In: 54th Hawaii International Conference on System Sciences (HICSS 2021), 5-8 January 2021, Maui, Hawaii, pp. 1958-1967
http://hdl.handle.net/10125/70853

Design Thinking for Cyber Deception

Debi Ashenden
University of Adelaide
debi.ashenden@adelaide.edu.au

Robert Black
Cranfield University
Defence Academy of the United Kingdom
r.black@cranfield.ac.uk

Iain D. Reid
University of Portsmouth
iain.reid@port.ac.uk

Simon Henderson
Deception by Design UK
simon@deceptionbydesign.com

Abstract

Cyber deception tools are increasingly sophisticated but rely on a limited set of deception techniques. In current deployments of cyber deception, the network infrastructure between the defender and attacker comprises the defence/attack surface. For cyber deception tools and techniques to evolve further they must address the wider attack surface; from the network through to the physical and cognitive space. One way of achieving this is by fusing deception techniques from the physical and cognitive space with the technology development process. In this paper we trial design thinking as a way of delivering this fused approach. We detail the results from a design thinking workshop conducted using deception experts from different fields. The workshop outputs include a critical analysis of design provocations for cyber deception and a journey map detailing considerations for operationalising cyber deception scenarios that fuse deception techniques from other contexts. We conclude with recommendations for future research.

1. Introduction

Current best practice for discovery of a network intrusion is 100 days, yet organisations that deploy cyber deception can experience a 91% reduction in dwell time in the network [1] by alerting system administrators to unauthorised access. For any military organisation not deploying cyber deception technology means an adversary potentially has a foothold on Defence networks and is able to extract information about military tactics, techniques and procedures for 100 days. This perimeter defence approach to cyber security is not enough. Nowhere else in the warfighting environment would it be acceptable for the enemy to operate inside our boundary without active measures being deployed once inside the perimeter. If we continue this comparison with the physical environment we can see that as the military move to a manoeuvrist approach [2], away from relying on a defeat level of force, success depends less on preparing solely for a zero- sum game through more efficient kinetic warfare and more towards understanding how to achieve a desired effect on the behaviour of the adversary. The same is true with cyber security measures and the use of cyber deception technology. While the current technology is increasingly sophisticated and is advancing as a technological capability, it is currently constrained primarily to issuing alerts and logging. It has so far failed to encompass the full range of deception tools and techniques that are used in other contexts. In this paper we use design thinking to explore ways to place cyber deception in the manoeuvre toolkit of network defenders and to offer solutions that have the potential to shape the will and behaviour of adversaries in our networks. Deception activities that aim to influence adversary behaviour have long been part of military planning and are critical to the success of physical military operations. With the network being seen as a domain of warfare we need to find ways to give cyber defenders the psychological tools to deny attackers the freedom to operate in cyberspace.

Deception has been defined as, 'deliberate measures to induce erroneous sensemaking and subsequent behaviour within a target audience, to achieve and exploit an advantage' [3]. The key elements here are erroneous sensemaking, an intentional act to bring the deceiver an advantage, a focus on the process of induction used by the victim, causing a change in behaviour, and targeting a specific audience. While some cyber deception technology arguably meets this definition, to a large extent the reach of such technology is contained within the network and does not consider the full scope of the attack/defence surface from the defender's cognitive processes to the attacker's cognitive processes.

Cyber deception has the potential, however, to restore the balance of power from the attacker to the defender, overcoming the asymmetric advantage that attackers currently have [4], [5]. Traditionally, outside of the network domain, deception is utilised to achieve a range of different effects beyond offering a physical honeypot to detect / distract an attacker. Such effects incorporate simulation such as mimicking, inventing and decoying, and dissimulation such as masking,

URI: https://hdl.handle.net/10125/70853
978-0-9981331-4-0

repackaging and dazzling [6]. If one looks beyond the technology-centric view and considers the experience of the attacker being targeted, many of the principles seen in other applications of deception could be used for cyber deception (for example, see [7]).

Our aim in this paper is to explore how we can synthesise deception techniques from deceptive concepts employed outside of the cyber domain with cyber deception technology and deploy them as part of a network defence strategy. We highlight design thinking as a methodology that can contribute to developing this fused approach. Design thinking is a philosophy as well as a practical process. It has creativity and innovation at its heart and is about translating an idea into a reality that creates value. It is inherently transdisciplinary bringing together different academic disciplines with practitioner experiences.

In this paper we detail the results from a design thinking workshop carried out with deception experts from academia, government, a cyber deception technology company and independent defence and security experts. Our participants have skillsets that encompass social and behavioural science expertise as well as computer science, software engineering and cryptography. The outputs from the workshop include a critical analysis of design provocations for cyber deception and a journey map of potential issues to consider when operationalising cyber deception scenarios. We conclude with our reflections on the workshop and recommendations for future research and practice.

2. Background and related works

Cyber defenders are required to understand usual network behaviour and activity, from which, deviations may suggest potential threats [5]. Cyber defenders operate in complex socio-technical systems and their responses are influenced by their shared awareness, organisational influences and level of expertise [8]. Nyre-Yu et al. [8] argue that too much emphasis has been placed on improving algorithms and cyber security analyst education while ignoring the integration and interaction between the human and technology.

