In these cases, abstract cybersecurity threats were translated into concrete consumer scenarios, where the primary risk became paying more or missing out, rather than facing direct technical vulnerabilities.

Narrative and metaphor strategies: how do influencers make VPNs feel relatable?

In this study, we observed that influencers simplified VPN's concepts and made them relatable to audiences through two main approaches. First, they used metaphors that translated abstract technical functions into familiar images. Second, they embedded VPNs within everyday scenarios, connecting the product to viewers' lived experiences. These approaches were adapted to different channel styles and audience profiles, resulting in distinct promotional framings.

Within the use of metaphors, we identified two subtypes. One subtype directly equated VPNs with concrete objects, such as '*an amulet for the digital world*' (Y05), '*a game item that grants rewards*' (Y07), or '*Harry Potter's invisibility cloak*' (Y19). The other subtype dramatized the experience of using VPNs through narrative scenarios, for example, '*we are now in Dubai, now in Egypt… a virtual world tour!*' (Y16), or '*at XX Travel Agency, we can pay Taiwan's price while enjoying food in New Zealand*' (Y08), which framed cross-regional browsing as travel. Some videos also drew on broader depictions of the internet, such as '*in the torrents of the web, you need to be as strong as a boxer… the VPN helps you cross the current*' (Y08), or '*shake off hackers like a race car breaking speed limits, as VPNs let you bypass restrictions and access global content*' (Y01). In these cases, the boxer or 'shake off' metaphors were especially telling: they implied not only endurance but also a kind of agency—the power to strike back or to escape. Such imagery suggested that VPNs enable users to actively respond to threats, even though in reality their functions are limited to encryption and IP masking.

In terms of everyday applications, the most common scenarios centered on entertainment uses. Examples included *accessing region-restricted videos* (Y03), *obtaining cross-border shopping discounts* (Y06), *playing games across regions* (Y18), or *using dating apps abroad* (Y10). Some influencers also drew on personal experiences to establish relatability, such as '*while gaming, an opponent suddenly messaged me my name, email, and address—it was terrifying!*' (Y20).

Audience engagement: how are viewers persuaded to take action?

In terms of audience engagement, we analyzed how the videos used language and framing to prompt action. First, most adopted assertive and definitive statements that directly linked VPN use with immediate benefits. These slogans, such as '*use a VPN and you can become the king of the internet*' (Y14), were less about literal claims than about stylistic flair—short, catchy phrases that fit influencers' personal styles and encouraged viewers to give the product a try. Within the brief span of a sponsorship segment, such calls created clear causal links and quickly established audience expectations.

Even when promoting the same features about VPN protection, influencers employed different persuasive framings. Some relied on positive appeals, emphasizing social or entertainment value, such as '*now I can cross-region to watch Taiwanese shows and join conversations with friends*' or *encouraging single viewers to try cross-border dating apps* (Y10). Others used negative pressure, implying that non-users would miss out, as in '*only with a VPN can you unlock all content*' (Y12) or '*why don't I get the perks? Because you didn't use a VPN!*' (Y07).

Finally, nearly all videos ended with calls to action, often combining discount codes or limited time offers to create urgency, e.g., 'download now' or 'click the link for exclusive deals.' Only one video encouraged viewers to *'learn more'* rather than purchase directly. Overall, the strategies primarily served marketing purposes, with limited relevance to the communication of cybersecurity knowledge.

Discussion

This study identifies several recurring rhetorical patterns in influencer-sponsored VPN advertisements on Taiwanese YouTube, including the simplification of technical information, the translation of cybersecurity risks into alternative domains, the use of metaphors and narratives,

and the reconstruction of trust. Together, these strategies illustrate how cybersecurity tools and risks are rhetorically framed within influencer advertising contexts. This qualitative content analysis focuses on the rhetorical presentation of cybersecurity tools and risks. Accordingly, the discussion offers analytical interpretation rather than empirical claims about creators' intentions or audience behavior.

