Human Factors in Cybersecurity, Vol. 91, 2023, 72–82
https://doi.org/10.54941/ahfe1003721

Social Engineering Penetration Testing Within the OODCA Cycle–Approaches to Detect and Remediate Human Vulnerabilities and Risks in Information Security

Erfan Koza[1], Asiye Öztürk[1], and Michael Willer[2]

[1]CLAVIS Institute for Information Security, Mönchengladbach, Germany
[2]Human Risk Consulting GmbH, Kassel, Germany

ABSTRACT

In more than 95% of all successfully conducted cyberattacks, the human factor is exploited as a vulnerability point. The following principle applies. Whenever a hacker uses external attack vectors and thus does not directly use the Internet as a medium, employees become the target of the attack. As a result, the current technical and intelligent defense mechanisms can only contribute to a limited extent to the increase the resilience of IT systems, as these technological approaches do not fully account for the behavioral, cognitive, and heterogeneous motivations that lead to human error in the security causal chain of information security using social engineering (SE) methods. In this paper, we present a strategic and iterative analysis tool to detect SE threats through systemic monitoring, to train and successfully defend against them. For this purpose, we use the so-called Course of Actions to practically check the security-compliant behavior of employees and to initialize the feedback processes for reactivating the human firewall based on the knowledge gained. This approach is already being applied to various types of organizations and critical infrastructure and can be seamlessly integrated into existing training and auditing programs.

Keywords: Social engineering, Human factors, Human firewall, Information security, Cyber threats

INTRODUCTION

The increasing system complexity, heterogeneity, and connectivity of IT and Operational Technology (OT) systems have led to an increasing number of internal and external vulnerabilities and attack surfaces, which can be attacked, especially by exploiting human vulnerability (Hughes-Lartey et al. 2021 | Widdowson et al. 2015). Owing to the deep penetration of ICT in almost all areas of a company's value chain, the human aspects of information security (IS) can be defined as core elements of security-related considerations. Given this interpretation, an appropriate level of IS can only be achieved if the

conceptual thinking and decision-making processes of strategists/decision-makers take up such a collective view and integrate the weighting of the sociological, psychological, and human characteristics of system users in terms of their expertise and risk awareness into their cognitive decision-making processes. As a result, human factors are an irreplaceable addition to the security chain and can help prevent incidents in terms of prevention and detection.

The focus is on role- or group-specific targeting of threat scenarios and awareness activities, as well as efficiently identifying optimization measures to effectively activate the human firewall (Koza, 2022a). Although a wealth of awareness and training models and offerings already exists, targeted and, above all, flexible decision-making and defense models are required in the sense of prioritization and decision-making strategies that can be dynamically tailored to companies. Approaches such as the 'man with the hammer syndrome' ultimately show that simple and artificially based simulation procedures, such as automated phishing or vhishing procedures (click and buy), are only of limited help because the explicit training of employees in the human emotional context and the knowledge factor are left out. This mindset, coupled with a lack of evaluation processes, namely whether IS training has led to visible and practicable security-compliant behavior, leads to decision makers repeatedly using the same tools, such as simple automated simulation gamification procedures and as a result, developing a deceptive sense of security and thus being unable to cope with the volatility, complexity, and dynamics of emerging attacks. For example, this results in the need for SE pen testing to discover potential vulnerabilities and root causes that could not have been discovered, and consequently optimized and eliminated without this set of tools. Causes include, for example, inadequate communication as an obstacle, mindset, lack of safety culture, low fault tolerance, organizationally unfavorably designed operational structure and process organization, which automatically lead to pressure situations and increased stress factor, misinterpretation of IS as not being a collective task of all employees in the organization, behavioral intention, ignorance etc. The objective of the present paper is to embed two practical interlocking strategic approaches in an OODCA (Observe-Orient-Decide-**Check**-Act) cycle defined for this purpose, which allows for sustainable, efficient, and transparent human threat awareness and mitigation in the context of IS (Koza, 2022b). Thus, the OODCA cycle represents a role- and issue-based practical approach to developing situational awareness. This approach was conceptualized based on Boyd's OODA-loop (Observe-Orient-Decide-Act) (Boyd, 1976). The defined OODCA-loop is used to identify and monitor person-specific threat scenarios and helps individual decision makers to monitor volatile threat vectors to design a targeted and sustainable strategic training and awareness program for employees. The OODCA cycle focuses on the role- and issue-based interactivity, dynamics, and diversity of human-based cyber threats such as fraud, vishing, phishing, pharming, scareware, tailgating, pretexting, Open-Source Intelligence (OSINT), and Social Media Intelligence (SOCMINT) and allows efficient mapping of existing roles to potential threats. As a check tool, the OODCA cycle uses social engineering pen-testing methods. Here,