Cyber deception, however, has been an active topic of consideration for some years. Cohen's Deception Toolkit (DTK) [9] dates back to 1998 and Spitzner's Honeynet Project has been in existence since 1999 [10]. In recent years there has been research on cyber deception to protect cyber-physical systems [11], cyber deception has been used in wargaming [12], and there has been research on active deception for cloud infrastructures [13]. While social engineering arguably focuses on deception at the interface between behaviour and technology it does not focus on the activity in a network. Stech et al. [14] argue that cyber tactics are not mapped onto the classic components of denial and deception tactics; that there is no conventional terminology to describe the phenomenon of deception in cyberspace; that classic deception domain terminology is rarely used; and that classic deception domain researchers are rarely cited within the field. In short, cyber deception researchers seem rarely to study, exploit or build on the body of research that already exists about deception in other domains.

As Shade et al. [15] have pointed out, most research on cyber deception tools tends to focus on honeypots [16], suggesting ways to improve them [17], deliver them as a service [18], or to recognise their deficiencies [18], [19]. Where cyber deception research extends beyond honeypots it still tends to build from a computer science or engineering perspective [18], [19], [20], [21] with a smaller number of examples of research that include the impact of humans on cyber deception through, 'cognitive models and experimental games' [22] and 'computational models of human cognition' [20]. The assumption in such research is one of rational decision-making with a focus on formal rules or models in how decisions are made [23]. As cyber deception research has highlighted, however, we also need to understand the cognitive and behavioural processes of both the attacker and defender to improve cyber deception [22], [24], [25].

There is a limited but growing amount of research that demonstrates the value of bringing behavioural science and cyber deception together to deliver more innovative cyber deception techniques. Oppositional human factors for example, seek to flip ideas that are used to improve cyber security, to instead disrupt attacker cognition and behaviour and increase decision-making biases [25], [26], [27]. Attackers may not be familiar with a system they are penetrating and they should not expect a system to have good usability or design that would benefit them [25]. This presents opportunities to create deceptive systems, which disrupt the attackers' decision-making processes. With a vast cyber-attack surface any gain from disrupting attackers' decision-making is positive [25].

Creating decoy systems with large numbers of false assets as opposed to real ones can reduce the ability of an attacker to successfully target a real asset through reducing their chances, distracting them from real assets, and forcing them to switch attention and perform additional tasks which, in turn, slows them down [4][28]. Further benefits of decoy systems can be the improved detection of attackers from their engagement with false assets and increasing their level of confusion about the network's credibility [4]. Decoy networks have been extensively tested through red-teaming

experiments, including detailed exploration of red teamers' behaviour, personality and cognition, and physiological responses to cyber deception [28].

Use of host-based deception and deceptive messages has been shown in experiments to disrupt attacker decision-making, increase the time to conduct the deception and increase confusion within the attacker or attacking team [15]. Other cyber security testing includes Capture the Flag (CTF) exercises where attackers are required to exfiltrate assets from a network [29]. Such exercises, however, rarely measure actual human behaviour, performance and cognition, in response to active network defence using cyber deception [29].

Exploring the behavioural facets of cyber deception through experiment is valuable because as Shade et al. [15] point out, experiments have the 'necessary rigor' to deliver evidence-based results for the implementation of cyber deception techniques. Gutzwiller et al. [25], however, demonstrate the value of qualitative research in cyber deception by using exploratory 'think-aloud' techniques with red teamers to understand how it might be possible to exploit 'traditional decision-making biases' with cyber deception tools. Shade et al. [15], also highlight the value of collecting the qualitative 'emotional experiences' of those experiencing cyber deception. Such methodological approaches facilitate the gathering of rich data that, while it may not have the generalisability of experimental results, arguably make up for this by taking a naturalistic approach to cyber deception and providing a range of complementary insights.

The argument for a more naturalistic approach is heightened with the knowledge that humans perform poorly compared to algorithms in low validity environments [23]. In naturalistic settings decision-making is often conducted in conditions of varying uncertainty, time pressure, and cognitive load, which affects the ability to assess options [23]. The key here is to understand attacker and defender decision- making across socio-technical contexts in the field, and using input from cyber deception experts to inform how deceptive systems are designed. In taking this approach we build on the work of Reid and Black
[30] and their qualitative, inductive study. This research includes a discussion about the connection between technology, and the real-world context (referencing the work on warrants and digital footprints by [31], [32]) and highlights the need for dynamic cyber deception systems and methods [17], [22] and for deception tools that are both 'generic and expressive' [28].

3. Design thinking

In the field of software development, design thinking has been used in requirements elicitation [33], to improve scrum and lean start up [34] and for innovation in agile software development [35]. The value of design thinking is also well recognised for designing military systems [36] where Zweibelson [37] states that design thinking facilitates 'creating what is needed but does not yet exist'. The aim of design thinking is to create new ways of seeing, thinking and acting to develop not one solution but 'high-quality interventions to bring the whole system forward into a more desired state' [38]. As Buchanan [39] suggests design thinking aims to 'connect and integrate knowledge from many different specializations into productive' results'.

Design thinking is both a philosophy and a practical process. It has innovation at its heart and is participatory and transdisciplinary in practice. The basic steps to the design thinking process [40] are:

(i) to empathise with the end user and their understanding from their perspective;
(ii) to focus on exploring the problem in depth and detail and from multiple perspectives rather than rushing to fix on a solution;
(iii) to ideate by using divergent thinking to open up creative possibilities;
(iv) to prototype those creative possibilities by developing low fidelity models to start to refine ideas through convergent thinking;
(v) to test the results.

