CyberGreen

A CYBER BELIEF MODEL

Technical Report [2023-01]

The Health Belief Model (HBM) is a longstanding family of models to explain why people don't act on health advice. Adaptation of the HBM to cybersecurity provides insight and explanations as to why cybersecurity advice is not consistently acted upon. This technical report presents motivation, a first Cyber Belief Model, results of an interview study and an interview coding scheme. The interview study with 9 participants analyzed enterprise responses to the log4shell crisis, and indicates that awareness and prompts to action are well addressed, but barriers to action remain. It may be that the overall cybersecurity investment could be rebalanced in ways that increase the rate of taking preventative actions. This Cyber Belief Model may be a useful way to identify and address inhibitors to action, leading to improved security globally.

CyberGreen Institute
January 2023

1. INTRODUCTION

CyberGreen is a non-profit organization focused on improving the state of cybersecurity public health, including defining a new science of cyber public health. This report documents our work in bringing an important model from public health to cybersecurity. Early in our exploration of public health, we discussed the question of "why is DNSSec not more widely deployed?" The intuitive answers included that deploying DNSSec carries risk to the organization, while benefits accrue mainly to customers. This is discussed further in "Cybersecurity background".

Those discussions of DNSSec led to a deeper conversation about why people do things for security. Our investigations of public health led to the discovery of a Health Belief Model ("HBM") which seemed relevant. We adapted the HBM to a version we call the Cyber Belief Model ("CBM").

This report presents the CBM, and presents an interview study that tested if the CBM helped analysis of discussions of a recent high-profile incident, and the results of that study.

The CBM works as an analytic framework to inform the likelihood that a person or organization will follow cybersecurity advice to engage in preventative action. This model was used to analyze a set of interviews about the Log4shell crisis.[1] The interviews were focused around the question of "what happened at your organization to respond to the crisis?" Between 66 and 85% of interview time was spent in ways that aligned with the model, giving us confidence that the model has explanatory and analytic power.

We believe that the CBM presents an important development in how we can motivate security improvements. As a society, we spend a great deal of energy alerting people to the need for action. But people hear a lot about the need for action, and don't rely on official sources. We spend less energy helping to prioritize between the various actions, helping overcome questions about the susceptibility to a problem, and addressing other barriers to action. Twitter came up in almost every interview and formal sources like CERTs, ISACs or ISAOs were almost never mentioned. However, quality of information from Twitter was a problem as people used an intentionally hobbled system, with its 280-character limit, to disseminate information about a complex problem with many variants. Even those interviewees who had threat intelligence staff and subscriptions made use of Twitter in the hopes of remaining up to date.

Our interviews also exposed the value of egress filters as a control. Several participants explicitly mentioned seeing partially successful attacks which broke because the exploit code needed internet access to continue. The interviews also cast light on the cost of patching and inefficiency in industry information flows.

Section 2 presents and discusses the model. Section 3 presents the interview study and its results. Section 4 presents an overall analysis of the work to date and suggests follow-ups. Specifics of the interviews and coding are included as appendices.

[1] For an overview of the Log4shell crisis, see the Cyber Safety Review Board Review (CSRB), 2022.

BACKGROUND

CYBERSECURITY BACKGROUND

This work is grounded in prior work around DNS and DNSSec. DNS is a building block protocol for the internet and is insecure by design. DNSSec is a protocol which prevents spoofing of or tampering with DNS responses. These improvements help ensure that those trying to reach an organization can do so with higher assurance. They don't directly address a risk of breach in, for example, the way a web application firewall might prevent SQL injection or other data extraction attacks. As such, the work to prevent a problem falls on the organization, but the costs of a problem probably accrue to their customers. This problem may lead to prioritizing DNSSec investments lower than other security investments. The prioritization problem is accompanied by a risk, namely that a DNSSec operational error may make it temporarily impossible for clients to reach an organization. This may result in leaders being even less willing to make an investment.

The result for society is that many organizations do not deploy DNSSec, and we are less secure than we could be.

The reader may be thinking "let's change that DNSSec protocol." Unfortunately, the risk of unreachability is an inherent property of name resolution protocol security. If a name resolution protocol can fail and clients can be convinced to use insecure information, an attacker may be able to invoke those conditions. Therefore, security engineers must make a design choice about the insecurity properties or failure conditions.

We initially considered these factors to be "inhibitors" of deployment, and have been investigating them in the hopes of learning to overcome them.

HEALTH BELIEF MODEL BACKGROUND

The CBM is derived from a version of the Health Belief Model presented by Rosenstock. Rosenstock describes the origin of the model as work by a set of social psychologists working for the Public Health Service, and looking to create a theory, rather than "merely solving practical problems one at a time" (Rosenstock 1974). This results in a qualitative model with varying specificity. Rosenstock's model is shown in Figure 1, and is referred to as "the" HBM.[1]

[1] There are a variety of Health Belief Models in use. For this work we are ignoring the variations.

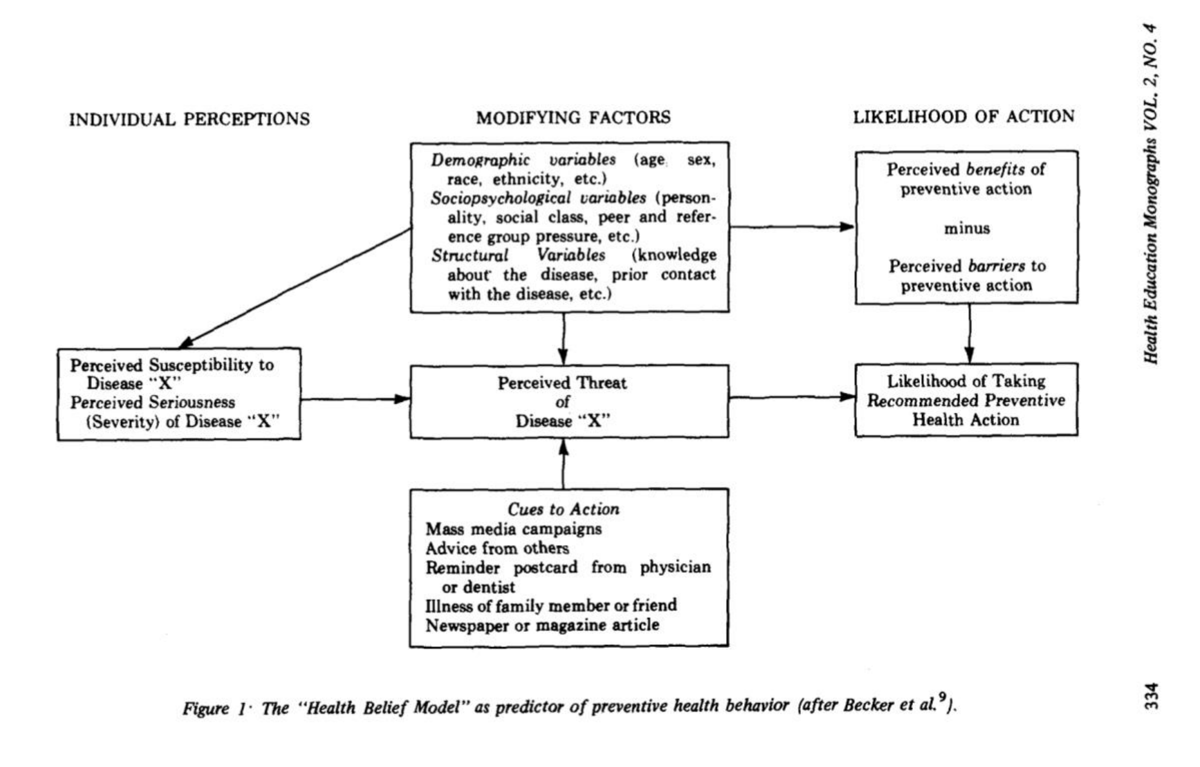

INDIVIDUAL PERCEPTIONS
MODIFYING FACTORS
LIKELIHOOD OF ACTION
Demographic variables (age sex, race, ethnicity, etc.)
Sociopsychological variables (personality, social class, peer and reference group pressure, etc.)
Structural Variables (knowledge about the disease, prior contact with the disease, etc.)
Perceived benefits of preventive action
minus
Perceived barriers to preventive action
Perceived Susceptibility to Disease "X"
Perceived Seriousness (Severity) of Disease "X"
Perceived Threat of Disease "X"
Likelihood of Taking Recommended Preventive Health Action
Cues to Action
Mass media campaigns
Advice from others
Reminder postcard from physician or dentist
Illness of family member or friend
Newspaper or magazine article
Health Education Monographs VOL. 2, NO. 4
334

Figure 1· The "Health Belief Model" as predictor of preventive health behavior (after Becker et al.[9]).