social engineering pen testing serves as a review instance and audit tool for identifying human vulnerabilities and reviewing sensitization measures. Furthermore, how to use the discovered vulnerabilities in the post ex phase by taking appropriate measures to the act phase to eliminate the discovered vulnerabilities. Thus, in this context, the human firewall is understood as an integral and indispensable part of the holistic approach to increasing the resilience of IT systems.

Methodology

To enable scientific viability and better classify the presentation of the achieved results, we define two successive research branches.

The first research branch involves the creation of a systematic and iterative strategic approach model for locating and tracking human vulnerabilities and hazards and disseminating this information to relevant stakeholders. Here, the focus is on dynamism and precision to enable aspects of resilience as well as aspects of flexibility and rapid action and response by decision makers. These aspects, when executed adequately, lead to the desired resilience to both act and react to changing and volatile security situations. Thus, based on the identified awareness needs, which are simultaneously derived from the identified cyber threat intelligence data in the Observe phase, to assign the detected threats to individual employees or even employee clusters (e.g., tailgating to security personnel in facility management) with effective awareness training measures specifically related to the knowledge factor. This means that, by using the modelled framework, CISOs can train employees on real-world threats instead of abstract and unrealistic scenarios.

The second branch of research deals with continuous improvement and defines the verification tool "Social Engineering pen test." Staying alert to the current threat situation and sensitizing employees to social engineering attacks is therefore not effective from a practical point of view if they are only trained unilaterally without determining whether the efforts and activities carried out and campaigns in IS awareness have achieved their objectives, and if not, what is the cause of this non-achievement in terms of root cause analysis. To pay attention to these possible impacts, the SE human auditing and maturity pen test is used to verify the effectiveness of training and awareness activities according to the detected potential and prioritized attack types and methods. The results (findings) of the SE pen test will be analyzed later and implemented based on inferences in terms of post-ex activities to optimize the achieved maturity level. In the following section, the two research branches of the OODCA cycle are presented concisely with their respective instruments.

Observe-Orient-Decide

To reduce the complexity of the representation of the OODCA cycle and to assign the individual artifacts to the corresponding model phases, the first

three phases "Observe-Orient-Decide" followed by "Check-Act" are listed and assigned to the corresponding artifacts and methodology.

For the concretization of the Observe phase, the roles and corresponding threat scenarios, including digital and analog threat vectors, must first be identified. In this process step, all roles that have a direct and indirect influence on the basic values of the IS availability, integrity, and confidentiality must be embedded in the scope according to the failure criticality (e.g., executives, managers, IT security personnel, system/network administrators as super/normal users, external parties with access to critical assets).

In the second step, observation takes place to assign the selected roles to the real requirement scenarios This merging of hazards and roles is documented in the role threat table. The role-threat table is used as a cyber threat intelligence tool in the Observe phase and assessed in the orientation phase. Each volatile and dynamic change at the observation level triggered a new observation and orientation. For example, hybrid warfare, disinformation campaigns, and information warfare may be necessary on a daily or even hourly basis. This makes it possible to visualize the threat situation in a transparent and dedicated way, so that awareness and training activities can be carried out not based on abstract threat vectors but based on real threat vectors that can be executed in the wild. This visualization gives cybersecurity engineers an order in their perception and orientation processes to better interpret their knowledge.

