Ethical Principles for Designing Responsible Offensive Cyber Security Training

Fredrik Heiding(✉) and Robert Lagerström

KTH Royal Institute of Technology, 100 44 Stockholm, Sweden
fheiding@kth.se
https://www.kth.se/

Abstract. In this paper we present five principles for designing ethically responsible offensive cyber security training. The principles can be implemented in existing or new study plans and target both academic and non-academic courses. Subject matter experts within various cyber security domains were consulted to validate and fine tune the principles, together with a literature review of ethical studies in related domains. The background for designing the principles is the continuous popularity of offensive cyber security (penetration testing, ethical hacking). Offensive cyber security means actively trying to break or compromise a system in order to find its vulnerabilities. If this expertise is placed in the wrong hands, the person can cause severe damage to organizations, civilians and society at large. The proposed ethical principles are created in order to mitigate these risks while maintaining the upsides of offensive cyber security. This is achieved by incorporating the ethical principles in offensive cyber security training, in order to facilitate the practitioners with ethical knowledge of how and when to use their acquired expertise.

Keywords: Ethical principles · Offensive cyber security training · Ethical hacking · Penetration testing · Privacy · Security training · Ethical guideline · Ethical framework

1 Introduction

The world is continuously becoming more digital and the digitalization brings a plethora of advantages such as increased connectivity, improved societal functions, and more accessible information [5,11]. However, the rapid digitalization is followed by an increase in digital threats [15,26]. Things that historically have been analogue are getting connected to networks and thus becoming exposed to malicious tampering. In 2017 the WannaCry malware caused damages estimated to more than 4b$ and soon after, the NotPetya malware caused damages estimated to over 10b$ [8].

Published by Springer Nature Switzerland AG 2021
M. Friedewald et al. (Eds.): Privacy and Identity 2020, IFIP AICT 619, pp. 21–39, 2021.
https://doi.org/10.1007/978-3-030-72465-8_2

The increased exposure is equally true for information as for physical machines and devices. Since personal information is often just as exposed (or even more exposed) as organizational information, our relation to privacy has been severely hurt during the last decades. This is further amplified by the rapid increase of information gathering performed by both organizations and states [23]. The rapid growth of data breaches and privacy infringements are likely to keep increasing in line with the continued digitalization, so we must act to stop or at least mitigate the situation [4].

As a response to the general increase in digital threats, offensive cyber security is becoming more popular. Offensive cyber security aims to increase the security of a system by actively trying to compromise and break the system. By doing this we can find the weaknesses of the system before an attacker has time to exploit them. This can help us stay ahead of attackers and create more secure environments. However, by educating people in offensive cyber security, we help them gain expertise and knowledge that could be used to break system. If placed in the wrong hands this information can cause severe damage. It is possible that we mistakenly train a person with malicious intentions, or that we train a person with good intentions, whose motivations later become malevolent due to some circumstance of life. Company layoffs, disputes, personal grudges or a number of other reasons may motivate a person to cause harm, and with the right technical education, it is possible for the hostile person to succeed. The weapons we are discussing is knowledge which is more elusive than a traditional firearm, but as recent history shows it is more than enough to cause severe damage, both to privacy and physical safety [8].

In order to mitigate these risks, we propose five principles for designing ethically responsible offensive cyber security training. The principles can be implemented by cyber security educators and officials, in order to create the best possible outcome of offensive cyber security training, maximizing the benefits while minimizing potential harm. The rest of the article describes the creation, foundation and finalization of these principles.

Throughout the article the terms security and cyber security are used synonymously. If the authors just write "security" the reader can assume they mean cyber security, if the context does not specifically state otherwise. Further, there is no clear consensus on what differentiates an ethical guideline and an ethical framework. For simplicity, we name the ethical suggestions proposed in this article as "principles" and consider them to be, to a large part, a building block for both ethical guidelines and ethical frameworks.

1.1 Outline

The remaining article is structured as follows: Sect. 2 analyzes ethical principles from other technical domains and those available within cyber security. Section 3 describes the creation of the principles, focusing on consulting subject matter experts and drawing conclusions from relating ethical principles. Section 4 presents the result from the consultations with the subject matter experts and Sect. 5 presents the five principles in detail. Section 6 discuss surrounding

thoughts on the principles and future possibilities for continued work, and Sect. 7 concludes the article.

2 Literature Review

A literature review was done to analyze the use of ethical studies in cyber security and other technical domains. A number of different search queries were used in five major databases *(Scopus, Web of Science, IEEE Xplore, ACM Digital Library and Google Scholar)*. Initially we searched in the abstract, title, and keywords of the articles. Some selected queries were later expanded to search in the entire article, in order to capture a wider range relevant papers. In addition, some papers were found via snowballing (reading references and cited works of the current related work) and by modifying the search phrases to include more general words such as "ethics" and "technology". When searching the full text of the article or when searching with the more general keywords such as "security", the resulting matches contained several false positives and were manually scanned for relevant articles. The queries are further described in Table 1 and Table 2.