Figure 1: Rosenstock's Health Belief Model

MOTIVATION

People don't act on security advice, in either their personal lives or as organizations. There are many reasons for this, including the actionability of that advice, the effort needed to take action, or the return on that invested effort. Available advice is a confusing jumble, sometimes within the advice from a single organization, but certainly across the set of organizations we interact with. Try comparing the security advice of your employer, your bank, and others with whom you interact. You will almost certainly see inconsistencies, and possibly even contradictions. Contradictions are most obvious in password requirements (shorter than/longer than; special characters allowed or forbidden), but often show up in areas such as resisting phishing (should I read the URL? Do you ensure that the URLs you send are all readable?)

This failure to take security advice can have direct impacts on them, and it can have deleterious effects on others. For example, if credentials are leaked, then the person's accounts are at risk, and organizations may need to take responsive action. If someone's computer is broken into, its contents may be rummaged, encrypted (as in ransomware), or the computer may be used to send spam.

This mimics the relationship between medicine and public health, where problems impact an individual and may lead to others becoming sick.

It is less obvious than the authors would hope that if we want to change organizational behavior, studying the reasons why organizations act as they do is a key step towards addressing the behavior. We initially considered the reasons that organizations don't deploy DNSSec to be "inhibitors" to action. However, that framing has important limits. For example, it doesn't consider tradeoffs that a firm might make. This limit is shared with the HBM, and possibly with the CBM (see discussion in Section 4).

ORGANIZATION OF THIS REPORT

We created initial versions of the CBM by adapting the HBM as faithfully as reasonably possible. The process and result are described in detail in Section 2. We then created a survey to test it, focusing in on the Log4shell set of vulnerabilities. That survey is described in Section 3 and the results are described and analyzed in Section 4.

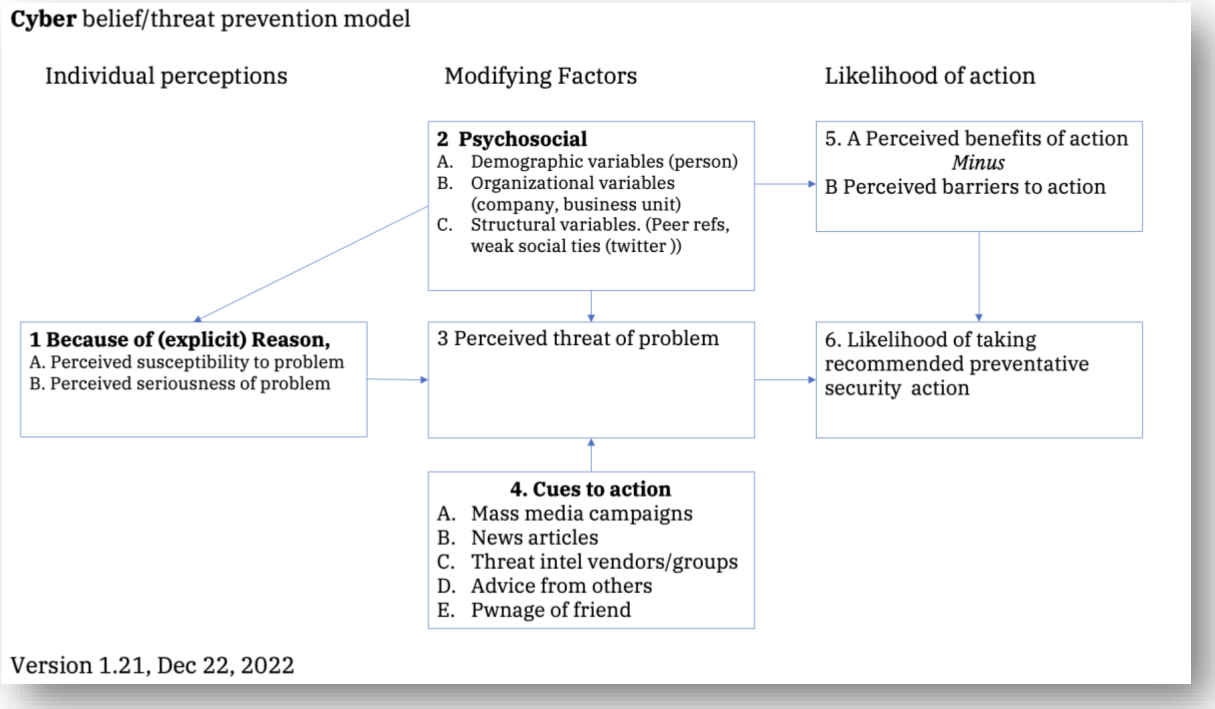
2. A CYBER BELIEF MODEL

This section presents version 1.2 of the Cyber Belief Model. The initial version was tied closely to Rosenstock. It has evolved as a result of both interviews and clarification needed to allow us to code interviews. This section presents the current model and then discusses differences to the HBM.

THE CYBER BELIEF MODEL

As shown in Figure 2, the model consists of three families of information (Individual perceptions; modifying factors; and likelihood of action). We added numbers and letters for ease of discussion and coding. These are further broken into:

1. Individual perceptions and reasons for action regarding a specific problem
    a. Perceived susceptibility
    b. Perceived seriousness
2. Psychosocial factors (modifying)
    a. Demographic variables (about the person)
    b. Organizational variables (about the organization)
    c. Structural variables, such as peer references
3. Perceived threat level
4. Cues to action (not specific to the receiver)
5. Likelihood of action
    a. Modified by perceived benefits and barriers
6. A resultant likelihood of taking recommended preventative actions.



Cyber belief/threat prevention model

Individual perceptions

Modifying Factors

Likelihood of action

2 Psychosocial

A. Demographic variables (person)
B. Organizational variables (company, business unit)
C. Structural variables. (Peer refs, weak social ties (twitter))

5. A Perceived benefits of action
Minus
B Perceived barriers to action

1 Because of (explicit) Reason,
A. Perceived susceptibility to problem
B. Perceived seriousness of problem

3 Perceived threat of problem

6. Likelihood of taking recommended preventative security action

4. Cues to action

A. Mass media campaigns
B. News articles
C. Threat intel vendors/groups
D. Advice from others
E. Pwnage of friend

Version 1.21, Dec 22, 2022

Figure 2: Our Cyber Belief Model

The individual perception and reason for action are fairly straightforward, and close to the HBM. Similarly, box 3, the "perceived threat of problem [X]" comes nearly directly from the HBM, substituting only "problem" for "disease", and box 6 similarly replaced "health" with "security". Box 5 is unchanged.

The psychosocial variables and cues to action were somewhat more complex to adapt both because we're dealing with organizations and because the *sociopsychological* and *structural* components are somewhat jargony. The first barrier was that people get diseases, organizations don't, but they do suffer security problems. We replaced sociopsychological ("personality, class, peer and reference group pressure") with organizational variables. For example, some of our interviewees worked for organizations with threat intelligence groups, and one works for a company which sells threat intelligence. Others worked for major retailers, or had other factors where the nature of the organization influenced the response (discussed in detail in Section 4).

Similarly, the cues to action have changed since 1974. The rise of the internet and social media have an impact on where we might be cued to action. In the 1970s, "advice from others" would be face-to-face communication, and the illness of a family member or friend might be communicated by phone or letter, not broadcast. In 2022, Twitter came up in most interviews as a source of information.

Structural variables (2c) represent the social structure, or social graph, in which people are embedded, while cues to action represent communication intended for many recipients.

DISCUSSION

Like any model, the CBM involves tradeoffs, and we discuss a set of these including:

- Medical advice is often more precise than cybersecurity advice
- Cybersecurity threats can often be addressed by a variety of defenses
- The CBM and HBM are focused on preventative action
- Neither model is currently quantitative
- The "cues to action" area reveals several tensions

PRECISION AND EVIDENCE

One important difference between health and cybersecurity is the frequent precision of medicine. If you have a specific health problem, the recommendation (a drug, surgery, etc.) may be exceptionally precise. Evaluating if someone had a surgery is a binary choice – either they did or they didn't. Even though there may be nuance, complications, or failure of a surgery, it clearly happened. Similarly, there exists a test for strep throat which definitively indicates whether a patient has it or doesn't.

We can differentiate between people taking a specific action (say, surgery) and taking general health-improvement actions (stop smoking, exercise, eat better).

Many cybersecurity preventative recommendations are closer to the general health improvement advice. But even those cybersecurity actions that are more precise are often less precise than medical advice. In part, this is because we have one body and many computers.[2] So even relatively simple advice such as "patch" can be complex.