The relevance of observation and orchestration follows from the following considerations. The art of espionage by human sources is sometimes one of the primal instincts of human species to protect themselves through information gained or to gain their own advantage from it. A modern human hacker still uses some of the same psychological manipulation as his predecessors did hundreds of years ago. However, with a significant difference. The information overload to which CISOs and cybersecurity engineers are exposed daily through numerous channels makes it difficult to separate the supposedly "important" from the "unimportant" information. This selection takes place through the highly subjective human perceptual filter. However, the perceptual filter can be trained specifically to train the necessary salience through proper selective perception, and ultimately, to be able to perceive the most important stimuli from the environment. However, the flood of information channels is also growing with the digitization and interconnectivity of the system via the medium of the Internet. The relevance of the targeted training of the perception filter can now also be better described with the following example: Intensive exposure to a topic has an essential effect on the perception filter. For example, searching for a new car is convertible. Less hours later, one perceives a multitude of convertibles on the streets. Of course, there were not fewer convertibles on the streets before, but for selective perception and especially for the perceptual filter this information and stimuli were not important before; therefore, one did not perceive them consciously (Willer, 2022).

In addition to the perception filter, time also plays an important role. Employees are driven by their tasks, targets, goals, expectations, and wishes.

Owing to the abundance of information that one must process and consume daily; one does not have much choice but to decide quickly. Therefore, questions such as, what is important for me and what is not? What is safe, and what is not? play an essential role. This creates a desire for more certainty in action to make the right decisions. However, this paradigm has a direct link to the dangers of social engineering. Anyone who does not take the time to check the perceived stimuli and information for their true content and manipulation content easily becomes a victim. To be able to protect oneself from manipulation in the context of human hacking, one needs basic knowledge of what human hacking is and what explicit danger vectors exist and how to classify them for my organization.

Therefore, observation and orientation take place at both the analog and digital levels. If one believes the quote of Bruce Schneier, the U.S. American author and expert for cryptography and computer security: "If you think technology can solve your security problems, then you don't understand the problems and you don't understand the technology," then the logical conclusion follows that at the beginning and at the end of the security consideration, the human factor is always the decisive factor (Willer, 2022).

In the Decision phase, IS awareness training can be derived based on the identified information and prioritization levels (Decide). Depending on relevance, different IS awareness plans can be developed and their effectiveness and efficiency validated using SE pen tests.

Check-Act

An SE pen test identifies vulnerabilities in a socio-technical environment (human-machine interaction) and puts the defined IS and IT security concepts, as well as the IS awareness measures and their effectiveness and efficiency in the test. Thus, SE pen test serves as an additional measurable indicator of the IS awareness level of an organization. From a technical point of view, SE pen test represents a selective vulnerability analysis of the human factor in the IS and serves as a set of tools for auditing and simulating realistic SE attack types. The main logical attack types are phishing, vhishing, pharming, SOCMINT, OSINT, and pretexting, as well as analogous physical attack vectors of Face2Face communication such as tailgating. Regardless of the two basic types of SE pen test (physical/purely logical digital SE pen test), the focus is on the following 4-modular process steps.

Pre-Preparation Phase: Detailed Planning

In this phase, the framework conditions for the SE pen test were determined. Clear objectives are defined, such as enticing employees to click on a link in a phishing e-mail, entering access data on simulated login pages (pharming), disclosing information, recording misinformation and passing it internally, handing over or sending goods without authorization, transferring money, changing master data, gaining access to the customer's property, and penetrating any special protection zones and (high) security areas unnoticed in order to simulate a physical or logical attack there. In addition to the objectives,

procedures must also be clearly discussed in advance, such as the behavior of pen tester in the event of reconnaissance by employees or the security service. In the case of purely digital SE pen tests, the authorities responsible for the IS should always be informed in advance of exactly when an attack will be carried out, so that in the event of an emergency they can always distinguish between the simulated SE pen test and a possible real attack. In further detailed planning, other actors, and stakeholders from cross-cutting areas such as data protection, compliance, and the works council should also be integrated into the procedure in an informative manner to reconcile any questions and expectations.

