CRASTE: Human Factors and Perception in Cybersecurity Education*

Vita Santa Barletta[1,*,†], Danilo Caivano[1,†], Miriana Calvano[1,†], Antonio Curci[1,2,†] and Antonio Piccinno[1,†]

[1]University of Bari Aldo Moro, Department of Computer Science, Via Edoardo Orabona 4, Bari, Italy
[2]University of Pisa, Department of Computer Science, Largo B. Pontecorvo, 3 56127 Pisa, Italy

Abstract

Human interaction plays a key role in the achievement of cybersecurity goals. Addressing cyber threats necessitates an emphasis on human behavior and cognitive models, not merely relying on technical details, since individuals are the weakest link in the cybersecurity context. Thus, the design of cybersecurity-related training programs should be carried out accordingly to increase their effectiveness. The following research proposes a framework, called "CRASTE", which maps human factors and perception, the Red and Blue Team simulation and the Cyber Kill Chain to improve cybersecurity education with respect. The introduction of Artificial Intelligence (AI) in this process can foster the proper employment of the MITRE ATT&CK, which is the most used knowledge base in cybersecurity, to present how the Large Language Models (LLMs) can support both Red and Blue Teams during attacks and their defense.

Keywords

Cybersecurity, Education, Human Factors, Perception, Kill Chain

1. Introduction

Cybersecurity plays a crucial role in today's era of technological innovation and the burgeoning digital economy. Threat actors pose risks to individual safety in terms of the integrity of their intellectual property by conducting attacks on several levels, such as illicit sales on the dark web or leveraging for ransom demands [1]. In the domain of cybersecurity, there are several frameworks that encompass the phases, methodologies, and techniques of attacks that aim at providing a standardized approach to facing attacks and solving issues. For example, the Cyber Kill Chain (CKC) gathers all the processes of an attack, grouping them in 7 phases. A more technical approach, instead, is provided by the MITRE ATT&CK which is a globally-accessible knowledge base of adversary tactics and techniques based on real-world observations [2]. Employing these frameworks, which will be explored below, supports facing cyber threats with a holistic approach that goes beyond technical aspects, emphasizing human behavior, cognitive models, and awareness [3].

This implies that, in this field, human's characteristics cover an important role since most of the security challenges comes from their behaviours and skills. Thus, it is necessary to explore human behavior when considering cyber threats to the mission, the mission-enabling infrastructure against which attacks occur, the human defenders' operational processes, and the roles that humans play in cyberspace operations [4, 5].

The integration of Artificial Intelligence (AI) in any field can boost productivity while providing humans with enhanced skills and abilities thanks to the high computational power [6, 1]. In recent years, Large Language Models (LLMs) have quickly and significantly spread, being used by millions of

DAMOCLES'24: First International Workshop on Detection And Mitigation Of Cyber attacks that exploit human vuLnerabilitiES, AVI '24, Arenzano (Genoa), Italy, June 3rd 2024
*Corresponding author.
†These authors contributed equally.
vita.barletta@uniba.it (V. S. Barletta); danilo.caivano@uniba.it (D. Caivano); miriana.calvano@uniba.it (M. Calvano); antonio.curci@uniba.it (A. Curci); antonio.piccinno@uniba.it (A. Piccinno)
0000-0002-0163-678 (V. S. Barletta); 0000-0001-5719-7447 (D. Caivano); 0000-0002-9507-9940 (M. Calvano); 0000-0001-6863-872X (A. Curci); 0000-0003-1561-7073 (A. Piccinno)

individuals. The motivation behind this lies in the fact that they are particularly effective in humanizing technology and addressing the mechanization of bottlenecks, implying an improvement of human factors and perception in the interaction of humans with any kind of technology [7]. These elements become relevant in cybersecurity because it heavily relies on humans, their cognitive skills, and their attitude toward reality. The goal of this research work is to investigate how to improve cybersecurity education through an adequate design of the CKC that highlights human factors and perception. Additionally, to understand how to further improve this process the integration of LLMs in the MITRE ATT&CK is also explored.

2. Background

Cybersecurity is an intricate domain and it is influenced by multiple factors. In this section, background concepts and additional context is provided concerning human-factors and perception, the Red and Blue teams, and the CKC.