Relatedly, many health issues are backed by solid testing approaches, and treatments are tested in rigorous ways which exceed the standards in cybersecurity.

DEFENSIVE OPTIONS

There may be a variety of defensive options, rather than a single "recommended preventative security action." Recommendations frequently involve a risk analysis of some form. It is unclear if that analysis is part of the recommendation or part of forming the recommendation. That is, if the recommendation is "Perform a risk analysis, and use its output to decide on a patching schedule" and the risk analysis says "never patch", the organization has probably completed (taken) the recommended preventative security action. In contrast, if the risk analysis is separate, then perhaps there is no "recommended preventative security action." This complexity may be inherent in the HBM, and may not be relevant unless we try to quantify the model and assess outcomes in a quantified way. That quantification would increase the value of precision in these issues, which may not matter for every use of the CBM.

PREVENTATIVE VERSUS REACTIVE

Both the HBM and the CBM formally focus on preventative action. The HBM starts with "perceived susceptibility to disease 'X'" and ends with "likelihood of taking *preventative* action" (emphasis added). We have followed that with the CBM. Many organizations are overwhelmed by

[2] The author has a desktop, a laptop and a phone nearby, and in the office there are at least half a dozen more IP-addressable systems.

alerts, and take very limited action on them, if any. It may be that a model like the CBM could be useful in understanding what alerts bubble to the top, but we have not investigated that.

QUANTIFICATION

Both the HBM and CBM are presented and designed as qualitative models, which draws the question of "could the CBM be turned into a quantitative model, where the effects of various interventions could be modeled to assess their impact on likelihood of taking action?" Such models might lead to testable predictions. At a high level, it seems feasible to quantify the CBM. Assessing susceptibility and seriousness are already being done, with tools like CVSS[3] or the "Known Exploited Vulnerabilities" list maintained by CISA[4]. Quantifying demographic variables might be harder, but relative scales might help. Similarly, assessing the intensity of cues to action with proxies such as advertising budget or social media metrics could provide for quantification there. Assessing benefits of action could use either a quantified impact or a type of impact scale. For example, the perceived benefits of DNSSec accrue to others, while a barrier might be that the firm's internet presence breaks. Some of the difficulties in quantification of the HBM are surveyed and discussed in Jones et al. 2015.

Work to scan the internet at scale, including but not limited to CyberGreen's Internet Infrastructure Health Metrics Framework (IIHMF)[5], could be combined with the CBM. The IIHMF produces ongoing measurements of a set of indicators of cybersecurity by country. Given a list of cybersecurity investments by a country, it would be possible to use the CBM to organize those investments. With several such lists, and the outcomes by country, it might be possible to see if different sorts of investment lead to different outcomes. If they do, then that might inform future investment choices.

CUES TO ACTION: NATIONALISM

In the HBM and in the initial CBM, the "cues to action" set of influences makes no mention of national origin. But there's evidence that it matters in both public health and Cyber Public Health.

Responses to COVID-19 have included mention of its apparent origin in China. Responses to other viral outbreaks have included mention of their origin in Africa, or in African countries, and of course, the 1918 influenza outbreak is frequently called the "Spanish Flu." We are not aware of research on this facet of people's willingness to engage in preventative action. There is some research on preference for vaccine by origin (see AlShurman et al. 2021; Dong et al. 2020).

As "cyber" grows in importance as a tool of statecraft and a domain of conflict, the national origin of software, hardware and services is being used as either a direct factor, or a way to influence opinions in ways that may leverage stereotypes, prejudice, or opinion rather than technical facts.

For example, the data gathering by TikTok, Facebook, and others. The mechanisms, content, and use are all different, and that complexity means that expert (and "expert" and influencer) opinions are more highly valued (Huddleston 2022; Lin 2021). Firms like TikTok are the subject of influence

[3] CVSS is a is a standard method used to supply a qualitative measure of cybersecurity severity.

[4] https://www.cisa.gov/known-exploited-vulnerabilities

[5] https://cybergreen.net/research/#iihmf

campaigns that have included specific reference to TikTok's Chinese origins (Lorenz and Harwell 2022).

CUES TO ACTION: "AWARENESS", SOCIAL MEDIA

Cybersecurity, like many diseases, has an awareness month. A focus on awareness may cue action better for an acute disease than a chronic condition, and it may cue action better for a chronic health condition than an apparently inherent and unchangeable aspect of technology.

Understanding the magnitude and qualities of these differences seems potentially important, especially as "awareness" is not just a marketing activity, but a specific requirement of many compliance programs. If awareness is present, but other aspects of belief dominate decision making, then effort to raise awareness may be taking energy from more productive activity, leaving us less secure.

Separately, the rise of social media clearly changes how health-impacting social cues are delivered and amplified: the impact on teenage anorexia is well documented (Jan, Soomro, and Ahmad 2017).

3. INTERVIEW STUDY DESIGN

We conducted 9 recorded, semi-structured interviews, lasting 45 minutes to an hour each, to explore the CBM and test its structure. We provided each interviewee with a participation form, informed them that interviews would be anonymized, and that they could choose to not answer or to terminate the interview without penalty. We did not provide any compensation for their time. Participants were employed by enterprises in retail, software, banking, and cybersecurity. Two participants came from one very large organization.

The nine interviewees were recruited via personal connections and some social media outreach. Each was responsible, in some fashion, for their organization's response to Log4shell issues. The interviews were semi-structured, and both the interview question list and the CBM evolved in response to both participant feedback and lessons as we interviewed. Interviews were conducted via Zoom and recorded. The recordings were transcribed by an automated service. Transcription errors were generally minor and did not impact our data analysis. Quotations herein have been cleaned up and are cited only for internal traceability; we will not release transcripts. Time markers are generally to what the transcription service marked as a paragraph.

The transcription service provided a time-marked CSV, which was manually coded by a single coder. Some remarks fit multiple codes, which is largely due to where the automated transcription put sentence breaks. We applied multiple codes to those statements rather than splitting them. This introduces slight double counting, which we believe is under 10%, a reasonable level for this preliminary work.

LIMITATIONS

The study was limited to nine participants because of the time and effort involved in recruiting. A larger study, or one with different participants, could have shown different results. Both the interview questions and the model evolved during the study. It's possible that early participants would have responded meaningfully differently to the different prompts. We are not releasing transcripts because of the difficulty of fully anonymizing them. We did not formally collect or analyze participant demographics.

The automated transcription would have impacted searches, especially for brands, software names and technical terminology where we assume the transcription tool is weaker. We did not rely on automated searches.

ETHICS

This work was funded by a grant from a private company and CyberGreen does not have an Institutional Review Board (IRB). Therefore, we were unable to get formal IRB review, but considered ethical factors as we designed and ran the study. First, we believe the risk to the participants was low: we were forthright about the goals and processes for the study, we did not seek to (nor did we) learn any operational security information. Again, we started each interview by informing participants we would respect their desire to not answer any or all questions; we did not mention without penalty because the interviews were uncompensated. The study was performed months after the issues to ensure that the study did not interfere with operational response. Quotes in this report are fully anonymized, and participants have been given a chance to review the report. Additionally, we worked to meet current standards such as providing both prior written information about our goals and data collection practices, and re-stating them at the start of each interview. We plan to delete the audio records and the full transcripts to further protect the privacy of the participants.

The funder maintained an arms-length relationship, and did not provide any feedback to the study's main author on drafts of the report.

4. INTERVIEW RESULTS

The CBM provided an interesting frame for creating and analyzing interviews. Several participants commented explicitly on the questions being an interesting set.

For this work, we felt it was useful to categorize and code statements. This allowed us to produce a breakdown of time spent on each element of the model (Figure 3). This is useful insofar as it shows that these semi-structured discussions of Log4shell had both alignment with the model and distribution across the various parts of the CBM.

In Figure 3, the CBM elements match the numbers shown in Figure 2, and are:

1. Explicit reasons
2. Psychosocial factors
3. Perceived threat
4. Cues to Action
5. Perceived benefits and barriers
6. Likelihood of taking action

These are augmented by additional codings:

7. Interview setup or administration

8. Responses to CBM
9. Contradictory points
10. Challenges/extensions to the model

A-I are the participants.

