Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

Complex yet attainable? An interdisciplinary approach to designing better cyber range exercises

Magdalena Glas [*], Gerhard Messmann, Günther Pernul

University of Regensburg, Universitätsstr. 31, Regensburg, 93047, Germany

ARTICLE INFO

Dataset link: https://github.com/InstruCRX

Keywords:
Cyber range
Authentic learning
Cybersecurity exercise
Instructional design
Cyber defense exercise

ABSTRACT

The global shortage of cybersecurity professionals poses a daunting challenge for organizations seeking to protect their assets and data. To counteract this workforce shortage, cyber range exercises (CRXs) can equip individuals with the necessary knowledge and skills to become security professionals. However, the complexity of CRXs tends to overwhelm trainees with little prior cybersecurity experience, resulting in ineffective learning experiences. To address this issue, we take an interdisciplinary approach, leveraging established models on learning and motivation for cybersecurity. In this pursuit, we propose a literature-based framework of six design principles that aim to facilitate CRX designers in creating more effective CRXs. To illustrate the framework's utility, we introduce a CRX for incident response built upon these principles. To evaluate the effectiveness of this principle-driven CRX design, we conducted a user study with $N = 89$ participants. The results of this study showed that the design provided an engaging learning experience that enabled participants to effectively acquire incident response knowledge and skills.

1. Introduction

In the light of an ever-changing threat landscape, organizations require a highly skilled cybersecurity workforce capable of effectively defending their infrastructure against cyberattacks. However, the growing demand for cybersecurity professionals outpaces supply ((ISC)^2, 2023). Commonly known as the cybersecurity workforce shortage, this issue is often reported as the main obstacle preventing organizations from achieving cybersecurity resilience (Bueermann and Doyle, 2023; ISACA, 2022). Organizations attribute this issue to the fact that applicants for open cybersecurity positions lack practical skills (Oltsik and Lundell, 2021; Zan and Di Franco, 2019). Aligning the requirements of organizations with the pool of aspiring cybersecurity professionals necessitates the development of training methods that equip novices with practical cybersecurity skills (Blažič, 2021). *Novices*, in this context, are individuals who seek to take on a role in cybersecurity but are yet lacking practical experience. This encompasses students enrolled in academic cybersecurity or computer science programs who have yet to gain practical work experience, as well as professionals who have a general technical background and are seeking to specialize in cybersecurity. In the past years, cyber ranges exercises (CRXs) have proven to be an effective way for participants to gain cybersecurity skills hands-on (Yamin et al., 2020). Cyber ranges replicate parts of digital infrastructures to provide an environment for realistic cybersecurity training in which participants can engage with complex real-world attack scenarios. This presents a valuable opportunity for cybersecurity professionals to apply and enhance their cybersecurity skills, as well as for cybersecurity teams to evaluate established processes (Kim et al., 2019). For novices, however, this experience can feel overwhelming and challenging to navigate (Owens et al., 2019; Mirkovic et al., 2015b). In this regard, studies on the effectiveness of CRXs indicate that trainees in a CRX often feel bombarded with an overload of information (Brilingaitė et al., 2020; Kick, 2014), finding themselves unsure of how to tackle given problems. This, in turn, inhibits their progress in the exercise (Weiss et al., 2016; Tobey et al., 2014), leading to frustration and possibly even resignation.

This issue can be partly attributed to the fact that the literature on CRX design primarily focuses on technological advances. Considerations related to instructional design – the discipline of applying theories on learning and motivation to create effective learning activities – are only marginally addressed (Maennel et al., 2023; Mirkovic et al., 2015a). Cyber ranges are large and complex technical infrastructures, often consisting of both virtual machines and physical components. Developing such infrastructures is time-consuming and cost-intensive (Vykopal et al., 2017). If the very goal of a CRX, which is for participants to acquire knowledge and skills, fails because of a lack of instructional rigor, these efforts are in vain. With our research, we want to investigate how to create attainable CRXs that provide an

[*] Corresponding author.
E-mail address: magdalena.glas@ur.de (M. Glas).

https://doi.org/10.1016/j.cose.2024.103965
Received 29 May 2024; Accepted 18 June 2024
Available online 25 June 2024

effective learning experience. This encompasses both that participants achieve the learning goals of the exercise and perceive the exercise as an engaging and rewarding experience. Consequently, this leads us to the following research question:

RQ. What are the characteristics of a CRX design that fosters effective learning for cybersecurity novices?

Contribution. We address this research question through an interdisciplinary approach by leveraging insights from *instructional design* for cybersecurity. To this end, we propose six instructional principles to create CRXs that provide a complex yet attainable learning environment. These principles are derived from established instructional models that foster the acquisition of complex knowledge in authentic learning environments, such as Cognitive Apprenticeship (Collins et al., 1991) and Cognitive Flexibility Theory (Spiro et al., 1994). To the best of our knowledge, this is the first approach to integrating instructional theory in CRX design. To demonstrate the principles' utility, we apply them to create a CRX in which trainees learn to detect and respond to real-time attacks against an industrial network. To evaluate the effectiveness of this design, we present a user study involving ten CRX events with $N = 89$ novice participants. We assessed both the participants' learning outcomes and their subjective perceptions of the learning experience. The evaluation results show that aligning the technological infrastructure of the CRX design with instructional principles creates a learning experience that is motivating for participants, fosters effective collaboration, and, ultimately, leads to effective learning outcomes.

2. Background and related work

This section briefly accounts the background of CRXs before discussing previous research efforts to examine the instructional aspects of CRXs.

2.1. Cyber Range Exercises (CRXs)

Cyber ranges simulate or emulate real-world systems, networks, and applications to provide a safe and legal environment for cybersecurity training and testing (National Initiative for Cybersecurity Education (NICE), 2020). Participating in a CRX allows trainees to experience real-time attacks against a training environment that closely resembles reality. The concept of cyber ranges emerged from the military sector, adopting the idea of a shooting range to prepare trainees for digital warfare (Ferguson et al., 2014). Today, the concept has been adopted to improve cybersecurity competencies in academia (Čeleda et al., 2015; Hatzivasilis et al., 2021), the industry (Airbus, 2023; Accenture, 2023; IBM, 2023), and public institutions (Leitner et al., 2020; Brilingaitė et al., 2017). In addition to technical skills, such as incident response or system hardening, CRXs can also help trainees practice soft skills such as teamwork skills and stress resilience (Beuran et al., 2018; Švábenský et al., 2018). In the literature, the terms cyber defense exercise and cyber ranges exercise are commonly used interchangeably (Brilingaitė et al., 2020).

2.2. CRXs for cybersecurity novices

To contextualize our research, we discuss related work that investigates instructional aspects of CRXs. That is literature that either explicitly examines the design of CRXs for novices or offers methods and insights that are applicable to novice training. First, we identified several frameworks that focus on feedback as a key design element of CRXs that helps participants achieve effective learning outcomes (Maennel et al., 2017; Vykopal et al., 2018; Švábenský et al., 2022; Andreolini et al., 2020; Braghin et al., 2020). The authors of these works propose methods for monitoring trainees' actions during an exercise to provide trainees with constructive feedback on their actions throughout and after an exercise and evaluate the effectiveness of a CRX. While these frameworks offer advanced techniques for integrating trainee monitoring into the architecture of a CRX, it is essential to recognize that monitoring and feedback modules, in isolation, are not sufficient to create immersive learning experiences. Consequently, our research builds upon these findings by considering feedback as one aspect of a holistic CRX design.

With respect to the learning experiences of novice trainees in CRXs, the body of extant literature is limited. Vykopal and Barták (2016) report experiences from a CRX conducted on KYPO cyber range with participants with various skill backgrounds, including novices. For novice trainees, the authors highlight the importance of providing sufficient guidance (e.g., a demonstration of the most important functions of a tool), especially at the beginning of the exercise. Furthermore, they emphasize the importance of embedding all relevant information for participants (e.g., study materials, hints, solutions, etc.) directly in the CRX environment instead of providing it externally (via email or in-class explanations of a trainer) to keep trainees' attention on the CRX. Brilingaitė et al. (2020) report on their experience with large-scale CRXs and propose a CRX design framework based on their findings. While the authors describe that novice participants encountered considerable difficulties in following the exercise, this issue is only marginally accounted for in their framework. That is, the authors propose conducting a pre-training for novice participants, however, without specifying its design or execution. Moreover, Maennel et al. (2023) propose a multi-dimensional approach that aims at better incorporating social, emotional, and cognitive considerations, such as participants' psychological safety and relatedness to their team members, into the design of a CRX. This approach aims to enhance learning experiences for both novices and more experienced participants. An implementation of this approach, however, is not described. The authors' findings in Vykopal and Barták (2016), Weiss et al. (2016), Maennel et al. (2023) provide far-reaching insights into which learning designs can improve trainees' learning in a CRX, which we carefully incorporated in the design principles we propose in this paper. However, the discussed studies are exploratory in nature and base their research primarily on participants' anecdotal feedback. In contrast, our research takes a more rigorous approach and aims to provide empirical evidence of the effectiveness of our proposed design, e.g., by systematically assessing participants' skills and knowledge before and after the exercise. In essence, the majority of works that address the improvement of learning effectiveness in CRXs concentrate on isolated elements of CRX design rather than presenting holistic approaches. Previous works that do investigate CRX designs more comprehensively either insufficiently address the needs of novices or lack empirical evidence of the usefulness of the proposed designs. Consequently, there is a lack of sound approaches to guide CRX designers in creating meaningful learning experiences that are accessible to novices. In this paper, we aim to fill this gap by presenting a CRX design explicitly tailored to introduce novice learners to practical cybersecurity. The theory-driven approach we followed to establish this design is described in Section 3.