2.1. Human Factors & Perception

The way that individuals interact with computers and make decisions is a dynamic and intricate issue, encompassing numerous factors [8]. Software solutions are not a one-size fits-all solutions because individual differences, personality traits, and cognitive abilities play pivotal roles in the requirements of any systems; this aspect influences cybersecurity as well: biases and heuristics shape risk perception. Both risk perception and individual differences are also affected by the environment in which they occur; thus, it is necessary to emphasize the critical role of human factors in cybersecurity training [9]. Corradini and Nardelli, for example, stress the need for tailored training programs and digital awareness interventions, respectively, to address the human element in cybersecurity [10, 11]. Neigel et al. further underscore the importance of individual differences, such as trust in technology and intrinsic motivation, in shaping cyber hygiene knowledge and behavior [12]. Instead, Thackray et al. suggest that integrating social psychology into cybersecurity education can enhance communication and understanding of cyber risks [13]. These findings collectively highlight the need for a holistic approach that considers the human element and the perception in cybersecurity education.

2.2. Cyber Kill Chain (CKC)

The CKC is one of the defense models designed to help companies and large organizations mitigate the most advanced cyber-attacks. The goal is to compromise a particular asset that contains specific data or valuable information [14, 15]. A characteristic to keep in mind of this attack is that the attacker's actions are performed over a considerable time.

The phases of the Kill Chain improve visibility into an attack and enrich the analyst's understanding of the adversary's tactics, techniques and procedures. It consists of 7 phases [16, 17]:

1. *Reconnaissance*: the goal of this phase is to collect information about the victim and to understand which are the most appropriate actions to perform during the attack; This process is also called "footprinting" since at the end it is possible to obtain a detailed "snapshot" of the target.
2. *Weaponization*: it is the preparation and staging phase which has the objective to define a penetration plan utilizing the information gathered from the previous stage.
3. *Delivery*: it is the malware transmission and delivery phase. It is expected that the victim downloads and/or executes malicious files or visits malicious web pages.
4. *Exploitation*: in this phase the vulnerability previously found are exploited by the attacker to obtain access to the target system and conduct the attack.
5. *Installation*: in this phase the backdoor or any equivalent systems is installed on the victim's device to guarantee the attacker persistent access in time.
6. *Command & Control (C2)*: in this phase the victim is manipulated by the attacker through the malicious code previously installed.

7. *Actions on Objectives*: in this phases the attacker execute the attack reaching their objective. Typically, the attackers aim to perform data exfiltration which involves collecting, encrypting and extracting information from the victim environment.

2.3. Red Team vs Blue Team Approach

The cybersecurity field can be described referring to two different and opposite perspective: *Red Team* and *Blue Team* which represent the attack and defence side respectively. The *Blue Team* is "the group responsible for defending an enterprise's use of information systems by maintaining its security posture against a group of mock attackers" [18]; the *Red Team* is "a group of people authorized and organized to emulate a potential adversary's attack or exploitation capabilities against an enterprise's security posture" [19].

This approach is employed from organizations to deeply understand how to face a cyber threat being able to "think like an attacker"; it brings benefits in terms of risk assessment, detection, and evolution of threats [20, 21].

The success of the *Blue Team vs. Red Team* approach hinges on interaction and mutual feedback. The primary goal is to refine an organization's detection and response capabilities. Through this collaboration, it is possible to increase awareness of attack techniques and bring to light the vulnerabilities in the attacker's defense infrastructure. It is important to highlight that when a Security Operations Center (SOC) fails to detect an intrusion, it is not always due to operator preparation or technological inefficiency. Instead, the attacker's success may result from the ineffectiveness of controls against sophisticated, previously unknown techniques. Therefore, the actions of the Red Team are functional in exposing these control deficiencies, preventing them from being exploited to cause real damage [22]. For this reason, a final report must be created to highlight the details concerning all the aspects related to the attack (e.g. how the breach occurred, the timeline of the attack, the details of the vulnerabilities that were exploited to gain access, the business impact to the company, etc.) [23].

3. CRASTE: Human Aspects in Cyber Kill Chain

In this section, CRASTE is presented. It is a framework concerning the integration of human factors and perception at various stages of the CKC. The goal is to understand how such human aspects can improve training in cybersecurity and the Red/Blue simulation.

Human Factors They *refer to environmental, organisational and job factors, and human and individual characteristics, which influence behaviour at work in a way which can affect health and safety* [24]. This concept includes three main aspects that must be considered:

- *Job*: the nature of the task, workload, the working environment, the design of displays and controls, and the role of procedures (**What** people are being asked to do).
- *Individual*: this aspect includes his/her competence, skills, personality, attitude, and risk perception. Individual characteristics influence behaviour in complex ways (**Who** is doing it).
- *Organization*: the work patterns, the culture of the workplace, resources, communications, leadership and so on (**Where** they are working).

