Spidey Sense: Designing Wrist-Mounted Affective Haptics for Communicating Cybersecurity Warnings

YOUNGWOOK DO, Georgia Institute of Technology, USA
LINH THAI HOANG, Georgia Institute of Technology, USA
JUNG WOOK PARK, Georgia Institute of Technology, USA
GREGORY D. ABOWD, Georgia Institute of Technology, USA and Northeastern University, USA
SAUVIK DAS, Georgia Institute of Technology, USA

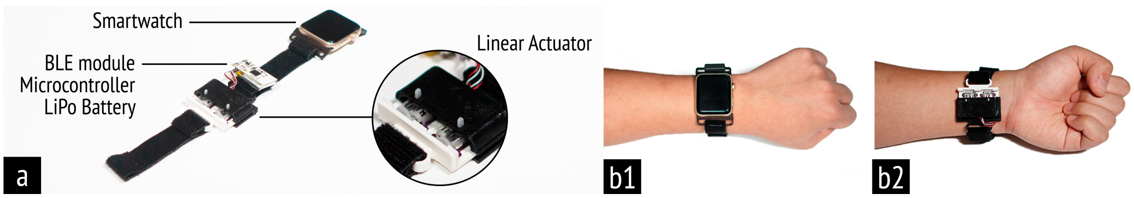

Fig. 1. Spidey Sense is a smartwatch wristband that produces squeezing sensations to alert people to "urgent" cybersecurity warnings. (a) The latest version of the Spidey Sense wristband, compatible with an Apple Watch. This version was developed after our evaluation. (b1) Top view of an Apple Watch with Spidey Sense; and (b2) Bottom view of an Apple Watch with Spidey Sense.

Improving end-users' awareness of cybersecurity warnings (e.g., phishing and malware alerts) remains a longstanding problem in usable security. Prior work suggests two key weaknesses with existing warnings: they are primarily communicated via saturated communication channels (e.g., visual, auditory, and vibrotactile); and, they are communicated rationally, not viscerally. We hypothesized that wrist-based affective haptics should address both of these weaknesses in a form-factor that is practically deployable: i.e., as a replaceable wristband compatible with modern smartwatches like the Apple Watch. To that end, we designed and implemented Spidey Sense, a wristband that produces customizable squeezing sensations to alert users to urgent cybersecurity warnings. To evaluate Spidey Sense, we applied a three-phased 'Gen-Rank-Verify' study methodology with 48 participants. We found evidence that, relative to vibrotactile alerts, Spidey Sense was considered more appropriate for the task of alerting people to cybersecurity warnings.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; *Haptic devices*; User studies.

Additional Key Words and Phrases: Wearable Device; Haptics; Usable Security Privacy; Security Warnings.

ACM Reference Format:
Youngwook Do, Linh Thai Hoang, Jung Wook Park, Gregory D. Abowd, and Sauvik Das. 2021. Spidey Sense: Designing Wrist-Mounted Affective Haptics for Communicating Cybersecurity Warnings. In *Designing Interactive Systems Conference 2021 (DIS '21), June 28-July 2, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3461778.3462027

1 INTRODUCTION

Cybersecurity warnings remain an essential component of human-in-the-loop secure systems, yet they often fail to capture people's attention and motivate action [8]. Prior work suggests there are two underlying problems that help explain why.

First, existing security warnings are communicated through the same overloaded channel as all other interruptions, warnings and notifications: via primarily visual modals, occasionally accompanied by a short audio and/or vibrotactile cue. Thus, with each non-essential notification users receive — and there are many [43] — they are trained to ignore the behavioral cues that are meant to bring their attention to security warnings [14, 49, 51]. Second, unlike threats in the physical world, cyber threats are communicated rationally but not viscerally [14, 39]. For example, an SSL/TLS certificate encrypts a user's sensitive data (e.g., passwords, credit card numbers) and help them securely to be sent to a host server. Once the certificate is expired, such data cannot be securely sent [52]. People might be able to *see* that they are navigating to a website with an expired SSL/TLS certificate the way they might see a red light on a pedestrian crossing. However, they may not *feel* that they should not proceed to that website when they see an expired certificate warning as they would if a friend had grabbed their wrist before they jaywalked.