2.3. Instructional design in cybersecurity exercises

A research domain in which considerations on the instructional design of cybersecurity exercises have been investigated more deeply is the field of Capture the Flag (CTF) challenges in an educational context. CTFs are cybersecurity exercises in which participants solve puzzles, crack codes, and exploit vulnerabilities in commonly static computer systems to capture "flags", usually strings of text or files, to earn points. In contrast to CRXs, the focus of CTFs is not for trainees to deal with scenarios that are as realistic as possible but to teach general, mostly offensive, cybersecurity concepts (Votipka et al., 2021). Nonetheless, the insights into how participants learn to solve problems effectively in a CTF challenge are also informative for the design of CRXs, which is why we want to point out some significant findings of prior works in this field. Weiss et al. (2016) conducted a study on balancing guidance and independence in a CRX for college students. The authors argue that a cybersecurity exercise for novices needs to be broken down into clear steps to focus novices' attention on what is important for achieving the learning goals of the exercise. Backman (2016) reports about a CTF challenge that specifically targets undergraduate students with little prior experience in cybersecurity. To provide better accessibility for this target group, the author emphasizes the roles of collaboration, especially pairing inexperienced students with more senior participants. Similarly, Mirkovic et al. (2015b), Mirkovic and Peterson (2014) report on their experiences with Class CTFs (CCTFs), CTFs especially designed to be part of an academic curriculum. In particular, the authors describe the importance of providing CCTF participants with sufficient knowledge resources, which, in the case of CCTFs, are provided in preparation for the challenge. Owens et al. (2019) present the CTF challenge picoCTF. In the design of the challenge, the authors focus primarily on gradually increasing the difficulty of tasks within a CTF so as not to overtax novices. They also report on the positive influence of competitive elements when participants can measure themselves directly against their peers in their course. Votipka et al. (2021) provide a review-driven approach and compare different CTF challenges regarding the instructional principles they employ. The authors take the analysis rather than the design perspective. For this reason, the design principles, with regard to the CTFs, show many overlaps with our framework (rf. Section 3.3), but are not congruent. These studies (Weiss et al., 2016; Backman, 2016; Mirkovic and Peterson, 2014; Mirkovic et al., 2015a; Owens et al., 2019; Votipka et al., 2021) provide far-reaching insights into characteristics of learning designs that improve trainees' learning in cybersecurity exercises, which we consequently considered when developing and implementing the design principles we propose in this paper. However, the discussed studies are mostly exploratory in nature and base their research primarily on participants' and trainers' anecdotal feedback. In contrast, our research takes a more rigorous approach and aims to provide empirical evidence of the effectiveness of our proposed design by systematically assessing participants' knowledge and skills before and after the exercise.

3. A framework for leveraging instructional design for CRXs

In this section, we propose a framework of six instructional principles for designing effective CRXs for cybersecurity novices. In the following, we first describe how instructional models can help with the integration of instructional design into CRX design (Section 3.1). Subsequently, we outline the method we followed to derive the six design principles from the literature (Section 3.2) before describing the individual principles in detail (Section 3.3).

3.1. Deriving the core characteristics of CRXs: Authenticity and complexity

Instructional models stem from the field of instructional design and represent prescriptive, normative models that seek to make descriptive scientific insights into human learning, motivation, and interaction accessible to educational practitioners in schools, universities, and organizational training (Glaser, 1976). In an attempt to synthesize different streams in the field of instructional design (i.e., traditional instructional design, educational and cognitive psychology, and social-constructivist views), instructional models should provide insight into the goals that can be achieved when employing the model and clarify how the gap between designed learning environments and target application contexts can be bridged (van Merriënboer et al., 2002). As described above, cyber ranges as infrastructures for training (in contrast to other cybersecurity exercises such as CTF competitions) strive to offer trainees the most realistic training environment possible. We therefore argue that, from an instructional design perspective, CRXs must reflect two overarching principles in particular: authenticity and complexity. *Authenticity* refers to the degree to which learning environments capture the essential characteristics of the real-life context in which learning outcomes are intended to be used later on Honebein et al. (1993). Authentic learning environments help trainees to understand how they can put the knowledge and skills they acquire to use in the future, which, in turn, may foster their intrinsic motivation to learn (Collins et al., 1991). To this end, learning goals should match competencies that are necessary for effective performance in the context of the target application. Likewise, learning tasks should resemble the tasks that trainees will encounter in the real-life context for which they are training. *Complexity*, the second overarching principle, is closely linked to authenticity. As tasks in real-life contexts bear a natural complexity, learning environments can only be meaningful and effective if they capture this real-life complexity in the learning tasks that trainees are confronted with. A meaningful, complex learning environment exposes trainees to this complexity, however, without overwhelming them (Spiro et al., 1994). In the following section, we describe the method we applied to identify instructional models that foster those two principles.

3.2. Selection method

The overall method employed in this paper follows the hypothetico-deductive research approach (Edgar and Manz, 2017) with the goal of identifying instructional design principles that foster meaningful learning for cybersecurity novices and empirically evaluate their effectiveness. How this research approach was adopted in this paper is depicted in Fig. 1. Recognizing the diversity of models within the field of instructional design, solely relying on a systematic literature search would not be sufficient to capture relevant models. Instead, we initiated a literature search grounded in well-known publications that provide an overview of established instructional models (Reigeluth, 1999; Merrill, 2002; Kolodner et al., 2003). Our selection of instructional models from these works was based on their promotion of authenticity and complexity as central principles (selection criteria). From this initial set of applicable models, we followed a snowballing approach to include (and select) further models based on a forward search. This approach led us to the following set of instructional models which are applicable for designing authentic and complex learning environments:

- **Cognitive Apprenticeship** (Collins et al., 1991): confronting trainees with increasing complexity by giving them more and more responsibility for carrying out tasks while supportive measures are simultaneously fading out.
- **Cognitive Flexibility Theory** (Spiro et al., 1994): fostering the development of trainees' ability to solve tasks in complex domains by applying acquired knowledge and skills in a flexible and creative way.
- **Collaborative Problem Solving** (Nelson, 1999): fostering social interactions between trainees, thereby capturing the increasing importance and presence of teamwork in real-life work contexts.
- **Four-Component Instructional Design (4C-ID)** (Kirschner and van Merriënboer, 2008): enabling trainees to solve complex tasks by providing them with sufficient knowledge resources and the purposeful organization of tasks.
- **Goal-Based Scenarios** (Schank et al., 1999): providing trainees with a complex and intrinsically motivating story and a corresponding mission they need to accomplish.
- **Problem-Based Learning** (Hmelo-Silver, 2004): enabling trainees to collaboratively solve a problem through self-directed knowledge acquisition and teamwork.

In an iterative process, we derived six key principles from this set of models. We want to note that these principles do not claim to be exhaustive (i.e., other less important principles might be useful in certain environments) but that we captured those we deemed most

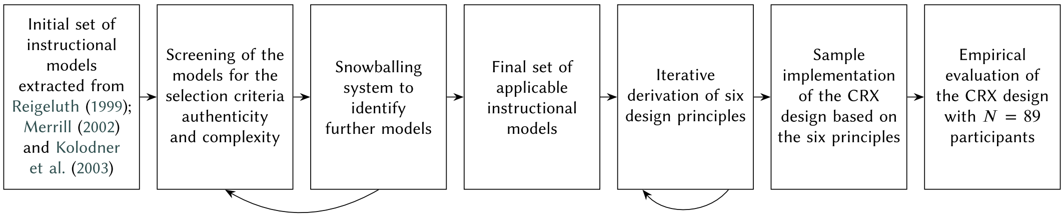

Initial set of instructional models extracted from Reigeluth (1999); Merrill (2002) and Kolodner et al. (2003)
Screening of the models for the selection criteria authenticity and complexity
Snowballing system to identify further models
Final set of applicable instructional models
Iterative derivation of six design principles
Sample implementation of the CRX design based on the six principles
Empirical evaluation of the CRX design with $N = 89$ participants

Fig. 1. Methodological approach: Seven steps to establishing design principles, implementing them in a CRX design, and evaluating its effectiveness.

relevant for creating an authentic, complex, and thus effective CRX design. This concise framework aims to provide a practicable guideline for CRX designers to create more effective learning experiences for cybersecurity novices.

3.3. Design principles

In the following, the six design principles (P1–P6) are described in detail.

P1: Active Participation. The principle of active participation foresees trainees not as passive recipients of knowledge but as active participants in their own learning processes (Spiro et al., 1994). By letting trainees self-determine how they acquire knowledge to solve given problems, this approach shifts away from traditional one-way teaching methods and promotes self-directed learning. This enables trainees to make use of their existing knowledge when acquiring new knowledge and, by this means, to improve both the quantity and the quality of their knowledge base. This, in turn, aids in easier retrieval and reuse of knowledge in future situations (Schank, 1999). Active participation aims to prepare trainees to solve tasks in complex real-world domains — domains in which effective problem-solving does not require following predefined procedures but applying the acquired knowledge in a situation-specific manner. Cybersecurity operations necessitate individuals with this kind of flexibility and adaptability. Consequently, we view active participation as a central principle in the design of CRXs — both for novices and experienced practitioners.

P2: Realistic Environment. Realistic environments embed learning activities in relevant real-world contexts (Honebein et al., 1993). This approach is meant to ensure that trainees do not acquire knowledge in an abstract and isolated manner but in the context of meaningful tasks that demonstrate how they can apply these competencies in real life. Using realistic tasks as a starting point for learning is motivating because trainees immediately experience the purpose of their learning activities (Eccles and Wigfield, 2002). Furthermore, through such realistic activities, knowledge acquisition is contextualized, and trainees acquire episodic knowledge, which is richer and easier to integrate than purely abstract semantic knowledge. According to Schank et al. (1999), realistic learning environments can be created by structuring exercises around "goal-based scenarios". This involves the creation of an engaging and relatable "mission" that trainees must undertake. When this approach is effectively implemented, trainees become intrinsically motivated, driven by the significance of their mission rather than external factors like passing a post-exercise test. Within this mission, trainees take on the role of an expert in the domain (Collins et al., 1991). Making trainees "think like experts" helps to contextualize the competencies they acquire throughout the training with real-world tasks and responsibilities (Spiro et al., 1994). CRXs for novice participants aim to prepare trainees for future careers as cybersecurity experts. To this end, the CRX design should facilitate the seamless application of acquired competencies in organizational cybersecurity operations. We argue that this can be achieved through creating a CRX as a realistic learning environment.