The process for design thinking encourages movement between divergent and convergent thinking. Divergent thinking is about *creating* choices by opening up possibilities and potential solutions, while convergent thinking is about *making* choices by tempering these ideas with real-world expertise and knowledge. It is an approach that encourages constructive and generative dialogue between inductive, deductive and abductive reasoning. The innovation that potentially comes from design thinking occurs because participants are encouraged to step into different worlds and experience different worldviews. This experience encourages abductive reasoning by focusing on 'what might be' [41] in the gaps between these various worlds and different perspectives. The act of synthesising information from different sources is an abductive sensemaking process where we make a 'motivated, continuous effort to understand connections' [42]. The resulting synthesis of world views encourages innovation through a 'leap of inference or intuition' [41].

Given that our aim is to fuse deception techniques from different contexts and environments so that we can innovate cyber deception tools, design thinking offers an interesting approach. In terms of our research problem our end user is the attacker as it is the attacker who is the

individual experiencing the network environment and the deception technology. As our interest was primarily to generate new ideas for cyber deception and to start to understand the implications of operationalising those ideas, we focused on those parts of the design thinking process that seek to understand the problem, and on ideation and prototyping. We developed two tasks. The first was to use provocations from different types of deception activities as a jumping off point for understanding the problem and ideation. The second task was to use journey mapping as a way of prototyping the ideas generated by thinking through how they would be realised.

4. Data collection

We needed to take account of a number of potentially destabilising aspects in our design- thinking workshop. Due to the Covid-19 pandemic we had to run the workshop online. While the facilitators have experience of running design thinking workshops, online delivery was new to them. We decided to use Zoom for video communication and for creating two virtual break-out rooms for group work. For the workshop activities we used an outline whiteboard application called Miro. One of the benefits of delivering the workshop online meant we could include expert participants from across both the UK and Australia. The drawback was that this meant we were working across two time zones - for UK participants the workshop ran from 0830 to 1100 and for Australian participants the workshop ran from 1730 to 2000. We wanted diversity in our participants' occupations and our aim was to synthesise knowledge across disciplines, or at least bring it into generative dialogue. This meant for this initial workshop we wanted to invite participants who would be more likely to be comfortable engaging in this type of activity. To maintain privacy and encourage the sharing of creative ideas we chose not to record the workshop but to collect the outputs in the form of field notes and the information created on the Miro whiteboard with virtual post-its throughout the workshop activities.

4.1 Participants

While taking these factors into account we also needed to identify participants with expertise in deception, along with expertise in across social and behavioural sciences, cyber security, software engineering, computer sciences and cryptography, who would also be open to a design thinking approach. As design thinking is inherently transdisciplinary it is not a natural skill for everyone. It is often a journey of trust, risk and fear because as Bernstein [43] highlights there is 'pain inherent in abandoning one's intellectual comfort zone by working outside one's home discipline'. The lack of understanding of each other's methods leads to fragile trust relationships that can break down' [44]. Accordingly, our pool of potential participants was small and we invited six participants who comprised two commercial technology designers, one Government scientist from Defence, one independent behavioural deception expert, one academic researching cyberspace operations, and one cryptography engineer. Three of our participants have experience of thinking from an adversarial perspective in a cyber security environment. Two academics facilitated the workshop – one in the UK and one in Australia. Both have expertise in cyber security in a national defence and security context. A further academic with experience in the psychology of deception acted as a note taker throughout the workshop.

4.2. Procedure

The two facilitators set up the Miro board with the agenda, introductory remarks and the two workshop tasks. Given that none of the participants had experience using Miro we decided not to allow them user access to the whiteboard because it would have been a distraction, but to share our facilitator screens so that they could see the tasks as well as the facilitators adding virtual post-its of their contributions on the whiteboard as they worked their way through the design thinking tasks.

The workshop was kept to 2.5 hours because our experience of participating in online workshops is that they are more tiring than physical workshops and also because of the time zone difference. We started the workshop with introductions, setting the teams for the group work, discussing the agenda and setting workshop rules to ensure we created a safe space where participants felt comfortable to contribute freely. One of the facilitators gave a brief overview of the design thinking process.

Prior to the workshop each participant had received a set of six design provocations on slides. The aim of the provocations was to spark ideas and prompt different ways of thinking about cyber deception. Each slide comprised an image and brief explanatory text. The slides included the following: (i) Japanese castle defences; (ii) code smells in software engineering; (iii) sliding doors (iv) gang graffiti and tagging; (v) the plot of an Indiana Jones action adventure movie; and (vi) deception for physical safety.

The first task in the workshop took one hour and was a discussion of each provocation in turn, with discussion points recorded as ideas on virtual post-its. For the second task we also allocated one hour. We put participants into two break-out groups to discuss what a journey map for implementing the provocations would

look like.

5. Analysis and discussion

5.1. Provocations

5.1.1. Japanese castle defences. This provocation used the example of Japanese castle defences [45] where paths would be designed to canalise attackers. The path might be sloped with uneven steps so that attackers would be forced to slow down or consciously navigate the path. It might wind far longer than it needed to both increase the distance attackers had to travel, but also to make them more visible to defenders. The discussion that arose from this focused on whether we could guide an attacker through a network in a similar way, thereby increasing the mental load on the attacker while also slowing them down and making observation easier. Interestingly the idea of lengthening paths has been researched and, while still in early stages, changing configurations in a network can increase uncertainty and ambiguity [46], [47]. One participant made the point that if the path is too complex attackers may well simply find an easier way into the system and that might be via a different exploitation route such as exploiting someone with legitimate access. The other concern was that such a defence might impede the normal running of the network. The conclusion was that for cyber deception this was '*interesting but complex'* and the trade-off between normal running and network defence needed to be understood. There was discussion over whether this was a deception technique even though it focused on, '*influencing and shaping behaviour'*. While not using deception in the way we would normally expect, one participant pointed out that it does work deceptively in the form of a lure.