One way of addressing this problem is to have a secondary notification channel that: (i) is used only to communicate important cybersecurity warnings — so that users might learn to differentiate alerts for urgent cybersecurity warnings versus other types of alerts; and, (ii) can communicate threats affectively — so that users might "feel" a potential threat in addition to seeing pertinent information about the threat. More formally, we need a mechanism of delivering cybersecurity warnings that helps bridge the gap between the *presence* of a cyber threat and the *perception* of that threat. We hypothesize that affective haptics and interfaces [10, 16, 50] can help bridge that gap. To explore this hypothesis, we present Spidey Sense, a smartwatch wristband that alerts people to urgent cybersecurity warnings through affective squeeze haptics. In short, we designed Spidey Sense to be an affective, secondary channel to deliver urgent cybersecurity warnings to end-users in varied contexts: not just when a user is situated in front of a laptop screen, but also when they are on-the-go, in a meeting, or interfacing with a screenless device (e.g., a smart lock).

Historically, deploying a wearable device specifically for cybersecurity has been considered impractical. Cybersecurity is, for most people, a secondary concern [13, 31, 36], and it is unlikely that a typical end-user would purchase and wear a tangible apparatus specifically for the purposes of improving their cybersecurity behaviors. However, smartwatches are increasingly popular and many, like the Apple Watch, have switchable wristbands. By developing a smartwatch wristband that can produce an affective haptic effect independent of the standard vibrotactile cues, we should be able to improve the perception of cybersecurity warnings in a manner that is both practical and deployable.

In designing Spidey Sense, we explored a wide spectrum of haptics, e.g., providing electrical impulses and producing prickling sensations using a robotic finger. Ultimately, we settled on "squeezing" for three reasons. First, we drew inspiration from prior work in psychology that found that people perceive a wrist squeeze as being associated with the emotions of fear and surprise [12, 21, 22, 25]. Second, squeezing sensations are generally painless. Third, the mechanical and electrical components used to produce a squeezing sensation can unobtrusively fit into a smartwatch wristband.

To explore the design space of wrist-mounted squeeze haptics and converge on one that is empirically effective for urgent cybersecurity warnings, we employed a three-phased study methodology — "Gen-Rank-Verify" — adapted from the "Find-Fix-Verify" crowd programming pattern from prior work in HCI and haptics [5, 40]. For the "Gen" phase, we recruited a small set of participants to independently create a unique squeeze notification using a design GUI that we constructed for manipulating the Spidey Sense wristband. For the "Rank" phase, we recruited an additional 30 participants to rank the initial set of squeeze notifications created in the previous phase, through a randomly-bracketed tournament where they made a series of pairwise comparisons. Finally, for the "Verify" phase, we ran a within-subjects experiment comparing the top-ranked squeeze notification vis-a-vis a vibrotactile baseline control condition — i.e., the current state-of-the-art for communicating warnings through a haptic channel. Through this three-part study, we converged on a Spidey Sense cue — one with many large, rapid pressure pulses — that was considered more appropriate than the vibrotactile baseline for alerting users to urgent cybersecurity warnings. Examining how psychological events are affected by squeeze haptics is beyond the scope of this paper.

Our work has limitations: most notably, we do not account for habituation effects, nor do we test Spidey Sense's effect on user behavior in the field. Nevertheless, our contributions should help synchronize the research agendas of the affective computing and usable security communities—a partnership we expect will become increasingly important as computing, and the cybersecurity threats it entails, further physicalizes. More concretely, we offer the following contributions in this paper:

- We iteratively designed and developed a smartwatch wristband that produces affective squeeze haptics to alert people to urgent cybersecurity warnings.
- We employed a three-part study methodology, "Gen-Rank-Verify", to systematically explore an open-ended haptic design space for affective cybersecurity warnings. Through this study, we show that cybersecurity warnings accompanied with a Spidey Sense cue were rated as more appropriate for urgent cybersecurity warnings than when they were accompanied by a vibrotactile cue.