Simplification of cybersecurity information

This study shows that Taiwanese YouTube VPN advertisements generally present technical information in highly simplified forms. While this simplification is a pattern what we have observed, its underlying cause is less certain: it may reflect the limited depth of sponsor-provided materials, the time constraints of short-form sponsorship segments, or influencers' own choice not to engage with technical details. This pattern aligns with Akgul et al.'s (2022) analysis of English-language VPN ads on YouTube, which identified broad security assurances, cross-regional convenience, and expanded claims as recurrent strategies. Their subsequent study further found that frequent exposure increased brand familiarity and perceived threat awareness, but did little to improve technical understanding (Akgul et al., 2024).

We further break down this simplification into three types. First, strategic simplification, where influencers intentionally downplay technical detail and emphasize only the most understandable or attractive functions: for example, briefly and quickly display lists of VPN features in a near-instantaneous flash of on-screen text, making those text difficult for viewers to read. Second, resource-based simplification, where influencers working under constraints of time and technical expertise rely more heavily on sponsor-provided material, resulting in generalized claims instead of detailed explanation. Third, spontaneous simplification, in which influencers often improvising in the moment, attempt to lower barriers through colloquial language or everyday metaphors, but without realizing that such ad-hoc phrasing can leave gaps in technical accuracy and risk communication.

In Taiwan, a broader structural factor also plays a role: the ubiquity of VPN advertising across channels has fostered a collective assumption that audiences already know what VPNs are for, further reducing the perceived need to explain technical foundations. As a result, most influencers emphasize entertainment and economic benefits, such as cross-regional streaming or online shopping discounts, rather than the principles and limitations of VPNs. While this approach lowers the threshold of engagement and makes content more appealing, it simultaneously avoids essential background information, leaving knowledge gaps unaddressed.

Viewed through the Security & Privacy Acceptance Framework (SPAF; Das et al., 2022), these ads primarily stimulate motivation (via threats or discounts) and reduce action barriers, while contributing little to knowledge or ability. Consequently, audiences may adopt the simplified belief that '*installing a* VPN *equals complete online safety*,' without recognizing its actual scope and limits. In this case, VPN ads succeed in raising visibility and brand value but fall short in fostering deeper cybersecurity literacy, potentially reinforcing partial or distorted understandings.

Prior research underscores that self-efficacy—the belief in one's ability to take protective actions—is critical for translating threat perception into behavior. Protection motivation theory (PMT) highlights that unless individuals feel capable of executing the recommended response, perceived risks will not lead to protective action (Almansoori et al., 2023). Empirical evidence on security advice adoption shows that overly technical or marketing-heavy messages can undermine users' sense of efficacy, while clear and actionable guidance enhances confidence and uptake (Redmiles et al., 2016). Consistently, Bada et al. (2015) argue that awareness campaigns fail when they rely on fear appeals or assurances without offering concrete '*how-to*' instructions. In VPN advertisements, the lack of efficacy-enhancing information, such as practical usage steps, contextual

recommendations, or clear boundaries, means that even motivated viewers may abandon adoption when facing difficulties or misuse the tool in ways that reinforce misconceptions.

The communicative impact of risk translation and metaphor strategies

Our analysis shows that many influencers reframed cybersecurity threats as economic losses, suggesting that the real danger was *'missing out on a deal.'* While such narratives are rhetorically powerful, they are not entirely accurate: what viewers encounter when paying different prices or losing access abroad is primarily the outcome of contractual arrangements and licensing practices, not a cybersecurity vulnerability. To resonate with audiences' everyday concerns, this narrative strategy recast VPNs from tools for encrypting traffic and masking IP addresses into instruments for unlocking discounts or cheaper fares. VPNs were thus repositioned as consumer products, redirecting attention from *'how to protect oneself from attacks'* to *'how to save money or access more content.'* In the short term, this economic framing likely increases acceptance by appealing to immediate personal benefits. Over the longer term, however, it may risk diluting the seriousness of cybersecurity issues. Viewers may adopt VPNs to secure discounts without recognizing underlying threats, and when security measures are equated with shopping perks, their protective value may be overshadowed.

