Human Factors in Cybersecurity, Vol. 127, 2024, 115–126
https://doi.org/10.54941/ahfe1004771

Discovering Cognitive Biases in Cyber Attackers' Network Exploitation Activities: A Case Study

Palvi Aggarwal[1], Sridhar Venkatesan[2], Jason Youzwak[2], Ritu Chadha[2], and Cleotilde Gonzalez[3]

[1]The Department of Computer Science, El Paso, TX 79912, USA
[2]Peraton Labs Inc., NJ 07920, USA
[3]Carnegie Mellon University, Pittsburgh, PA 15213, USA

ABSTRACT

Understanding a cyber attacker's behavior can help improve cyber defenses. However, currently, there are no substantial publicly available datasets to learn about attackers' decision-making processes and associated cognitive biases. Recent research has significantly advanced our understanding of attackers' weaknesses, new research is needed to provide clear metrics of cognitive biases in professional red teamers, using testbeds that represent realistic cybersecurity scenarios. New studies should go beyond exploratory observations and rely on formal metrics of cognitive biases derived from actions taken by the attackers (i.e., rely on what attackers "do" rather than what they "say") and demonstrate how defense strategies can be informed by such biases. In this paper, we start to build upon existing work to demonstrate that we can detect and measure professional red teamers' cognitive biases based on the actions they take in a realistic Advanced Persistent Threat (APT) scenario.

Keywords: Cybersecurity, Cognitive biases, Cyberpsychology, Oppositional human factors

INTRODUCTION

Cyber defense is a complex field that requires domain knowledge and cognitive abilities to detect adversarial activities and determine potential adversarial intent within a dynamic and complex environment (Ben-Asher and Gonzalez, 2015; Gonzalez et al., 2014). To improve cyber defenses, studies are needed that leverage skilled individuals (attackers and defenders) interacting with high-fidelity cyber ranges to collect activity data. However, the availability and recruitment of skilled participants is a significant challenge for researchers seeking to study the cognitive aspects of cyber operations. In particular, human attackers and defender behavior have been studied for more than a decade (e.g., Aggarwal et al., 2016; Dutt et al., 2011) but research on attacker behavior is especially scarce. Studies are often limited to abstract testbeds where cyber expertise is not mandatory, studying general attack behaviors exhibited by participants recruited from anonymous crowd-sourcing platforms (Cranford et al., 2020). Exceptions to the above are studies performed using red teams in realistic cyber environments and

observational assessments of cyber defense behaviors (Buchler et al., 2018), exploratory experimental studies of attacker behaviors in the presence of deceptive strategies (Aggarwal et al., 2022), and field experiments with professional red teamers operating in realistic networks (Ferguson-Walter et al., 2018). Yet, most evidence of the role of cognition in cybersecurity activities relates to defenders and much less is known regarding attackers.

We expect that cyber defense can greatly improve if we can determine the cognitive biases of the adversary; this will enable the development of technology informed by the attacker's behavior and take advantage of the attacker's cognitive weaknesses. Recent work has begun to set the stage for the study of attackers' cognitive biases with the goal of implementing defenses that disrupt the goals of malicious cyber attackers (Gutzwiller et al., 2018). For example, in the Tularosa study (Ferguson-Walter et al., 2018), researchers recruited professional red teamers to participate in a network penetration test. They provide evidence of attackers' confirmation bias and framing effects that reflect contrasting risk tolerances based on whether a problem is presented as positive or negative (Ferguson-Walter, 2020). Other studies involving similar attack activities (e.g., traversing familiar attack paths) and emotional responses (e.g., frustration) have observed attacker biases under similar settings across groups of anonymous, student, and professional participants (Aljohani and Jones, 2022).

While past research represents a significant advancement in our understanding of attackers' weaknesses, a continuous research is needed to provide clear metrics of attacker cognitive biases in professional red teamers, using testbeds that represent realistic cybersecurity scenarios. New studies should go beyond exploratory observations and rely on formal metrics of cognitive biases that can use the actions taken by the adversaries (i.e., rely on what adversaries "do" more than what they "say") and be able to demonstrate how defense strategies can be informed by such attacker biases. In this paper, we start to build upon existing work to demonstrate that we can detect and measure professional red teamers' cognitive biases based on the actions they take in a realistic cybersecurity scenario.