5.1.2. Code smells. This provocation was based on 'code smells' [48] and the idea that software code may have a surface indication that 'corresponds to a deeper problem in the system'. For example, Martin Fowler [49] suggests that 'A long method is a good example of this - just looking at the code and my nose twitches if I see more than a dozen lines of java'. This provocation raised the question of whether you could introduce a code smell so that an attacker would conclude that, '*This doesn't feel right for what I understand about the machine'*. One participant noted that code smells have been looked at in forensic analysis, and they can cause people to look more closely at areas of code. Depending on the attacker there could be at least two outcomes from a deceptive code smell; it could either scare the attacker away or motivate them to explore further. Another participant pointed out this was, in effect, '*subconscious anomaly detection'* and dependent on cues and pattern recognition. To be useful the defender would need to consider the level of fidelity required to ensure that anomalies were recognised. This prompted discussion of what would need to be considered when using cues in a network because, in terms of behaviour and sense-making, they need to not be '*too thick, too big, too many*'. The conclusion from participants was that we could take this concept and break it out into more detail and that it could have lots of practical utility. The question was raised though about how to make this deception realistic for use against adversaries; for example, how could it be automated, scaled and tested?

5.1.3. Sliding doors. This was a second provocation based on Japanese castle defences and focused on the idea of sliding doors that would allow access to different parts of the building depending on which side they were opened from, or may even just be one- way doors so that attackers would be trapped. One participant noted that this is classic deception. Another pointed out that there are parallels in cryptography with one-way functions, but few examples in the network space. A comparison was made with '*tar pits'* as cyber deception tools where the attacker gets stuck and quitting the network is the only means of escape. Participants discussed whether it would it be possible to treat different types of network traffic in different ways. For example, could this idea be used to augment a tripwire approach to network protection so that if a tripwire is activated the system then changes how it looks to an attacker. Participants believed that this had value for cyber deception but some raised the issue of what would happen once the deception was known about. One participant countered this by linking the idea to cryptography where, '*design philosophies focus on the strength of the algorithm but knowledge of the algorithm doesn't mean that you have the key*'. Just because you know the mechanism is there doesn't mean it isn't still valuable. This has been addressed in recent research that demonstrates that deception tactics can still be effective even when known about [50].

5.1.4. Gang graffiti and tagging. This provocation used the idea of gang graffiti and the way it is used to symbolise a gang's turf and to act as a 'no trespassing' sign to rival gangs. When it came to considering how this could be used in cyber deception, participants felt it could act as a potential deterrent to attackers. It was observed though that what is more often seen on a network is that rather than acting as a deterrent an attacker would be more likely to either carry on with their attack and force other attackers out, or to simply use the network side-by-side with the other attackers. This raises the question posed by one participant, '*if you have*

multiple state actors and you see someone from your organisation there, does that cause a deconfliction issue back in the office, before you proceed further?' Could this be used to slow down an attacker's progress if they were a state actor? Another issue (similar to the one made about code smells) is how this would impact on the attacker's decision-making process because for some this would be a deterrent and for others a call to arms to attack. Again, it was felt that this required knowledge of the attacker and who or what they might be afraid of. This raised the question of what fear might look like and how it might be used in cyber space, causing one participant to ask *'Is it fundamentally different doing cyber deception then?*' A suggestion was made that it might be possible to increase fear in an attacker through the use of '*cold, generic info to make them feel that they have been doxxed'*. A discussion ensued about how we bridge the gap between cyber and physical to capture or disrupt an adversary. Cyber creates threats and opportunities for new forms of deception that do not occur (or that occur quite differently) in the real world, through automation, anonymity, impersonation, and digital footprints.

5.1.5. Indiana Jones narrative. This provocation took the narrative from the classic film Indiana Jones and The Last Crusade. In the film Indiana Jones has to deal with a series of protective booby traps to find the Holy Grail, which is hidden amongst hundreds of other potential grails, all of which fatally poison the person who drinks from it believing it to be the Grail. One participant pointed out that the idea of multiple potential grails is the equivalent of the cyber deception technique of hay-stacking (for example, hiding real database entries amongst vast numbers of fake entries). Hay-stacking has some weaknesses though if an attacker can see that the defender is only touching the real object and the defender has to ensure that they touch both the real and fake objects. Another participant highlighted the fact that in the film Indiana Jones has to escape from the system of traps once he has got the Grail. In the network this would be an attacker who wants to achieve their aim but also wants to leave the network without being caught. In the discussion it was pointed out that some of the obstacles Indiana Jones faces are binary; if you trip one of them you are out; if you are caught, you are out and just as in the film there are, *'lots of booby traps, just not working the way you would expect it'* so the question was raised about whether you *'could conceive a parallel to this in a cyber domain - booby traps tied with the network self-destructing'* If a network self-destructs this undermines availability inflicting damage on both defender and attacker, potentially creating a denial of service attack on oneself. If the attacker's objective is to deny the network this might make it easier for them because it could tie up SOC staff from defending other areas. The issue of exploits not working all the time was discussed with the conclusion that attackers getting a feel for their success rate could be a potential area for cyber deception. If everything is failing, or not working in expected ways, this increases confusion in the attacker.