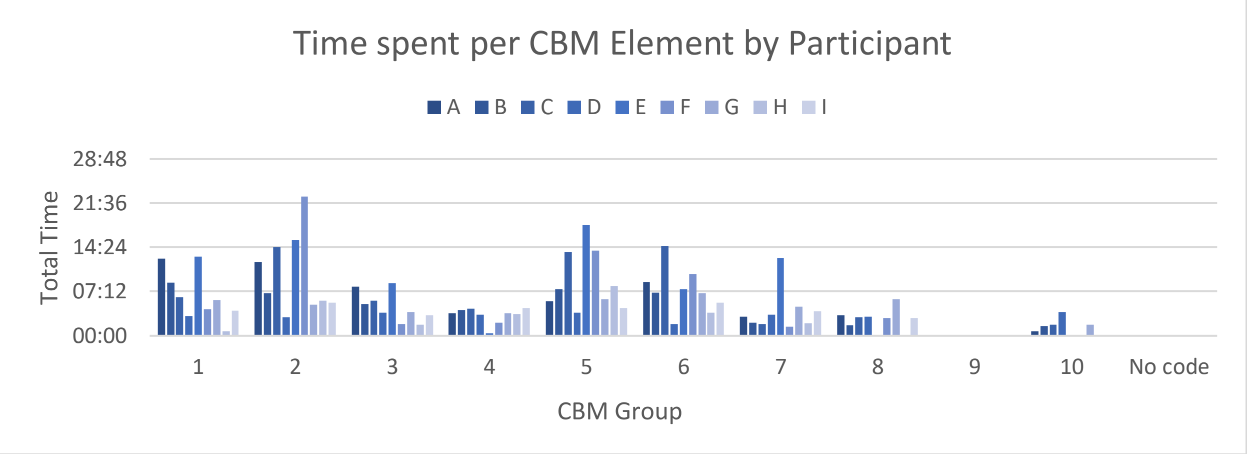


Figure 3

Time spent discussing cues to action (4) is consistently low, as are responses to the model (8). Contradiction (9) is very low, and discussion of challenges people see and ways to extend it (10) are also low. The lack of contradiction is reassuring. It is unsurprising that people don't see ways to extend the model on first exposure — there's a lot to take in.

Different participants found different parts interesting – for explicit reasons (1), psychosocial factors (2), perceived benefits and barriers (5), and likelihood of taking action (6), time spent varied widely. (And for one participant, there was extended discussion in the setup, 7.)

Participants spent relatively little time in perceived threat (3) or cues to action (4). It is likely that an interview focus (other than Log4j) would result in those discussions being less obvious and thus longer.

INTERVIEW ANALYSIS

This section presents analysis of the interviews and points which we found interesting that are centered on the CBM.

GROUP 1: EXPLICIT REASONS: SUSCEPTIBILITY AND SERIOUSNESS

Explicit discussions of reasons to be worried were frequent across interviewees, and most were thoughtful about why they were vulnerable, and how this issue related to other issues they experienced. For example:

> We looked at the specifics on how you could interact and create the problem that the vulnerability would allow. We saw that it was relatively easy to do and that you didn't have to have direct access to the platforms. In other words, we're always really concerned with if we have an internet-exposed application or service, those ones kind of get priority. Log4j didn't, but the way that the particular app works, it didn't necessarily have to be exposed directly to the internet. So we saw that and that was concerning. And then we started looking at, well, what's our

> level of exposure? Does our product use it? Does some of the infrastructure that our product runs on use it? And are there third-party apps that use it that we don't control? And the answer was yes to all of those. So that's when we were like, oh shit. Okay, this is going to be a nightmare. Right? *(Participant D, 12:16*)

The same participant, whose company sells software, continued a few minutes later:

> "...As this is happening, I'm getting deluged with emails from our own customers saying, can you attest? `You guys don't use Log4j anywhere.' And of course we couldn't attest to it."

Another participant discovered that their employer was vulnerable when "... we received report that one of our products had been affected publicly." (Participant G, 3:21) Participant G works for a Fortune 100 company and continued:

> Well, for an enterprise our size, I'm still not sure we're complete...I think we probably got to 90% coverage within a few days, and then 99% coverage within ten or 15 days. [...] But I am not yet convinced that the whole dependency chain is yet known or that there's not some unsupported bit of item somewhere that we haven't flushed out yet.

Not all code which contains vulnerabilities is exploitable. Participant C substantiated this point:

> Our scans told us that in a number of cases the library was present, but it wasn't used by the code base. So, okay, it's there, yes, it's critical, but if the code is not using it in this particular pipeline, for example, or in some package application, it's not going to be exploited because you just can't call it. Okay, well then we can note that and figure out a way to eliminate the library just for good hygiene. (7:31)

Another participant wasn't initially concerned, but the availability of proof-of-concept code helped them validate and reassess:

> I don't think we made the connection that this is really bad, initially. (Interviewer: "Do you remember if there was a thing which helped that tick over?") Yeah, it was a realization that we can validate very quickly what's vulnerable, we can validate third-party services being vulnerable and sort of green light to go do that validation and then it's hard to argue with the shell, like pretty much, so I would say a little bit to get there and then it was kind of nice that it was so simple to validate. You had the smoking gun, the dead body, you had everything there to reconstruct the entire crime. And then it became, hey, wait a second, this is probably a pretty bad threat, I think was probably when we figured out that we could test empirically very quickly. Then that sort of reiterated, hey, wait a second, other people can do the same thing. Right. I think that's when it got really bad. (Participant B, 6:48)

But they ultimately decided that it was impossible to exploit:

> But at the end of the day, it was impossible really to exploit it. You really had to know a lot about the target system and you did all this work and really at the end of the day, the threat wasn't really going to impact you and maybe you feel better about doing all that work. Maybe you've exercised all these muscles you used in your analogy, exercise or whatever, but the end of the day, you didn't improve security... (Participant B 36:44)

GROUP 2: PSYCHOSOCIAL FACTORS

Demographic factors include who you know. For example, one participant reported:

> An anecdote: a friend, a colleague of mine's significant other, also is in security and sort of the joke was, oh, yeah, they haven't even figured out — they don't even know anything about this. I'm like, well, why don't you tell her? You're in the same house. "I don't really want to ruin a weekend." And I'm like, okay, well, her week is

> going to be ruined or her month. I think there is kind of a little bit of that sort of social cues threat did happen in this particular one. (Participant B: 32:58)

Organizational factors include silos, definitions of responsibilities, prior experience, staff competence and more.

Responsibilities include operational systems without defined owners:

> ...where ... no one was even really sure who owned the apps. You can imagine in a big company they have what are typically called like white elephant networks or whatever, where they have apps that are running and doing stuff. And so as we started scanning, we found more and more of this and less and less interest in taking ownership. (Participant E, 15:42)

An executive's personal prior experience was brought up as a factor:

> So if there's a leader in the org who has had a negative experience responding to or maybe not responding to correctly a vulnerability and has been burned by it, they are more likely to respond in the future. [...] But learning your experience through pain is a good teacher and this doesn't seem to account for that either. So maybe that's part of the [...] psychological factors. (Participant E, 38:13)

While another explicitly criticized their organizational readiness:

> We lacked readiness, largely because we didn't have a proper playbook. Prior to the event, leadership was more concerned with who had ownership of the IR program than with how to effectively handle such an incident. When the incident dropped, those that claimed ownership were largely stunned and were either not responsive or did not provide sufficient direction. The early response was therefore left to a small group of engineers who risked upsetting their own management to make forward progress. We did get early buy-in from senior leadership to do this, but we had to go around our supervisors. Had the severity of the issue been less clear, bias for action might have been lost to bureaucracy. In a sense, we got lucky. (Participant [redacted])

Sectoral factors were raised. For example:

> All the law firms start panicking about the same thing at the same time and sometimes that's going to be because they've been talking to each other […] and one of them had a problem and now they're all going "do we have that problem as well?" (Participant A, 48:04)

GROUP 3: PERCEIVED THREAT

There is an interesting question of how to distinguish perceived threat, seriousness, and susceptibility. They are closely related. Many people define threat (or risk) as a function of probability and impact. We looked for specificity in matching, but see "Using the CBM" later in this section.

Most participants commented along the lines of "this was a pretty serious issue."

After the model was revealed, participant G said:

> I am staring at the perceived threat of disease, the middle box, and it seems clear to me that threat is a key variable in convincing someone to take action. How bad is it? And it's also clear to me that there's a limit to its usefulness, okay? Or a limit to the number of times you can go back to that. And I have come to think of that limit as normalcy bias, where although Log4j was a crystal-clear vulnerability, the fact that it resulted in relatively… not zero, but … and I don't mean to trivialize it, but the Internet did not, in fact, melt down. The

> damage was not commensurate with the press. It's hard to know if that is a cause or an effect. But the more you go back to that, well, the more often people think, yeah, this is business as normal. (30:57)

Participant I had somewhat critical words of perception of threat:

> We didn't even have the product that was vulnerable, but because it was in the news, now all of a sudden I have to jump through hoops to prove that we're okay. Specifically, the Confluence vulnerability. That was a month ago, maybe two at this point. Okay, we use Confluence internally at [my employer]. Some VP in the company saw it on the news. Confluence is bad. Okay, well now I'm getting asked all sorts of questions from VP, C suite teams. Are we okay? What's going on with this? We use the cloud version. Cloud is not susceptible. I said the same thing and then I also had to start telling clients that no, yes, we use Confluence, but we use the SaaS version. (35:09) [Note — the interviewee is discussing a different issue than Log4j, Confluence, a widely used package.]