DECISION-MAKING BIASES IN CYBERSECURITY

Humans make thousands of decisions every day. However, in complex environments, the human mind is incapable of paying equal attention to all the relevant attributes required to make decisions, and it often creates mental shortcuts, called "cognitive heuristics." These heuristics are fast, economical, and often effective. However, they can lead to systematic and predictable errors, called "cognitive biases" (Tversky and Kahneman, 1974). Such biases result in systematic decision-making errors that may be problematic in various environments. For example, the effect of cognitive biases has been demonstrated in a large number of practical decision tasks, including venture formation, clinical medicine, share price reversal, software engineering, information retrieval, and crowdsourcing (Azzopardi, 2021; Draws et al., 2021; Klein, 1990; O'Sullivan and Schofield, 2018; Simon et al., 2000). Similarly, in the context of cybersecurity, attackers, and defenders may exhibit

cognitive biases. However, only a few research efforts on understanding cognitive biases in the realm of cybersecurity have been conducted to date. This paper focuses on decision-making biases in cyber attackers.

Recent research provides initial evidence that cognitive biases can be induced and thus can be leveraged by cyber defenders to exploit attackers' behavior and impact their decisions. Ferguson-Walter et al. (2018) conducted *Tularosa study* with 138 professional red teamers and provides an evidence of attackers' biases derived from a two-days cyber exercise. With deception, defenders were able to induce negative affective states such as frustration, confusion, self-doubt, etc. which might have impacted attackers' performance in the experiment. By analyzing the "*Tularosa study*" dataset, Gutzwiller et al. (2018) provide evidence of attentional tunneling, the illusion of control, anchoring bias, and framing effects that reflect contrasting risk preferences according to whether a problem is presented as positive or negative. This paper postulates that deception could induce various cognitive biases which would otherwise be absent. Gutzwiller et al. (2019) further analyzed the dataset in Tularosa study and observed *confirmation bias*, *anchoring bias*, and evidence of the use of the *take-the-best heuristic* by red teamers. Dataset collected in the Moonraker study (Shade et al., 2020) showcased a preference for selecting the first or the last IP address in a list returned by network reconnaissance operations (*default effect bias*). Understanding attackers' cognitive limitations has been helpful in delaying and deterring attackers' activities. Cyber Deception research conducted under the MURI project used cognitive modeling and computational game theory to develop transformative advances in the science of security about attackers' behavioral processes and cognition. The project provides evidence of successfully exploiting human attackers' cognition using a combination of truthful and deceptive information (Aggarwal et al., 2023a; 2023b; Bao et al., 2023; Miah et al., 2023). Using the similar approach of oppositional human factors, Johnson et al. (2022) designed experiments to induce *Sunk Cost Fallacy* by manipulating uncertainty, project completion, and difficulty which resulted in effectively delaying the attacker's activities. Observation of cognitive biases is challenging; however, inducing these biases to manipulate the attacker's behavior in cyber operations is even harder. Aggarwal et al. (2020) deployed cyber deception strategies against adversaries and observed evidence of *certainty bias* and *risk aversion* in attackers' decisions. In a follow-up study, Aggarwal et al. (2022) exploited risk aversion using prospect theory and thereby reducing defender's losses significantly. Similar evidences of cognitive biases were provided in Aggarwal et al. (2023) where attackers exhibited sunk cost fallacy and default effect bias.

The initial work described above laid a foundation for conducting research on cognitive biases in cybersecurity. However, there is a huge gap in formulation of cognitive biases in the context of cybersecurity. Thus, there are numerous avenues for future research to detect, quantify and trigger cognitive biases in cybersecurity. Recognizing this research vacuum, we attempted to uncover some of the primary cognitive biases prevalent in the participants in a case study described in this paper. We hypothesize that attackers will exhibit cognitive biases as they progress through the attack kill chain in an

Advanced Persistent Threat (APT) scenario. We believe that it is valuable to gain an understanding of these biases in attackers so that the biases can be exploited to develop significantly stronger cyber defense measures.

CYBER SCENARIO IN CYBERVAN TESTBED

We used the CyberVAN testbed (Chadha et al., 2016) for conducting Human in The Loop (HILT) experimentation. We designed a network scenario in which an attacker would execute an APT-style attack campaign with a goal to obtain sensitive documents from the target host. To achieve this goal, human attackers were asked to perform network reconnaissance, laterally move to hosts and gain access to the relevant systems, and finally, perform data exfiltration as a post-exploitation task. The network scenario enable a multi-step attack campaign wherein participants were required to make several decisions as they progress toward the goal. Hosts were running Ubuntu 20.04 with either intentional misconfiguration or vulnerable services. The participant host was running Kali Linux 2022 which contained pre-installed attack tools such as Nmap and Hydra, and customized wrapper attack scripts to simplify the execution of the attacks.