The literature review is categorized in three parts: *ethical studies in other domains, ethical studies on cyber security*, and *ethical studies on offensive cyber security*.

Table 1. Searches made in the abstract, title and keywords of the articles.

Search phrase	Scopus	WoS	IEEE	ACM	GS
"ethical framework" AND "cyber security"	6	2	0	0	2
"ethical guidelines" AND "cyber security"	16	1	0	0	1
"ethical framework" AND "it security"	0	0	0	0	1
"ethical framework" AND ("Artificial intelligence" OR "AI")	54	18	0	0	32
"ethical framework" AND "technology"	291	118	19	2	44
ethics AND "cyber security"	103	60	44	5	68

Table 2. Searches made in the entire article, include its full length text.

Search phrase	Scopus	WoS	IEEE	ACM	GS
"ethical framework" AND "cyber security"	30	2	10	6	459
"ethical guidelines" AND "cyber security"	16	1	1	12	650

2.1 Ethical Studies in Other Domains

Ethical challenges are not isolated to cyber security or offensive cyber security, nor is the challenge of educating people in potentially dangerous techniques. New knowledge can be dangerous whether it is acquired in cyber security [20], artificial intelligence [21], CBRNE subjects (chemical, biological, radiological, nuclear or explosive) [12], and many other domains. Some frameworks exist within these areas as explained below, we have studied a number of them in order to gain a better understand of how to construct our principles on offensive cyber security education.

Before creating an ethical guideline it is good to investigate what problems the guideline can help solving. For example, if the guideline treats how to digitalize health care in an ethically responsible way, it should be clear what problems may arise from digitalizing health care [18]. This provides a reason for why the guideline is needed, which should recur as a common theme throughout the different parts of the guideline [18]. In the case of health care, problems that may arise when digitalizing it include reduction of autonomy, independence, quality of life, beneficence, non-maleficence, and justice of the patients. More specific issues of cyber security includes privacy and confidentiality as well as general cyber security in itself. With that background, a guideline for ethical digitalization of health care should aspire to mitigate these given problems [18]. This line of thinking is not directly targeting the ethics of offensive cyber security training, but it presents a foundation for how ethical guidelines can be designed. Parts of this methodology will inspire our ethical principles on offensive cyber security, as shown in Sect. 3.

The actual deliverable of an ethical guideline may be a set of principles that policy-makers can use in order to evaluate decisions and processes. The principles should be easy to interpret and apply, and clearly work for the benefit of the population as a whole [3].

When dealing with human subjects in cyber security research we should use three ethical hallmarks to ensure that the subjects are treated fairly. The study should always possess the consent of the studied part, debrief the subjects of the actual impacts of the study, and when it is not possible to gain consent from the actual participants, we must gain consent from so called surrogate participants who can represent the participant [2].

There exists a number of studies that create ethical guidelines for various technical domains, such as a guideline for biofuels [3], a guideline for bioprinting human organs [25] or guidelines for CBRNE subjects (chemical, biological, radiological, nuclear or explosive) as described above [12]. A substantial portion of the found guidelines target the ethics of artificial intelligence. Searching for *"ethical framework" AND ("Artificial intelligence" OR "AI")* gave 53 results, containing articles that discuss different takes on the ethics of implementing artificial intelligence. One noteworthy study also mentions possible ways to compensate for damages caused by artificial intelligence [14]. The study also discusses the ethics of cyber security but the main focus is given to artificial intelligence and the paper does not treat offensive cyber security nor the training of new security

employees. Compensation of damages caused by AI is an interesting topic that can be extended to compensation for damages caused by offensive cyber security. If it becomes evident that a hacker who caused injury has been trained at a certain university or educator, it can be asked whether the educator should take part in the responsibility (legally or ethically) of compensation.

The responsibility of damages can be further distinguished in different categories, such as accountability, liability, and culpability. In order to present a solid ethical guideline one must take note of the precautions from not following the suggested actions [21]. This can be hard since ethical values and even legal obligations vary between cultures and countries. An offensive cyber security course may host students from different origins, with different opinions of what is ethical, or even different experiences of what is legal. This problem is further described in Sect. 6.

Another set of ethical principles for artificial intelligence aims to be implemented and interpreted in machines, rather than used by humans. In order to make the framework machine readable its structure is adapted to machine answerable questions, often resulting in two-folded decision trees [24]. This offers some insight to our principles on offensive cyber security and could give rise to interesting future use cases in the area. When making frameworks for human interpretation, they often focus on offering assistance in difficult questions, rather than stating definite answers. The target audience could be any decision maker or leader, and the intended outcome is often to assist and incentivize them to choose ethical ways to deal with AI [6]. Although both of the above frameworks deal with ethics towards artificial intelligence, some of the concepts can be used as inspiration when designing ethical principles for offensive security.