GROUP 4: CUES TO ACTION

Twitter came up repeatedly, in ways that were an amalgam of contributors to perception of threat and as cues to action. We chose to treat the general discussion as a cue to action, unless there was a specific reason that the connection was strong enough to drive action. Participant C said "We don't track researchers typically, but the people who are legit, are. And so we will go and see, what are these guys saying" (11:15). This seems to be a clear cue, rather than a personal, organizational, or other "sociodemographic" factor. It may be useful to further investigate how public health professionals conceptualize social media.

Another interesting point that came up was linguistic:

> One of the engineers in a China office posted a link to a Mandarin-language blog on the issue in our company's Slack channel. After translating it, the severity, scope, and impact were evident. It was early morning in the US and at the time there was maybe one English-language reference to the bug I could locate. Between these sources and my knowledge of our environment it was clear that this was going to be a serious problem. At that point I notified my manager, the CISO, and a few other trusted colleagues on Signal, letting them know that this was going to be significant." (Participant F, 1:39)

They returned to the topic:

> I was the first in my (security-focused) social circle to become aware of it. I reposted the blog to our group's private Slack instance--I want to say eight hours before anyone I knew of in the US had seen it. I definitely benefited from being in the US with a China-based, Mandarin speaking colleague. We got a head start on it for sure.. (6:59)

GROUP 5: PERCEIVED BENEFITS AND BARRIERS

Engineering tradeoffs were a frequent element of perceived benefits and barriers discussions. For example, participant E spent quite a bit of time delving into this, discussing first a risk management decision to update a library within the elasticsearch system without updating it in full:

> There was one case where we made the decision to deploy the Log4j update without deploying a newer elasticsearch. We felt our use case was stable, it wouldn't break anything. But you've probably faced this in your[6] Microsoft days, is not a comfortable thing to do. You would much rather take the elasticsearch update as it is instead of trying to shoehorn a new DLL in there. (Participant E, 5:48)

[6] Referencing the author, who previously worked at Microsoft.

They continued to discuss how people perceived benefits (intermixed with susceptibility):

> But there was a kind of a work balance thing where I'm having to work with my peer, the VP of engineering, and saying, hey, those cool projects you're working on, [you] need to kind of stop all of those and work on this for a while. And so that ended up in a series of escalations where it's like, "well, we're not worried about it." I think I explained to people why Log4j, not being internet exposed, was still vulnerable maybe 100 different times over and over again, because the typical answer would be, oh, that's in a database and it's not on the internet, so you're okay, we don't have to patch it. And I was like, no, you do, and here's why. (Participant E, 7:49)

As well as perceived costs of patching systems over and over again:

> A couple of things that made this worse, as you well know, I'm sure, having listened to quite a few of these, is that there were multiple different patches, there were multiple different workarounds. The original patches and workarounds turned out to not work, and they had to do other ones. (Participant E, 9:30)

Another touched on perceived costs:

> However, did I do the minimal amount of work to improve my security? No. So I would say yes, but there was definitely the wasted time where the patch was in a location or otherwise. We didn't improve security based on the threat or potential threats or future threats. […] I would say we did waste a lot of time and a large amount of time and money and resources remediating, which didn't improve our security. (Participant B, 19:15).

A few minutes later, Participant B went on to say:

> There were plenty of services that used the vulnerable library that didn't have [the combination of] a low impact, low likelihood or even [the combination of] low likelihood, high impact, where frankly, I guess [evaluating the security impact of that situation] depends how you want to define security. I guess what I'm saying is very low likelihood anything would have ever had…to zero likelihood…that anything would have happened. But we did still spend the time and effort to patch the bug.

Participant C also touched on cost: "Yeah, a part that I don't see here [discussing the CBM] ... is the cost of action, right? A lot of times that's huge." (39:43)

Organizational costs were sometimes self-imposed:

> ...finding the right person, making sure that they were available, making sure that they understood the priority on this. Something that should have taken five minutes, took like 2 hours to get posted. It was just inefficient, and by the time they posted it, well, now shifts change. This is a rapidly developing situation, and you just took any lead time we had so it looked decent in front of our customers. I don't know. A lot of that has changed. One of the big things I pushed for after that whole thing was the customer support CSE, to actually have an on-call schedule and be able to be paged from pager duty and then all that happy, fun stuff. (Participant I, 24:44)

Another participant raised "getting back to work" as a benefit, and the cost of testing as a barrier:

> We (Security) attempted to be measured in our recommendations to Engineering. We estimated that we had one chance to get them to implement the changes we asked for. Initially, we didn't have verification that the possible mitigations were going to be effective in our environment. Given the urgency of the situation, we engaged Engineering to go over the options and collect their thoughts. Due to the holiday code freeze and their desire to return to Q1 feature development, Engineering was predisposed to select the fastest (and what they saw as the least intrusive) option without consideration for robustness. In one instance, despite urging them to help us test a potentially stronger mitigation, one group decided to deploy the runtime environment variable mitigation and then stopped communicating with us believing that the issue was fixed. Ultimately, the runtime environment

> variable mitigation was discovered not to be effective. There was a tendency, I think, to move too quickly on the product side. (Participant F, 16:56)

Checking work is important. One participant described having to throw away their incident response playbook and play it by ear and experience because of a lack of testing of the official playbook. (Participant F, around 23:00-25:00)

One participant used their bug bounty program to check their work:

> We actually used our bug bounty program. So we made a couple of very quick changes. Not right away, but once we had completed assessment and we started remediation, we had those folks helping us look for things that we may have missed. Just sort of say, hey, here's special bounty, go search, see what you can find. And it was a really good verifier of like one, are we really fixing things right by people who are going to come and tell us? And the answer generally was yes, we missed a couple of things here and there, but generally it was good. But also it gives you insight into how much of your external infrastructure old seat and we learned a little bit there as well. So good verification, a lot of false positives, obviously, but again, we try to muster all resources and throw them at these big problems. (Participant C, 12:39)

Leading to them offering a summary: "And so I think engineering's interest in patching this waned over time."

Regulation was reported as a barrier to action by participant I, who said:

> We had to restart [with the second patch] and then get permission to push it into the FedRAMP environment from our FedRAMP customer. And so it was just kind of like a debacle back and forth, back and forth. So [intensely frustrating].[7] (18:06)

Another surprising barrier to action was live events:

> I don't think anything necessarily [makes patching challenging] other than live sports. Unlike watching your favorite show rerun, it's valuable to a customer in a moment that the events happen. Right? So, otherwise you know what the score is and all that. But I think it's just more of the value of the content and the value of that particular event is high. And there are ways, I think, to do patching in sort of a very kind of safe way. So it's more of a question of how can we patch this over the next week or do we need to patch this over the next hour? (Participant B, 12:32)

Several participants were retailers. It's well-known that retailers freeze their code before Thanksgiving in preparation for the Christmas rush. Each reported that they were able to break the deployment freezes with relative ease given the severity of the issue.

GROUP 6: LIKELIHOOD OF TAKING ACTION

Many actions were decided by senior leadership teams. One participant, who provides software to a critical infrastructure sector, said:

> Internally we had good support for the feature teams, and for our leadership that they recognized that this was serious and I think some of that was down to them having confidence that when we are saying actually we're pushing the button, this is an emergency, that's not something we would do lightly. The fact that they could also

[7] The author is not a FedRamp expert. It may be that FedRamp doesn't require this, but certain customers add requirements that are grouped. Reviewing the recording, it is clear that the participant says "customer" in the singular. We also note a perceived need for permission to patch plagues medical devices, years after FDA has publicly clarified that patching security issues does not impact certification.

> see sort of everyone in the Java land screaming their heads off and then they were pretty convinced that this was really bad. I think what prevented having too much sort of flapping and difficulty was when it was the fact that we were able to quickly give them a list, saying this is very valuable. So I think that was critical to ensuring we did get a good response, particularly for my clients. (Participant A, 21:18)

It is notable that the work of government agencies was not brought up; the company is in a highly regulated sector with a globally well-regarded regulator.