Preparation Phase: Exploration

In the preparatory phase within a physical SE pen test, an on-site analysis is carried out to obtain a picture of the target company directly on site. Among other things, physical and environmental security perimeters, procedures, routines, responsibilities play a major role in the development of the actual "attack scenarios." In all digital SE pen tests, specific initial analyses are operationalized via the OSINT and SOCMINT methods to enable targeted and precise information gathering to form the basis for developing the attack scenarios. As a result of the preparation phase, Courses of Action (COAs) are defined, in which the nature of attack scenarios is specified and made more precise.

Attack Phase: Reconnaissance

In the attack phase, the actual attack is carried out based on the fundamentals and COAs developed in the previous phases. In a physical SE pen test, legends, i.e., identities, are developed for this purpose, which are intended to enable the SE-pen tester to penetrate the target object unnoticed and gain unrestricted freedom of movement. SE pen tester then attempts to gain access to the IT system by means of a physical attack vector. Depending on the agreement with the target organization, removable media such as USB sticks, physical keyloggers, WLAN sniffers and wiretaps (fuzzy) are used. Analog information is also sifted through by the SE pen tester on-site, recorded, and, if necessary, purloined (photographed). The physical SE pen test (variant I) is completed with an undetected and successful exfiltration of the target object. In the case of a digital SE pen test, the respective COA is always completed when it has either achieved one of the defined objectives, initiated or supported another COA, or the latter is repelled. Within variant II, the physical SE pen tests end after the actual attack simulation is executed with the initialization of impulses for IS awareness formation. For example, at the end of an SE pen test, one can move more subtly within the target company until finally being approached by an employee (deliberate tactical detection). Now the SE pen tester can persist in sticking to his legend, so that the employee ends up having to send a message and report an incident. Here, in addition to the employee reaction, the focus is on the functionality and efficiency of reporting paths, reporting chains, alerting plans and accessibility. However, the SE

pen tester can also resolve the situation, explain what is being performed to the employee, and transition into active real-time awareness training. This approach can also be performed in the same analogy for logical SE pen tests, in which, for example, the vishing attack is used to unauthorizedly determine the most neuralgic information possible or to force misbehavior. In variant II of a logical SE pen test, the SE pen tester poses as an employee of a known company, for example, and usually asks for help with an urgent operational problem. However, the SE pen tester promptly resolves the situation, briefly introduces himself, explains the purpose of the attack simulation, and thus becomes a social engineering awareness trainer.

Post Phase: Follow-Up and Discussion of Results

This phase forms the conclusion of an SE pen test and includes a detailed execution and results description of the scenarios performed, as well as the SE-pen testers approach. Each attack is documented in writing in a situation description and all identified vulnerabilities are evaluated through in-depth analysis. The report is anonymized to protect employees from discreditation and is transferred to the target organizations via a secure digital path in PDF format. After the in-depth analysis, the results from the written report are discussed in post-workshops with the IS decision-makers and the affected employees as part of the debriefing to analyze the human behavior shown, and to crystallize the possible individual causes, and to develop suitable mitigation options. Through dialog with employees, real problems often come to light, which were previously known at the operational level but rarely communicated beyond that. The following is a concise example of how SE pen test can be operationalized after the above-mentioned phase.

How Does Social Engineering Penetration Test Work?

Before the SE pen test is performed, it must be prepared. This is where it comes into play that, unfortunately, many decision-makers, employees, and companies no longer have a feeling for their external presentation on the Internet. Regarding social networks, it can ultimately be said that in the long term there will no longer be any way to completely avoid these digital communication channels without foregoing external social contacts and external representations. However, for IS reasons, it is generally not advisable to have a Facebook or Instagram profile, as the data can be arbitrarily sold or leaked to third parties without authorization. However, the perception of the information presented in networks is an essential difference. Thus, the human hacker can obtain information about the target person from every like, post, or comment. There are two approaches to achieving this.

Information is shared with the world forever and ever on the Internet, or one retains the information to oneself. If you are aware of the external presentation and the information published on the Internet, you can also deduce in the event of an SE attack how a person who may be a stranger managed to

build up similarities to the target person or target company so quickly. In this context, commonalities are the key to building rapport, that is a kind of relationship level with a person. In this context, humans love the common ground. They are important building blocks for the social environment and provide security. People with whom we have common ground are not only more likeable than others but also seem more credible, convincing, and competent. With business networks such as Indeed, Nexxt, and LinkedIn one naturally recognizes the networking approach and the benefits that derive from it for the individual. However, does a business profile really have to include one's entire CV, all further education measures, training courses, one's own contact data, date of birth, and much more?