2.2 Ethical Studies on Cyber Security

Ethical guidelines for cyber security can target organizations, individuals, or researchers [13]. An ethical question that has been debated in cyber security research is the collection and usage of research data. Several studies have used data from controversial sources, such as exposed data from embedded devices that has collected user data without consent, and later been exposed. These types of publicly available "found data" presents a foundation for research that is ambiguous and not necessarily ethical nor legal [13,23]. In addition to finding data one may ask the ethical validity of entering other systems (even hostile ones) without consent. It is possible to argue that unlawfully entering or compromising devices of criminal systems, such as an illegal botnet, is similar to enter the house of criminal without a search warrant [13].

When an ethical guideline is created, the subsequent task is to determine how to spread it and to whom it shall be offered. This is not necessarily a straight forward task and it can be determined in different ways depending on the guideline. Some guidelines even target this meta question, analyzing what types of leadership is most feasible for cyber security. Cyber security leadership can be divided in three branches depending on the how cyber security is being used. One branch for groups that implement cyber security and work with it

practically. This could be divisions in organizations (an IT security department) or entire organizations (dedicated security companies). The second branch is people who innovate and research cyber security. This could be cyber security companies (innovating their products) or cyber security researchers. The third branch is people who regulate cyber security, this could be governmental agencies or ethical review boards [13].

Some parties question the capacity of institutions, governments or even ethical review boards to lead the progress of cyber security ethics, and rather propose more distributed and decentralized peer groups [13]. This is interesting to consider when creating an ethical foundation for offensive cyber security education. Especially regarding questions on how to spread and implement the guideline. If decentralized ethical regulators are more trusted than a unified regulating body, the guideline must be adapted for this cause, being flexible and adaptable to different nuances.

It is often said that the most susceptible part of an organization is its employees. The list of social cyber attacks is long and there are several ways to exploit the people of an organization in order to perform a cyber attack [1]. Therefore, it is not surprising that organizations are spending resources to test and train their employees on cyber security. This training is often valuable, but it can be ethically questionable, especially if the employees suffer consequences from failing a test. A test could be sending intentional phishing emails to ensure that employees do not press suspicious links or hiring physical testers who aspire to tailgate (follow someone who just entered a locked door) an employee into the office. In order to ensure these tests are ethically justified, we should keep the tests legal, always consider the consequences of our actions towards the employee, business owners, and clients, avoid actions that could cause productivity loss, and keep the tests confidential and honest [1].

2.3 Ethical Studies on Offensive Cyber Security

Searching for *"'ethical framework" AND "cyber security"'* gave scarce results from all major databases. Scopus only resulted in six matching articles, of which [14] and [21] were more focused on the ethics of artificial intelligence, as mentioned above. However, some studies exist from which we can learn something useful.

When assessing the ethics of offensive cyber security it is good to start by defining which actions are acceptable and which are to be denounced, as well as defining how to categorize different hostile actions [9]. Some say a hostile cyber activity should be considered as an act of crime [10], others see it more as an act of warfare, which could justify a military response [20].

A problem with the legal approach is that countries tend to have their own laws, which makes global reinforcement of laws problematic. Although international laws and treaties have been conducted on several occasions, aligning such works from global leaders on cyber security may be challenging [10]. This could be an argument for making the ethical guideline on cyber security adaptable, allowing local variations in its implementation.

Of all the reviewed literature, one paper was found targeting how to ethically conduct offensive cyber security education [22]. The paper struggles to find relevant information from research databases, so the authors used three senior information security professionals as experts on the matter, and took their input on how to conduct ethically responsible offensive cyber security education. The interviews focused on analyzing how to make students more inclined to apply the ethics they already knew (and the ethics they learned) to hacking tasks. The questions focused solely on college students and did not try to analyze other education sources such as employers or online courses. Four areas were identified as important for making offensive cyber security education more ethical: social interaction/support system, competition, recognition, and ongoing skills development. In addition, three smaller areas were chosen to be of further importance for keeping students ethical: interaction with cyber security-related law enforcement, cyber security internships and student attendance at meetings and conferences of professional cyber security organizations [22].

The questions posed by the related studies are challenging and relevant. They offer some inspiration for how an ethical framework on offensive cyber security education can be created, but they also present a research gap for an up to date paper, specifically written on the ethics of offensive cyber security education.

2.4 Hacker Motivations

Peter Grabonsky's book *Computer Crime: A Criminological Overview* states that computer criminals are driven by a large number of different motivators depending on each individual case. The main motivators are believed to be the same as for regular criminals (and individuals), namely: greed, lust, power, revenge, adventure, and a desire to try forbidden technologies [7].

Schwartz's Theory of Motivational points out that human behavior can be motivated by three human requirements: *"biological needs (for organic survival), social interaction (for interpersonal coordination), and social institutional demands (for group survival)"* [17]. The human requirements can be translated into values such as striving for world peace in order to ensure survival of the species. The value thus refers to ideals about how we are supposed to act or how the world is supposed to be, these beliefs are what determines our actions. Ten main values of motivation are further determined: *Universalism, Benevolence, Conformity, Tradition, security, power, achievement, hedonism, stimulation and self direction.* These motivational factors were studied in combination with hacking activities in order to determine the most prominent motivators for hackers. The result suggested that the correlation between motivational factors and hacking activities is not always straightforward. The most relevant motivators for hackers were determined as intellectual challenge and curiosity [17].