Another participant talked of executives wanting to "do the right thing," a phrase the participant used repeatedly, along with "upfront":

> So I know what we did was not only a benefit for all the customers from a security perspective, but it was the right thing to do. I do feel there's a lot of companies out there that don't necessarily want to do the right thing because it's going to cost them more time, labor, money, et cetera. That's one thing I can say [my company] was very good at. They pushed and wanted to be upfront about everything. Even before the new SEC dictation on disclosure for cyber incidents, they were still trying to be very upfront with what happened or what will happen. I had quite a few conversations with our legal departments on that, because what you're obligated to does not necessarily always mean what is the right thing. (Participant I, 22:03)

Another talked about the importance of keeping perspective on security:

> What comes to mind about the issue besides what we've talked about so far is the old xkcd cartoon that you probably know. And I'm talking about all modern infrastructure and block and little block down at the bottom, says one critical component maintained by a volunteer in Nebraska since 2013 that's been shared infinitely many times in the context of this incident and brings the same trouble every time anyone shares it. The other thing, that…two other things, actually… one thing is that the headlines and the bandwidth that everybody had to extend to resolve this incident created a great opportunity to remind people of the importance of security. So we did internal talks featuring many people, you know, myself included, on what Log4j did and from different perspectives. And that was a very timely way to get the engineering community back together after a long COVID enforced, [video-only] sort of training sessions. There's a limit to how much impact, attitude adjustment you can have in a lecture over video. (Participant G, 21:56)

USING THE CBM

Above, we discussed the usefulness of coding interviews as an analytic technique. However, many statements were somewhat hard to categorize, and can fit in multiple places within the model. This may appear to be a weakness, but it is not.

If the CBM is seen as a conceptual framework for considering reasons that people act, then it is not a substantial weakness if their reasoning is not expressed in ways which perfectly align with the model. Complexity in categorizing a statement may reflect that people's thoughts are organic.

In threat modeling, people often refer to STRIDE[8] as a categorization system. It is relatively easy to find threats that don't fit perfectly (for example, is deleting the content a file tampering or denial of service?) However, if we use STRIDE to prompt us to think about factors that might be relevant, the categorization complexity is simply not relevant. The CBM may share that property.

This differentiation between a prompting model and a decision tree leads to complexity in putting labels on some statements. Are they descriptions of threats or of susceptibility? For example, in Group 1 (Susceptibility) we quoted Participant B:

[8] STRIDE stands for Spoofing, Tampering, Repudiation, Denial of Service and Elevation of Privilege.

> Yeah, it was a realization that we can validate very quickly what's vulnerable, we can validate third-party services being vulnerable and sort of green light to go do that validation and then it's hard to argue with the shell, like pretty much, so I would say a little bit to get there and then it was kind of nice that it was so simple to validate. *You had the smoking gun, the dead body, you had everything there to reconstruct the entire crime. And then it became, hey, wait a second, this is probably a pretty bad threat*, I think was probably when we figured out that we could test empirically very quickly.

In this instance, should we treat the discussion of threat (in italics) as separated from the surrounding discussion of susceptibility? This is a challenge created by treating the HBM or CBM as a decision tree.

TAKEAWAYS

This section presents themes and discoveries that are broader than the CBM. They are:

- The participants are not highly confident that their work improved security
- Architectural investments pay off, notably "egress filtering"
- Perceived costs of patching are high
- There are issues with industry information flows

LACK OF CONFIDENCE IN SECURITY

Question 9 was "What's your confidence that your work improved your security" (Later we added, "or that of your customers?") The answers were frequently that people were not confident. This is surprising because the participants were generally leaders charged with addressing the problem, and we expect that, after an 'all hands on deck' response, they would uniformly report confidence.

ARCHITECTURE PAYS OFF

Several participants mentioned egress filtering as a particularly helpful control. For example:

> From what I remember, it was interesting because you could see the attempts, but a lot of the places where we saw the attempts did not have egress access enabled. You only got about halfway through the exploit chain and then just you couldn't really complete it. (Participant C, 14:13)

> Because we are egress filtering, once people launched their first stage attack they wouldn't be able to download the payloads to ask it for code execution... [which only allows communication to] hosts that we have explicitly allowed. (Participant A, 4:48)

Outbound or egress filtering seems to be under-represented in control frameworks. For example, it is not obviously in CISA's 2022 Cross-sector Performance Goals or the NIST CSF's PR.PT, or references such as CIS CSC 8, 12, 15. [9]

Additionally, Participant G pointed out that log infrastructure was an invisible part of their Trusted Computing Base (TCB):

[9] Version 8 of the CIS controls includes 12.3, "Establish and maintain a secure network architecture. A secure network architecture must address segmentation [...]" and 13.4, "Perform traffic filtering between network segments, where appropriate." Either or both can be read to incorporate egress filtering, but do not explicitly call it out.

> [You combine the need for logging and that] we've underappreciated the complexity of logging systems which are big, complicated pipelines…complicated ETL jobs. And it's a critical attack surface and it's big and complicated and it's sort of necessarily in your TCB then what? ... We're paying more attention to logging systems as a source of attack surface than we have previously as a result of Log4j. (12:39)

Another commented:

> One of the things we learned, and it was serendipitous, was, like, we had way underestimated the amount of time it would take us to collect and process all the logs that we needed at our scale in order to determine whether exploitation happened. (Participant C, 14:13)

PERCIEVED COSTS OF PATCHING ARE HIGH

Many participants talked about the high cost in terms of effort to patch their systems. While we did not investigate that, it seems that further research into the actual costs and reasons for those costs may be productive.[10]

INDUSTRY INFORMATION FLOWS

Timeliness of information was a recurring theme, both in getting information about the vulnerabilities, but also about products. One participant said:

> The biggest challenge was understanding what our suppliers were going to do [...] One of our vendors pulled down their product until they could release something that was known safe [...] Another supplier - it was really an exercise in frustration – that they said 'the mitigations are applied' and they did not commit to when that new version went live. (Participant D, 12:16)

Inaccurate information and curation was a consistent problem. The cost of collecting, deduplicating, and verifying information was borne by many organizations, despite the existence of funded capabilities for such intelligence gathering and dissemination. We did not investigate the motivations for this activity.

Even once the facts are established, information flows about system vulnerability are inefficient. Several participants discussed the barrage of "are you vulnerable?" requests, and how those interacted with suppliers, as well as with the set of issues. Several organizations were in the midst of a complex web, taking in software and delivering more than one product to others, as packaged software, SaaS, or both. The metaphor of a chain is insufficient. The set of issues were collectively referred to as "Log4shell," and so when asked to "attest you're not vulnerable to Log4shell," there was an issue of clarity: what request is being made? Is it about CVE-2022-44228? The set of issues at the time of the request? The set of issues at the time of the response?

That inefficiency is compounded by the cost and delay inherent in generating reliable information:

> [We threw a lot of people] at this issue right now to try to gather as much information. And obviously they're doing what intel people do, right? They're gathering data, deconflicting probably 80% to 90% of what you find out there is just amplification of a handful of sources. Like people seem to be sort of rebuilding even threat intel vendors, like the same blogs, the same posts. So it's easy to see something and think, wow, there's a lot of noise. But it's like, no, everybody's just repeating the same thing. It's amplified noise. (Participant C, 11:15)

[10] It's conceivable that software creators don't spend money on ensuring their systems can be easily patched because they can shift those costs to their customers, who have a hard time evaluating the costs before they purchase, or because most vendors don't invest in easily patched systems.

Similarly:

> [We use X for a threat feed]. It was not helpful, to be honest. Partly for some good reasons and some lots of good reasons. One is because they're curating information, what they're doing is taking the raw data from what we're seeing in terms of Twitter, one alike trying to validate it and then posting that curated information. So for some necessary delay there. By that process, I would have hoped for some kind of urgent notification of our confidence level on this is reasonable, but perhaps not of our later standards and telling us about it, given the impact, but we didn't get that. And that's when it's continued threat. They produced some dashboards and things that were, I guess, kind of useful, but way after we finished. So it was not effective for us in any way. (Participant A, 17:48)

Another participant said they expect this trend to continue:

> I don't know if [we're] going back to how we had to deal with all our other vendors and their lack of transparency and really having to hold their feet to the fire. That's something that I think now the industry has seen a lot more because of this vulnerability, or at least that's what I'm positing. We are going more and more to vendors and say, hey, where are you at with this particular patch? That's not something we would have said five or ten years ago or even maybe two years ago. (Participant B, 15:31)

This is not simply a concern between organizations. Our two participants from a single organization appear to have contradicted each other. One said it came from a Twitter post, the other said they believed they got advanced notice somehow.[11] There are many possible explanations for the disagreement, including failures of internal alerting mechanisms.

Investments in information sharing are high. Many governments maintain Computer Emergency Response Teams (CERTs). The United States also encourages the creation of sectoral organizations. These organizations seem to have charters to perform information gathering, verification and analysis, and to disseminate the results. However, they were not mentioned spontaneously in our interviews. It may well be that they would have been mentioned if the author had asked.
It may be worthwhile for those funding such organizations to assess if they are performing to goal.

