Teaching Adversarial Thinking for Cybersecurity

Seth T. Hamman †, ‡
sethhamman@cedarville.edu

Kenneth M. Hopkinson †
kenneth.hopkinson@afit.edu

† Air Force Institute of Technology
Wright-Patterson AFB, OH 45433

‡ Cedarville University
Cedarville, OH 45314

Abstract - The academic discipline of cybersecurity is still in its formative years. One area in need of improvement is teaching cybersecurity students adversarial thinking—an important academic objective that is typically defined as “the ability to think like a hacker.” Working from this simplistic definition makes framing student learning outcomes difficult, and without proper learning outcomes, it is not possible to create appropriate instructional materials. A better understanding of the concept of adversarial thinking is needed in order to improve this aspect of cybersecurity education. This paper sheds new light on adversarial thinking by exploring it through the lens of Sternberg’s triarchic theory of intelligence. The triarchic theory’s division of the intellect into the analytical, creative, and practical components provides a comprehensive framework for examining the characteristic thought processes of hackers. This exploration produces a novel, multidimensional definition of adversarial thinking that leads immediately to three clearly defined learning outcomes and to some new ideas for teaching adversarial thinking to cybersecurity students.

Categories and Subject Descriptors
K.3.2 [Computers and Education]: Computer and Information Science Education
General Terms

Computer science education, Curriculum

Keywords

Adversarial Thinking Definition, Cybersecurity Education, Triarchic Theory of Intelligence

1. INTRODUCTION

It is widely acknowledged that teaching adversarial thinking to cybersecurity students is important. In a recent editorial highlighting the state of cybersecurity education in colleges and universities, Fred Schneider writes, “Can adversarial thinking for cybersecurity even be taught, or is it an innate skill that only some can develop? The answer, which is neither known nor aggressively being sought by those who study cybersecurity education, seems central to the development [emphasis added] of an effective cybersecurity course” [1, p. 4].

A team of subject matter experts convened by the Association of Computing Machinery (ACM) to identify cybersecurity curricular guidelines agrees that teaching adversarial thinking is vital. Their summary report states, “To protect systems…we need to temporarily adopt the thinking process of the malevolent hacker…Developing this way of thinking must be part of [emphasis added]…educating cybersecurity professionals” [2, p. 16].

While there is a consensus that adversarial thinking should be taught in higher education settings, current cybersecurity curricular guidelines, both from academia and industry, seemingly omit this aspect of cybersecurity education. The recent “CS Curricula 2013” [3], which made headlines for its new emphasis area on cybersecurity, does not explicitly mention the term adversarial thinking, nor does the National Security Agency (NSA) in their National Centers of Academic Excellence (CAE) curricular guidelines [4]. What explains this disconnect between the acknowledged importance of teaching adversarial thinking and the apparent lack of curriculum support for doing so? The hypothesis of this paper is that the root of the
problem is an imprecise understanding of the concept of adversarial thinking. Constructing curriculum guidelines is predicated on having identified measurable student learning outcomes, and it is difficult to identify learning outcomes for adversarial thinking when working from the common and simplistic definition that adversarial thinking means “the ability to think like a hacker.”

This paper sets out to rigorously define the concept of adversarial thinking by viewing it through a lens provided by the discipline of cognitive psychology, an appropriate place to turn to for guidance for exploring the minds of hackers. Specifically, this paper homes in on Sternberg’s triarchic theory of intelligence as an anchor for understanding the characteristic thought processes of skilled hackers. Then, with new insights gained from this exploration, a novel, multidimensional definition of adversarial thinking is proposed that leads immediately to three clearly defined learning outcomes and to some new ideas for teaching adversarial thinking to cybersecurity students.