P3: Scenario Operations. The competencies trainees acquire throughout the exercise should be embedded in activities that let them actively engage with the scenario. This fosters the acquisition of "knowledge as knowledge in use" (Spiro et al., 1994) (cf. P1). In a scenario-based exercise, this is achieved through Scenario Operations (SOs), that is, single tasks or problems that trainees must accomplish or solve in order to successfully complete the overall mission (Schank et al., 1999). To succeed in SOs, trainees must acquire knowledge and put this knowledge into practice. The complexity of SOs must be tempered for practicality without undue simplification to preserve the realistic character of the learning tasks (cf. P2) (Honebein et al., 1993). This means that SOs should resemble real-world tasks but can be simplified to what is necessary to address the learning goals of an exercise (Schank et al., 1999). Each SO sets a realistic goal that trainees must attain successively. This helps to structure the learning process, thereby enabling trainees to keep track of their progress and not get cognitively overwhelmed (Sweller, 2010). Likewise, this facilitates trainees' goal striving and helps them to stay motivated by providing a sense of accomplishment after completing SOs (Ryan and Deci, 2000; Votipka et al., 2021). In this respect, we argue that novice participants need more guidance in a CRX than experienced practitioners do. To address this need, structuring an exercise around well-defined SOs helps to guide novice trainees through complex scenarios and prevent confusion or frustration.

P4: Knowledge Resources. In SOs, trainees face complex tasks they are not familiar with. To this end, SOs need to be accompanied by knowledge resources that provide trainees with the information they need to accomplish the task (Owens et al., 2019; Mirkovic and Peterson, 2014). By this means, the level of intrinsic cognitive load that is imposed onto the trainees can be managed, and trainees are enabled to master tasks without experiencing cognitive overload (Sweller, 2010). In addition, managing the complexity of learning tasks by providing knowledge resources that are tailored to trainees' level of competencies bears the motivational benefit that trainees will feel optimally challenged, experiencing learning tasks neither as too difficult nor as too easy. We classify knowledge resources into three distinct types of information: Supportive information, procedural information, and feedback (Kirschner and van Merriënboer, 2008). The subsequent sections describe the significance of these three information types for CRX learning.

- **Supportive Information.** Supportive information provides trainees with guidance on problem-solving in a specific domain and aims at fostering the development of knowledge trainees need for accomplishing tasks in their domain (e.g., job tasks) (Kirschner and van Merriënboer, 2008). In addition, and closely connected to P1, supportive information aims to enable trainees to connect such new knowledge to their existing knowledge about solving particular domain-specific problems. In scenario-based exercises, this supportive information helps trainees understand the background and context of SOs.
- **Procedural Information.** Procedural Information refers to instructions on how to perform routine aspects of authentic tasks, respectively, SOs (Kirschner and van Merriënboer, 2008). This type of resource should be presented in a manner that empowers novices to approach SOs as independently as possible, only

accessing instructions if they need it. This can be facilitated by applying the instructional concepts of “fading” and “scaffolding” (Collins et al., 1991). Scaffolding refers to providing trainees with the right amount of individual assistance they need to complete a SO. Fading refers to the gradual removal of this support throughout the exercise, enabling trainees to become increasingly self-sufficient, ultimately reaching a point where they can complete SOs with minimal external assistance by the end of the exercise.

- **Feedback.** Feedback refers to information provided to trainees in response to their actions and aims at providing trainees with insight into the effectiveness of their actions both from a process and an outcome perspective. Within the learning process, feedback is designed to enable trainees to understand which aspects of the set learning goals they have already accomplished and which steps they need to take in order to further improve their knowledge and skills. For instance, feedback on a mistake may serve as a trigger for promoting learning from mistakes. That is, the feedback on the mistake may encourage trainees to reflect on the cause of the error by reviewing the information provided or by discussing things they have not yet fully understood, reconsidering their action strategy (Schank et al., 1999).

P5: Competition. The principle of competition enables trainees to compare their own learning progress with that of their peers. This can serve as a powerful source of motivation to succeed in an exercise (Collins et al., 1991). By involving comparisons with others, trainees will more easily identify information and relevant steps of action that are relevant to achieving the goals of the learning activity at hand. Similarly, comparisons offer trainees valuable insights into their strengths and weaknesses, thereby potentially enhancing their perceived competence (Ryan and Deci, 2000). Competition is a common element in CRX designs, typically in combination with point-based rewards and leaderboards to showcase trainees’ scores, as observed in previous studies (Yamin et al., 2020; Brilingaitė et al., 2020). While we acknowledge the potential competition has to raise trainees’ motivation in a CRX, we emphasize the need for careful implementation of the concept. Firstly, because competition might encourage hasty problem-solving only for the sake of winning, and secondly because some trainees might feel inhibited rather than motivated when comparing themselves to others.

P6: Collaboration. In collaborative learning environments, trainees actively engage with their peers to share knowledge and collaboratively solve problems. This has the potential to enhance learning outcomes and boost creativity when tackling complex problems (Nelson, 1999). Working collaboratively is efficient for several reasons. First, when individual team members struggle with the accomplishment of a task, other members of the group can provide various forms of guidance, such as additional information or explanations (Backman, 2016). Similar to the resources outlined above, such social resources can prevent cognitive overload (Sweller, 2010). Moreover, collaborating in a stable group enables trainees to develop a sense of relatedness and belonging. That is, when the members of a group feel related to each other, there is a greater chance that they might go beyond their comfort zone and voice challenging ideas or point out problems (Messmann, 2022). This can not only enable creative problem-solving but can mitigate the potential negative consequences of a competitive approach. Finally, working together with other trainees during an exercise enables individual trainees to gain experience in collaboration and, through this experience, improve their teamwork skills (Hmelo-Silver and Barrows, 2006).

4. A principle-driven CRX design for incident response

In this section, we describe how the six principles were integrated into the design and development of a CRX in which novice participants gain hands-on experience in incident response. We want to explicitly clarify that this implementation serves as an example and does not represent the only approach to how the six proposed principles can be implemented. CRX designs, like any complex system, involve a spectrum of considerations, including technological feasibility, resource constraints, and the evolving landscape of requirements. Our design represents a trade-off between these factors, aiming to strike a balance that aligns with our effectiveness goals (rf. Section 4.1) while acknowledging that alternative approaches may exist, each with its own set of advantages and limitations.

4.1. Effectiveness goals

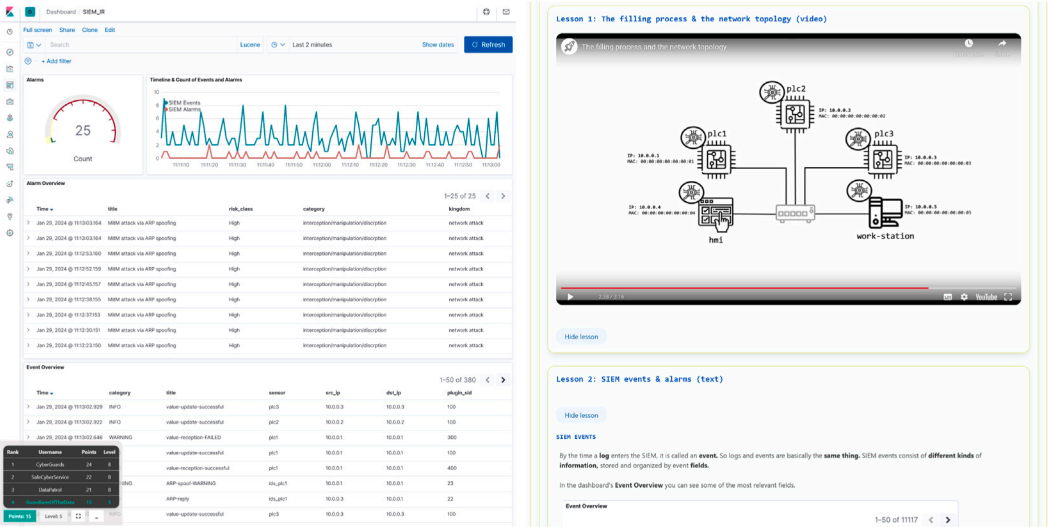
As introduced earlier, we consider a CRX effective if trainees perceive it as a positive *learning experience* and if the CRX facilitates the achievement of a defined *learning outcome*. Within our CRX design, we therefore define goals referring to learning experience (G1–G2) and goals referring to learning outcomes (G3–G6). This categorization is in accordance with the first three of Kirkpatrick’s interconnected levels at which the effectiveness of a training exercise can be assessed (i.e., *reaction*, *learning*, and *behavior*) (Kirkpatrick, 2005).

Concerning trainees’ immediate *reaction* to the training, we used trainees’ learning experience as an indicator of how favorably they responded to the learning process they participated in. Specifically, we took into account the factors of Keller’s ARCS model, that is, whether the learning environment was engaging in the sense that it guided trainees’ attention, was perceived as relevant by the trainees, bolstered their confidence, and contributed to a satisfying learning experience (Keller, 1987) (G1). Furthermore, we considered trainees’ experience of the social aspects of the learning environment (G2), specifically whether they felt related to their team members (Ryan and Deci, 2000) and were satisfied with the collaboration experience (Tseng et al., 2009). The former aspect encompasses whether the trainees were well integrated into their team. The latter aspect refers to the extent to which trainees think they benefit from collaborating with their team members in comparison to learning individually. Focusing on these two indicators of a favorable learning experience is in accordance with Maennel et al. (2023), who describe the quality of collaboration as a possible indicator of overall engagement in a CRX. Moreover, concerning the connection between learner’ *reactions* and subsequent training outcomes (see below), it is clarified in Kirkpatrick’s model that a favorable reaction of trainees is crucial for the chance that subsequent learning outcomes can be attained. This is because a positive reaction of trainees to the characteristics of the training increases trainees’ learning motivation and, thus, their active and persistent participation in learning activities (Ryan and Deci, 2000). In summary, the CRX pursued the following two goals related to the learning experience:

G1: Engaging learning experience. The training captures trainees’ attention, is perceived as relevant, and leads to feelings of confidence and satisfaction.

G2: Positive collaboration experience. Trainees experience a sense of relatedness and satisfaction when collaborating with others during the training.

At the next level of Kirkpatrick’s model, immediate *learning* outcomes need to be assessed. It needs to be determined whether trainees acquired the intended knowledge and skills as a result of their participation in the training. As our CRX aims at introducing novices to practical incident response in an industrial environment, we took into account whether participating in the exercise leads to the acquisition of knowledge in the areas of industrial control system (ICS) security (G3) and incident response (G4). Since collaborative problem-solving and teamwork is crucial for incident responders, we also examined whether the collaboration in the exercise helped participants to improve their teamwork skills (G5).