In terms of rhetorical strategies, influencers often relied on metaphors to reduce technical barriers and make abstract processes more relatable. Metaphors inherently simplify reality, and when overextended, they risk fostering inaccurate mental models. VPNs serve two core functions, encrypted tunnelling, and IP masking, which provides an essential baseline for evaluating metaphorical extensions. For example, some videos described VPNs as a *'battle tool'* or *'universal internet cleaner'* (Y06, Y07, Y17), metaphors that not only suggest active defensive power but also implies a kind of all-purpose protection. Such framing risks overstating VPNs' role: cybersecurity defences are always conditional, shaped by specific threats, contexts, and configurations, whereas VPNs serve limited functions related to traffic confidentiality and concealment rather than universal security, aligning more closely with an *'invisibility cloak'* metaphor than an offensive weapon. Similarly, to market cross-regional functions, videos often invoked metaphors of *'one-click travel'* or *'world tours'* (Y06, Y08, Y16), suggesting constant mobility, whereas VPNs merely route traffic through a chosen server to simulate location. These exaggerated portrayals may obscure the technology's actual scope.

Beyond the use of metaphors that may mislead audiences' mental models of VPNs, our observations also suggest that some influencers' expressions fall within specific metaphorical systems of cybersecurity. In our dataset, one influencer employed two contrasting metaphors: a defensive one, portraying cybersecurity as *'an amulet for the digital world'* (Y05), and an offensive one, drawing on a boxing scenario in which cybersecurity was likened to *'the gloves in the digital arena,' needed to confront the torrents of the internet* (Y02). This juxtaposition resonates with prior research on Chinese-language cybersecurity metaphors, which identified a *'warfare system'* of metaphors (Chen et al., 2025). Within this system, cybersecurity is conceptualized as an ongoing battle, commonly framed through references to combat, weapons, and adversarial roles. While such rhetoric effectively amplifies a sense of threat and urgency, it may also confine public imagination to conflict-oriented framings, thereby overlooking other important dimensions of cybersecurity such as resilience, cooperation, and preventive governance. Consistent with expert evaluations in the same study, warfare metaphors may succeed in capturing attention and heightening tension but are less effective than metaphorical systems such as public health or risk management, which emphasize prevention and capacity-building, in fostering long-term and constructive cybersecurity practices.

As Fischhoff (1995) noted, oversimplified or poorly framed risk communication can confuse audiences and undermine understanding. Our findings echo this concern: when influencers privilege vividness over accuracy, audiences may gain entertainment but lose precision in their

understanding. Ultimately, metaphor operates as a double-edged sword—its value lies in balancing accessibility with fidelity, a challenge that future cybersecurity marketing or awareness promotion must confront.

Trust transfer and constructed assurance

VPN sponsorship videos also highlight the reconstruction of trust. When advertisements claim that '*using a VPN makes you safer*,' they essentially encourage users to shift their trust in the online environment from existing intermediaries to VPN providers. For instance, many videos stress the dangers of public Wi-Fi to emphasize the necessity of VPN use. From a security expert perspective, however, this is a form of trust transfer: users no longer rely on Wi-Fi providers but instead entrust their traffic to VPN companies, believing they will not intercept or misuse the data. Crucially, this transfer often arises not from rational evaluation but from subtle cues and persuasive framing in the advertisements.

Only a small minority of videos provide information directly related to building trust, such as assurances about '*no-log policies*' or third-party security audits—measures that function as rational guarantees of reliability. The vast majority, however, omit such details. influencers rarely discuss the reputation or safeguards of the VPN provider, nor do they remind viewers to critically assess the trustworthiness of the service itself. Whether intentional or not, this omission leaves audiences unaware that adopting a VPN entails transferring their online security to another third-party entity. If viewers simply follow slogans like '*just use it*,' their decisions risk being based more on affective impressions than on informed judgment.