3 Methodology

In this section we describe how the ethical principles were created. It was primarily done by gathering information from subject matter experts and performing a literature review on related ethical guidelines.

3.1 Literature Review

In order to create the ethical framework a systematic literature review was conducted to analyze ethical guidelines in related domains, as further explained in Sect. 2. The combined search results received a total of 2074 articles. However, a large number of these came from the Google Scholar full text searches (1059 papers) or the more general searches on "ethical framework" AND "technology" (474 papers) or ethics AND "cyber security" (280 papers). For these more general searches the search results were sorted for relevancy and the first pages were manually scanned to include relevant papers. The inclusion and exclusion criteria were thus intentionally flexible and based on the article's connection to ethics, ethical principles, and offensive cyber security education. An article with a clear and well planned ethical principle was considered relevant even though it came from another field than offensive cyber security, such as the ethical principles posed by the UN peace keeping unit discussed in Sect. 3.4. More information of the searches are found in Table 1, Table 2 and Sect. 2. The more specific search phrases generated a smaller set of matching articles so these were scanned more thoroughly, as further described in Sect. 2.

3.2 Ethical Guidelines from Computer Science Societies

There exist a number of ethical guidelines proposed by computer science societies, such as those from the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE). These were analyzed in order to gain a better understanding of ethical principles in related areas and to gain inspiration for the principles on offensive cyber security education.

The ACM Code of Ethics and Professional Conduct targets general computer science practitioners and does not have specific information about offensive cyber security nor cyber security education. As described on the web page of the ethical code *"the code is designed to inspire and guide the ethical conduct of all computing professionals, including current and aspiring practitioners, instructors, students, influencers, and anyone who uses computing technology in an impactful way"*[1].

Some of the principles in the code is highly relevant for our principles on offensive cyber security education. Principle 1.1 (*Contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.*), 1.2 (*Avoid harm.*), 1.3 (*Be honest and trustworthy.*) and 1.6 (*Respect privacy.*)

[1] https://www.acm.org/code-of-ethics.

concerns topics that should be treated during offensive cyber security education. Point 1.1 and 1.2 can be especially linked to principles 5 of our proposed ethical framework.

The IEEE Computer Society Code of Ethics consists of eight ethical principles for software engineer professional. Similarly to the above principles from ACM, these principles have a high level of abstractions and target general computer science rather than cyber security or offensive cyber security education. Three principles were found especially relevant for our study and used as inspiration for the created ethical principles: principle 1 (*Software engineers shall act consistently with the public interest.*), principle 6 (*Software engineers shall advance the integrity and reputation of the profession consistent with the public interest.*), and principle 4 (*Software engineers shall maintain integrity and independence in their professional judgment.*)[2].

3.3 Ethical Principles from Standardization Bodies - NIST

The authors could not locate a clear set of ethical principles or a code of conduct from the National Institute of Standards and Technology (NIST). Some documents containing information to employees were found and contained minor parts with ethical instructions, but nothing that was concrete for computer science or cyber security. A document regarding the ethical conduct for employees of the executive branch states the importance of acting in a way that serves society, which is in line with main takeaways from other frameworks [19].

3.4 Ethical Principles Within the Military, Using Violence to Keep Peace

The United Nations (UN) has an ethical code of conduct for its military peace keeping forces. There exists an interesting connections between this and offensive cyber security since both fields use (some form of) violence in order to reduce violence. There is a further similarity since UN forces operate in different nations and the peace keeper is not necessarily familiar with the laws and customs of the country they assist. Of the UN rules of conduct for UN peace keepers, rule 2 (*Respect the law of the land of the host country, their local culture, traditions, customs and practices*) and 5 (*Respect and regard the human rights of all*) are especially related to our work on offensive cyber security education. It is not uncommon for operations within offensive cyber security to cross many digital borders (such as a server or a cloud service being located in another country than that of the target), it is not always clear whether to follow the laws of the country where the target company is located or the laws of other countries encountered on the way, it is especially difficult if different laws contradict one another [16].

[2] https://www.computer.org/education/code-of-ethics.

3.5 Ethical Principles Within Professional Offensive Security Courses

A number of professional offensive security courses exists such as the Certified Ethical Hacking course[3] and the Certified Penetration Testing course[4] offered by the EC Council, the Global Information Assurance Certification offered by the SANS institute[5], and the Offensive Security Certified Professional course[6]. None of these seemed to contain substantial information about ethical principles so they were not investigated further. However, since only parts of the content can be viewed without purchasing the courses, it is possible that the complete material contains ethical principles or similar instructions for how to use the hacking material.

3.6 Subject Matter Experts

In order to solidify the ethical principles, expert opinions were gathered from security professionals and educators within cyber security. These sessions were partly constructed as semi-structured interviews, direct interviews or surveys based on seven questions, as displayed below. In total 24 people with various positions in cyber security were questioned in order to design and validate the ethical guidelines. The prearranged questions were created in order to give the interviews the right direction, while still being open enough to let the participants come with their own ideas and suggestions to the framework. The result of the consultations is explained in Sect. 4 and the complete list of questions are shown below.