Perception It is a subjective process influenced by various factors which shape how individuals interpret and comprehend their surroundings. These factors influence the selection, organization, and interpretation of sensory information, resulting in diverse and unique perceptions among people. For this reason, it is essential to design a learning and playful environment that integrates different aspects of cybersecurity education; in this way individuals can be allowed to expand their knowledge beyond technical aspects and consider human factor risks [25, 3].

Table 1 presents the human factors and perceptions identified in the previous work [3] considering which phase of the kill chain they affect and which perspective (i.e. Blue and/or Red Team). Each

element and its description is provided in Section 4 referring to a broader context, which is represented by the MITRE ATT&CK framework, along with the influence of LLMs in this context.

Table 1
CRASTE: Cyber Aspects in Kill Chain

Element	Aspect	Kill Chain Phase	Team
E1	Human Factor/Perception	Reconnaissance/Delivery	Red
E2	Perception	Weaponization/Installation	Red
E3	Human Factor	Exploitation/Installation	Red
E4	Human Factor	Reconnaissance/Delivery/Action	Blue
E5	Perception	Reconnaissance/Weaponization/Delivery	Blue
E6	Perception	Installation/C2/Actions	Blue
E7	Human Factor/Perception	C2/Action	Red/Blue
E8	Perception	Weaponization/Exploitation	Red/Blue
E9	Perception	All phases	Blue

4. LLMs role in the MITRE ATT&cCK

This section aims at gathering insights into the influence of AI in cybersecurity by considering the MITRE ATT&CK framework, human factors and perception. The difference between the MITRE ATT&CK and the CKC lies in the fact that the first is a knowledge base that encompasses all the elements, tools, techniques, and procedures that belong to cyber-attacks [2]; the second instead, is a more general framework that gathers all the processes that lead to a successful execution of an attack. The MITRE ATT&CK diverges from this sequential model, since it prioritizes aiding security professionals in identifying and addressing individual adversary tactics and techniques as they manifest in various contexts.

The integration of AI can bring significant advantages because it can provide details, explanations, and suggestions concerning aspects that humans might not detect by themselves. It can strongly influence human factors and perception of users while being involved in an attack from both perspectives (i.e., Red and Blue Team). The correlation between the elements presented in Table 1 with the specific role of LLMs is discussed below.

(E1) Social Engineering This element of the framework exploits human psychology to manipulate individuals in order to make them perform actions that are in favor of the attacker. The integration of LLMs in this component becomes crucial in the *Reconnaissance* phase, since it can allow humans to receive easily-understandable information concerning the threat target and suggest a suitable delivery method for the attacker based on their expertise; it influences their perception of the attack's success and their own abilities.

Human factors are influenced in terms of individual skills, in fact, when a red team consists of multiple individuals, they are all influenced by where the operation takes place and the tasks assigned to others to carry out the attack.

(E2) User Awareness It refers to the perception of risks and humans' understanding of safe practices. In this case, for *Weaponization*, a LLM can suggests techniques for the attack design and tools. During *Installation*, an AI agent can help in the implementation of a strong and long-lasting backdoor. Perception is influenced by the success or failure of implementing these suggestions. Humans can utilize AI to understand ongoing situations (both from blue and red team perspectives) - user awareness.

(E3) Usability and Security Trade-Offs Lack of usability can lead to issues in the effective communication between humans and any kind of system. This implies that users might attempt to find

workarounds to ignore security protocols or warnings that could prevent them from falling into dangers. In this case, a LLM tool can suggest exploit and malware installation methods, leveraging usability and security. It influences red team human factors in terms of skills and support for conducting an attack.

(E4) Organizational Culture This element encompasses the array of values, anticipations, and behaviors that direct and shape the behavior of every team member. An organization that allows the employment of LLMs when it comes to cybersecurity can allow individuals to improve their understanding of dangerous situations, providing support in obtaining explanations regarding notions or techniques. The blue team can improve its understanding reconnaissance, delivery, and action methods. Organizational culture is considered because individuals within the blue team behave according to their skills, workplace environment, and organizational influences.

(E5) Risk Assessment Being able to appropriately assess risks and threats implies possessing the right judging skills in order to find the proper danger levels. The blue team, with the help of AI, understands how attacks are executed to perform risk assessments. Perception is influenced in terms of the ability to manage, confront, and evaluate situations.