2. BACKGROUND

2.1 Hacker Definition

Given the starting point that adversarial thinking means “thinking like a hacker,” the first question that should be addressed in any attempt to define adversarial thinking is, “What kind of a hacker?” For example, the following hacker activities differ substantially: email spear phishing, writing worms and viruses, circumventing digital rights management (DRM) protection, coding buffer overflow attacks, and password cracking. Additionally, there are various different broad categories of hackers, ranging from script kiddies to highly trained professionals and from insider threats to hacktivists. For the purposes of this paper, because the emphasis is on exploring adversarial thinking in the context of cybersecurity practice, all references to hackers refer to individuals whom cybersecurity personnel are hired to prevent from breaching their networks and computer systems.
2.2 Definitions of Adversarial Thinking

To date, no commonly accepted definition of adversarial thinking exists. When
the term is used in the literature, in many cases it is not defined at all, taking it for
granted that adversarial thinking merely means “thinking like your cyber adversary
(i.e., a hacker).” However, this raises the obvious question, “What is unique about
the way hackers think?”

Two notable attempts to unpack the concept further have been made in recent
editorials promoting teaching adversarial thinking in cybersecurity education.
Melissa Dark, Education Editor for *IEEE Security & Privacy*, proposes the following
definition of adversarial thinking: “Let’s say that adversarial thinking is the ability to
look at system rules and think about how to exploit and subvert them as well as to
identify ways to alter the material, cyber, social, and physical operational space” [5,
p. 78]. Another definition comes from Schneider, who writes that adversarial
thinking is “the very essence of game theory. In it, actions by each player are
completely specified; for cybersecurity and safety-critical systems, identifying
possible player actions is part of the central challenge” [1, p. 4].

On the surface these two definitions are very different, but what they have in
common is the identification of some of the salient objects of a hacker’s attention.
For Dark, these are “system rules” and “operational spaces,” and for Schneider it is
“player actions.” Hackers undoubtedly bring a unique perspective to system rules,
they strive to alter operational spaces to their advantage, and they carefully consider
possible player actions. Combining these two definitions in a concise way might
lead to the following definition: *adversarial thinking is the ability to approach system rules,*
*operational spaces, and player actions from a hacker’s perspective.*

This is certainly more helpful than the simplistic “thinking like a hacker”
definition. However, this paper takes the exploration a step further in that it orients
the term not around the objects of a hacker’s focus, but around the primary
structures of his intellect. In other words, the goal of this paper is to provide a more
fundamental definition of adversarial thinking that transcends overly specific details.
2.3 Cognitive Psychology

According to the American Psychological Association, cognitive psychology is the study of “higher mental processes such as attention, language use, memory, perception, problem solving, and thinking” [6]. Because of its focus on the human mind, and in particular on the structures of thought, cognitive psychology is a natural place to turn to for guidance for exploring the minds of hackers.

Well-known psychology professor Robert Sternberg proposed a cognitive model called the *triarchic theory* that breaks the intellect down into three component parts: the analytical, the creative, and the practical [7]. While there are many competing cognitive models, Sternberg’s is appreciated for its simplicity and explanatory power. Long before Sternberg, Aristotle developed a roughly parallel three-pronged model of the intellect, which may have provided some of the inspiration for Sternberg [8].

Sternberg’s analytical area captures the popular conception of intelligence, and coincides with the notion of IQ. It includes mathematical ability and logical reasoning. The creative area of the intellect includes the ability to make unique connections and to see the world in original ways. Artists, authors, and musicians excel in this aspect of the intellect. And lastly, practical intelligence includes the ability to plan, strategize, and accomplish goals. CEOs and military leaders typically have high degrees of practical intelligence (see Table 1).

The three areas of the triarchic theory are meant to capture different modes of intelligence that all people possess to a greater or lesser extent. The three areas are not necessarily correlated with one another—a person might be above or below average in any given area independent of the other areas. The model is useful to help explain why some people succeed in some arenas and fail in others. Sternberg notes that many students with high analytical intelligence do very well in the highly structured world of undergraduate education, but they struggle as graduate students because they lack creative and practical intelligence, both of which are paramount for conducting and completing original research [7].
### Table 1
Summary of Sternberg’s Triarchic Theory of Intelligence