Interview questions

1. Q1: Do you think there is a need to include offensive security training in the curriculum of cyber security education?
2. Q2: Do you think there is a need to incorporate ethics in offensive security education?
3. Q3: What do you think is the best way to incorporate ethics in offensive security education?
4. Q4: What do you think is the most influencing factor in making students use their knowledge for good, such as regular penetration testing work (several points can be selected if desired)?
   (a) Moral values.
   (b) Reputation.
   (c) Ideological gains.
   (d) Fear of repercussions from doing something illegal.

[3] https://www.eccouncil.org/programs/.
[4] https://www.eccouncil.org/programs/certified-penetration-testing-professional-cpent/.
[5] https://www.giac.org/.
[6] https://www.offensive-security.com/pwk-oscp/.

   (e) Their past education.
   (f) Their current peers or social group.
   (g) Other (feel free to specify any motivator not listed above).
5. Q5: What do you think is the most influencing factor in making students use their knowledge for illegal purposes (several points can be selected if desired)?
   (a) Curiosity.
   (b) Financial gains.
   (c) Ideological gains.
   (d) Reputation.
   (e) Their past education.
   (f) Their current peers or social group.
   (g) Other (feel free to specify any motivator not listed above).
6. Q6: What do you think is the best way to ensure that students maintain their ethical mentality and use their knowledge for good, years after they have finished their training?
7. Q7: If you have any other feedback or points you wish to share, please feel free to do so here.

4 Result

In this section the proposed ethical principles are explained. Each principle of the guideline is described in detail, and the findings of the expert consultations are visualized.

4.1 Result from the Expert Consultations

The answers to the questions were categorized in different segments, in order to better classify and arrange them. The sections for each question, and their corresponding answer percentages can be seen in Table 3. For some questions more than one answer could be selected, the displayed percentage is then the fraction of votes from all consulted parts, on this question.

In general we can see a strong trend towards the belief that a peer group with solid ethical values is important to make the students stay within legal bounds. Furthermore it was common to believe that a stable job market for white hat penetration testing reduced the risk of students turning to illegal actions. Developing sound ethical values were also considered important, as well as giving the students a chance to earn positive reputation by legal means. It was deemed important to have strict application criteria, both to ensure the students remain ethical during the training and during their own time after they had graduated. This makes sense since the course can not replace all the students past experiences and if they are too far down an unwanted path, such as having a strong criminal record, it may not be feasible to change their ethical values during the time of the course. Some interesting suggestions appeared in the open discussions such as the idea to keep track of students after they have graduated and hold alumni gatherings where they can meet and share their past experiences (hopefully earning positive reputation from their legal penetration testing work).

Table 3. Result from interviews with Subject Matter Experts.

Question	Answer segments
Q1: Shall we include offensive security training in the cyber security courses?	Yes: 88% No: 4% Maybe: 8%
Q2: Shall we include ethics in offensive security training?	Yes: 96% No: 0% Other/don't know: 4%
Q3: What is the best way to incorporate ethics in offensive security education?	Working with practical cases: 25% Discussions and reflections: 25% Integrate ethics in ordinary coursework: 38% Informative material (legal, ethical etc.): 8% Other/don't know: 4%
Q4: What makes students use their knowledge for good causes?	Moral values: 30% Reputation: 25% Ideological gains: 8% Fear of repercussions: 25% Their past education: 21% Their current peers or social group: 46% Other: 21%
Q5: What makes students use their knowledge for malicious causes?	Curiosity: 42% Financial gains: 42% Ideological gains: 8% Reputation: 13% Their past education: 4% Their current peers or social group: 38% Challenge/amusement: 13%
Q6: What is the best way to ensure that students maintain their ethical mentality	Teach students to always consider and know the consequences of their actions: 25% Help the students develop a solid reasoning (ethical and otherwise) about the actions they do: 29% Have a good white hat job market, help the students develop a professional network and help them prepare to get a legal pentesting job: 17% Other/don't know: 21%
Q7: Other points	Incorporate bug bounty programs in the education Show how ethics can used to grow professionally

5 Ethical Principles for Offensive Cyber Security Education

The final five principles are designed to assist in creating more ethically responsible offensive cyber security training. The principles are motivated by the information gathered during the literature review of related ethical principles and from analyzing other ethical codes, such as *The ACM Code of Ethics and Professional Conduct* and the military principles for ethical behavior proposed by the UN peace keeping unit, as described in Sect. 3. The principles were further validated and fine tuned by the input obtained in the consultations with subject matter experts. Together the principles serves to minimize the harm while maximizing the benefits of offensive cyber security education. The principles are:

1. Include ethics in offensive cyber security education.
2. Inform the students of lucrative and legal ways of applying their hacking knowledge.
3. Introduce ways for students to earn positive reputation from their hacking skills.
4. Have selective application criteria for joining an offensive cyber security course.
5. Offensive cyber security should rarely teach hacking tools that can threaten people's essential rights.