Another notable strategy is the use of fictionalized testimonial stories. While user experiences or expert endorsements are longstanding methods of securing consumer trust, some influencers dramatize or fabricate scenarios, such as claiming to have received personal data threats while gaming (Y20), to underscore VPN necessity. Whether such stories are genuine or rhetorical is difficult to verify, raising the possibility that viewers may conflate commercial persuasion with credible risk information. As Stubb et al. (2019) note, influencer marketing depends heavily on credibility and authenticity; once audiences perceive content as overly commercialized or insincere, trust in both influencer and brand may erode.

This concern is amplified in the era of generative AI (GAI), where text, images, and video materials can be produced rapidly, cheaply, and at scale, enabling influencers to construct seemingly authentic '*testimonials*.' Such fabricated assurances, however, often lack accountability. If audiences accept them uncritically, they may further undermine accurate understanding of cybersecurity risks.

Distributing responsibility: providers, platforms, policymakers, and creators

This study underscores the need for clearer standards in influencer VPN advertising, where responsibility is distributed across four levels: providers, platforms, policymakers, and creators.

Providers and Platforms. Our analysis, together with prior research on English-language content (Akgul et al., 2024), indicates that influencer VPN advertisements often contain vague, misleading, or overstated security claims. Analytically, this observation draws attention to the role of providers in shaping the informational materials that circulate through influencer campaigns. Platforms such as YouTube constitute an additional layer, as platform policies and affordance structure how sponsorship disclosure, visual elements, and metaphors appear within advertising content.

Policymakers. Regulators form the third layer of accountability. These government interventions show that regulators already recognize the risks posed by misleading digital advertising. Similar oversight could be extended to VPN advertisements, where exaggerated security claims risk distorting public perceptions of cybersecurity. These government interventions demonstrate that when risk perception is involved, exaggerated claims are restricted.

Creators. While creators are not positioned as cybersecurity experts in this analysis, they emerge as key mediators of trust within sponsored content. From a rhetorical perspective, creators' choices reflect negotiations among provider materials, platform norms, and audience expectations. This multi-layered view frames responsibility for cybersecurity risk communication as structurally distributed across actors, aligning the discussion with the interpretive scope of the analysis rather than normative or policy-oriented claims.

Conclusion

This exploratory study examined how Taiwanese YouTube influencers frame VPN advertisements and communicate cybersecurity risks. Compared with prior research that has largely focused on formal discourses such as news reporting or policy documents, our analysis concentrated on what influencers actually say and how they say it in everyday promotional contexts. We found that, likely under conditions of limited time and resources, influencers often lacked a clear grasp of the boundary between what VPNs can and cannot do. To add entertainment value and build intimacy with audiences, they simplified technical descriptions, reframed risks beyond cybersecurity, and deployed metaphors and everyday scenarios to lower barriers of understanding. These strategies enhanced accessibility and attention to a cybersecurity tool but also risked obscuring the conditional and limited nature of VPNs. For instance, reframing VPNs as tools for saving money or dramatizing protection through metaphors of warfare redirected attention from preventive practices toward consumption and conflict. Although these observations are drawn from a specific sample, the recurring rhetorical patterns may point to broader cultural narratives surrounding cybersecurity. Future research could extend this analysis through larger datasets or comparative cross-cultural approaches.

As influencer advertising has become one of the most prominent ways the public encounters VPNs in Taiwan, it plays a central role in shaping cultural meanings of cybersecurity and deserves closer scholarly and policy attention. In future research, we aim to extend this work by comparing how narratives vary across sponsors, analyzing the influence of sponsor-provided materials, and systematically examining non-verbal elements such as visual design and imagery. Interviews with both influencers and audiences could further clarify how these narratives are produced, interpreted, and circulated.