OPPORTUNITIES TO IMPROVE THE CBM

As this project draws to a close, there are several apparent opportunities to improve the model, including consideration of more reliable coding approaches, quantification, and deeper investigation into how cues to action work.

CUES TO ACTION

There are several opportunities to refine how we treat cues to action. One question is, should they be more broadly connected?

For example, Figure 4 shows a version where cues to action can drive perceptions of susceptibility or directly influence likelihood of taking action.[12]

[11] The interviewer did not press on this claim early in the interview on the assumption that it would have cost trust.
[12] It is reasonable to assume the author sees and uses models differently than sociologists.

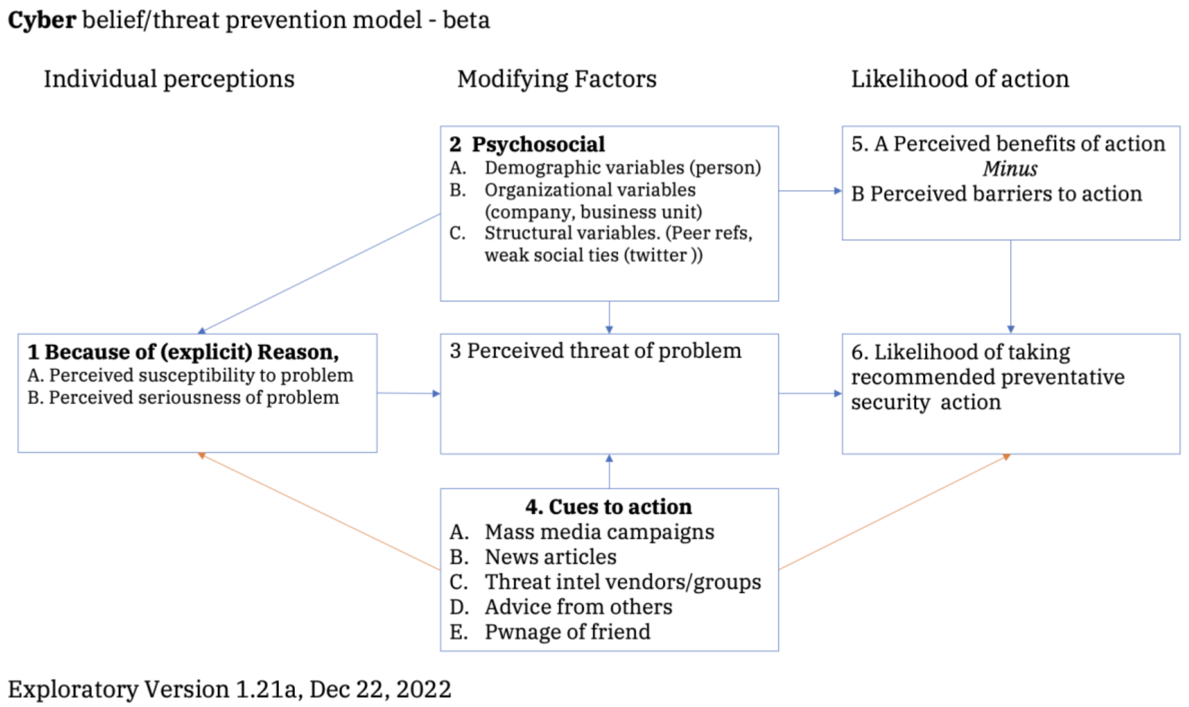

Cyber belief/threat prevention model - beta
Individual perceptions
Modifying Factors
Likelihood of action
2 Psychosocial
A. Demographic variables (person)
B. Organizational variables (company, business unit)
C. Structural variables. (Peer refs, weak social ties (twitter))
5. A Perceived benefits of action
Minus
B Perceived barriers to action
1 Because of (explicit) Reason,
A. Perceived susceptibility to problem
B. Perceived seriousness of problem
3 Perceived threat of problem
6. Likelihood of taking recommended preventative security action
4. Cues to action
A. Mass media campaigns
B. News articles
C. Threat intel vendors/groups
D. Advice from others
E. Pwnage of friend
Exploratory Version 1.21a, Dec 22, 2022

Figure 4

Another is the question of "what exactly does pwnage of a friend mean?" Participant A reported:

> So I had the one side at work where people were we had our own idea of what was happening and I was able to take that into a completely I mean, it was a similar context, but it was completely different. Eyes on it and say, this is our response strategy. What's your response strategy? And take the best ideas out of all of it, bring it back here, take the best ideas, just do this back and forth. Yeah, absolutely. It would have been harder had I not had that network. For sure, for sure. Others were seeing different things depending on their workers. Some of them, like I said, we got lucky. Were really not hit with script kiddies. Some people were getting pounded through people active attacks on their systems. Right. So I was able to use that information to guide our blue team to be like, hey, these are the patterns. Some of my friends were just dumping the path. This is what we're seeing. Like, here's the raw traffic. (1:09)

There are several things to note here. First, attacks are not pwnage unless they succeed, and so perhaps our label should be broader to incorporate attacks.

We might move "Pwnage[13] of friend" to an explicit part of box 2, but no one reported exactly this, and so we have been unable to explore it except as a hypothetical. For example, if Alice hears over beer with Bob that Bob's company was compromised, she might want to share that with her coworkers to impel action. However, this raises questions of discretion. May she share? What can she share, in what form?

In this case, the sharing seems to be of attack information, which is simpler than breach information.

[13] "pwn" is hacker slang for taking over a system; pwnage is a noun form.

5. CONCLUSIONS

The CBM is a new tool we can use to understand why people don't act on cybersecurity advice. Because it adapts a powerful and well-established model from public health, there is a wealth of example uses and deep reflections that are not present when models are created from a blank slate.

Having created, exercised, and validated the model, we can turn attention to its use by enterprises (end users of software), cybersecurity vendors, and policy makers.

Enterprises, and especially their security groups, can use the CBM to assess the deployment of new tools, and possible reasons their organizations are not acting on their recommendations.

Cybersecurity vendors may find the CBM useful to assess if there are factors, such as lack of benefit or excessive barriers to action. Many of these barriers may be addressable by creating new software or training capabilities.

Policy makers can use the CBM to understand why organizations are not taking hoped-for defensive measures, and assess if policy interventions can influence relevant factors. In particular, the quantification approach discussed in Section 2 may lead to improved understanding of the impacts of policy choices.

Beyond use of the model, we hope that the work here inspires scientists and researchers to extend and adapt the model.

REFERENCES

AlShurman, Bara' Abdallah, Amber Fozia Khan, Christina Mac, Meerab Majeed, and Zahid Ahmad Butt. 2021. "What Demographic, Social, and Contextual Factors Influence the Intention to Use COVID-19 Vaccines: A Scoping Review" *International Journal of Environmental Research and Public Health* 18, no. 17: 9342. https://doi.org/10.3390/ijerph18179342

Bals, Fred, Why developers need a supplemental source to NVD vulnerability data, Synopsys blog, June 1, 2020 https://www.synopsys.com/blogs/software-security/developers-need-more-than-nvd-vulnerability-data/

Center for Internet Security, CIS Controls, Version 8, May 2021.

Cyber Safety Review Board, Review of the December 2021 Log4j Event, July 11, 2022, https://www.cisa.gov/sites/default/files/publications/CSRB-Report-on-Log4-July-11-2022_508.pdf

Cybersecurity Infrastructure Security Agency, Cross-Sector Cybersecurity Performance Goals (2022), https://www.cisa.gov/sites/default/files/publications/2022_00092_CISA_CPG_Report_508c.pdf

Dong, Dong, Richard Huan Xu, Eliza Lai-yi Wong, Chi-Tim Hung, Da Feng, Zhanchun Feng, Eng-kiong Yeoh, and Samuel Yeung-shan Wong. "Public preference for COVID-19 vaccines in China: A discrete choice experiment." *Health Expectations* 23, no. 6 (2020): 1543-1578.

Jan, Muqaddas, Sanobia Soomro, and Nawaz Ahmad. "Impact of social media on self-esteem." *European Scientific Journal* 13, no. 23 (2017): 329-341.