Now, let us take a closer look at a LinkedIn profile through the eyes of a pen tester and, in bad cases, a human hacker, and exploit this information for the pre-preparation phase (Willer, 2022):

- Employer and current activity
- What company knowledge is the target person having access to?
- What are the target's areas of responsibility?
- Does the target person also have responsibility for personnel?
- How long has the target been with the target company?
- Where was the target person before?
- Are the contact details freely accessible?
- Further education, qualifications and events attended.

How can you draw conclusions about the software or processes used by the target company? Can you meet the target person in person during a training session or event? Are the contact lists and details available? If so, through which of the contacts one can approach the target person? Now that the pen tester has found out who their target is networked with via the OSINT and SOCMINT methods, they can identify connections and play the individuals against each other if needed. To get into the contact list of a target person, it is advisable to first ask for known contacts, who are regular contact collectors (500+). As a rule, they hardly check new requests. The target person, however, receives the impression that we already have a common contact network if we write to them now. Finally, this is exactly what is visually represented - remember the point above: People love and need common ground. Or is the date of birth visible? Just in time for the birthday, the target person receives a personalized and individual phishing email with a little surprise or maybe with a Trojan in the attachment. In this context, the date of birth also helps the human hacker to verify himself as the target person on the telephone. For some authorities, offices, or other agencies, a surname, first name, date of birth, and convincing appearance on the phone are often sufficient to successfully pose a target person. Once the pen tester identified his target person and gathered the initial information, profiling begins. In this type of profiling, the aim is to identify personality traits and theoretical behavioral patterns based on the information obtained from OSINT and SOCMINT to derive a response to manipulation. Which stimulus is the person most likely to respond to which reaction? Thus, after profiling, the pen tester

can access the possibility of legend formation and the use of psychological reinforcers. Tradesmen, suppliers, maintenance companies, telecommunication providers, partners and customers, external service providers, facility management, and catering companies, as well as consultants, sales trainers, auditors, and many more, can be used as legends. Decisive for a successful SE attack but also for the SE-pen test is the authenticity of the attacker as well as the timing and right communication channel at the right time. Psychological enhancers are small details that suggest an image to the victim or target person that reinforces the impression of the situation just experienced as well as the expectations arising from it. These details can be identification cards, clipboards with logos, the right clothes (e.g., a suit or even a jumpsuit), well-known signal colors on a pen or key fob, electronic devices, a uniform, a dialect, or a large luxury limousine.

For a better understanding, a fictitious COA is now listed (Willer, 2022). Three colleagues have met at the back entrance of a company to enjoy a cigarette break outside the building. A white delivery van pulls up and a man frantically gets out, approaches three colleagues with packages on his arm and politely asks if someone could open the door for him. The parcel carrier wears the typical colors of a well-known delivery service, and his van also has a logo on the door. The probability that one of the colleagues will open the door to the parcel delivery man is already quite high. However, now the three colleagues are sensitized, curious or arrogant and ask why he would not use the front entrance as usual? In the second row, the parcel delivery driver apologizes and explains that a car had just parked at the front entrance and that he did not want to completely block the company entrance. The three colleagues are satisfied with the answer conditionally and give the messenger with view of the own cigarette the instruction to announce itself, but immediately at the receipt, since none of them is ready its cigarette break prematurely to terminate. However, the package-delivery man never arrived.

Amplifier:

- Motor vehicle with logo, work clothes in well-known signal colors,
- Parcels, stressed, friendly, rather submissive behavior,
- Group dynamics, triggering of helpfulness and perhaps even triggering a sense of superiority through subjective perception of one's position/status in society.

Timing: The moment the cigarettes are lit, the attack begins.

Group dynamics: Each individual shifts perceived responsibility to others via access control. TEAM (Great Someone Else Is Doing It).