<table>
<thead>
<tr>
<th>Area</th>
<th>Description</th>
<th>Popular Conception</th>
<th>Exemplar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical</td>
<td>Mathematical ability and logical reasoning</td>
<td>Book smarts</td>
<td>Einstein</td>
</tr>
<tr>
<td>Creative</td>
<td>The ability to make unique connections and see the world in original ways</td>
<td>Creativity</td>
<td>Van Gogh</td>
</tr>
<tr>
<td>Practical</td>
<td>The ability to plan, strategize, and accomplish goals</td>
<td>Street smarts</td>
<td>Napoleon</td>
</tr>
</tbody>
</table>

3. TRIARCHIC THEORY APPLIED TO HACKERS

Applying Sternberg’s framework to the minds of hackers provides some valuable insights and a more thorough understanding of how they think. This section of the paper views hacker behavior through each of the three lenses provided by the triarchic theory with an emphasis on explaining how each category of the intellect contributes to success in hacking.

3.1 Analytical

In the popular culture hackers are typically portrayed as highly intelligent “computer wizards.” Hackers in television shows and movies sometimes seem like aliens to those around them because of their uncanny technical abilities. Typically, these characters are irresistibly drawn to computing from their youth. While these
portrayals are fictitious, there is support for this popular hacker stereotype; hackers do seem to have an unusual affinity and knack for technology.

Hacking involves detailed knowledge of many highly technical aspects of computing, including computer networking protocols, assembly language programming, and operating systems. In Sternberg’s paradigm, this technical knack exhibited by computer hackers ties into the analytical component of their intellect. In this case, their analytical gifts translate into an unusual facility with computers and technology. Having strong technical abilities is vital to hackers because many kinds of cyber attacks involve overcoming significant technological hurdles. Here are a few examples: to infiltrate a computer network, a hacker may need to construct precisely malformed network packets; to exploit a programming flaw, a hacker may have to tediously code a buffer overflow attack; and to remain undetected on a system, a hacker might need to modify an operating system’s libraries.

One real-life example of a hacker who leveraged his analytical intelligence is Robert Tappan Morris. At the age of 19, Morris published a technical paper on a major vulnerability in a trust protocol used in the BSD Unix operating system [9]. A few years later, freshly graduated from Harvard, he used his advanced understanding of computer networking protocols and operating systems to write a software worm that infiltrated an alarming percentage of the computer systems on the Internet at that time [10]. Another example of a hacker who excelled in this area is Elias Levy (a.k.a. Aleph One) who wrote the seminal paper on buffer overflow attacks [11]. Both of these individuals used their analytical gifts to dissect software and network and security protocols, and this enabled them to identify exploits.

In summary, to think like a hacker in terms of the analytical component of his intellect is to embody his technological capabilities, which includes low-level programming skills and a deep familiarity with operating systems and computer networking protocols.

3.2 Creative
Sternberg cites “lack of conventionality” as one of the markers for creative intelligence [7, p. 245]. This is similar to the way cybersecurity guru Bruce Schneier describes what he calls the “hacker mindset.” Schneier writes that a hacker is a person who “discards conventional wisdom,” and who by “thinking differently,” is able to uncover security vulnerabilities that had not occurred to the system’s designers [12]. This aspect of adversarial thinking may be what Dark is referring to in her definition (quoted above) when she mentions the ability to subvert system rules.

Creativity is at the core of the “hacker mindset.” While fiction writers excel at creating original stories that capture the imagination, hackers excel at creating original exploits that bend technology in unexpected ways. Both are manifestations of the same root – they involve seeing the world in a unique way, and the ability “to put old information together in a new way,” as Sternberg puts it [7, p. 245]. While most technologists are concerned with making systems work, hackers are obsessed with pushing the limits of systems and exploring possibilities that many people would never consider. This aspect of hacking is the main connection between the pejorative way the term hacker is today, and the original, complimentary term from a previous era which connoted being highly skilled in the art of computer programming.

IP fragmentation attacks provide a good illustration of the way hackers apply their creativity to bend technology and protocols. This class of attacks is where IPv4 packets are intentionally fragmented by hackers for purposes ranging from crashing computers to circumventing firewalls [13]. All computer network students learn that routers are programmed to automatically fragment IPv4 packets that are too large to traverse the next hop link, but the creative and unconventional mind of a hacker realizes that packets could also be fragmented by programmers, intentionally, and in unusual ways. This opens up a world of possible attacks, many of which have exposed unsafe security assumptions made by system designers.