Alarm Overview

1–25 of 25

Time	title	risk_class	category	kingdom
Jan 29, 2024 @ 11:13:03.164	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:13:03.164	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:12:53.160	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:12:52.159	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:12:45.157	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:12:38.155	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:12:37.153	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:12:30.151	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack
Jan 29, 2024 @ 11:12:23.150	MitM attack via ARP spoofing	High	interception/manipulation/discrption	network attack

Event Overview

1–50 of 380

Time	category	title	sensor	src_ip	dst_ip	plugin_sid
Jan 29, 2024 @ 11:13:02.929	INFO	value-update-successful	plc3	10.0.0.3	10.0.0.3	100
Jan 29, 2024 @ 11:13:02.922	INFO	value-update-successful	plc2	10.0.0.2	10.0.0.2	100
Jan 29, 2024 @ 11:13:02.646	WARNING	value-reception-FAILED	plc1	10.0.0.1	10.0.0.1	300
[illegible]	[illegible]	value-update-successful	plc1	10.0.0.1	10.0.0.1	100
[illegible]	[illegible]	value-reception-successful	plc1	10.0.0.1	10.0.0.1	400
[illegible]	NING	ARP-spoof-WARNING	ids_plc1	10.0.0.1	10.0.0.1	23
[illegible]	[illegible]	ARP-reply	ids_plc1	10.0.0.3	10.0.0.1	22
[illegible]	INFO	value-update-successful	plc3	10.0.0.3	10.0.0.3	100

Rank	Username	Points	Level
1	CyberGuards	24	8
2	SafeCyberService	22	8
3	DataPatrol	21	8
4	GuardiansOfTheData	15	5

Lesson 2: SIEM events & alarms (text)

Hide lesson

SIEM EVENTS

By the time a **log** enters the SIEM, it is called an **event.** So logs and events are basically the **same thing.** SIEM events consist of **different kinds** of **information,** stored and organized by event **fields.**

In the dashboard's **Event Overview** you can see some of the most relevant fields.

Event Overview

1–50 of 11117

Fig. 2. User interface of the CRX: The user interface is divided in the dashboard of the SIEM system (left) and the LMS (right), including the trainee scoreboard (bottom left corner).

G3: ICS security knowledge. Trainees understand the general functionality of an ICS system and a Man-in-the-Middle attack based on ARP-spoofing.

G4: Incident response knowledge. Trainees understand the steps of a typical incident response process and the functionality of a SIEM system.

G5: Teamwork skills. Trainees improve their ability to collaboratively solve complex problems.

Moreover, a training exercise should only be considered effective if, at the next level, changes in *behavior* occur in the sense that trainees can apply their newly acquired knowledge and skills, for instance, by being able to perform relevant performative tasks in training. We thus used trainees' performance in the incident response tasks they had to perform throughout the training as an indicator of effectiveness at this evaluation level (G6):

G6: Incident response performance. Trainees are able to transfer their incident response knowledge to the performance of certain practical incident response tasks.

In the following, we give a brief overview of the design before describing it in light of the proposed design principles.

4.2. Design overview

The proposed CRX design aims to introduce novices to practical incident response. The CRX scenario is located in an industrial setting. On the one hand, this scenario was chosen because the shortage of cybersecurity specialists can have particularly drastic effects in the operational technology (OT) environment, e.g., in the case of attacks on critical infrastructures. On the other hand, because cybersecurity training for this domain is particularly complex, as future incident responders need to understand both the attacks and the functioning of the physical processes of the involved industrial systems. The CRX is designed to be completed within the duration of a regular lecture, with a maximum completion time of 90 min.

In the CRX, trainees take on the role of incident responders who oversee the secure operation of a simulated ICS, which is part of a filling line that fills a liquid into bottles. The ICS is composed of a tank, a pipe, and a filling module. Three sensors are integrated into the system, each controlled by a Programmable Logic Controller (PLC), measuring the liquid level in the tank (PLC1), the flow rate in the pipe (PLC2), and the liquid level in the bottles (PLC3). In addition, PLC1 controls an actuator that opens and closes the tank's motorized valve. An intrusion detection system (IDS) runs on each PLC to detect potentially malicious activities within the network. The PLCs generate logs on both the physical process' operational events (e.g., when a sensor measures data) and the IDS monitoring activities. The log data is normalized as security events are correlated in a SIEM system, with which the trainees interact throughout the CRX. In the scenario, an attacker gains access to the network through a workstation serving as the human–machine interface of the filling plant. The attacker disrupts the filling process by executing a Man-in-the-Middle attack that interferes with the communication between PLC1 and PLC3, resulting in an overfill of the bottles. The trainees participate in the CRX via a user interface (UI) that consists of the dashboard of the SIEM system for monitoring the CRX and a Learning Management System (LMS) that provides trainees with information about their operations within the scenario and background information on these operations (rf. Fig. 2).

The ICS simulation and SIEM system are implemented as a microservice infrastructure using Docker. The ICS simulation is implemented with Mininet and Ettercap. The LMS is implemented with VueJS, connected to the Docker infrastructure via a REST API, implemented with Flask. User data management is implemented with Google Firebase. For a detailed description and documentation of the CRX, we want to refer to the GitHub project[1] on which we made the source code of the CRX publicly available.

4.3. Implementation of the design principles

In the following, we describe how the six design principles proposed in Section 3.3 were implemented in our CRX design.

P1: Active Participation. In the presented CRX design, trainees progress in the exercise by autonomously acquiring knowledge that enables them to solve practical tasks (P3). This self-directed approach is facilitated by the LMS, which provides trainees with task assignments and associated background information (P4). This enables trainees to

[1] https://github.com/InstruCRX

navigate through the exercise at their own pace, while the scenario automatically adapts to each individual's or group's progress. In this setup, the trainer serves as a facilitator of each trainee's individual learning journey (instead of taking on a front-of-class role) in order to support trainees when they face problems throughout the exercise.

P2: Realistic Environment. As described in Section 4.2, the scenario of the CRX is an attack against an industrial filling line. The trainees take on the role of an incident responder responsible for ensuring the secure operation of the ICS as their mission and, therefore, need to detect the attack, eliminate the attacker, and restore the initial network configuration. This scenario creates a relatable goal (i.e., defending the ICS) that embeds the knowledge and skills trainees acquire throughout the exercise in a realistic context.

P3: Scenario Operations. Although the simulated incidents replicate real-world incident response procedures, they are deliberately designed not to encompass a full incident response process. This choice ensures that the exercise remains manageable for novice trainees in terms of both scope and complexity. At the beginning of the scenario, trainees examine SIEM events to grasp the SIEM dashboard's features and security event log structure under normal operation of the ICS. This allows them to become familiar with the SIEM dashboard before the attack starts. After the attack triggers, the trainees investigate SIEM logs to detect and analyze the attack. In subsequent steps, the trainees eliminate the compromised ICS component and restore the network configuration using a simulated command line interface (CLI). The SOs are flag-based, i.e., the trainees either submit a solution for a SO (e.g., the attacker's IP address) or an artificial flag that, for instance, appears when trainees enter the correct CLI syntax. As outlined in P1, the scenario adapts to the pace of the trainees' individual learning process. This is achieved by automating the attacker's actions so that the stages of the attack are triggered corresponding to the trainee's progress in the SO.

P4: Knowledge Resources. Regarding knowledge resources in our CRX, *supportive information* aims at helping trainees to understand four aspects of the scenario: (1) the operation of the ICS, (2) the mechanics of a MiTM attack, (3) the generic steps of an incident response process, and (4) the capabilities of a SIEM system. The information is presented in the form of different knowledge units, referred to as "lessons", that are accessible over the LMS (rf. Fig. 2). Lessons come in short videos or textual descriptions. As lessons build upon each other, the design of the LMS allows trainees to revisit previous lessons if needed. Furthermore, *procedural information* supports trainees in the two areas of acquiring the knowledge and skills for (1) analyzing SIEM events and (2) using CLI utilities in an incident response process. The first aspect is addressed through the detailed description of the analysis of sample SIEM events as part of one knowledge unit. The second aspect is addressed via simplified help pages integrated into the simulated CLI. Before using a utility command, trainees can access these pages to understand how to use the respective command. The level of detail of the instructions given along the SOs is gradually reduced when an SO addresses skills similar to those practiced in a previous SO. Each SO includes hints available for trainees seeking additional guidance (i.e., scaffolding). Trainees have to trade these hints for points, challenging them to solve a SOs as independently as possible (P5). Finally, CRXs in the literature integrate various forms of *feedback,* such as qualitative feedback through trainers or after-exercise talks (Vykopal et al., 2017). For novice trainees, however, we propose to integrate feedback as a direct consequence of actions within SOs. In our CRX design, the flag-based submission format aims to provide trainees with instant feedback when submitting a solution. This feedback provides trainees with information on why a solution is (in)correct and illustrates its impact on the overall scenario (e.g., which effect taking down the network interface of the attacked device has on the operation of the ICS).

P5: Competition. We aimed to exploit the potential of competition to act as a motivator while ensuring that trainees approached the SOs with sufficient rigor. The user interface of the CRX includes a scoreboard on which trainees see their own scores and those of their peers (rf. Fig. 2). Trainees can accumulate up to three points for each SO, with one point deducted for each incorrect attempt, resulting in zero points for three wrong attempts. The time a trainee takes to complete a SO does not affect this score. This approach is designed to encourage trainees to spend sufficient time on each SO, to review associated resources, and to consider different strategies before submitting a solution. If trainees feel constrained by viewing their peers' scores, they have the option to temporarily hide the scoreboard, giving them greater control over their individual learning process.

P6: Collaboration. In the initial version of our CRX design, we opted for a design mode in which participants could take part in the exercises individually from home. We call this initial design mode, in which the principle of competition (P5) is implemented without a collaborative element (P6), the competitive-only design mode. Collaboration was later on implemented in an updated design, which we refer to as the competitive–collaborative design mode. Here, trainees collaborate in small teams while competing against the other participating teams. The implementation of the collaborative design mode is illustrated in Fig. 3. Each trainee within a team can individually access the CRX platform through their own device. This setup enables trainees to interact with the SIEM system and explore instructional materials autonomously without the need to coordinate with other team members. The collaborative aspect comes into play during the submission of SO solutions, which occurs on a designated submission platform on a separate device for each group. This design aims to facilitate individual immersion in the scenario, allowing trainees to independently engage with it before coming together to collaboratively analyze the scenario and tackle the solution of a SO. The user study described in Sections 5 and 6 comprises both design modes. One group of trainees used the competitive-only design mode of the CRX, while the other group used the competitive–collaborative design mode. In addition to examining the effectiveness of the overall CRX design, the availability of data sets for participants in each design mode allowed us to examine the effect of integrating collaboration in the CRX design by comparing the learning experience and learning outcomes for the two groups of trainees.