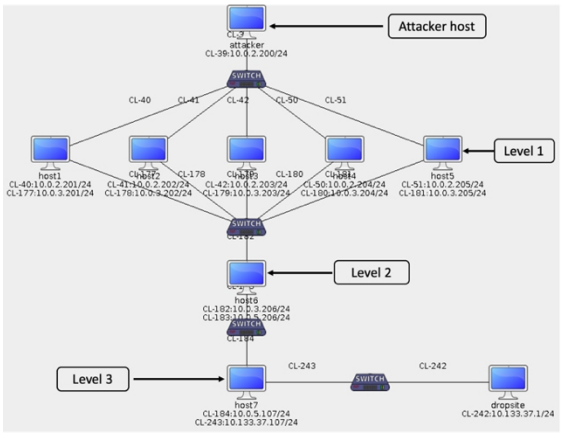

Figure 1: Network architecture.

Attack Workflow. The activities in the network (shown in Figure 1) were divided into three levels. Attackers start their activities from the attacker host. At level 1, they encounter five hosts, and their goal is to gain unauthorized access to one of these hosts by cracking the passwords of valid users on the system. Specifically, they scan port 22 (for SSH) in an IP address range between 10.0.2.201-250 using the Nmap scanning tool. After *choosing a host* with port 22 open, they use a password-cracking tool, Hydra, to crack the passwords over SSH. Next, they *choose a credential* to login to the host following which they enumerate additional credentials from the local shadow file – a credential store for Linux systems – using a pre-installed tool called John the Ripper. The next step is to choose a set of credentials different from the ones that were identified by the Hydra tool to access the host

at level 1. The network was configured such that some of the unauthorized login attempts would be detected and the corresponding session would be terminated. Attackers are informed through prompts on the terminal when their login attempts are detected by the defender. Once attackers successfully log in to a host at level 1, they pivot to the host at level 2 by choosing one of vulnerabilities to exploit that was present on that host. The host was configured to be vulnerable to CVE-2014-6271, CVE-2023-23752, CVE-2017-12636, and CVE-2023-28432.

Similar to level 1, some of the unauthorized login attempts would be detected and the corresponding session would be terminated. At level 2, the attacker's goal is to gain access to the *host7* machine at level 3 and exfiltrate as many files as possible from the target machine. From level 2, attackers are given two options to execute the attack: (i) an open-source tool that is reliable but requires additional effort to set up and execute, and (ii) a prepared shell script that is unreliable (small probability of success) but easy to execute. Upon compromising the *host7,* the final action is to exfiltrate as many files as possible from the host to an external drop site. For exfiltration, attackers choose between standard file transfer applications such as SCP and FTP. Attackers were periodically informed that the network defenders might be monitoring the network and that they might be detected at any stage of the task. If detected, attackers were returned to the previous step and had to perform the task again by choosing a different host/credential/exploit.

CASE STUDY DESIGN AND PROCEDURE

In this case study, we aimed to identify specific biases such as default bias, recency bias, and availability heuristics. A hypothesized mapping of the target cognitive biases to various phases of the cyber kill chain process is shown in Figure 2. Participants played the role of attackers and performed the attack workflow in the CyberVAN testbed. Initially, participants that agreed to participate provided demographic information. Next, they were presented with task instructions regarding the goal of the task and the general procedure.

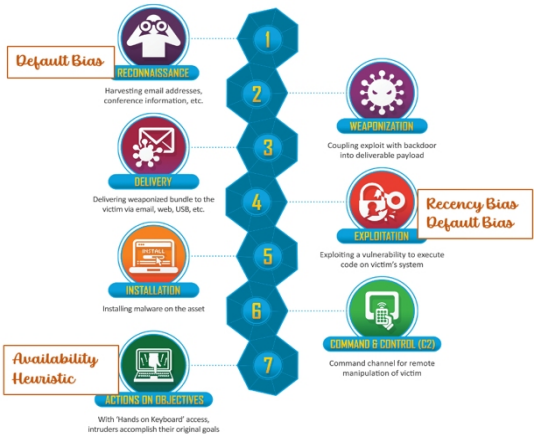

Figure 2: Cyber kill chain and bias mapping.