In summary, the creative aspect of adversarial thinking involves embodying the unconventional perspectives of hackers which enable them to manipulate technology in unexpected ways.
3.3 Practical

The practical component of Sternberg’s triarchic theory is the aspect of the intellect that involves planning, strategizing, and overcoming obstacles to accomplish goals. While script kiddies are known to indiscriminately fire point-and-click exploits at random in hopes of finding unpatched systems, more highly skilled hackers select targets, conduct reconnaissance, carefully plan their attacks, and meticulously cover their tracks [14]. In general, hackers attempt to use their time and resources wisely, and they strive to outwit security personnel. A researcher who interviewed hackers recorded, “One [hacker] described how he attempted to anticipate the moves of his adversary [security personnel] by stating, ‘how can I predict, how can I anticipate what they’re going to do?’” [15, p. 23]. The researcher concludes that strategizing is an essential aspect of hacking. Schneider, in his definition of adversarial thinking (quoted above), probably has the practical component in mind when he compares adversarial thinking to game theory – the study of strategic reasoning.

A good example of a real-life hacker who excels in the area of practical intelligence is the famous social engineering expert Kevin Mitnick. While Mitnick is undoubtedly very intelligent, his intellectual gifts can be better described as street smarts than book smarts. Mitnick had a knack for thinking on his feet, and he was rarely denied the prizes he sought. During his hacking days, he routinely employed strategic maneuvering to evade detection and capture. For example, during his years on the lam from the FBI, he routinely hacked into his pursuers’ phone lines, voicemails, and email accounts, which enabled him to stay one step ahead of them for years [16]. Interestingly, it was not until the FBI enlisted the help of another hacker, Tsutomu Shimomura, that they finally caught him. Because practical intelligence is associated with success in business, it is no coincidence that Mitnick was able to parlay his hacking infamy into the lucrative career as a cybersecurity consultant that he enjoys today.

In summary, adversarial thinking positioned in the light of the practical component of the intellect is embodying a hacker’s ability to think strategically. It
is captured in the ways hackers plan their attacks, outmaneuver security personnel, and overcome obstacles.

3.4 Summary

Having outlined all three areas of the hacker’s intellect separately, it may be helpful to take a real-world example of a cyber attack and see how each of the three aspects contributed to the hacker’s success. Clifford Stoll published the first detailed account of a computer hacker in the research literature in 1988 [17]. (He later turned the paper into a bestselling book [18].) Although today’s cybercrime is worlds apart from the hacking of the 1980’s in terms of motivation, scale, and organization, the fundamental techniques of hacking have not changed.

Stoll describes how his hacker was deeply familiar with the Unix operating system and computer networks in general (on the level of a professional systems and network administrator), and was adept at cracking passwords, writing scripts, and modifying operating system utilities to act as Trojan horses. These strengths can be attributed to the analytical component of the hacker’s intellect. Stoll also describes how the hacker was able to escalate his privileges on systems from a regular user to root level with Gnu-Emacs, a popular text editor with a built-in mail feature which enabled users to communicate with one another by moving files into each other’s home directories. The hacker had the key insight that it was also possible to use the mail utility to move files (like a simple shell script programmed to change user permissions when executed by Cron) into the systems directory. This possibility likely never occurred to the Gnu-Emacs developers because there was no legitimate reason to send “mail” to the systems folder. This insight can be attributed to the creative component of the hacker’s intellect. And lastly, the paper describes how the hacker installed backdoors so that he could gain access to systems even after they had been patched, how he modified logs and audit trails to avoid detection, and how he employed many shrewd tactics for identifying new login credentials, including searching in emails and files, installing Trojan horses to capture login attempts, and password cracking and guessing. These strategies can be attributed to the practical component of the hacker’s intellect.
This short example illustrates that in the case of a skilled hacker, all aspects of his intellect may contribute to his success. While not all areas are strictly necessary, a hacker without analytical intelligence (i.e., technical expertise) is a nonstarter, one lacking creative intelligence never discovers novel vulnerabilities and is fully dependent on recycled, and likely widely known, hacks, and one without practical intelligence has little chance of successfully evading detection or of overcoming unexpected hurdles.