The cigarette in the hand makes one's own intention, namely the need for nicotine or social exchange with colleagues, appear as a priority. After a successful COA, the results achieved must now be analyzed. Therefore, the following exemplary results can be derived from this post-pen test phase in the totality of the observations (Tab. 1).

Table 1. Example of results from the depth analysis.

Empirical observation (Findings)	Findings from the post-pen test phase
The SE pen tester is perceived by employees as an unknown and apparently external person, but is not approached, even though he sometimes behaves in a very conspicuous manner.	{Failure or non-existence of error culture: insecurity and fear of doing something wrong} {Failure to have a culture of safety or IS awareness as a collective responsibility: that's not my responsibility, that's what we have a receptionist and security for} {Missing knowledge and skills - Perceived behavioral control: I wouldn't even know what to say}. {Human traits-Self-efficacy expectation: What if I approach this one and it's an auditor or other outside consultant, even worse a supervisor?}
The SE pen tester is simply taken in by a coworker through a side entrance, although the two do not know each other. (Tailgating)	{Human characteristics - Injunctive Perceived Norm: I did not mean to be rude}. {Missing communication and inadequate staff training and awareness: I did not realize that was my job, to pay attention to who was allowed in and who wasn't}. {Missing knowledge - behavioral intent and salience: why would anyone want to come in here, what do we have to hide, we're just a local government}.
The SE pen tester can overcome a singulation facility by passing an employee's badge.	{Human traits - Injunctive Perceived Norm: I didn't mean to be rude}. {Human traits - Behavioral Intention and Habit: We always do this at lunch, we always have just one card with us when we go out to eat together}.
Employees continue to click on links in training mail despite repeated phishing training.	{Inefficiently designed operational structure and process organization: I process about 200 mails a day - I simply don't have time to take a closer look at every mail}.
Employees fall for a vishing pen test with an IT support scenario in the rows.	{Human characteristics - behavioral intention and habit: I have never had contact with our IT support in the home office.} {Missing knowledge and skill - perceived behavioral control: How am I supposed to recognize our outside support, I don't even know the name. {Missing knowledge and skills - perceived behavioral control: I did not realize that IT support would never ask us for our password}

CONCLUSION

Human factors are an important addition to the security chain and can help prevent incidents by preventing and detecting them. In this context, it should not be considered as a weak link, but as the strongest in the IS chain. The objective of the work was to embed two practicals, interlocking strategic approaches into an OODCA cycle defined for this purpose, which enables sustainable, efficient, and transparent human threat awareness and mitigation in the context of IS. The cycle focuses on role- and issue-based interactivity of human-originated cyber threats such as fraud, v(p)ishing, pharming, scareware, tailgating, pretexting, OSINT, and SOCMINT, and enables efficient mapping of existing roles to potential threats. As a tool, the cycle uses SE pen testing methods. SE pen testing serves as an additional security consideration and directs the focus to the most important entity in the security chain. The human factor, which operates technical systems, follows organizational measures, and is significantly involved in maintaining IS.

REFERENCES

Boyd, J. R. (1976): "Destruction and Creation," 1976, pp. 1–8.

Hughes-Lartey, K., Li, M., Botchey, F. E., Qin, Z. (2021): "Human factor, a critical weak point in the information security of an organization's Internet of things," Heliyon, Vol. 7, I. 3, 2021, pp. 1-13.

Koza, E. (2022a): "Information Security Awareness and Training as a Holistic Key Factor – How Can a Human Firewall Take on a Complementary Role in Information Security?" 13th International Conference on Applied Human Factors and Ergonomics (AHFE 2022), pp. 49–57.

Koza, E. (2022b): "OODA Loop as a Decision Support Model to Continuous and Dynamic Vulnerability Management and Incident Response Management of Critical Infrastructures" Proceedings of the 32nd European Safety and Reliability Conference, pp. 2859–2866.

Willer, Michael (2022): Insight Human Hacking, Recognize and understand social engineering, pp. 1–58.

Widdowson, A. J., Goodliff, P. B. (2015): "CHEAT, an approach to incorporating human factors in cyber security assessments," 10th IET System Safety and Cyber-Security Conference, Bristol UK, 2015, pp. 1–5.