2 BACKGROUND AND RELATED WORK

2.1 Improving Cybersecurity Warnings through Design

Cybersecurity warnings have been extensively studied. To date, the dominant thrust of this background work has focused on improving visual communications. Akhawe et al. ran a field study to evaluate the effectiveness of widely-used web browsers' security warnings. The authors found that while malware and phishing warnings helped prevent some users from visiting malicious websites, many users ignored the warnings outright [2]. Similarly, Sunshine et al. found that people learned to ignore the visual cues of SSL security warnings [49]. Egelman et al. examined the effectiveness of visual phishing warnings in web browsers and found that even though security indicators did help some people heed phishing warning, those warnings were still often ignored [14]. Beyond web browsers, Felt et al. surveyed users on the effectiveness of Android's permissions system, finding that end-users often did not pay attention to and had low comprehension of permission screens [18].

Prior work has also demonstrated the need for non-visual warning modalities. Micallef et al., for example, explored combinations of modalities for smartphone notifications and found that people preferred privacy notifications to be non-auditory and distinct from other types of notifications [35]. Beyond user preference, Vance et al. [51] found neurological evidence that non-essential notifications can blur with security warnings when distributed through the same communication channels, further motivating the need to communicate security and privacy information through distinct notification channels. Yet, existing smartphones and smartwatches are limited to visual, auditory, and conventional motor-based vibrotactile feedback—these channels are already overloaded and associated with non-essential notifications.

Overcoming habituation — i.e., people's tendency to be desensitized to warnings over repeated exposures — is another key challenge in improving security and privacy warnings. Prior work describes that people develop an automated response to ignore security warnings as they repeatedly experience the same warning cue [46, 47, 49]. Anderson et al. identify that providing variations of feedback patterns—polymorphic warnings— could reduce people's *habituation* to security warnings [3].

While prior work has studied and improved the visual and information communication of security warnings, there remains significant room for improvement. Specifically, the combination of this prior work suggests that existing warnings are often ignored by end-users for two key reasons: (i) they are communicated via over-saturated communication channels that users learn to ignore; and, (ii) they are communicated rationally, not viscerally. We explore affective haptics to help address both of these problems. To address these two gaps, we explore the use of affective haptics to improve security warning communication — a design space inspired but under-explored by prior work.

2.2 Haptic Feedback in Wrist-worn Devices

Prior work in HCI has explored the incorporation of expressive haptic feedback in wrist-worn devices. One recent approach is to use tactile displays: Ion et al. implemented a tactile display in a wrist-worn form factor to communicate a variety of information via tactile messages [26], while Huang et al. use tactile displays to afford new gaming and video experiences [24]. Moreover, Lee et al. explored the use of air to provide non-contact tactile stimulation for situations when contact between skin and a tactile display may not be possible. [32].

Beyond tactile displays, other prior work has explored mechanically-powered wrist-worn haptics. Leigh et al., for example, developed a wrist-mounted robotic finger for mechanical hand augmentation that enables synergistic interaction between a hand and machine joints (e.g., notification with a gentle tap on a user's hand) [34]. Other researchers have utilized wrist-mounted squeeze haptics for different application areas: e.g., to enhance virtual reality experiences [41] and help novice surgeons hone their surgical skills [9].

Additionally, researchers have developed wrist-worn haptic devices that provide on-skin electrotactile feedback. Pohl et al. designed a wrist-worn haptic device that provides electric stimulation on a skin to replicate itching sensation [42]. Withana et al. developed an on-skin tactile interface to provide electric stimuli to a finger [54].

While prior work illustrates the promise of wrist-worn haptics for a variety of application areas, the exploration of wrist-mounted haptics to better communicate cybersecurity warnings remains unexplored. In this work, we connect these two promising but disparate thrusts of research.