In the following subsection the principles are explained in more detail. Each principle is described with the intent of the principle, a potential use case of how the principle can be implemented, and possible criticism against the principle.

5.1 Principle 1 - Include Ethics in Offensive Cyber Security Education

Intent. It seems clear that hacking knowledge can be used for destructive purposes. However, it is equally clear that offensive cyber security is becoming an important tool within cyber security. Thus, we ought to educate students in offensive cyber security but the education must incorporate ethics and teach students how to use their knowledge in an ethically responsible way.

Implementation. Ethical material of how to use ones hacking skills should be included in the study plan of the course. The education of ethics can for example be incorporated in the normal education slides, discussed in dedicated seminars, practiced in case studies or practically displayed by letting students use their knowledge for bug bounty hunting or other legal hacking activities.

Potential Criticism. Educational programs may be more incentivized to focus the course on raw hacking skills. Being forced to incorporate ethics can take time away from this practical hacking. Incorporating the ethical material may present additional costs for the educators and if this principle is not legally enforced, actors may ignore it. If the principle is legally enforced, it will take freedom away from course administrators and it will create bureaucratic costs for controlling the organizations.

5.2 Principle 2 - Inform the Students of Lucrative and Legal Ways of Applying Their Hacking Knowledge

Intent. In order to help the students develop an ethical compass and use their knowledge for good, it is important to demonstrate how the knowledge could be used for good, and give plenty of opportunities for them to do so. This includes

capitalizing on their hacking skills by earning money, but it can also include non-monetary benefits such as ideal causes. One student may be driven by income, another may be driven by an ethical cause such as protecting the privacy of citizens. We want to teach the students how they can achieve what they want within a legal and ethical context.

Implementation. The principle can be implemented in a number of ways, one example is to make bug bounty programs part of the education. The students could even be taught how to capitalize from bug bounty challenges and earn money during the course. This principle also encourages courses to use realistic and stimulating hacking infrastructures, so the students can get practical hands on hacking experience without needing to look for challenges outside the classroom.

The implementation of this principle also includes the legal aspects of offensive cyber security. The course should inform students of the legal boundaries of offensive cyber security and how these can differ across countries and what they should keep in mind to stay legal while they perform penetration tests.

Potential Criticism. Introducing bug bounty programs and monetizing may act as a double edged sword. It is possible that this could motivate the students curiosity to keep searching for ways to earn money from hacking, which could lead them to the black market. This is especially the case if the black market offers larger monetary rewards for similar tasks.

5.3 Principle 3 - Introduce Ways for Students to Earn Positive Reputation from Their Hacking Skills

Intent. Reputation appears to be a major motivator for hackers. Educators can benefit from this by letting the students use their hacking knowledge to earn reputation in legally and ethically responsible ways. This can help the students develop a healthy attitude towards applying their hacking knowledge.

Implementation. An ethical hacking course may introduce scoreboards where the top students get a prize (the honor of being seen in top can be prize enough). It is probably best to only show the score for the top students. Publicly displaying the worst students can create resentment from those with slower learning curves.

Capture the flag (CTF) competitions is another good way to help students earn positive hacking reputation in an ethically responsible way. In these competitions the participants solve computer security problems in order to obtain "flags", the flags are often a sequence of characters. Many of the competitions are public and open for everyone to participate in. CTF competitions can be integrated in the standard curriculum of the course or shown as a place where the students can apply their skills after the course is completed.

Potential Criticism. If some students are far better than their peers (perhaps from earlier experience), they may seek new scoreboards or challenges with more equal competition. It is important that the teachers are attentive to this and present further material, such as online communities. If not, the student may be inspired to seek fame from less ethical challenges, such as tasks found on the black market. Also, in order to implement the scoreboards or ranking systems educators will have to take time and energy away from developing the raw hacking material.

5.4 Principle 4 - Have Selective Application Criteria for Joining an Offensive Cyber Security Course

Intent. Hacking knowledge can be used for malicious causes. If someone has ethical warning flags such as an extensive criminal record, it may not be suitable to accept them to a hacking course.

Implementation. Ethical hacking courses can perform a background check on the applicants, the same way one usually goes through a number of tests in order to purchase a weapon, open a bank account, apply for a job, or apply for many other courses. It is possible that a person has malicious intentions, even if the person does not have a criminal record or other ethical warning flags. Because of this, it may be relevant to do some personality assessment before accepting students to an offensive cyber security course.

Potential Criticism. From an organizational point of view it may be difficult or even impossible to grant the teacher/course responsible mandate to determine who can join the course. If it would be possible to implement, the selection can give rise to discrimination if each educator is allowed to create their own admission rules. If the admissions are instead determined by a global standard, the educators will loose freedom and bureaucratic costs will arise from controlling the admissions and arranging potential assessment tests. Furthermore, equal information can often be found online. A banned student may be fueled with stronger motivation to learn the material on their own. If that happens, anger from not being admitted may place the student at an even larger risk of using the knowledge for destructive causes.