3.5 Adversarial Thinking Definition

A concise summary of the above exploration leads to the following multidimensional definition of adversarial thinking: *adversarial thinking is the ability to embody the technological capabilities, the unconventional perspectives, and the strategic reasoning of hackers* (see Table 2). The word *embody* used in the definition is intended to capture the sense in which actors embody the characters they play. It connotes “becoming one” with hackers and seeing the world through their eyes. To the extent that cybersecurity students can acquire this ability, in their future careers they will be able to identify the digital fingerprints of hackers in their systems and compete with them on a level playing field (the analytical component), identify and fix security vulnerabilities before hackers have the opportunity to exploit them (the creative component), and anticipate future attacks, thwart attacks in progress, and help track down hackers (the practical component).
Table 2
The Triarchic Theory Applied to Adversarial Thinking for Cybersecurity

<table>
<thead>
<tr>
<th>Area</th>
<th>Adversarial Thinking Application</th>
<th>Example Attack</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical</td>
<td>Understanding technology at a deep level, including computer networking protocols, programming languages, and operating systems</td>
<td>Buffer Overflow</td>
<td>Technological capabilities</td>
</tr>
<tr>
<td>Creative</td>
<td>Identifying unsafe security assumptions through manipulating and stretching technology in unexpected ways</td>
<td>IP Fragmentation</td>
<td>Unconventional perspectives</td>
</tr>
<tr>
<td>Practical</td>
<td>Reasoning strategically to plan and execute attacks, evade detection, and overcome obstacles</td>
<td>Social Engineering</td>
<td>Strategic reasoning</td>
</tr>
</tbody>
</table>

4. ADVERSARIAL THINKING FOR CYBERSECURITY EDUCATION

As explained in the introduction of this paper, the reason for developing a more precise definition of adversarial thinking is to help identify appropriate learning outcomes around which curricula can be built. This section of the paper briefly examines current educational practices in terms of each of the three dimensions outlined in the definition. For each area, three aspects in particular are addressed:
1) Awareness – how aware is the educational community of the importance of this area?

2) Progress – how well is the educational community currently addressing this area?

3) Potential – how much potential is there for developing students’ skills and abilities in this area?

4.1 Technological Capabilities

Although it is typically not associated with adversarial thinking, in order to think like a hacker, cybersecurity students must understand his technological capabilities. This cybersecurity learning objective has been understood for a long time, and teaching students the details of technology and the tricks of the hacking trade is the primary emphasis of cybersecurity education today. For example, the NSA’s CAE in Cyber Operations curriculum stresses low level programming, software reverse engineering, operating systems theory, computer networking, and many other highly technical topics [4].

Not only is this area of cybersecurity well established, it is also particularly effective at accomplishing its ends due to the fact that most computer science students (i.e., the typical cybersecurity student) enjoy a knack for technology that is on par with hackers.

4.2 Unconventional Perspectives

Because it is widely recognized as being important, helping cybersecurity students develop the unconventional perspectives of hackers is the subject of much active research. One recent innovative approach to achieving this involves encouraging students to cheat on an otherwise impossible-to-pass exam. The authors explain, “For it is only by learning the thought processes of our adversaries that we can hope to unleash the creative thinking [emphasis added] needed to build the best secure systems” [19, p. 51]. Another cybersecurity educator attempts to teach students this type of creative thinking by assigning hacking labs. He writes, “We find students truly learn when challenged with defeating a computer protocol” [20,
p. 2]. Others have written about how Capture the Flag (CTF) exercises also may contribute to developing this type of creativity [21].

Unlike the technological capability area above, computer science students do not necessarily have strong innate creative abilities. On the contrary, most technically minded people are predominately “left-brained,” meaning that they resonate with logic, rigidity, and rules to the detriment of “outside-the-box” thinking. Therefore, teaching this aspect of adversarial thinking may prove to be an uphill battle. It is not yet known how effective approaches like the ones mentioned above are at developing (as opposed to merely revealing) cybersecurity students’ abilities in this area.