5.1.6. Deception for physical safety. This was a photo of a perfectly flat floor tiled to look as if it was undulating to ensure that people walked rather than ran on it. Again, one participant raised the issue of the longevity of this technique and doubted it would continue to work once it was known but another participant compared this to Kerckhoff's Principle [51] noting that a cryptographic system should be secure even if everything about the system, except the key is public knowledge. This led to a discussion about what we gain, if anything, from the element of surprise. Widely publicising the use of cyber deception could mean that we cause our attacker to question everything which, in turn, could slow them down, as one participant commented, **'***Even though I know it's there - if I went there, I would still feel uncomfortable'.* While this form of deception may have diminishing returns, with the effect being greatest when first encountered, it could be relatively cheap to implement. We see these types of visual disruption in road flow control systems (e.g. illusory 3D crosswalks, anamorphic images of children playing in the street) to try to get drivers to slow down. Initially, such systems are effective, however, drivers soon become desensitized and are complacent in their presence, thereby increasing risk [52]. We would have to consider where we would get the most value from using this in the network and whether we design a cyber deception tool that was context aware. In response to this one participant noted, *'perceptual cues are very powerful but very brittle – this might not work from other end of corridor or at night, for example'* and there may be similar issues deploying this type of cyber deception in a network. The issue of timing of the use of deception was raised. If you deploy deceptive assets too early you can waste them, and if they are deployed too late then their value may have '*withered on the vine'*. Participants highlighted the need to consider both timing of surprise in cyber deception and link it to the deceptive outcome we want to achieve. Interestingly, this accords with research recently carried out that demonstrated that the late timing of deception in an intrusion reduces cyber attacks [53].

5.2. Provocations and ideation.

As was hoped, the provocations raised more questions than answers and participants demonstrated divergent thinking when it came to creating new ideas for implementing cyber deception. The discussion about

what deception is, and whether cyber deception is fundamentally different. Similarities and differences between cyber deception and deception in the physical world were raised especially around the idea of creating fear in an attacker and whether it was possible to do that on a network. There was discussion about the difference between deception, and influencing and shaping behaviours as well as the role played by more subtle deception techniques such as lures, cues and patterns and how these could be implemented on a network. The potential of cyber deception techniques to work two ways was also noted – either to deter an attacker or to provoke interest so that they investigate further. Finally, practical issues were highlighted such as how such cyber deception techniques could be automated, scaled and tested.

5.3. Journey mapping

A journey map is a way of designing a user experience for a new product. In our case the user is an attacker in the network and the product is cyber deception technology. For this task we broke the journey map down into five sections: (i) activities – this is where we focused most effort considering what needed to happen to develop innovative cyber deception technology; (ii) risks – what we should be concerned about through the development process; (iii) questions – what we need to know at each stage; (iv) success – defining how we would know if we had been successful; (v) opportunities and threats – what else could help or hinder us in achieving our goals? To select the provocations to journey map we asked participants to think about where they would place each the provocation on a PICK (Possible, Implement, Kill or Challenge) chart – a two by two matrix that plots ideas in terms of ease of implementation and potential benefits. There were three provocations that looked the most promising after the PICK activity, these were the Indiana Jones narrative, the deception for physical safety, and the gang graffiti and tagging. When we reflected back on the journey maps there was significant overlap between them and so we synthesised them into one overarching journey map for developing new cyber deception tools.

5.3.1. Activities. The journey that we mapped for developing new cyber deception exploits comprised the following activities. Firstly, we need an idea of whether the attacker is human or software and, from this, we also need to understand the motivation and what different attack strategies might be invoked. For example, does the attacker care if they are discovered or do they hope to carry out their attack undetected? Second, we need a clear idea of what we are protecting with cyber deception and to surmise the likely path through the network that the attacker might take. So far these two activities are very similar to the process that would be followed for any risk assessment on a network (understanding the threat actor, the vulnerabilities and the assets to be protected). The difference perhaps is the lens through which risk is understood. For cyber deception capability development, the focus should be on putting ourselves in the position of the attacker and seeing the network through their eyes, rather than ours as the defender, and focusing on the technological tools we can improve.

The next set of activities are different, however, and this is where, in the process of visualising the attacker's journey, our participants brought together their expertise of deception techniques in different contexts. The first step suggested by participants involved actively exploring analogues from other domains of deceptive practice. As a result of doing this a range of risks, question, measures of success, opportunities and threats arose so that the following additions to the journey map were added for consideration when developing innovative cyber deception techniques.

5.3.2. Questions. A range of questions central to the design process was raised by participants. The most important questions related to understanding how the network might look from the attacker's perspective. This would require some kind of persona or threat model. Participants raised questions about the level of detail needed and highlighted the risk of relying on a model of the attacker that turned out to be incorrect or flawed (for example relying on previous attacker behavioural data which is not always consistent with the present).

At the start of the process cyber defenders would also need to define what they want the attacker to do as a result of cyber deception. What is the desired behaviour and how might the attacker's sensemaking be shaped to generate that behaviour? Subsidiary questions stemming from this include:

- What perceptual cues do cyber defenders need to provide to attacker to shape their sensemaking? How can cyber defenders ensure that such cues will be detected and attended to by an attacker? And how can designers ensure that attackers will make desired sense of these cues?
- How can cyber defenders increase an attacker's dwell time in the network to deplete their resources and gather more information about their tactics, techniques and procedures? Would the return on being able to do this be of sufficient value for threat analysis to warrant the activity?
- How might cyber defenders force an attacker to follow a desired path through a network?
- How might cyber defenders instil fear in an attacker online? Would it be sufficient to startle them and, if so, how might this be achieved?