5. Method

In the following Section, we describe the methodological approach of the user study we conducted to evaluate our CRX design. To accommodate the exploratory nature of our study, we opted for evaluation questions instead of hypotheses. These evaluation questions refine the central research question of this paper raised in the introduction by evaluating if aligning a CRX design with our design principles helps to create a learning experience that is motivating for participants, fosters effective collaboration, and, ultimately, leads to effective learning outcomes. The first two evaluation questions address whether participants perceived the exercise as a positive learning experience and the extent to which they achieved defined learning outcomes. This corresponds to the effectiveness goals defined in Section 4.1.

EQ1. Do participants perceive the exercise as a positive learning experience?

EQ2. Does the CRX design enable participants to achieve the exercise's learning outcomes?

To understand the effect of the CRX design more deeply, we investigated what influenced trainees' learning outcomes. To this end, we assessed the impact of their learning experience on the achievement of learning outcomes.

EQ3. To what extent does participants' learning experience affect their achievement of learning outcomes?

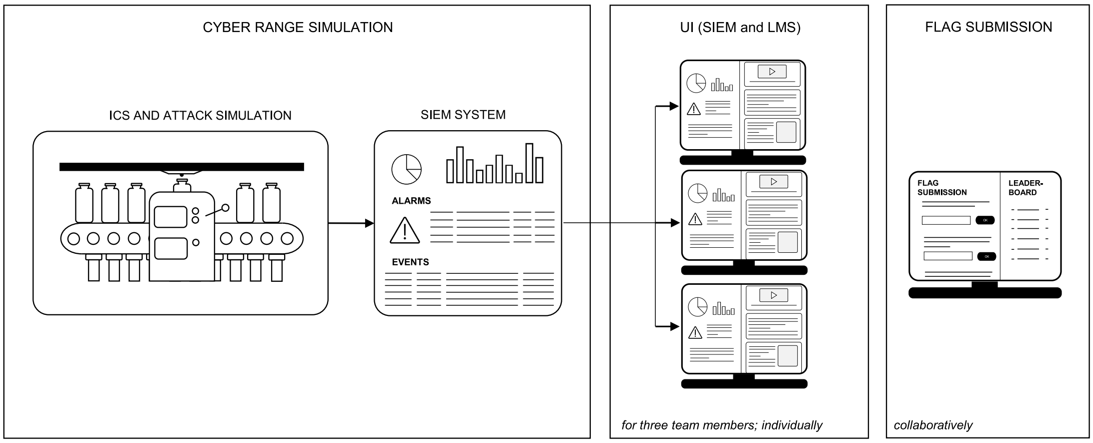


Fig. 3. Collaborative CRX design: The CRX design consists of three building blocks: (1) the cyber range simulation, (2) the CRX application with one access for each team member, and (3) the flag submission tool for collaborative flag submission.

Finally, as described above, we leverage the data collected from participants of the two design modes of the CRX (i.e., competitive-only vs. competitive–collaborative design mode) to explore the impact of the collaborative design.

EQ4. Is the CRX more effective in the competitive–collaborative design mode than in the competitive-only design mode?

5.1. Sample and data collection procedures

The research question of this paper was addressed in a study with $N = 89$ cybersecurity novices, with $n = 48$ participants using the competitive-only design mode of the CRX and $n = 41$ participants using the competitive–collaborative design mode. The study participants were drawn from students of undergraduate and graduate-level study programs in management information systems and human–computer interaction at a German university.

The data collection for the competitive-only CRX took part in five separate training sessions between December 2021 and May 2022. The assessment of the competitive–collaborative CRX took part in five sessions between December 2022 and May 2023. Students were recruited via online and in-class announcements about the upcoming CRX sessions. The announcement said the CRX gave students a chance to participate in a "hands-on cybersecurity training that gave them practical insights into the working procedures of a security analyst as part of a research study." Participants did not gain monetary compensation but actively signed up to take part in the training. Participants could end the exercise or withdraw their consent to participate in the study at any time. The study data was anonymized right after the collection of data. The study design was approved by the ethics committee of the first author's university. Informed consent was obtained from each participant at the beginning of the study. From initially 65 participants in the competitive-only CRX and 53 participants in the competitive–collaborative CRX, 17 participants in the competitive-only group and five participants in the competitive–collaborative group did not finish either the exercise or the evaluation questionnaires. The data from these participants were removed from the study, which resulted in an overall sample size of 89 participants. Learning outcomes (G3, G4) were assessed before and after the training to enable pre-post comparisons, and incident response performance was assessed throughout the training. All other variables were assessed cross-sectionally after the training. See Table 1 for an overview of the variables included in the questionnaires used in the data collections for the competitive-only and the competitive–collaborative groups, respectively. Due to COVID-19 restrictions at the time of the study, the participants in the competitive-only group attended the study remotely from home. The participants attended a video call to receive instructions for accessing the CRX platform and to have the possibility to speak to their trainer. For participants in the competitive–collaborative group, the training was conducted in-class at the university, where participants were randomly assigned to groups of two to three students, depending on the number of participants in the respective training session. On average, participants took $M = 01{:}02{:}24$ h to finish the exercise ($SD = 00{:}15{:}03$). Regarding background variables, 30.33% ($n = 27$) of participants defined themselves as female, and the rest as male. 69.67% ($n = 62$) of participants were undergraduate students; the rest were graduate students. With regard to the distribution in the two groups, the proportion of female participants was 25% ($n = 12$) in the competitive-only group and 36.59% ($n = 15$) in the competitive–collaborative group. The proportion of undergraduate students was 70.83% ($n = 34$) in the competitive-only group and 68.29% ($n = 28$) in the competitive–collaborative group. The study described in Glas et al. (2023) references a segment of the data of the ongoing user study, namely the data collected from participants of the competitive-only CRX. However, in this work, we solely utilized the description of the user study to demonstrate the application of the evaluation framework.

5.2. Measures

In the following, we describe the measures we used to address the evaluation questions. An overview of all measures is given in Table 1, listing the variables that were assessed for all 89 participants and the collaboration-related variables that were additionally assessed for the 41 participants of the competitive–collaborative CRX.

Learning Experience. Participants' engagement in the exercise (G1) was measured with an 8-item scale based on the ARCS model by Keller (Keller, 1987) that proposes four conditions (i.e., Attention, Relevance, Confidence, and Satisfaction) which a learning environment must meet to motivate trainees (e.g., 'In my opinion, the tasks I performed represent skills that are required in real incident response work'). Within the 8-item scale, hereafter referred to as *ARCS*, each condition was represented with two items ($\alpha = 0.78$). For the assessment of participants' teamwork experience (G2), the two variables *perceived relatedness* and *teamwork satisfaction* were used as indicators. For the measurement of *perceived relatedness*, which relates to participants' feelings of belonging to their team members, six items from the basic psychological needs satisfaction at work scale (Deci and Ryan, 2000; Johnston and Finney, 2010) were adapted to the training

Table 1
Overview of the evaluation measures and their allocation to the two design modes competitive-only (CO) and competitive–collaborative (CC).

Goal category	ID	Goal	Variable	CO	CC
Learning Experience	G1	Engaging learning experience	Attention, relevance, confidence and satisfaction (ARCS)	•	•
	G2	Positive collaboration experience	Perceived relatedness		•
			Teamwork satisfaction		•
Learning Outcomes	G3	ICS security knowledge	ICS network knowledge (NK)	•	•
			ICS attack knowledge (AK)	•	•
	G4	Incident response knowledge	SIEM knowledge (SK)	•	•
			Incident response process knowledge (PK)	•	•
			Incident response tool knowledge (TK)	•	•
	G5	Teamwork skills	Teamwork skills		•
	G6	Incident response performance	Mission performance	•	•

context (e.g., 'I collaborated actively with my teammates and did not keep to myself during the training', $\alpha = 0.89$). For the assessment of *teamwork satisfaction*, which captures participants' perceived quality of collaboration in their team and its effect on the learning process, a 6-item-scale developed by Tseng et al. (2009) was adapted. While the scale was initially developed to assess the quality of online collaboration, we adopted the scale to only include aspects that equally apply to on-site collaboration (e.g., 'Interacting with my teammates increased my motivation to learn', $\alpha = 0.93$). For the scales measuring *ARCS*, *teamwork satisfaction* and *perceived relatedness*, a 5-point Likert-type response format ranging from 1 = 'Fully disagree' to 5 = 'Fully agree' was used.

Learning Outcome. To measure cognitive learning outcomes, we subdivided the two cognitive effectiveness goals (G3, G4) into five knowledge categories: ICS network knowledge (NK), ICS attack knowledge (AK), SIEM knowledge (SK), incident response process knowledge (PK) and incident response tool knowledge (TK). The allocation of the knowledge categories to the outcome goals is shown in Table 1. Each knowledge category was assessed with three questionnaire items (e.g., 'How can you identify the attacking host during an ARP-based man-in-the-middle attack?') employing a single-choice format with four answers to choose from for every question (i.e., one correct answer, three distractors). In a pre-post design, participants answered questions directly before and after the exercise. The variables $exercise_{pre}$ and $exercise_{post}$ represent the mean percentage of correctly answered questions in the respective knowledge category in the pretest and posttest, and $exercise_{\Delta}$ indicates the difference between posttest and pretest results (i.e., $exercise_{post} - exercise_{pre}$). Regarding *teamwork skills* (G5), which refers to individuals' ability to effectively collaborate with others, we opted for a self-assessment to capture the context-dependent nature of the construct that cannot be objectively assessed by external raters. Specifically, a self-assessment instrument was adapted from Britton et al. (2015) captures the degree to which the CRX helped participants to improve their teamwork skills (e.g., 'Participating in the training together with my teammates has helped me strengthen the ability to participate actively and take a fair share of the group work', $\alpha = 0.94$). The response format was a 5-point Likert-type scale ranging from 1 = 'Fully disagree' to 5 = 'Fully agree'. Finally, participants' capacity to transfer their acquired knowledge and skills to the performance of practical incident response tasks (G6) was evaluated by assessing their *mission performance*, that is, their ability to succeed in the eight SOs encompassed in the CRX as indicated by an overall score that was calculated by summing up all individual SO scores. Participants could reach a maximum of 24 points, which decreases by one point for each wrong attempt or activated hint.