Participants were provided access to one of the machines on a network as shown in Figure 1 as the Attacker Host. They received information on the target network and their objective, which was to steal sensitive information from level 3. The participants were provided with a high-level description of vulnerabilities present at the host and a set of prepared shell scripts to exploit those vulnerabilities. Participants were asked to choose a vulnerability and run the corresponding script to, say, leak credentials from the target host and subsequently, use those credentials to log in to the host at level 2. Participants were asked to repeat the exercise 3 times and the environment was reset at the end of each run. After the completion of the main task, a feedback questionnaire was presented to the participants for them to provide feedback about the task and the strategies that they used during the task. Participants also responded to behavioral survey questions which collected information about their frustration, surprise, confusion state, etc.

Data Collection. In this experiment, we collected user asciinema i.e. a recording of the user's interaction with the terminal, including all the typed commands and the output produced by such commands, as well as timing information such as time to input a command, the time between output display and subsequent command, etc. Additionally, we collected gnome video recording of all the participant activities for additional validation. Network data and host data was also captured during the experiment session. Table 1 provide details about which specific data points were used for analysis.

Table 1. Data collected.

Level	Actions Recorded
Participant host	Chosen Host: 10.0.2.201-205 Chosen username from Hydra: Alice, Bob, and Carol Chosen username from shadow file: Alphabetically sorted list of users
Level 1 (Vulnerability Selection)	• CVE-2014-6271 – Shellshock: remote code execution • CVE-2023-23752 – Joomla: exposure of credentials • CVE-2017-12636 – CouchDB: remote code execution – Required most effort to setup and execute. • CVE-2023-28432 – Minio: exposure of credentials
Level 2 (Exploitation Method)	• Open-source tool - Required more effort but is reliable • Buffer Overflow shell script - Required less effort but is unreliable
Level 3	Data exfiltration method: SCP vs FTP

Participants. Six participants from a cybersecurity company with cyber expertise participated in this case study. Four participants identified as male, one participant identified as female, and the remaining participants did not

specify their gender. The participants had an average age of 40 years, with a standard deviation of 19.09. Two participants hold a Ph.D. degree, two possess a master's degree, one has a bachelor's degree, and one reported having a high school education. The average duration to complete this study was 120 minutes. We gathered data on participants' prior experience with computer usage, network operations, and programming. Five out of six participants indicated spending more than 8 hours on a computer per day while one participant reported 4–8 hours per day. Four participants were familiar with 5–10 programming languages, while two reported familiarity with 2–5 programming languages. Five participants possessed cybersecurity education, with three having completed more than 5 courses, one participant taking 3–5 courses, and another taking 1–2 courses in cybersecurity. Only one of the participants had no formal training in cybersecurity. Three participants reported over 10 years of experience, one participant reported 5–10 years of experience and two participants reported having 1–5 years of experience in cyber operations.

RESULTS

Actions performed by the participants were analyzed for evidence of cognitive biases based on the different network levels as shown in Table 1, which presents the recorded variables at each level. Results for each of the three studied biases (default bias, availability heuristic, and recency bias) are provided below.

Default Bias. When people are given a choice between several alternatives, there is a tendency to choose the default one (Johnson et al., 2020). The default could be decision made in last round, preference for a brand or a tool, or simply choosing first or last option. This is known as the default effect. At level 1, we calculated the selection of hosts in round 1 of the task. The assumption is that if the participants choose host 1 (10.0.2.201; first) or host 5 (10.0.2.205; last), that indicates the presence of default effect bias. Figure 3 presents the host selection behavior of participants. We observe that ~23% participants chose the first host 10.0.2.201 and ~38% participants chose the last available host i.e. 10.0.2.205. In total, ~60% of the participants selected either the first or the last IP address from the network scan result, representing an indication of default effect bias.

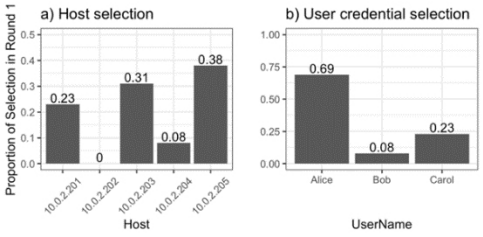

Figure 3: Evidence of default bias: a) participant's host selection in round 1 and b) participant's user name selection in all three rounds.