5.3.3. Measures of success. When considering measures of success participants raised the issue of what success might look like; whether it would look the same for all stakeholders and what it would look like to a customer. Participants felt that cyber defenders should also distinguish between measures of success at the level of individual cyber deception techniques as well as when they were used in combination and at steps in-between, such as observing cues that support deception techniques. One participant suggested we might want to look at a broader range of measures, perhaps even including stories about deception experiences in our network that were publicised on hacking forums. Another participant questioned whether measures of success might be overrated and that there could be benefits in blind application of techniques. This led to a discussion about whether this would be acceptable from a cost and ethical perspective.

5.3.4. Risks, opportunities and threats. When considering risks, opportunities and threats participants raised the opportunity or threat of advertising the presence of deception and considered whether this might increase or decrease its effectiveness against an attacker. The point was also made that there could be unforeseen consequences if cyber deception techniques were used in combination – for example, might we inadvertently risk launching a denial of service against ourselves as defenders?

Commercial considerations that could pose a risk to the development of cyber deception include the confidence cyber defenders have in cyber deception technology. One participant suggested that people would rather use cryptography than hay-stacking or covert paths. Cyber defenders need to increase confidence in the use of cyber deception technology as well as the integrity of the technology itself. Other points that participants believed organisations needed to consider were the legalities and ethics of using cyber deception technology and the cost of resources to develop and deploy it.

Finally, an interesting theme that emerged as a result of the journey mapping exercise was that of time. There are temporal costs in terms of the length of the lifecycle of cyber deception exploits as well as other temporal aspects to consider such as the necessary length of time for exposure to cues, dwell time in the network, the potential value of false temporal cues as well as the temporal aspects of the commercial development and deployment of cyber deception technology.

6. Limitations and reflections

The use of design thinking enabled us to bring different academic disciplines and practitioners' experiences into a generative dialogue to discuss ways of designing cyber deception tools and techniques that fuse deception techniques from different contexts and that bring together cognitive and technology effects. Running a design thinking workshop online across two countries and two time zones worked well and could be replicated with similarly sized groups of participants. Larger groups of participants would be difficult to manage in an online setting and to maintain organic group discussions. Running the workshop online was successful but resource intensive and demonstrates the need for skilled facilitators. This was a relatively small, qualitative, study but we believe this limitation was off-set by the expertise of our participants. Now we have demonstrated the potential of using design thinking for cyber deception our next steps will be to run further iterations of this workshop with a broader range of participants both in terms of occupation and academic disciplines. We would also like to trial other design thinking tasks within the online workshop and to prototype and test the outputs from this workshop further.

The potential benefits of using design thinking included a broad conversation about how to affect the attacker's behaviour and decision-making processes while in the network. Participants demonstrated divergent thinking by taking the provocations as a creative starting point and expanding the potential solutions from the technology through to the cognitive and behavioural effects on the attacker and how to achieve them. The journey map demonstrated convergent thinking with regard to operationalising cyber deception on a network. Participant discussions raised the positioning of a deceptive asset as part of an integrated and layered network defence strategy. The difficulty of evaluating and assessing the effectiveness of cyber deception technology against different attacker types and under a range of commercial pressures is an important consideration.

From a research perspective a range of interesting areas for future work were identified. We identified a need for research that looks at fear, surprise and startle effects online; whether this is possible to achieve and what these effects might look like. The issue of time needs further research, specifically the effect of tempo on designing, developing and delivering cyber deception. Risks and unanticipated effects of cyber deception techniques, such as triggering a denial of service attack and the psychological impact on both attackers and cyber network defenders, needs to be explored. Each of these topics requires researchers to lift cyber deception research into a space that considers the socio-technical dimension of both the attacker and defender in cyber deception.

7. Conclusion

In conclusion, we have demonstrated the potential benefits of using design thinking to synthesise deception techniques from a range of contexts and illustrated the potential this could offer for developing new cyber deception tools and techniques. Such tools and techniques may offer a more subtle and nuanced approach to cyber deception and could become an important part of the cyber defender's toolkit. In a military environment this could complement other manoeuvrist approaches to warfare. The substantive contribution has been to highlight areas for future research and practice in cyber deception. The methodological contribution has been to demonstrate the potential value of design thinking for cyber deception tools. Overall the study demonstrates the value of seeing the network through the eyes of the attacker and understanding how their experience could be shaped as they move through the network.

8. Acknowledgements

We would like to thank our participants for the time, energy and creativity that they contributed to this research.

9. References

[1] D. Brody, "Industry Analysts Now Fully Endorse Deception Technology" Online. Available: https://www.illusivenetworks.com/blog/industry- analysts-offer-full-throated-deception-technology- endorsement/

[2] Ministry of Defence, UK Defence Doctrine: Joint Declaration Publication 0-01, 5th ed. Shrivenham: Ministry of Defence, 2014.

[3] S. Henderson, "Deceptive Thinking Workshop," Paper presented at the 1st MilDec Military Deception Symposium, 2nd-3rd November 2011, Defence Academy of the United Kingdom, Shrivenham.