We made all questionnaires and the full data set of the user study are available in the aforementioned GitHub project.

5.3. Analyses

Regarding participants' perceived learning experience (EQ1) a descriptive analysis (i.e., means and standard deviations) was carried out for *ARCS*, *perceived relatedness*, and *teamwork satisfaction*. To address EQ2, we conducted five separate ANCOVAs (analyses of variance including covariates) to examine the overall effect of participating in the CRX (i.e., *exercise*) on participants' knowledge gain in the five knowledge categories *NK*, *AK*, *SK*, *OK*, and *TK*. This was accomplished by comparing pre- and posttest scores (i.e., $exercise_{pre}$ and $exercise_{post}$) while controlling for the covariates *design mode*, *education*, and *gender*. As no longitudinal data were gathered for participants' perceived improvement of *teamwork skills* and their *mission performance*, merely a descriptive analysis was conducted for these variables. Regarding EQ3, we examined the effect of participants' learning experience (i.e., *ARCS*, *perceived relatedness*, and *teamwork satisfaction*) on their learning outcomes as represented by pre-post differences for *NK*, *AK*, *SK*, *OK*, and *TK* as well as by *mission performance* and *teamwork skills*. In this regard, multiple linear regression analyses were performed to test the predictive effect of the learning experience variables on the learning outcomes variables in a multivariate setting. while additionally controlling for the background variables *gender* and *education*. For EQ4, the ANCOVAs conducted for EQ2 were analyzed with a focus on the impact of the within-variable *design mode* (i.e., competitive–collaborative vs. competitive only) on participants' learning outcomes. Two additional ANCOVAs were conducted to examine the effect of *design mode* on *ARCS* and *mission performance*, respectively, again controlling for the covariates *gender* and *education*.

6. Results

The following section outlines the key findings of our study by presenting the results of our analyses in order of the four proposed evaluation questions.

EQ1. Regardless of the particular CRX design mode (i.e., competitive-only and competitive–collaborative), the mean rating of all 89 participants for *ARCS* was notably high at 4.13 ($SD = 0.47$). In addition, the 41 participants of the competitive–collaborative CRX rated the *perceived relatedness* to their team members ($M = 4.49, SD = 0.52$) and their *teamwork satisfaction* ($M = 4.48, SD = 0.57$) very positively. These findings indicate that the participants perceived the CRX design both as engaging and conducive to a positive collaboration experience.

EQ2. The mean difference between pre- and posttest in the five knowledge categories was between 18% (SK) and 33% (TK) correctly answered questions. The ANCOVAs showed a significant main effect of participating in the CRX (*exercise*) for all knowledge categories while controlling for the background variables *gender* and *education* (rf. Table 2). This indicates that the CRX design is highly effective in facilitating participants' acquisition of knowledge. Furthermore, participants' mean *mission performance* was at 18.63 points ($SD = 5.5$),

Table 2
Results of the ANCOVAs analyzing the effect of CRX participation (*exercise*) on learning outcomes and the interaction with *design mode* ($N = 89$).

	NK				AK				SK				PK				TK			
	M	SD	F	η_p^2	M	SD	F	η_p^2	M	SD	F	η_p^2	M	SD	F	η_p^2	M	SD	F	η_p^2
exercise$_{pre}$	.47	.30			.59	.27			.62	.32			.59	.26			.58	.31		
exercise$_{post}$	.67	.24			.82	.24			.80	.24			.78	.25			.95	.12		
exercise$_{\Delta}$	.20	.33	10.54**	.11	.23	.30	27.83***	.25	.18	.31	23.34***	.21	.19	.27	26.24***	.37	.33	.34	27.83***	.25
exercise$_{\Delta}$ * design mode			.59	.01			1.45	.02			4.98*	.06			3.65	.04			1.98	.02
exercise$_{\Delta}$ * gender			.43	.01			1.24	.01			.01	.00			.18	.00			1.05	.01
exercise$_{\Delta}$ * education			.72	.01			4.67	.05			2.33	.03			.42	.01			.77	.01

Note. $df_b = 1$, $df_w = 83$; NK: ICS network knowledge, AK: ICS attack knowledge, SK: SIEM knowledge, PK: incident response process knowledge, TK: incident response tool knowledge; means scores for exercise$_{pre}$, exercise$_{post}$, and $exercise_{\Delta}$ refer to the mean percentage of correctly answered questions in the respective knowledge category.

** $p < .01$.

*** $p < .001$

which means that participants, on average, needed less than six wrong attempts or activated hints to solve the eight SOs of the exercise, indicating that they were generally able to successfully apply their acquired knowledge to practical tasks. Finally, the mean rating of the 41 participants of the competitive–collaborative group for *teamwork skills* was 4.12 ($SD = 0.64$), suggesting a strong positive response towards the CRX's potential to improve one's teamwork skills.

EQ3. For EQ3, we analyzed the impact of learning experience on the achievement of learning outcomes while controlling for *gender* and *education*. The results of the regression analyses showed that participants' *education* level had an effect on *mission performance* and learning outcomes. For learning outcomes, however, this effect was only observed for one knowledge category (AK). Moreover, the analyses revealed a significant positive effect of *ARCS* on *mission performance* ($\beta = .48, p < .001$) (rf. Table 3 in the appendix). This finding indicates that a positive learning experience positively affects participants' ability to solve practical incident response tasks within the exercise. An effect of *ARCS* on learning outcomes was not observed.

With regard to the variables exclusively measured within the competitive–collaborative group ($n = 41$), we first conducted a correlation analysis for *perceived relatedness* and *teamwork satisfaction* in order to account for the strong theoretical connection between the two constructs. As the results showed that the variables were indeed strongly correlated ($r = .83$) we combined them into the composite variable *teamwork experience* ($M = 4.49, SD = 0.52$). The subsequent regression analysis showed that *teamwork experience* had a significant positive effect on trainees' improvement of*teamwork skills* ($\beta = .63, p < .001$) (rf. Table 4 in the appendix). An effect of *teamwork experience* on ARCS on learning outcomes was not observed.

EQ4. As already discussed with respect to EQ1, the participation in the CRX led to a significant increase in knowledge across all knowledge categories. However, when comparing the competitive-only and the competitive–collaborative group, descriptive statistics for pre- and posttest scores (Table 5 in the appendix) and the results of the corresponding ANCOVAs (see Table 2) showed that a significant difference between the two design modes was only observable for one knowledge category (SK), where participants in the competitive–collaborative group demonstrated significantly better results. In addition, *mission performance* in the competitive–collaborative group ($M = 21.54, SD = 2.41$) was significantly better than in the competitive-only group ($M = 16.15, SD = 6.21$), $F(1, 85) = 34.86, p < .001$.

7. Discussion

The results of the user study show that the CRX design is both capable of shaping trainees' perception of the CRX as a positive learning experience (EQ1) and of facilitating the achievement of learning outcomes (EQ2). The self-assessed scores relating to EQ1 show a very positive perceived learning experience participating in the CRX (G1, G2). This indicates that the introduced design principles were able to foster authentic learning, which, in turn, created a positive learning experience for participants. Regarding *perceived relatedness* and *teamwork satisfaction* in particular, it is noteworthy that participants rated these aspects of the exercise very positively despite being randomly assigned to their team. In terms of cognitive learning outcomes (G3, G4), the significant increase in participants' knowledge across all knowledge categories demonstrates the ability of the CRX design to enable trainees to effectively acquire knowledge through participation in the CRX. Albeit the self-assessed improvement of *teamwork skills* (G5) does not allow to determine an objective improvement in skill levels, the participants' high scores still serve as an indicator that the CRX design fostered the development of teamwork skills. The interpretation of participants' *mission performance* is complex because of the exercise's scoring mechanism. The CRX design does not aim at optimizing *mission performance* but at optimizing learning. This is because making mistakes, learning from them, and using hints for guidance lowers trainees' scores respectively *mission performance*; however, these actions are seen as an integral part of trainees' learning process. Nonetheless, *mission performance* still indicates whether participants were generally able to apply their new knowledge and skills effectively (G6). In summary, the user study results demonstrate that the CRX design fulfills all six predefined goals established to assess its effectiveness. This indicates that the design principles upon which the CRX design is built facilitate effective learning.

Investigating the relationship between learning experience and learning outcomes in detail (EQ3), our study shows that *ARCS* had a positive effect on *mission performance*. This indicates that trainees who feel engaged by the exercise are better able to translate acquired knowledge and skills into practical performance. This finding advocates for the continued integration of motivational aspects in the design of CRXs. Regarding collaboration, the analysis of EQ3 showed a positive effect of *teamwork experience* on the CRX's capability to let trainees practice and, in turn, improve their *teamwork skills*. This finding indicates that simply including collaboration in a CRX design does not automatically lead to the improvement of *teamwork skills*. Only if a CRX promotes respectful and constructive teamwork through its design will trainees effectively collaborate with their team members. This is consistent with Maennel et al. (2023), who propose to integrate considerations on how to ensure a high quality of teamwork into CRX designs to achieve better learning outcomes.

The evolution of the CRX design from a competitive-only to a competitive–collaborative design mode allowed us to investigate the effect integrating collaborative learning into the CRX design exerts on the achievement of learning outcomes (EQ4). In this respect, the results of the study showed that although the competitive–collaborative design mode hardly had an impact on participants' knowledge and skills acquisition, participants in the competitive–collaborative group showed a significantly better *mission performance*. Concerning the latter finding, we argue that this superior *mission performance* in the competitive–collaborative group can be attributed to the collaborative nature of the design, which allows trainees to exchange ideas before submitting solutions. We may thus conclude that this finding is an indication of effective collaboration rather than evidence for an inherently better transfer of knowledge and skills in this group. To summarize, the two design modes met the exercise goals comparably well. On top of the acquisition of cognitive learning outcomes, though, the competitive–collaborative design mode effectively enables trainees to have a positive experience with collaboration in a team (G2) and subsequently improve their teamwork skills. Consequently, we highlight the instructional value of collaboration and advocate to include this design principle in CRXs for novice trainees.