Jones, C. L., Jensen, J. D., Scherr, C. L., Brown, N. R., Christy, K., & Weaver, J. (2015). The Health Belief Model as an explanatory framework in communication research: exploring parallel, serial, and moderated mediation. *Health communication*, *30*(6), 566–576. https://doi.org/10.1080/10410236.2013.873363

Huddleston, Tom Jr., TikTok shares your data more than any other social media app — and it's unclear where it goes, study says, CNBC, 2022 https://www.cnbc.com/2022/02/08/tiktok-shares-your-data-more-than-any-other-social-media-app-study.html

Lin, Pellaeon, TikTok vs Douyin A Security and Privacy Analysis, Citizen Lab Research Report #127, March 22, 2021 https://tspace.library.utoronto.ca/bitstream/1807/123974/1/Report%23137--TikTok.pdf

Lorenz, Taylor and Drew Harwell, Facebook paid GOP firm to malign TikTok, The Washington Post, March 30, 2022. https://www.washingtonpost.com/technology/2022/03/30/facebook-tiktok-targeted-victory/

National Institute of Standards and Technology, Framework for Improving Critical Infrastructure Security, Version 1.1, April 16, 2018 https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf

Rosenstock, Irwin M. "Historical origins of the health belief model." *Health education monographs* 2.4 (1974): 328-335.

ACKNOWLEDGEMENTS

This report was written by Adam Shostack for the CyberGreen Institute.

We would like to thank the study participants, who gave generously of their time and attention. We would like to thank the sponsor who made this work possible.

APPENDIX A: INTERVIEW QUESTIONS

The interviews had three phases: introduction, interview and debrief.

INTRODUCE THE PROJECT

- CyberGreen is a non-profit, working on a new science of cyber public health.
- Our plan is to produce freely available science without specific attribution of anything you say here or a list of participants
- We are recording this call, and will aim to transcribe and then destroy the recordings
- We expect that this call will take about an hour.
- If any question makes you uncomfortable, you're free to either ask me to skip it or to end the interview
- To avoid biasing the research, I'd prefer to not talk about where we're going until the end.

THE INTERVIEW QUESTIONS

1. What was your role in handling Log4shell and related vulnerabilities?
2. How did you personally first become aware of one of the issues?
3. How did you perceive its seriousness?
4. How did you think about susceptibility?
5. if personal awareness proceeded formal organizational plans:
    a. Did you work to make others aware of the issue?
    b. Were there things they brought up that made it harder to take action?
6. If organizational awareness proceeded personal awareness:
    a. How did you become aware?
    b. What was your role?
7. Did you set up a formal approach to hearing about variants?
8. Were you aware of any question from leadership about the benefit of the work? (if so, what?)
9. What's your confidence that your work improved your security (or that of your customers)?
10. Are there things you did that didn't seem effective?
11. What else comes to mind about the issue?
12. Is your confidence in your ability to handle future such crises improved/neutral/diminished?
13. What are you doing differently as a result of your experience?
14. What else should I ask?

DEBRIEF

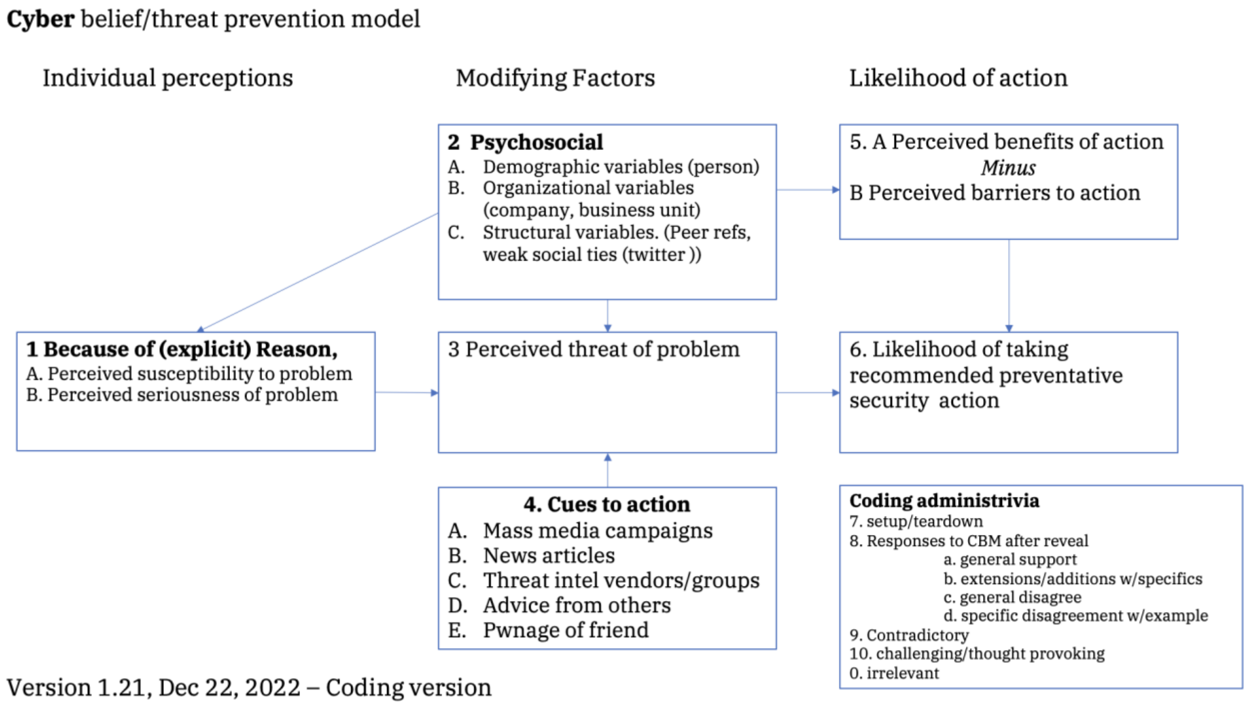
The interviewer presented the CBM, explained it, and asked for feedback.

APPENDIX B: CODING SCHEME

Cyber Belief Model Interview Coding instructions
(Sept 28, 2022 Draft)

1. Transcribe the interview, and put the output in a spreadsheet.
2. Break up long lines
3. Assign codes based on the chart on the next page.
4. For "via twitter" focus on the impact on belief, not the channel
    a. Awareness from people you follow/interact with is 2c
    b. "There's a lot of chatter on twitter" is 4
5. When an issue could be either a 1 or a 3 – default to coding as "3" in august
    a. Consider Imminence, Specificity

- HBM description of Demographic:
    - Demographic variables include age, sex, race, ethnicity, and education, among others. Psychosocial variables include personality, social class, and peer and reference group pressure, among others. Structural variables include knowledge about a given disease and prior contact with the disease, among other factors.
- CBM:
    - Demographic variables relate to a person including "I've been at this a white"; I wasn't able to convince someone (5b) because I'm a woman and leadership disrespects women (2a)
    - Organizational variables (new to CBM) relate to a business, or a business unit - is it a tech company, a software producer. video provider not patching because of a big game coming up.. the choice is a specific thing because of *our business*
    - Psychosocial 2c includes peer references, weak social ties people we follow talking about it (twitter, slacks, github)
    - Perceived susceptibility is about do I believe my org is vulnerable? Some cues to action are of the form of "you may be vulnerable even if you don't use Java"
    - Exhortations about the issue being large/real are cues to action when tied to a specific (specified, completable) task.
    - That information shows up on twitter is secondary to what sort of information it is
        - We treat awareness from someone we follow as cue to action (#4); "I was popped" is awareness because of a friend's susceptibility



Cyber belief/threat prevention model

Individual perceptions

Modifying Factors

Likelihood of action

2 Psychosocial

A. Demographic variables (person)
B. Organizational variables (company, business unit)
C. Structural variables. (Peer refs, weak social ties (twitter))

5. A Perceived benefits of action
Minus
B Perceived barriers to action

1 Because of (explicit) Reason,
A. Perceived susceptibility to problem
B. Perceived seriousness of problem

3 Perceived threat of problem

6. Likelihood of taking recommended preventative security action

4. Cues to action

A. Mass media campaigns
B. News articles
C. Threat intel vendors/groups
D. Advice from others
E. Pwnage of friend

Coding administrivia
7. setup/teardown
8. Responses to CBM after reveal
a. general support
b. extensions/additions w/specifics
c. general disagree
d. specific disagreement w/example
9. Contradictory
10. challenging/thought provoking
0. irrelevant

Version 1.21, Dec 22, 2022 – Coding version

Decisions:

Prompt	Issue	Decision	notes
"We're a company that creates alerts based off of vulnerabilities."	This could be either an organizational variable (2b) or a cue to action (4c)	Remove "group" from 4c	An org that creates a threat intel *group* has an investment in that group succeeding in a way that's probably distinct from "a vendor told us this.
"So that's when we were like, oh shit. Okay, this is going to be a nightmare. Right?"	Is this perceived seriousness of the problem (1b) or perceived threat (3)?	We've combined susceptibility and seriousness to what it means to us, so perceived threat	