[4] K. J. Ferguson-Walter, D. S. LaFon, and T. B. Shade, "Friend or Faux: Deception for Cyber Defense," Journal of Information Warfare, vol. 16, no. 2, pp. 28– 42, 2017.

[5] R. S. Gutzwiller, S. M. Hunt and D. S. Lange, "A task analysis toward characterizing cyber-cognitive situation awareness (CCSA) in cyber defense analysts," 2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), San Diego, CA, 2016, pp. 14-20, doi: 10.1109/COGSIMA.2016.7497780.

[6] N. Rowe and J. Rrushi, Introduction to Cyberdeception. Springer International Publishing, 2016.

[7] S. Henderson, R. Hoffman, L. Bunch, and J. Bradshaw, Applying the Principles of Magic and the Concepts of Macrocognition to Counter-Deception in Cyber Operations. Paper presented at the 12th International Naturalistic Decision Making Conference, 9–12 June, 2015, McLean, Virginia, USA.

[8] M. Nyre-Yu, R. S. Gutzwiller, and B. S. Caldwell, "Observing cyber security incident response: Qualitative themes from field research," in Proceedings of the Human Factors and Ergonomics Society 2019 Annual Meeting, 437-441.

[9] F. Cohen and Others. "The Deception Toolkit". Risks Digest, 19, 1998.

[10] https://www.honeynet.org/about/ accessed 20th November, 2020

[11] T. Vollmer, and M. Manic. "Cyber-physical system security with deceptive virtual hosts for industrial control networks," IEEE Transactions on Industrial Informatics, 10: 1337-1347, 2014.

[12] K. E. Heckman, M. J. Walsh, F. J. Stech, T. A. O'boyle, S. R. DiCato, and A. F. Herber, "Active cyber defense with denial and deception: A cyber-wargame experiment," Computers & Security, 37: 72-77, 2013.

[13] A. Brzeczko, A. S. Uluagac, R. Beyah, and J. Copeland, "Active deception model for securing cloud infrastructure," in 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), April 2014, 535-540

[14] F. Stech, K. Heckman, P. Hilliard, and J. R. Ballo, "Scientometrics of Deception, Counter-deception, and Deception Detection in Cyber-space," PsychNology Journal, 9: 79-122, 2011.

[15] T. B. Shade, A. V. Rogers, K. J. Ferguson-Walter, S. B. Elson, D. K. Fayette, and K. E. Heckman, "The Moonraker study: An experimental evaluation of host-based deception," Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 20201875-1884.

[16] N. C. Rowe, T. D. Nguyen, M. M. Kendrick, Z. A. Rucker, D. Hyun, and J. C. Brown, "Creating effective industrial-control-systems honeypots," in Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 2020,1845-1854.

[17] A. Niakanlahiji, J. H. Jafarian, B-T Chu, and E. Al- Shaer, "HoneyBug: Personalized Cyber Deception for Web Applications," in Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 2020, 1895-1904.

[18] J. H. Jafarian and A. Niakanlahiji, "A Deception Planning Framework for Cyber Defense," in Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 2020, 1905- 1914.

[19] M. S. Miah, M. Gutierrez, O. Veliz, O. Thakoor, and C. Kiekintveld, "Concealing Cyber-Decoys using Two-Sided Feature Deception Games," in Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 2020, 1915- 1924.

[20] E. A. Cranford, P. Aggarwal, C. Gonzalez, S. Cooney, M. Tambe, and C. Lebierre, "Adaptive Cyber Deception: Cognitively Informed Signalling for Cyber Defense," in Proceedings of the 53rd International Conference on System Sciences, Jan 2020, 1885-1894.

[21] G. Ayoade, F. Araujo, K. Al-Naami, A. M. Mustafa, Y. Gao, K. W. Hamlen, and L. Khan, "Automating Cyberdeception Evaluation with Deep Learning," in Proceedings of the 53rd International Conference on System Sciences, Jan 2020, 1925-1934.

[22] C. Gonzalez, P. Aggarwal, E. A. Cranford, and C. Lebiere, "Design of Dynamic and Personalized Deception: A Research Framework and New Insights for Cyberdefense", in Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 2020, 1825-1834.

[23] D. Kahneman, and G. Klein, "Conditions for intuitive expertise: A failure to disagree," American Psychologist, 2009, 64, 515-526. doi: 10.1037/a0016755.

[24] K. Ferguson-Walter, S. Fugate, and C. Wang, "Introduction to the Minitrack on Cyber Deception for Defense," in Proceedings of the 53rd International Conference on System Sciences, Jan 2020.

[25] R. S. Gutzwiller, K. J. Ferguson-Walter, and S. J. Fugate, "Are cyber attackers thinking fast and slow? Evidence for cognitive biases in red teamers reveals a method for disruption" Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Oct. 2019.

[26] R. Gutzwiller, K. J. Ferguson-Walter, S. Fugate, and A. Rogers, "'Oh, Look, A butterfly!' A framework for distracting attackers to improve cyber defense," Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Oct. 2018.

[27] C. Wang, and Z. Lu, "Cyber deception: Overview and the road ahead," IEEE Security and Privacy, March/April 2018, 80-85.

[28] K. J. Ferguson-Walter, T. B. Shade, A. V. Rogers, E.M. Niedbala, M. C. Trumbo, K. Nauer, K. M. Divis, A. P. Jones, A. Combs, and R. G. Abbott, "The Tularosa Study: An Experimental Design and Implementation to Quantify the Effectiveness of Cyber Deception," in Proceedings of the 52nd Hawaii International Conference on System Sciences, p. 10, Jan. 2019.