From amulets that promise magical protection to boxer metaphors that dramatize combat with unseen adversaries, influencer narratives demonstrate the multiple ways cybersecurity is imagined. Unlike prior studies that mainly examined how news or policy texts frame security, our analysis recovers metaphors directly from influencers' speech—expressions close to audiences' everyday language. These simple metaphors reveal both opportunities for engagement and risks of misrepresentation: they attract attention but can also obscure the conditional and limited scope of VPNs. Future research can extend this focus by tracing how metaphors not only reflect but also shape public awareness and the communication of cybersecurity policy. Risk communication is always both imaginative and practical—anchored in cultural symbols as much as in technical facts.

Acknowledgements

This work was financially supported by the National Science and Technology Council (NSTC) in Taiwan, under NSTC 113-2627-M-002-001-, and the Center for Research in Econometric Theory and Applications (Grant no. 114L900202 & 114L910509) which is under the Featured Areas Research Center Program by Higher Education Sprout Project of Ministry of Education (MOE) in Taiwan.

About the authors

Yu-Wen Huang is a doctoral student in Library and Information Science at National Taiwan University, Taiwan. She holds a bachelor's degree in Library and Information Science from National Taiwan University, with a minor in design from National Taiwan University of Science and Technology. Her research interests include visual rhetoric, information design, and science communication, with a particular focus on the communication of cybersecurity, privacy, and risk. She can be contacted at f11126004@ntu.edu.tw

Yu-Jie Lin is a master's student in Library and Information Science at National Taiwan University, Taiwan, where she also received her B.A. Her research focuses on misinformation studies through the lens of humour and ambiguity. She can be contacted at r13126008@ntu.edu.tw

Kai-Hsiang Chou is a master's student in Computer Science at National Taiwan University, Taiwan, where he also received his B.S. in Computer Science. His research interests include human-centered security and technology policy. He can be contacted at b07705022@csie.ntu.edu.tw

An-Jie Li is a master's student in Cyber Security at ETH Zurich, Switzerland. He received his B.S. in Computer Science from National Taiwan University. His research interests include applied cryptography, privacy, and human-centered security. He can be contacted at lianli@student.ethz.ch

Tsai-Hsuan Hsieh is a master's student in Cybersecurity at Georgia Institute of Technology, USA. Hsieh received her bachelor's degree in National Taiwan University, and her research interests include cybersecurity policy and user privacy. She can be contacted via email thsieh61@gatech.edu

Li-Fei Kung is a researcher in Taiwan's National Institutes of Cyber Security. Her research interest focus on cybersecurity tabletop exercises and knowledge interpretation between expert and non-expert. She can be contacted at lfkung@nics.nat.gov.tw

Wei-Jeng is an Associate Professor in the Department of Library and Information Science at National Taiwan University. She received her PhD from the iSchool at the University of Pittsburgh. Her research focuses on open science and data practices, and more recently on cybersecurity communication, including metaphors, visual narratives, and public-facing cybersecurity awareness practices. She also has experience in public-sector cybersecurity governance and currently serves as Deputy Director-General at the Administration for Cyber Security, Ministry of Digital Affairs, Taiwan. She can be contacted at wjeng@ntu.edu.tw

References

Akgul, O., Roberts, R., Namara, M., Levin, D., & Mazurek, M. L. (2022). Investigating influencer VPN ads on YouTube. 2022 IEEE Symposium on Security and Privacy (SP), 876–892. https://doi.org/10.1109/SP46214.2022.9833633

Akgul, O., Roberts, R., Shroyer, E., Levin, D., & Mazurek, M. L. (2024). As advertised? Understanding the impact of influencer VPN ads. arXiv. https://doi.org/10.48550/arXiv.2406.13017

Almansoori, A., Al-Emran, M., & Shaalan, K. (2023). Exploring the frontiers of cybersecurity behavior: A systematic review of studies and theories. Applied Sciences, 13(9), 5700. https://doi.org/10.3390/app13095700

Álvarez-Monzoncillo, J. M. (Ed.). (2023). The dynamics of influencer marketing: A multidisciplinary approach. Routledge. https://doi.org/10.4324/9781003134176

Bada, S., Sasse, M. A., & Nurse, J. R. C. (2015, February). Cyber security awareness campaigns: Why do they fail to change behaviour? In International Conference on Cyber Security for Sustainable Society (CSSS 2015).