4.3 Strategic Reasoning

Unlike the previous two areas, there is very little awareness of the need to teach strategic reasoning to cybersecurity students. One hypothesis for this blind spot is that because cybersecurity education was born out of a technical discipline (i.e., computer science), it has tended to stay revolved around technology to the neglect of the human element inherent in cybersecurity. But the fact of the matter is, without cyber adversaries, cybersecurity would not exist. At least one educational researcher has noted this weakness in cybersecurity education. He writes, “These topics [the technical aspects of the curriculum] must be augmented with large doses of ethics, legal studies, behavioral science, and military strategic studies” [22, p. 2].

As for potential, this area of adversarial thinking is particularly promising because it is believed that, in general, a person’s ability to engage in strategic reasoning can be improved. Colin Camerer, author of the seminal text on behavioral game theory, writes, “Strategic thinking seems to be more like learning to windsurf, ski, or fly an airplane, activities that require people to learn skills which are unnatural but teachable, and less like weight-lifting or dunking a basketball, where performance is constrained by physical limits” [23, p. 244].
5. RECOMMENDATIONS

There are at least three helpful observations that emerge from this brief analysis (see Table 3). First, any attempt to teach adversarial thinking to students with little technical aptitude could prove futile, because in order to understand how hackers think, a student must have some baseline level of innate technical ability. This argues for cybersecurity to continue being taught as a sub-discipline of computer science.

Second, associating what Schneier calls the “hacker mindset” with the creative component of the intellect could lead to novel approaches for teaching the “unconventional perspectives” of hackers. For example, it may be possible to adapt practices used to stimulate creativity in other disciplines (e.g., creative writing) to cybersecurity education.

Third, the strategic dimension of adversarial thinking is not being adequately addressed in the classroom. This observation has already led to progress in cybersecurity education. The authors of this paper conducted an educational experiment where basic game theory concepts were taught to cybersecurity students. The results show that learning game theory had a statistically significant impact on the students’ abilities to anticipate the strategic actions of others. This study demonstrates that with the proper educational support, students can learn how to better compete in the “battle of wits” that sometimes plays out in the practice cybersecurity.
## Table 3
Summary of Adversarial Thinking Instruction in Cybersecurity Education

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Learning Outcome</th>
<th>Awareness</th>
<th>Progress</th>
<th>Potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technological Capabilities</td>
<td>Understand computer networking protocols, low-level programming languages, and operating systems.</td>
<td>[ ][ ][ ]</td>
<td>[ ][ ]</td>
<td>[ ][ ]</td>
</tr>
<tr>
<td>Unconventional Perspectives</td>
<td>Identify unconventional uses of software and protocols that could be exploited as attack vectors by hackers.</td>
<td>[ ][ ][ ]</td>
<td>[ ][ ]</td>
<td>[ ][ ]</td>
</tr>
<tr>
<td>Strategic Reasoning</td>
<td>Anticipate the strategic actions of hackers, including where, when, and how they might attack, and their tactics for evading detection</td>
<td>[ ][ ]</td>
<td>[ ][ ]</td>
<td>[ ][ ]</td>
</tr>
</tbody>
</table>

**Key:**
- High [ ][ ]
- Medium [ ]
- Low [ ]

### 6. CONCLUSION

In conclusion, by defining more precisely and rigorously what it means to “think like a hacker,” this paper has shed new light on how adversarial thinking can be
addressed in the classroom. Perhaps its most beneficial contribution is its observation that, despite the fact that it has been overlooked, strategic reasoning is a vital aspect of adversarial thinking.

Future work could build on this research by potentially expanding the definition to include other aspects of a hacker’s mind, such as his motivations and unique personality traits (see [24]). It would be interesting to study whether these types of insights could prove beneficial to the practice of cybersecurity and how they should be addressed in the cybersecurity classroom.
DISCLAIMER

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the U.S. Government.

REFERENCES