[29] K. J. Ferguson-Walter, M. M. Major, D. C. van Bruggen, S. J. Fugate, and R. S. Gutzwiller, "The world of CTF is not enough data: Lessons learned from a cyber deception experiment", IEEE Workshop on Human Aspects of Cyber Security, 2019, 346-353.

[30] I. D. Reid, and R. Black, "Toward a holistic model of deception: Subject matter expert validation," in Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 2020, 1865- 1874.

[31] P. K. Forster, "Countering individual jihad: Perspectives on Nidal Hasan and Colleen LaRose". Counterterrorism Exchange, 2: 1-11, 2012.

[32] A. Sandham, T. Ormerod, C. Dando, R. Bull, M. Jackson, and J. Goulding, "Scent trails: Countering terrorism through informed surveillance," Paper presented at Engineering Psychology and Cognitive Ergonomics – 9th International Conference, Orlando, Florida, USA. 2011.

[33] E. Canedo, A. Vinhadelli Papadópolis, A., Cerqueira, and A. P. F. De Araujo, "Application of Design Thinking for Elicitation Requirements in Mobile Applications," in Proceedings of the 53rd Hawaii International Conference on System Sciences, Jan 2020, 6651-6660.

[34] F. Dobrigkeit, and D. de Paula, "The best of three worlds-the creation of innodev a software development approach that integrates design thinking, scrum and lean startup," in DS 87-8 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 8: Human Behaviour in Design, Vancouver, Canada, 21-25.08. 2017, 2017, 319-328.

[35] L. Corral, and I. Fronza, "Design thinking and agile practices for software engineering: an opportunity for innovation," In Proceedings of the 19th Annual SIG Conference on Information Technology Education, September 2018, 26-31.

[36] A. P. Jackson. (2019). Design Thinking: Applications for the Australian Defence Force [Online]. Available https://video.army.gov.au/ADC/Publications/documents/joint_studies/JSPS_3_Design_Thinking.pdf.

[37] B. Zweibelson, "Fostering deep insight through substantive play," in Design Thinking: Applications for the Australian Defence Force [Online], A. P. Jackson, Ed. retrieved from https://video.army.gov.au/ADC/Publications/documents/joint_studies/JSPS_3_Design_Thinking.pdf, 2019, pp. 105-120.

[38] K. Dorst, "Design beyond Design," She Ji: The Journal of Design, Economics, and Innovation, 5, 2, pp. 117- 127, 2019.

[39] R. Buchanan, "Design research and the new learning," Design issues, 2001, 17: 3-23.

[40] S. Doorley, S. Holcomb, P. Klebahn, K. Segovia and J. Utley, Design Thinking Bootleg [Online], Stanford d.school, 2018. Available: https://static1.squarespace.com/static/57c6b79629687fde090a0fdd/t/5b19b2f2aa4a99e99b26b6bb/1528410876119/dschool_bootleg_deck_2018_final_sm+%282%29.pdf, accessed 30 July 2018.

[41] J. Kolko, Well-Designed: How to Use Empathy to Create Products People Love. Harvard Business Review Press, 2014.

[42] G. Klein, B. Moon, and R. Hoffman, "Making Sense of Sensemaking 1: Alternative Perspectives," Intelligent Systens, 21, 4, pp. 70-73, 2006.

[43] J. H. Bernstein, " Transdisciplinarity: A review of its origins, development, and current issues," Journal of Research Practice, 2015 11, Article R1.

[44] M. MacLeod, "What makes interdisciplinarity difficult? Some consequences of domain specificity in interdisciplinary practice," Synthese, 2018 195: 697- 720.

[45] S. Henderson, S. (2020, Feb. 29). The Talented Hawk Hides Its Claws [Online]. Available: https://deceptionbydesign.com/the-talented-hawk- hides-its-claws/

[46] V. E. Urias, W. M. Stout, and C. Loverro, "Computer network deception as a moving target defense," in 2015 International Carnahan Conference on Security Technology (ICCST), September 2015, 1-6.

[47] G. S. Ahn, K. J. Kwak, A. Bogaevskiy, J. Li, G. Briskin, and R. Vaeth, "NetShifter: A Comprehensive Multi-Dimensional Network Obfuscation and Deception Solution," in Autonomous Cyber Deception, Cham, Springer, 2019, pp. 125-146.

[48] M. Fowler, Refactoring: Improving the Design of Existing Code 1st ed. Reading Massachusetts: Addison-Wesley, 1999.

[49] M. Fowler. (2006, Feb. 9). *CodeSmell* [Online]. Available: https://www.martinfowler.com/bliki/CodeSmell.html

[50] K. J. Ferguson-Walter, "An Empirical Assessment of the Effectiveness of Deception for Cyber Defense," PhD Thesis, 2020. [Online]. Available: https://scholarworks.umass.edu/dissertations_2/1823/

[51] A. Kerckhoff, "La cryptographie militaire,". Journal des Sciences. Militaires, IX 5-38, 1883.

[52] T. Vanderbilt, Traffic: Why We Drive the Way We Do (And What It Says About Us). London: Penguin, 2009.

[53] P. Aggarwal, C. Gonzalez, and V. Dutt, "HackIt: A Real-Time Simulation Tool for Studying Real-World Cyberattacks in the Laboratory," in Handbook of Computer Networks and Cyber Security. Cham: Springer, 2020, pp. 949-959