(E6) Risk Detection It gathers the processes and actions to identify concealed threats inside a network or system and responding to them. AI assists the blue team in risk detection to understand how attacks are executed. It affects how the blue team interprets and evaluates the attack, consequently affecting their ability to counter it. For example, an individual can ask a LLM model to interpret or analyze an email in order to understand if it is phishing. The LLM can provide human-like explanations, fully understandable even by non-experts.

(E7) Incident Response It refers to how individuals respond to security incidents and how to report them. AI helps the red team understand how the blue team might counter the attack, and the blue team in actually countering it. Human factors and perception are influenced. Human factors include personal skills and their impact on their work and organization (from both perspectives); perception is influenced by the success of the attack (red team) and the response skills (blue team). Perception in terms of difficulty.

(E8) Response to Threat It refers to security threats and incidents that have actually happened. Referring to the timing of the attack, AI can signal real-time attacks for the blue team. For the red team, AI can monitor the attack's progress and assess if their objectives are met. It influences the blue team's perception of how the attack is interpreted and the stress it causes, while the red team focuses on the attack's success and effects.

(E9) Compliance It ensures an organization's security measures meet regulatory standards and guidelines. These standards are designed to protect the integrity, confidentiality, and availability of sensitive data from various cyber threats. With AI, the blue team can evaluate if security measures comply with regulations. Thus, the blue team's perception refers to how they perceive their security measures based on whether the attack was successful or not.

In conclusion, the employment of LLMs with the MITRE ATT&CK can bring positive implications on cybersecurity education. Through the explanations and indications provided by LLMs, individuals are enabled to receive real-time and personalized feedback that can be adapted to specific threats and dangerous situations. In this way, individuals can avoid feeling lost when facing cybersecurity risks, especially in case of low levels of expertise.

5. Conclusion

This paper analyzes how human factors and perception, integrated in the CKC, can support the activities concerning the attack and defence perspective. This aspect is further investigated considering a broader context concerning the MITRE ATT&CK framework along with the impact that LLMs can have on human's behaviors and feelings. By integrating AI functionalities it is possible to improve the training process of both perspective having, in some cases, a real-time feedback and suggestion. The *CRASTE* framework maps the elements of human factors and perception at the various stages of the CKC to improve cybersecurity training.

Future work concerns the execution of experiments of this approach in university courses to compare how the integration and analysis of these aspects can improve the knowledge about attacks (Red Team) and defense (Blue Team). It is also intended to introduce gaming elements in the learning process to increase student's knowledge and skills by recreating red and blue team simulations through the application of gamification and serious games.

Acknowledgments

This study has been partially supported by the following projects: SSA - "Secure Safe Apulia - Regional Security Center" (Codice Progetto 6ESURE5) and SERICS - "Security and Rights In the CyberSpace - SERICS" (PE00000014) under the MUR National Recovery and Resilience Plan funded by the European Union - NextGenerationEU. The research of Antonio Curci and Miriana Calvano is supported by the co-funding of the European Union - Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 – Partnerships extended to universities, research centers, companies, and research D.D. MUR n. 341 del 15.03.2022 – Next Generation EU (PE0000013 – "Future Artificial Intelligence Research – FAIR" - CUP: H97G22000210007).

References

[1] V. S. Barletta, F. Cassano, A. Pagano, A. Piccinno, New perspectives for cyber security in software development: when end-user development meets artificial intelligence, in: 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2022, pp. 531–534. doi:10.1109/3ICT56508.2022.9990622.

[2] T. M. Corporation, Mitreatt&ck, https://attack.mitre.org/, 2024.

[3] V. S. Barletta, M. Calvano, F. Caruso, A. Curci, A. Piccinno, Serious games for cybersecurity: How to improve perception and human factors, in: 2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 2023, pp. 1110–1115. doi:10.1109/MetroXRAINE58569.2023.10405607.

[4] N. L. Crabtree, J. A. Orr, Cyber red/blue and gamified military cyberspace operations, Lincoln Lab. J. 23 (2019) 1–12.

[5] M. Calvano, F. Caruso, A. Curci, A. Piccinno, V. Rossano, A Rapid Review on Serious Games for Cybersecurity Education: Are "Serious" and Gaming Aspects Well Balanced?, in: Proceedings of the 9th International Symposium on End-User Development (IS-EUD 2023), volume 3408, CEUR Workshop Proceedings, Cagliari, Italy, 2023. URL: https://ceur-ws.org/Vol-3408/short-s3-05.pdf.