2.3 Designing Haptic Feedback Patterns

Exploring the design space of haptic feedback patterns best suited to a specific purpose remains a largely open problem, though prior work provides pointers. Saket et al. experimented with various smartphone vibration patterns to see which vibration pattern was the most effective in alerting end users [44]. The authors designed ten vibration patterns created with four signal types (e.g., short vibration on/off, long vibration on/off) and asked participants to choose the pattern best capturing urgency of notification. However, Saket et al. did not do a systematic design space exploration of possible vibration patterns — they designed the tested patterns themselves. Obrist et al. introduced a method to systematically explore and evaluate an open-ended design space of haptics with distinct groups of participants [40]. We adapt and extend Obrist et al.'s methodology to systematically explore the design space of affective squeeze haptics for cybersecurity warnings.

2.4 Affective Haptics and Applications in Cybersecurity

Physiological theories of emotion suggest that the sense of touch communicates emotion — e.g., the sensation of being squeezed can communicate anger, fear, sadness, and surprise to a person [4, 21, 22]. Affective haptics is the study and design of devices that take advantage of this fact to elicit, enhance, or influence users' emotional state via their sense of touch [16, 50]. The design space for affective haptics has been explored in various areas such as gaming [28], anxiety and depression treatment [6], and assistive communication technologies for children with autism [11]. Perhaps most relevant for our purposes, the sensation of a wrist-squeeze is highly linked to eliciting emotions of anger and fear [25].

To date, there has been relatively little cross-exploration of affective haptics with efforts in cybersecurity. The most related prior work come from works-in-progress that explore thermal feedback as a way to improve a end-user's security awareness. Wilson et al. introduced thermal feedback as a new way to communicate web browser security warnings to a user by examining the association between people's perception about temperature and the level of security [53]. Furthermore, Napoli et al. discuss that the effect that thermal stimulation has on an end-user's cognition can help improve their security awareness. Their preliminary findings suggest that thermal stimulation could help communicate the security of TLS certificate to end-users. However, they also note that this thermal feedback might confuse end-users because the heating pad used to deliver the thermal warning may contain residual heat after users navigate away from an insecure website [38]. Additionally, Song et al. implemented a wrist wearable that provides thermal and squeeze haptic feedback, and studied that thermal feedback was more confusingly recognized than squeeze feedback [48].

Inspired by background literature in affective haptics, we explore an alternative mechanism to deliver haptic cybersecurity warnings: a wristband that produces programmable squeeze effects.

3 TARGET DESIGN SPACE OF SPIDEY SENSE

In the lexicon of Schaub et al.'s design space for privacy notifications, we can think of Spidey Sense as a just-in-time (urgent), haptic, non-blocking notification through a secondary channel readily differentiable from other commonly used alert modalities [45]. More generally, Spidey Sense helps address an old problem — that people ignore cybersecurity warnings — with a new solution: a functioning, deployable haptic notification apparatus that produces a novel, squeezing haptic effect. In designing and evaluating Spidey Sense, we build on prior work in haptics, tangible interaction, and usable security to construct a deployable and practical form factor.

4 EXPLORATORY FORM FACTOR EVALUATIONS

Inspired by its namesake,[1] our top-level design consideration for Spidey Sense was to create a practically deployable device that produces affective haptic effects that alert people to urgent cybersecurity warnings. We consider cybersecurity warnings to be "urgent" when the risk of harm is severe and immediate *user action* is necessary to mitigate harm.

We targeted "urgent" scenarios in which outright prevention of potentially risky behavior may be paternalistic and inappropriate (and doing so can result in users finding insecure workarounds or becoming frustrated [1, 13]). For example, a user may be about to install software with predatory data collection practices, or may have navigated to a website that is known to distribute misinformation. In these situations, we want the user to be clearly aware of the risk and facilitate a safer course of action, but want users to retain agency.