Chen, W.-N., Huang, P.-P., Huang, Y.-W., Lin, Y.-J., & Jeng, W. (2025). Decoding cybersecurity for civic defense: Exploring real-world cybersecurity metaphors from digital frontlines [Conference poster]. iConference 2025 Proceedings. iSchools. https://hdl.handle.net/2142/126203

Das, S., Faklaris, C., Hong, J. I., & Dabbish, L. A. (2022). The security & privacy acceptance framework (SPAF). Foundations and Trends® in Privacy and Security, 5(1–2), 1–143.

Fischhoff, B. (1995). Risk perception and communication unplugged: Twenty years of process. Risk Analysis, 15(2), 137–145. https://doi.org/10.1111/j.1539-6924.1995.tb00308.x

Goldhammer, J. (2015, March 16). Cybersecurity metaphors: How they shape national cyber policy, technical research and the future of US national security. Lawrence Livermore National Laboratory, California, U.S. https://www.youtube.com/watch?v=yrGLGqB6y9U

Hsieh, T.-H., Lin, Y.-J., Huang, Y.-W., Chou, K.-H., Li, A.-J., Kung, L.-F. and Jeng, W. (2025), Conversations Reimagined: Human-AI Collaboration in Analyzing How Creators Explain Security and Privacy Tools. Proceedings of the Association for Information Science and Technology, 62: 1478-1480. https://doi.org/10.1002/pra2.1440

Internet World Stats. (2020, March 31). Top ten Internet languages in the world – Internet statistics (Number of Internet users by language). Retrieved August 31, 2025, from https://www.internetworldstats.com/stats7.htm

Malcolm, J., Hodkinson, P., & Colley, H. (2003). The interrelationships between informal and formal learning. Journal of Workplace Learning. https://doi.org/10.1108/13665620310504783

Ministry of Digital Affairs. (2024). National Cyber Security Status Report 2023. Ministry of Digital Affairs, Taiwan. https://moda.gov.tw

Nurse, J. R. C., Creese, S., Goldsmith, M., & Lamberts, K. (2011, September). Trustworthy and effective communication of cybersecurity risks: A review. In 2011 International Conference on Availability, Reliability and Security (pp. 19–26). IEEE. https://doi.org/10.1109/ARES.2011.11

Ollis, T. (2011). Learning in social action: The informal and social learning dimensions of circumstantial and lifelong activists. Australian Journal of Adult Learning, 51(2), 248–268.

Redmiles, E. M., Kross, S., & Mazurek, M. L. (2016). How I learned to be secure: A census-representative survey of security advice sources and behavior. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16) (pp. 666–677). Association for Computing Machinery. https://doi.org/10.1145/2976749.2978307

Reuters. (2025, January 6). Chinese cyberattacks on Taiwan government averaged 2.4 mln a day in 2024, report says. Reuters. https://www.reuters.com/technology/cybersecurity/chinese-cyberattacks-taiwan-government-averaged-24-mln-day-2024-report-says-2025-01-06

Strauss, A., & Corbin, J. M. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Sage Publications, Inc.

Wash, R. (2010). Folk models of home computer security. In Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10) (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/1837110.1837125

Yuan, Y., Du, Y., Yao, F., Zhu, Y., & Lv, Z. (2020). Influencers in intercultural communication: Analysis of opportunities and risks. Frontiers in Economics and Management, 1(12), 52–57.