8. Limitations and future work

There are some limitations to our work that we would like to acknowledge in this section and outline how they will be addressed in future work. Our user study focuses on participants' experience within the exercise and their immediate learning outcomes. Consequently, the study does not provide conclusive evidence regarding the CRX's long-term effectiveness. Given that learning experience and knowledge and skills acquisition are essential for long-term learning (Kirkpatrick, 2005), the results of the study can still provide valuable information about the quality and effectiveness of the CRX design. Nevertheless, assessing the long-term effectiveness of the CRX design will be the subject of future research. Specifically, we want to examine the impact of the CRX design within an organizational context, assessing whether it can contribute to positively shaping the working practices of emerging cybersecurity professionals.

Moreover, the design of the user study was based on the goal of evaluating the effectiveness of the CRX design. While it can be concluded that the CRX design contributed to the achievement of these goals, the study design does not allow for the effectiveness of the overall design to be related to our design principles (except for collaboration). The subject of future studies, therefore, should be to examine the effectiveness of the individual design principles in the context of a CRX design.

Finally, the design principles we propose only encompass those we selected to be most relevant for creating complex and authentic CRXs for cybersecurity novices. Thus, we make no claim to completeness but aim to offer an initial framework for CRX designers that can and should be refined and extended. Crucially, our work aims to encourage CRX designers to foster the potential that instructional models have to offer for crafting more effective learning environments in the field of cybersecurity. A possible extension of our framework, for instance, contemplates larger-scale CRXs spanning multiple days. For this purpose, it would be promising to consider the integration of multiple representations (Spiro et al., 1994) as a design principle. This principle advocates that trainees should approach a problem from multiple perspectives and in multiple contexts — an undertaking that was not feasible within the relatively short duration of our current CRX design.

9. Conclusion

This paper demonstrates the potential of an interdisciplinary approach to CRX design. We leverage design principles rooted in the field of instructional design to create a CRX for cybersecurity novices. Our study shows that this design effectively promotes the acquisition of new knowledge and skills and facilitates their practical application in an authentic attack scenario. With this approach, we aim to demonstrate how to create CRXs that are better attainable to cybersecurity novices and, thus, have the potential to prepare more individuals with the knowledge and skills to take on a position in cybersecurity. Regarding practical implications, we aim to encourage CRX designers in academia and industry to acknowledge CRXs as complex learning environments that require a nuanced understanding of how individuals learn. Incorporating expertise from other disciplines helps to gain this understanding and shape new foundations for cybersecurity training and education.

Availability of source code and study data

A detailed description of the implementation of the CRX and the full source code (structured in four repositories) are available at the following GitHub project: (https://github.com/InstruCRX). Furthermore, the project contains all questionnaires and the full data set of the user study.

CRediT authorship contribution statement

Magdalena Glas: Conceptualization, Methodology, Software, Writing – original draft. **Gerhard Messmann:** Conceptualization, Methodology, Writing – review & editing. **Günther Pernul:** Supervision, Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The full source code and study data is shared over GitHub: https://github.com/InstruCRX.

Declaration of Generative AI Technologies in the Writing Process

During the preparation of this work, the authors used ChatGPT (Model GPT-3.5) in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Acknowledgments

First, we want to thank the participants of the user study for their interest in our research and their participation in the study, without which this research would not have been possible. Furthermore, we want to thank Muhammad Mudassar Yamin for his valuable comments on this manuscript. This work was partly performed under the INSIST project, which is supported under contract by the Bavarian Ministry of Economic Affairs, Regional Development and Energy, Germany (DIK0338/01).

Appendix

A.1. Regression analyses (EQ3)

See Tables 3 and 4.

A.2. Comparison of the two design modes (EQ4)

See Table 5.

Table 3
Effects of learning experience (*ARCS*) on learning outcomes (*NK*, *AK*, *SK*, *PK*, *TK*), and *mission performance* (MP) ($N = 89$).

	NK	AK	SK	PK	TK	MP
gender						
education		−.25*				.20*
ARCS	−.13	.12	.10	.05	−.01	.48***
R^2	.03	.09	.04	.01	.01	.34

Note. NK: ICS network knowledge, AK: ICS attack knowledge, SK: SIEM knowledge, PK: Incident response process knowledge, TK: Incident response tool knowledge, MP: mission performance.

* $p < .05$.

*** $p < .001$.

Table 4
Effects of learning experience on learning outcomes including variables additionally assessed for the competitive–collaborative group (i.e., *teamwork experience* and *teamwork skills*) ($n = 41$).

	NK	AK	SK	PK	TK	MP	TS
gender						−.37*	
education							
ARCS	−.13	−.03	.17	.11	.04	.47**	.14
TX	.09	.29	.15	.06	−.04	.10	.63***
R^2	.04	.09	.12	.05	.03	.40	.48

Note. NK: ICS network knowledge, AK: ICS attack knowledge, SK: SIEM knowledge, PK: Incident response process knowledge, TK: Incident response tool knowledge, MP: mission performance, TS: teamwork skills, TX: teamwork experience.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 5
Results of average of correctly answered questions in pre and post test among the two groups.

		NK		AK		SK		PK		TK	
		M	SD	M	SD	M	SD	M	SD	M	SD
CC	exercise$_{pre}$	.38	.26	.56	.29	.50	.30	.49	.21	.61	.33
	exercise$_{post}$	.61	.21	.75	.28	.75	.26	.75	.27	.93	.14
	exercise$_{\Delta}$	.23	.33	.19	.33	.25	.35	.26	.26	.32	.39
CO	exercise$_{pre}$	.54	.30	.62	.25	.73	.30	.67	.28	.56	.28
	exercise$_{post}$	.71	.25	.88	.19	.84	.22	.81	.23	.97	.09
	exercise$_{\Delta}$	.17	.33	.26	.29	.11	.27	.15	.27	.41	.29

Note. CC: Competitive–collaborative design mode, CO: competitive-only design mode, NK: ICS network knowledge, AK: ICS attack knowledge, SK : SIEM knowledge, PK: incident response process knowledge, TK: incident response tool knowledge.

References

Accenture, 2023. Accenture Security ICS Cyber Range. https://www.accenture.com/us-en/services/security/cyber-resilience. (Accessed 28 May 2024).

Airbus, 2023. Airbus CyberRange: An advanced simulation solution. https://www.cyber.airbus.com/cyberrange. (Accessed 28 May 2024).

Andreolini, M., Colacino, V.G., Colajanni, M., Marchetti, M., 2020. A framework for the evaluation of trainee performance in cyber range exercises. Mob. Netw. Appl. 25 (1), 236–247.

Backman, N., 2016. Facilitating a battle between hackers: Computer security outside of the classroom. In: Proceedings of the 47th ACM Technical Symposium on Computing Science Education. SIGCSE '16, ACM.

Beuran, R., Tang, D., Pham, C., ichi Chinen, K., Tan, Y., Shinoda, Y., 2018. Integrated framework for hands-on cybersecurity training: Cytrone. Comput. Secur. 78, 43–59.

Blažič, B.J., 2021. The cybersecurity labour shortage in europe: Moving to a new concept for education and training. Technol. Soc. 67, 101769.

Braghin, C., Cimato, S., Damiani, E., Frati, F., Riccobene, E., Astaneh, S., 2020. Towards the monitoring and evaluation of trainees' activities in cyber ranges. In: Proceedings of the 2020 Conference on Model-Driven Simulation and Training Environments for Cybersecurity. MSTEC 2020. Springer International Publishing, pp. 79–91.

Brilingaitė, A., Bukauskas, L., Juozapavičius, A., 2020. A framework for competence development and assessment in hybrid cybersecurity exercises. Comput. Secur. 88, 101607.

Brilingaitė, A., Bukauskas, L., Kutka, E., 2017. Development of an educational platform for cyber defence training. In: Proceedings of the 2017 European Conference on Cyber Warfare and Security. Academic Conferences International Limited, pp. 73–81.

Britton, E., Simper, N., Leger, A., Stephenson, J., 2015. Assessing teamwork in undergraduate education: a measurement tool to evaluate individual teamwork skills. Assess. Eval. Higher Educ. 42 (3), 378–397.

Bueermann, G., Doyle, S., 2023. Global Cybersecurity Outlook 2023. Technical Report, World Economic Forum, Geneva, Switzerland.

Čeleda, P., Čegan, J., Vykopal, J., Tovarňák, D., et al., 2015. Kypo–a platform for cyber defence exercises. In: M&S Support to Operational Tasks Including War Gaming, Logistics, Cyber Defence. NATO Science and Technology Organization.

Collins, A., Brown, J.S., Newman, S.E., 1991. Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In: Idol, L., Jones, B.F. (Eds.), Educational Values and Cognitive Instruction: Implications for Reform. Routledge, pp. 453–494.

Deci, E.L., Ryan, R.M., 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychol. Inq. 11 (4), 227–268.

Eccles, J.S., Wigfield, A., 2002. Motivational beliefs, values, and goals. Annu. Rev. Psychol. 53 (1), 109–132.

Edgar, T.W., Manz, D.O., 2017. Hypothetico-deductive research. In: Research Methods for Cyber Security. Elsevier, pp. 215–249.

Ferguson, B., Tall, A., Olsen, D., 2014. National cyber range overview. In: Proceedings of the 2014 IEEE Military Communications Conference. pp. 123–128.

Glas, M., Vielberth, M., Pernul, G., 2023. Train as you fight: Evaluating authentic cybersecurity training in cyber ranges. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. CHI '23, Association for Computing Machinery, New York, NY, USA.

Glaser, R., 1976. Cognitive psychology and instructional design. In: Klahr, D. (Ed.), Cognition and Instruction. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 304–315.

Hatzivasilis, G., Ioannidis, S., Smyrlis, M., Spanoudakis, G., Frati, F., Braghin, C., Damiani, E., Koshutanski, H., Tsakirakis, G., Hildebrandt, T., Goeke, L., Pape, S., Blinder, O., Vinov, M., Leftheriotis, G., Kunc, M., Oikonomou, F., Magilo, G., Petrarolo, V., Chieti, A., Bordianu, R., 2021. The THREAT-arrest cyber range platform. In: Proceedings of the 2021 IEEE International Conference on Cyber Security and Resilience. CSR, pp. 422–427.