These cybersecurity warnings typically occur on what Hong calls a *Tier 1* device [23] — i.e., devices that afford high interactivity and require high computation and high user attention. For example, laptops and smartphones are Tier

[1]Spider-Man, a Marvel Comics superhero, has the ability to discover nearby danger with his "spideysense."

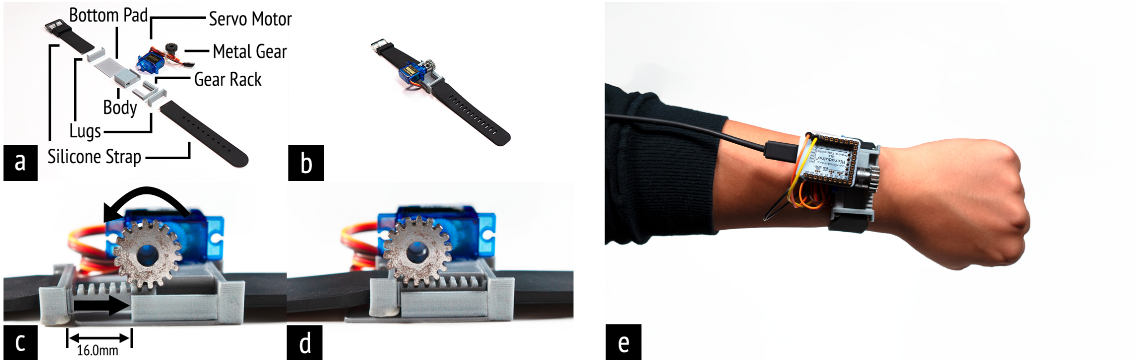

Fig. 2. Assembly of Spidey Sense: (a) before assembly; (b) after assembly. Spidey Sense actuation: (c) initial status; (d) Spidey Sense provides squeezing sensation when a gear on a servo motor rotates shortening the length of the wristband by 16.0mm at the maximum; (e) Top view of the Spidey Sense device with a microcontroller (Microduino Core and Microduino BLE)

1 devices, while a Nest thermostat is not. These Tier 1 devices typically collect large amounts of sensitive data and offer federated access to other important accounts (e.g., bank accounts). If these devices were compromised, a large proportion of an individuals' personal data and devices would be compromised, in turn. Therefore, heeding urgent cybersecurity warnings on these devices is critical.

4.1 Selecting the Right Affective Haptic Type and Form Factor

Tier 1 devices are typically not wearable, and thus have limited ability to provide direct haptic feedback. Historically, wearable devices have been impractical for day-to-day cybersecurity applications — they can be socially inappropriate and/or prohibitively expensive, both of which are damning for any application that primarily benefits a secondary concern, like cybersecurity [13, 31, 36]. However, smartwatches with interchangeable wristbands, like the Apple Watch, are increasingly popular and socially acceptable [37]. Moreover, smartwatch wristbands are relatively cheap and are placed directly on the skin. Accordingly, our key design insight is that if we can create a smartwatch wristband that can deliver a noticeable haptic effect and that is compatible, out-of-the-box, with commercial smartwatches, we should be able to create a practically deployable affective haptic system for cybersecurity warnings.

Inspired by prior work in affective haptics, we started by implementing early prototypes of four competing designs that we hypothesized would elicit a visceral response: (i) *Symbiont*: a robotic finger that produces a programmable scratch [34]; (ii) *Goosebumps*: an electrical impulse generator [42] that produces localized goosebumps; (iii) *ModiFiber*: a thread-based, artificial muscle actuator that produces a light squeeze[19]; and, (iv) *Mechanical*: a linear actuator that produces a programmable squeeze. To choose among these four early form-factors we evaluated each prototype, in turn.