Participants were asked to use the Hydra tool to get login credentials to the selected host. The hydra output provided the list of user names and passwords in alphabetical order (Alice, Bob, and Carol). Similar to the host selection, we observed that participants chose the first username (i.e., Alice) from the list of cracked passwords more than 70% of the time (see Figure 3b). These observations provide strong evidence of the default bias.

Availability Heuristic. The availability bias occurs when people rely on readily available information or examples that come to mind easily when making judgments or decisions (Tversky and Kahneman, 1973). In this experiment, we measured the availability bias at level 1 and level 2. At level 1, the availability bias is attributed to the participant if they preferred simple/easy-to-execute options over complex options (options are different CVEs). This availability of easily accessible information can lead them to overestimate the effectiveness or likelihood of the simple option, thereby influencing their decision-making. Among the presented list of vulnerabilities, CVE-2017-12636 required the highest efforts, leading to more mental effort. As shown in Figure 4a, at level 1, we observed that participants rarely chose the vulnerability CVE-2017-12636 that required the highest effort to exploit i.e. (only chosen ∼13% of the time). Other vulnerabilities were exploited more, as required less effort and reduced complexity levels. Similarly, at level 2, we observed whether participants used a less reliable but easily available shell script over reliable open-source code that required additional effort to perform the attack. Participants chose the simpler buffer overflow shell script option more frequently (∼77% of the time) although it was less reliable than its alternative (see Figure 4b). These observations suggest the existence of availability bias in the attacker's decisions.

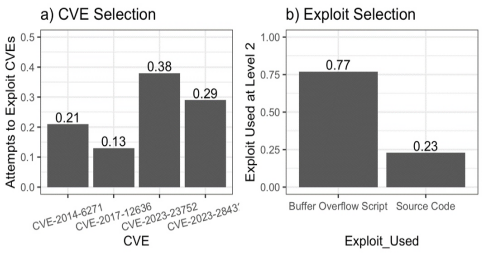

Figure 4: Evidence of availability heuristic: a) participant's selection of CVE in all three rounds and b) participant's selection of exploit at level 2.

Recency Bias. In this experiment, we measure the recency bias at level 2 and level 3. At level 2, the recency bias is attributed to the participant if they exploit recently discovered vulnerabilities in preference to old vulnerabilities. At level 3, participants are attributed with the recency bias if they choose a previously used command (e.g., SCP was used multiple times during level 1 and level 2) rather than a command that was not used recently (e.g., FTP). We observed that recently discovered vulnerabilities were exploited 67% of the time although they only made up 50% of the available vulnerabilities (2 out

of 4; see Figure 4a). These observations suggest the presence of recency bias. We also observed the likely effects of recency on the participants' actions. Specifically, at level 3, participants chose SCP over FTP (~70% of the time) as their preferred method for exfiltration (see Figure 5). Although SCP is the stealthier choice of the two methods, participants used SCP multiple times during level 1 and level 2 which may have primed their memory. Furthermore, few participants reported that they were more comfortable with SCP.

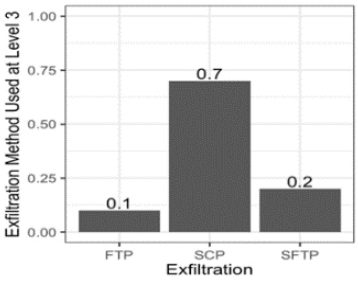

Figure 5: Evidence of recency bias: participant's selection of exfiltration method at level 3.

DISCUSSION AND CONCLUSION

Detection and identification of cyber attackers' cognitive biases is a challenging task. Such biases can manifest in different ways; e.g., an attacker who is loss-averse may exhibit behaviors such as tenacity (sticking to a target), while one with choice overload may exhibit an aggressive scanning style, and so on. This paper provides a case study to observe attacker's cognitive biases. We use a dataset collected in the network and provide objective rather than subjective analysis. We observed that the *strength* of the presence of a bias was related to whether the participants found a task to be interesting or not. We define the strength of the presence of a bias as the difference between the topmost and second-to-topmost choices made by the participants. If the difference was high, then we claim that there was a strong presence of bias. Based on the post-exercise questionnaire, the expert participants were curious about experimenting with different exploits for level 2 and different IP addresses for level 1 (both between rounds and within a round). Hence, we see that the difference between the topmost and next highest choices made by the participants for those tasks is small (Figures 3a and 4a). Curiosity has been found to battle cognitive biases (Kahan et al., 2017). On the other hand, mundane or common tasks such as copying files to a different system seem to indicate strong evidence of the presence of the corresponding bias (Figure 5). In other words, for cyber settings, it could be that mundane or common day-to-day operations may provide opportunities for observing System 1 biases while more interesting/rarely-practiced tasks such as exploitation may draw more curiosity and attention, and thus, provide opportunities for System 2 biases. Upon analyzing the responses from the post-exercise questionnaire, one of the participants reported specific patterns that were not part of the experiment which could indicate representativeness bias (representing another example of system 1 biases). According to the participant weak