Hmelo-Silver, C.E., 2004. Problem-based learning: What and how do students learn? Educat. Psychol. Rev. 16 (3), 235–266.

Hmelo-Silver, C.E., Barrows, H.S., 2006. Goals and strategies of a problem-based learning facilitator. Interdiscip. J. Probl. Based Learn. 1, 21–39.

Honebein, P.C., Duffy, T.M., Fishman, B.J., 1993. Constructivism and the design of learning environments: Context and authentic activities for learning. In: Duffy, T.M., Lowyck, J., Jonassen, D.H. (Eds.), Designing Environments for Constructive Learning. Springer, Berlin, pp. 87–108.

IBM, 2023. IBM Security X-Force Cyber Range. https://www.ibm.com/services/security-operations-center. (Accessed 28 May 2024).

ISACA, 2022. State of Cybersecurity 2022: Global Update on Workforce Efforts, Resources and Cyberoperations. Technical Report.

(ISC)^2, 2023. (ISC)^2 cybersecurity workforce study 2023 - how the economy, skills gap and artificial intelligence are challenging the global cybersecurity workforce. Technical Report, International Information System Security Certification Consortium, pp. 1–84.

Johnston, M.M., Finney, S.J., 2010. Measuring basic needs satisfaction: Evaluating previous research and conducting new psychometric evaluations of the Basic Needs Satisfaction in General Scale. Contemp. Educ. Psychol. 35 (4), 280–296.

Keller, J.M., 1987. Development and use of the ARCS model of instructional design. J. Instruct. Dev. 10 (3), 2–10.

Kick, J., 2014. Cyber Exercise Playbook. Technical Report, MITRE Corporation.

Kim, J., Maeng, Y., Jang, M., 2019. Becoming invisible hands of national live-fire attack-defense cyber exercise. In: Proceedings of the 2019 IEEE European Symposium on Security and Privacy Workshops. EuroS&PW, pp. 77–84.

Kirkpatrick, D.L., 2005. Evaluating Training Programs, third ed. Berrett-Koehler, p. 379.

Kirschner, P., van Merriënboer, J., 2008. Ten steps to complex learning: A new approach to instruction and instructional design. In: Good, T. (Ed.), 21st Century Education, Vol. 1, first ed. SAGE Publications Ltd, United Kingdom, pp. 244–253.

Kolodner, J.L., Camp, P.J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., Puntambekar, S., Ryan, M., 2003. Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by design(tm) into practice. J. Learn. Sci. 12 (4), 495–547.

Leitner, M., Frank, M., Hotwagner, W., Langner, G., Maurhart, O., Pahi, T., Reuter, L., Skopik, F., Smith, P., Warum, M., 2020. AIT cyber range: Flexible cyber security environment for exercises, training and research. In: Proceedings of the 2020 European Interdisciplinary Cybersecurity Conference. EICC 2020, Association for Computing Machinery, New York, NY, USA.

Maennel, K., Brilingaitė, A., Bukauskas, L., Juozapavičius, A., Knox, B.J., Lugo, R.G., Maennel, O., Majore, G., Sütterlin, S., 2023. A multidimensional cyber defense exercise: Emphasis on emotional, social, and cognitive aspects. SAGE Open 13 (1), 215824402311563.

Maennel, K., Ottis, R., Maennel, O., 2017. Improving and measuring learning effectiveness at cyber defense exercises. In: Lipmaa, H., Mitrokotsa, A., Matulevičius, R. (Eds.), Proceedings of the 2017 Conference on Secure IT Systems. NordSec 2017, Springer International Publishing, Cham, pp. 123–138.

Merrill, M.D., 2002. First principles of instruction. Educ. Technol. Res. Dev. 50 (3), 43–59.

Messmann, G., 2022. Fostering proactive behaviour: The role of work-related reflection, psychological empowerment, and participative safety for innovative behaviour and job crafting. Int. J. Train. Dev. 27 (1), 99–116.

Mirkovic, J., Dark, M., Du, W., Vigna, G., Denning, T., 2015a. Evaluating cybersecurity education interventions: Three case studies. IEEE Secur. Privacy 13 (3), 63–69.

Mirkovic, J., Peterson, P.A.H., 2014. Class Capture-the-Flag exercises. In: 2014 USENIX Summit on Gaming, Games, and Gamification in Security Education. 3GSE 14, USENIX Association, San Diego, CA.

Mirkovic, J., Tabor, A., Woo, S., Pusey, P., 2015b. Engaging novices in cybersecurity competitions: A vision and lessons learned at ACM tapia 2015. In: 2015 USENIX Summit on Gaming, Games, and Gamification in Security Education. 3GSE 15, USENIX Association, Washington, D.C..

National Initiative for Cybersecurity Education (NICE), 2020. The Cyber Range: A Guide. Technical report, National Initiative for Cybersecurity Education (NICE).

Nelson, L.M., 1999. Collaborative problem solving. In: Reigeluth, C.M. (Ed.), Instructional-Design Theories and Models. Routledge.

Oltsik, J., Lundell, B., 2021. The Life and Times of Cybersecurity Professionals. Technical report, The Enterprise Strategy Group (ESG) and Information Systems Security Association International (ISSA).

Owens, K., Fulton, A., Jones, L., Carlisle, M., 2019. pico-boo!: How to avoid scaring students away in a ctf competition. In: Proceedings of the Colloquium for Information System Security Education.

Reigeluth, C.M., 1999. Instructional-Design Theories and Models. Lawrence Erlbaum, p. 728.

Ryan, R.M., Deci, E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 55 (1), 68–78.

Schank, R.C., 1999. Dynamic Memory Revisited. Cambridge University Press.

Schank, R.C., Berman, T.R., Macpherson, K.A., 1999. Learning by doing: Goal based scenarios. In: Reigeluth, C.M. (Ed.), Instructional-Design Theories and Models. Routledge, pp. 161–181.

Spiro, R., Coulson, R., Feltovich, P., Anderson, D., 1994. Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In: Singer, H., Ruddell, R.B. (Eds.), Theoretical Models and Processes of Reading, fourth ed. International Reading Association, Newark, NJ, pp. 544–557.

Švábenský, V., Vykopal, J., Cermak, M., Laštovička, M., 2018. Enhancing cybersecurity skills by creating serious games. In: Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education. ITiCSE 2018, Association for Computing Machinery, New York, NY, USA, pp. 194–199.

Švábenský, V., Weiss, R., Cook, J., Vykopal, J., Čeleda, P., Mache, J., Chudovský, R., Chattopadhyay, A., 2022. Evaluating two approaches to assessing student progress in cybersecurity exercises. In: Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1. SIGCSE 2022, Association for Computing Machinery, New York, NY, USA, pp. 787–793.

Sweller, J., 2010. Element interactivity and intrinsic, extraneous, and Germane cognitive load. Educat. Psychol. Rev. 22 (2), 123–138.

Tobey, D.H., Pusey, P., Burley, D.L., 2014. Engaging learners in cybersecurity careers: Lessons from the launch of the national cyber league. ACM Inroads 5 (1), 53–56.

Tseng, H., Ku, H.-Y., Wang, C.-H., Sun, L., 2009. Key factors in online collaboration and their relationship to teamwork satisfaction. Q. Rev. Distance Educ. 10, 195+.

van Merriënboer, J.J.G., Seel, N.M., Kirschner, P.A., 2002. Mental models as a new foundation for instructional design. Educ. Technol. 42 (2), 60–66.

Votipka, D., Zhang, E., Mazurek, M.L., 2021. Hacked: A pedagogical analysis of online vulnerability discovery exercises. In: 2021 IEEE Symposium on Security and Privacy. SP, IEEE.

Vykopal, J., Barták, M., 2016. On the design of security games: From frustrating to engaging learning. In: Proceedings of the 2016 USENIX Workshop on Advances in Security Education (ASE 16). USENIX Association, Austin, TX.

Vykopal, J., Ošlejšek, R., Burská, K., Zákopčanová, K., 2018. Timely feedback in unstructured cybersecurity exercises. In: Proceedings of the 49th ACM Technical Symposium on Computer Science Education. SIGCSE '18, Association for Computing Machinery, New York, NY, USA, pp. 173–178.

Vykopal, J., Vizvary, M., Oslejsek, R., Celeda, P., Tovarnak, D., 2017. Lessons learned from complex hands-on defence exercises in a cyber range. In: Proceedings of the 2017 IEEE Frontiers in Education Conference. FIE, pp. 1–8.

Weiss, R., Turbak, F., Mache, J., Nilsen, E., Locasto, M.E., 2016. Finding the balance between guidance and independence in cybersecurity exercises. In: Proceedings of the 2016 USENIX Workshop on Advances in Security Education. ASE 16, USENIX Association, Austin, TX.

Yamin, M.M., Katt, B., Gkioulos, V., 2020. Cyber ranges and security testbeds: Scenarios, functions, tools and architecture. Comput. Secur. 88, 101636.

Zan, T.D., Di Franco, F., 2019. Cybersecurity Skills Development in the Eu: The Certification of Cybersecurity Degrees and Enisa's Higher Education Database. Technical Report, Publications Office, European Union Agency for Network and Information Security.

Magdalena Glas is a research assistant and Ph.D. candidate at the University of Regensburg, Germany. Her research interests include authentic learning environments in organizational cybersecurity and other people-centric domains of cybersecurity. She received her master's degree in management information systems from the University of Regensburg, the University College Dublin (Ireland), and the Ionian University (Greece). Contact her at magdalena.glas@ur.de.

Gerhard Messmann is an assistant professor and lecturer at the Faculty of Human Sciences at the University of Regensburg, Germany. His research interests include proactivity at work (e.g., innovative work behavior and job crafting) as well as reflection and learning from experience in formal and informal settings. He received his master's and Ph.D. degree in educational science from the University of Regensburg. Contact him at gerhard.messmann@ur.de.

Günther Pernul is a professor with the Department of Information Systems at the University of Regensburg, Germany. His research interests include different aspects of cybersecurity and advanced data-centric applications. He received his diploma and Ph.D. degree (Hons.) in business informatics from the University of Vienna, Austria. Previously, he held positions at the University of Duisburg–Essen, Germany; the University of Vienna, Austria; the University of Florida, Gainesville; and the College of Computing, Georgia Institute of Technology, Atlanta. Contact him at guenther.pernul@ur.de.