5.5 Principle 5 - Offensive Cyber Security Courses Should Avoid Techniques for Hacking Critical Infrastructures and Other Industries that Are Crucial for Society to Function

Intent. Critical infrastructures are getting more connected and thus more susceptible to cyber attacks. It is essential to maximize the cyber security of these industries and offensive cyber security is a great tool to do so. However, since penetration testing of power plants and similar industries are sensitive and must be done with care, it is best to treat this information in dedicated and highly controlled offensiveness security courses.

Implementation. General hacking courses can avoid teaching material that is especially used for targeting critical industries. Specific hacking courses can be arranged for penetration testing targets within sensitive areas, such as power plants or hospitals.

Potential Criticism. Much of the information acquired from general hacking courses can be used for attacks against critical infrastructures as well.

6 Discussion

We have presented a set of ethical principles for offensive cyber security education. The principles aim to mitigate the risks of educating many new practitioners in offensive cyber security. In order to fully gage the usefulness of the principles, they must be implemented and used in a real life scenario. After its implementation the result must be validated. A method of validation would be to analyze actions of students who has taken an offensive cyber security courses that follow the ethical principles. These students could then be compared with students who took regular offensive cyber security courses (courses that does not actively follow the ethical principles) and the two groups would then be compared. This comparison will not take place for some time, since we first need to implement the principles and let them run for an evaluation period. However, this article serves as the first step, designing principles that are now ready to be implemented.

The results showed that a peer group with strong ethical values is important for the student to develop healthy ethical values of their own and a strong job market further increase the chance that students remain legal. This makes sense since we often tend to be influenced by our peers, and a healthy job market will reduce the risk of students turning to the black market for monetary reasons.

Another discussion is whether to have static or dynamic principles. Currently the principles are not built to be dynamic but nor are they built to be unchangeable. As briefly discussed in previous sections, it may not be feasible to implement global laws and regulations for cyber security, due to cultural and legal differences between countries and regions. This speaks in favor of having a adaptable principles that can be tweaked for different regions. Although the principles do not currently have built in support for being adjusted, they should be relatively straight forward to change in order to meet local requirements. The obvious benefit of having customizable principles is an easier integration in different organizations and cultures, since it can be adapted to suit specific requirements. A downside is that alterations may tamper with the underlying ethical values and standard of the principles. Allowing alterations would also make it more difficult to oversee the principles, if it became a legal demand to use them. Local adaptations could be made to make the principles easier to implement and use while reducing its ethical impact. This could potentially undermine the value of the principles and each investigator would need to make a manual analysis, to judge if the principles meets the standard. This problem

could be mitigated by having cornerstones that could not be tampered with, but dynamic principles would still create more complicated enforcements and supervisions.

Cultural differences in ethics can also cause problems when people from different backgrounds participate in the same course. It is possible that an offensive cyber security course hosts students from a range of different cultures, these cultures may have different ethical values as well as different laws. This makes it important for the teacher to be attentive to each student in order to make sure they understand and agree with the ethical principles described. The educator should also make sure that students with different ethical views get the chance to describe and communicate their individual views.

With this said, it is not easy to ensure that students of offensive cyber security courses will use their knowledge for ethical and legal purposes. Including ethical principles in the education is no guarantee that students will behave accordingly, but it may increase the chances. The result of the interviews and literature reviews gives reason to believe that ethical principles can have a positive impact on the behavior of the students, but it is hard if not impossible to eliminate the risk of future criminal acts altogether. In light of this we must judge whether the potential benefit of educating people in the area outweighs the potential risks. Since cyber security expertise is in high demand the benefits appears to be strong enough for the trend to continue in the near future.

6.1 Future Work

The study has given rise to a number of ideas for future research in the area. One such area is adapting the principles to target offensive cyber security education for critical infrastructures. Certain requirements should be fulfilled when auditing systems that can cause severe damage to society.

Another research area of interest is the motivations of illegal hacking. There seems to be a research gap for analyzing what motivates hackers to use their knowledge for illegal causes. This could be expanded to target ethical hackers as well, analyzing if they are incentivized by certain causes such as money/career, ideological gains, reputation etc. Then a comparison could be made between the motivations for legal and illegal hacking.

A third area of future interest is how to compensate for damages caused by offensive cyber security. This includes investigating if the educator is responsibility for damages caused by the hacker and if so, how to legally reprimand the educators if the hacker commits a crime.

It would also be relevant to research adaptation of the principles. We need to assess how to best implement them and how to best guide others on how to implement them. This includes further research on whether the principles should be customizable or static.

Perhaps the most pressing future task is to implement the principles in actual offensive cyber security courses. This will provide the valuable feedback to fine tune and adjust the principles according to their practical usability.