passwords were associated with accounts that were being monitored when in reality no such monitoring policy was in place. This case study provides preliminary evidence suggesting the presence of cognitive bias in the attacker's actions based on observable data available to the network defenders. Certain behaviors noted in the initial case study exhibit overlapping biases. For instance, the selection of first and last host, attributed to default bias, also aligns with indicators of position bias. Similarly, opting for easy over the complex is linked to availability heuristic, as recalling simpler events or experience is easier compare to complex ones. This pattern of behavior could also indicate aversion to complexity. Moving forward, we aim to broaden the scope of this study to further explore cognitive biases influencing attackers' behavior and gain deeper insights.

ACKNOWLEDGMENT

The work reported in this paper was partially performed in connection with contract number W911NF-14-D-0006 with the U.S. Army Research Laboratory and Cooperative Agreement Number W911NF-13-2-0045 (ARL Cyber Security CRA).

REFERENCES

Aggarwal, P., Rubaiyet Nowmi, S., Du, Y., & Gonzalez, C. (2024). Evidence of Cognitive Biases in Cyber Attackers from An Empirical Study. Proceedings of the 57th Hawaii International Conference on System Sciences, 934.

Aggarwal, P., Jabbari, S., Thakoor, O., Cranford, E. A., Vayanos, P., Lebiere, C., Tambe, M., & Gonzalez, C. (2023a). Human-subject experiments on risk-based cyber camouflage games. Cyber Deception: Techniques, Strategies, and Human Aspects, 89, 25.

Aggarwal, P., Cranford, E. A., Tambe, M., Lebiere, C., & Gonzalez, C. (2023b). Deceptive signaling: Understanding human behavior against signaling algorithms. Cyber Deception: Techniques, Strategies, and Human Aspects, 89, 83.

Aggarwal, P., Thakoor, O., Jabbari, S., Cranford, E. A., Lebiere, C., Tambe, M., & Gonzalez, C. (2022). Designing effective masking strategies for cyber defense through human experimentation and cognitive models. Computers & Security, 117, 102671.

Aggarwal, P., Du, Y., Singh, K., & Gonzalez, C. (2021). Decoys in cybersecurity: An exploratory study to test the effectiveness of 2-sided deception. arXiv preprint arXiv:2108.11037.

Aggarwal, P., Thakoor, O., Mate, A., Tambe, M., Cranford, E. A., Lebiere, C., & Gonzalez, C. (2020). An exploratory study of a masking strategy of cyber deception using cybervan. Proceedings of the human factors and ergonomics society annual meeting, 64 (1), 446–450.

Aggarwal, P., Gonzalez, C., & Dutt, V. (2016). Looking from the hacker's perspective: Role of deceptive strategies in cyber security. 2016 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), 1–6.

Aljohani, A., & Jones, J. (2022). The pitfalls of evaluating cyber defense techniques by an anonymous population. HCI for Cybersecurity, Privacy and Trust: 4th International Conference, HCI-CPT 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings, 307–325.

Azzopardi, L. (2021). Cognitive biases in search: A review and reflection of cognitive biases in information retrieval. Proceedings of the 2021 Conference on Human Information Interaction and Retrieval, 27–37. https://doi.org/10.1145/3406522.3446023

Bao, T., Tambe, M., & Wang, C. (2023). Cyber deception techniques, strategies, and human aspects (1st ed., Vol. 1). Springer International Publishing.

Ben-Asher, N., & Gonzalez, C. (2015). Effects of cyber security knowledge on attack detection. Computers in Human Behavior, 48, 51–61.

Buchler, N., Rajivan, P., Marusich, L. R., Lightner, L., & Gonzalez, C. (2018). Sociometrics and observational assessment of teaming and leadership in a cyber security defense competition. computers & security, 73, 114–136.