[6] V. S. Barletta, F. Cassano, A. Pagano, A. Piccinno, A collaborative ai dataset creation for speech therapies, volume 3136, 2022, p. 81 – 85. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85130701005&partnerID=40&md5=e5349d9fedc09b849dc4fc1c8debf3be, cited by: 7.

[7] P. Kumar, Large Language Models Humanize Technology, 2023. URL: http://arxiv.org/abs/2305.05576, arXiv:2305.05576 [cs].

[8] V. S. Barletta, F. Caruso, T. Di Mascio, A. Piccinno, Serious games for autism based on immersive virtual reality: A lens on methodological and technological challenges, in: M. Temperini, V. Scarano, I. Marenzi, M. Kravcik, E. Popescu, R. Lanzilotti, R. Gennari, F. De La Prieta, T. Di Mas-

cio, P. Vittorini (Eds.), Methodologies and Intelligent Systems for Technology Enhanced Learning, 12th International Conference, Springer International Publishing, Cham, 2023, pp. 181–195.
[9] K. Parsons, A. McCormac, M. A. Butavicius, Human factors and information security : Individual , culture and security environment executive summary, 2011.
[10] I. Corradini, E. Nardelli, Building organizational risk culture in cyber security: The role of human factors, Advances in Intelligent Systems and Computing (2018).
[11] I. Corradini, E. Nardelli, Developing digital awareness at school: A fundamental step for cybersecurity education, in: International Conference on Applied Human Factors and Ergonomics, 2020.
[12] A. R. Neigel, V. L. Claypoole, G. E. Waldfogle, S. Acharya, G. M. Hancock, Holistic cyber hygiene education: Accounting for the human factors, Comput. Secur. 92 (2020) 101731.
[13] H. Thackray, J. McAlaney, H. Dogan, J. Taylor, C. Richardson, Social psychology: An under-used tool in cybersecurity, in: British Computer Society Conference on Human-Computer Interaction, 2016.
[14] T. Yadav, A. M. Rao, Technical aspects of cyber kill chain, in: Security in Computing and Communications: Third International Symposium, SSCC 2015, Kochi, India, August 10-13, 2015. Proceedings 3, Springer, 2015, pp. 438–452.
[15] V. S. Barletta, D. Caivano, M. De Vincentiis, A. Pal, F. Volpe, Automotive knowledge base for supporting vehicle-soc analysts, in: 2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), 2023, pp. 960–965. doi:10.1109/MetroXRAINE58569.2023.10405622.
[16] E. M. Hutchins, M. J. Cloppert, R. M. Amin, et al., Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains, Leading Issues in Information Warfare & Security Research 1 (2011) 80.
[17] M. J. Assante, R. M. Lee, The industrial control system cyber kill chain, SANS Institute InfoSec Reading Room 1 (2015) 2.
[18] N. I. of Standards, Technology, Red team/blue team approach, 2012. URL: https://csrc.nist.gov/glossary/term/red_team_blue_team_approach.
[19] N. I. of Standards, Technology, Red team team, 2015. URL: https://csrc.nist.gov/glossary/term/red_team.
[20] G. White, A. Conklin, The appropriate use of force-on-force cyberexercises, IEEE Security Privacy 2 (2004) 33–37. doi:10.1109/MSP.2004.58.
[21] T. G. Malone, R. E. Schaupp, The" red team": forging a well-conceived contingency plan, Air & Space Power Journal 16 (2002) 22.
[22] M. T. Baldassarre, V. S. Barletta, D. Caivano, D. Raguseo, M. Scalera, Teaching cyber security: The hack-space integrated model, in: Italian Conference on Cybersecurity, 2019. URL: https://api.semanticscholar.org/CorpusID:59615981.
[23] Y. Diogenes, E. Ozkaya, Cybersecurity-attack and defense strategies: Infrastructure security with red team and blue team tactics, Packt Publishing Ltd, 2018.
[24] Health, S. Executive, Human factors in the management of major accident hazards. Introduction to human factors, Technical Report, 2005.
[25] P. van Schaik, D. Jeske, J. Onibokun, L. Coventry, J. Jansen, P. Kusev, Risk perceptions of cyber-security and precautionary behaviour, Computers in Human Behavior 75 (2017) 547–559. doi:https://doi.org/10.1016/j.chb.2017.05.038.