7 Conclusion

Offensive cyber security is becoming an important part of making systems more resilient. Training people in offensive cyber security introduces a problem since it enables people to break systems, which could be used for illegitimate causes. In order to mitigate the risks associated with offensive cyber security training we propose a set of ethical principles describing how to conduct ethically responsible offensive cyber security training. The principles work to minimize the potential harm from offensive cyber security education, while maximizing its potential benefits. This article presents the background for the principles and describes each principle in detail, validated by a systematic literature review and consultations with subject matter experts. Subsequent research will implement the principles and fine tune them for real life applications.

Acknowledgments. This work has received funding from the Swedish Centre for Smart Grids and Energy Storage (SweGRIDS).

References

1. Archibald, J.M., Renaud, K.: Refining the pointer "human firewall" pentesting framework. Inf. Comput. Secur. **26**(4) (2019)
2. Bravo-Lillo, C., Egelman, S., Herley, C., Schechter, S., Tsai, J.Y.: You needn't build that: reusable ethics-compliance infrastructure for human subjects research. In: Cyber-Security Research Ethics Dialog and Strategy Workshop (2013)
3. Buyx, A., Tait, J.: Ethical framework for biofuels. Science **332**(6029), 540 (2011)
4. Durnell, E., Okabe-Miyamoto, K., Howell, R.T., Zizi, M.: Online privacy breaches, offline consequences: construction and validation of the concerns with the protection of informational privacy scale. Int. J. Hum.-Comput. Interaction **36**, 1834 (2020)
5. Eichhorst, W., Hinte, H., Rinne, U., Tobsch, V.: How big is the gig? assessing the preliminary evidence on the effects of digitalization on the labor market. Manag. Revue **28**(3), 298 (2017)
6. Floridi, L., et al.: AI4people-an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach. **28**(4), 689 (2018)
7. Grabosky, P.: Computer crime: a criminological overview. In: Tenth United Nations Congress on the Prevention of Crime and the Treatment of Offenders (2000)
8. Greenberg, A.: Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers, Doubleday (2019)
9. Holzer, C.T., Lerums, J.E.: The ethics of hacking back. In: 2016 IEEE Symposium on Technologies for Homeland Security, HST 2016 (2016)
10. Iasiello, E.: Hacking back: not the right solution. Parameters **44**, 105 (2014)
11. Jeske, T., Weber, M.-A., Würfels, M., Lennings, F., Stowasser, S.: Opportunities of digitalization for productivity management. In: Nunes, I.L. (ed.) AHFE 2018. AISC, vol. 781, pp. 321–331. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-94334-3_32

12. Jillson, I.A.: Ethical frameworks for CBRNE crises: toward shared concepts and their practical application. In: O'Mathúna, D.P., de Miguel Beriain, I. (eds.) Ethics and Law for Chemical, Biological, Radiological, Nuclear & Explosive Crises. TILELT, vol. 20, pp. 53–64. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11977-5_5
13. Kenneally, E., Bailey, M.: Cyber-security research ethics dialogue and strategy workshop. Comput. Commun. Rev. **44**, 76 (2014)
14. Khisamova, Z.I., Begishev, I.R., Sidorenko, E.L.: Artificial intelligence and problems of ensuring cyber security. Int. J. Cyber Criminol. **13**(2), 564 (2019)
15. Kolias, C., Kambourakis, G., Stavrou, A., Voas, J.: DDoS in the IoT: Mirai and other botnets. Computer **50**(7), 80 (2017)
16. Langholtz, H.J.: Ethics in peace operations. In: Peace Operations Training Institute (2008)
17. Madarie, R.: Hackers' motivations: testing Schwartz's theory of motivational types of values in a sample of hackers. Int. J. Cyber Criminol. **11** (2017)
18. Magnusson, L., Hanson, E.J.: Ethical issues arising from a research, technology and development project to support frail older people and their family carers at home. Health Soc. Care Community **11**(5), 431 (2003)
19. U.S. Office of Government Ethics. Standards of ethical conduct for employees of the executive branch. J. Int. Technol. Inf. Manag. (2001)
20. O'Connell, M.E.: Cyber security without cyber war. J. Conflict Secur. Law **17**(2), 187 (2012)
21. O'Sullivan, S., et al.: Legal, regulatory and ethical frameworks or standards for AI and autonomous robotic surgery. Int. J. Med. Robot. Comput. Assisted Surg. (2018)
22. Pike, R.E.: The "ethics" of teaching ethical hacking. J. Int. Technol. Inf. Manag. **22**(4), 4 (2013)
23. Schrittwieser, S., Mulazzani, M., Weippl, E.: Ethics in security research which lines should not be crossed? In: 2013 IEEE Security and Privacy Workshops (2013)
24. Tonkens, R.: A challenge for machine ethics. Minds Mach. **19**(3), 421 (2009)
25. Vermeulen, N., Haddow, G., Seymour, T., Faulkner-Jones, A., Shu, W.: 3D bioprint me: a socio ethical view of bioprinting human organs and tissues. J. Med. Ethics **43**(9), 618 (2017)
26. Zhou, J., Cao, Z., Dong, X., Vasilakos, A.V.: Security and privacy for cloud-based IoT: challenges. IEEE Commun. Mag. **55**(1), 26 (2017)