Chadha, R., Bowen, T., Chiang, C.-Y. J., Gottlieb, Y. M., Poylisher, A., Sapello, A., Serban, C., Sugrim, S., Walther, G., Marvel, L. M., Newcomb, E. A., & Santos, J. (2016). CyberVAN: A cyber security virtual assured network testbed. MILCOM 2016–2016 IEEE Military Communications Conference, 1125–1130.

Cranford, E. A., Gonzalez, C., Aggarwal, P., Cooney, S., Tambe, M., & Lebiere, C. (2020). Toward personalized deceptive signaling for cyber defense using cognitive models. Topics in Cognitive Science, 12 (3), 992–1011.

Draws, T., Rieger, A., Inel, O., Gadiraju, U., & Tintarev, N. (2021). A checklist to combat cognitive biases in crowdsourcing. Proceedings of the AAAI Conference on Human Computation and Crowdsourcing, 9 (1), 48–59. https://doi.org/10.1609/hcomp.v9i1.18939

Dutt, V., Ahn, Y.-S., & Gonzalez, C. (2011). Cyber situation awareness: Modeling the security analyst in a cyber-attack scenario through instance-based learning. Data and Applications Security and Privacy XXV: 25th Annual IFIP WG 11.3 Conference, DBSec 2011, Richmond, VA, USA, July 11-13, 2011. Proceedings 25, 280–292.

Ferguson-Walter, K., Shade, T., Rogers, A., Trumbo, M. C. S., Nauer, K. S., Divis, K. M., Jones, A., Combs, A., & Abbott, R. G. (2018). The tularosa study: An experimental design and implementation to quantify the effectiveness of cyber deception. (tech. rep.). Sandia National Lab. (SNL-NM), Albuquerque, NM (United States).

Ferguson-Walter, K. J. (2020). An empirical assessment of the effectiveness of deception for cyber defense.

Gonzalez, C., Ben-Asher, N., Oltramari, A., & Lebiere, C. (2014). Cognition and technology.

Gutzwiller, R., Ferguson-Walter, K., Fugate, S., & Rogers, A. (2018). "Oh, look, a butterfly!" A framework for distracting attackers to improve cyber defense. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 62 (1), 272–276.

Gutzwiller, R. S., Ferguson-Walter, K. J., & Fugate, S. J. (2019). Are cyber attackers thinking fast and slow? Exploratory analysis reveals evidence of decision-making biases in red teamers. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 63 (1), 427–431.

Johnson, C. K., Gutzwiller, R. S., Ferguson-Walter, K. J., & Fugate, S. J. (2020). A cyber-relevant table of decision-making biases and their definitions. DOI: https://doi. org/10.13140/RG, 2(14891.87846).

Johnson, C. K. (2022). Decision-making biases in cybersecurity: Measuring the impact of the sunk cost fallacy to delay attacker behavior (tech. rep.). Arizona State University.

Kahan, D. M., Landrum, A., Carpenter, K., Helft, L., & Hall Jamieson, K. (2017). Science curiosity and political information processing. Political Psychology, 38, 179–199.

Klein, A. (1990). A direct test of the cognitive bias theory of share price reversals. Journal of Accounting and Economics, 13 (2), 155–166. https://doi.org/https://doi.org/10.1016/0165-4101(90)90028-3

Miah, M. S., Aggarwal, P., Gutierrez, M., Thakoor, O., Du, Y., Veliz, O., Singh, K., Kiekintveld, C., & Gonzalez, C. (2023). Diversifying deception: Game-theoretic models for two-sided deception and initial human studies. Cyber Deception: Techniques, Strategies, and Human Aspects, 89, 1.

O'Sullivan, E. D., & Schofield, S. (2018). Cognitive bias in clinical medicine. Journal of the Royal College of Physicians of Edinburgh, 48 (3), 225–232.

Shade, T., Rogers, A., Ferguson-Walter, K., Elsen, S. B., Fayette, D., & Heckman, K. E. (2020). The moonraker study: An experimental evaluation of host-based deception. HICSS, 1–10.

Simon, M., Houghton, S. M., & Aquino, K. (2000). Cognitive biases, risk perception, and venture formation: How individuals decide to start companies. Journal of Business Venturing, 15 (2), 113–134. https://doi.org/10.1016/S0883-9026(98)00003-2

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. science, 185 (4157), 1124–1131.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive psychology, 5(2), 207–232.