Symbiont. We first attempted to replicate a"creepy-crawly" effect — i.e., the acute repulsion that accompanies the sensation of an insect crawling on one's skin. Leigh et al. introduced the concept of a "robotic symbiont": i.e., a programmable robot finger as an interactive agent [33, 34]. We adapted and modified the symbiont to produce a rubbing or scratching sensation on one's hand or wrist to approximate the "creepy-crawly" effect. Ultimately, this form factor was impractical for two reasons: first, it required bulky mechanics that made it uncomfortable to wear; second, the robot finger is intentionally obtrusive, reducing the situations in which it might be socially acceptable.

Goosebumps. Goosebumps are involuntary skin bumps that develop when a person experiences strong emotions. We next explored light electrical pulses to produce localized "goosebumps", inspired by prior work [42, 54]. Electrical pulse generators can be space efficient: they require only conductive electrode pads that make direct contact with the skin to provide noticeable electric stimulation. However, in our early user tests, we found two mitigating concerns. First, if the electrode pads lose contact with skin, the generated pulses are markedly less noticeable. Placed on the wristband of a smartwatch, it is likely that the electrodes will shift position and frequently break direct contact throughout the day. Second, we found that the produced effect was inconsistent within a user and across users: each individual has unique skin resistivity and this resisitivity varies with perspiration[42, 54]. Skin resistivity, in turn, impacted the noticeability of the electrical pulses.

ModiFiber & Mechanical Squeezes. Prior work on the communication of emotion via touch suggests that people naturally associate sudden "squeezes" as being associated with "anger", "fear" and "surprise." [22]. In addition, Baumann et al. studied that such emotions can be expressed through squeeze-haptics on a wrist [4]. Accordingly, we next explored two different types of actuators to produce a wrist squeeze—an artificial muscle and a servo motor. The underlying principle behind both actuation methods is the same: an adjustable length wrist band is constricted to produce a "squeeze" and expanded to relax the "squeeze." Recent prior work introduced ModiFiber: an artificial muscle in the form of thin silicone-laced threads that shrink when powered and return back to their original shape without power [19]. We implemented ModiFiber and interlaced the shrinking threads in a silicone wristband, but found that actuation was too slow and weak to be noticeable for our purposes (5% thread length shrinkage rate for around 20 seconds). We then tried a mechanical approach that we adopted and modified prototypes introduced in prior work [4, 48]: specifically, we designed an adjustable-length wristband that can be programatically shrunk or expanded through a linear servo motor. Ultimately, we proceeded with this fourth design because the mechanical servo motor can produce a reliable and strong squeeze effect in a comfortable form-factor that is small and unobtrusive.

5 SPIDEY SENSE — AN AFFECTIVE WRISTBAND FOR URGENT CYBERSECURITY WARNINGS

In order to produce a noticeable squeeze, Spidey Sense dynamically adjusts the length of a smartwatch wristband. We implemented a mechanical linear actuator using a servo motor to adjust the wristband length. For the Spidey Sense device (Figure 2), we attached a SG90 servo motor to a metal gear (McMaster-Carr, 14-1/2 Degree Pressure Angle) that traverses over a PLA-based 3D-printed gear rack. The motor is controlled via a microcontroller placed on the servo motor. We tried two types of microcontrollers, Arduino Uno and Microduino Core. Both microcontroller setups work similarly in terms of motor control. The microcontroller, in turn, is powered by a USB cable connected to an external power source (in practice, this would be the smartwatch). (Figure 2 (e)) At the ends of both sides, we affixed a silicone smartwatch wristband (width 22mm) that users could wear on their wrist. We used this model for our evaluations, but will show how we improved our design for practical deployability and wearability in the Discussion section.

The length of the wristband is adjusted by the motor. As the motor rotates clockwise or counter-clockwise, it pushes or pulls the gear rack that is, in turn, attached to a lug that holds a silicone strap connected to the full wristband. The gear rack's actuation range is 16.0mm (see Figure 2). Since each user has their own wrist circumference and might wear their watch differently, the squeezing pressure that Spidey Sense can produce can be different from individual-to-individual. To reduce this variation, when putting on the Spidey Sense wristband, the wristband should be fully fitted.

Note that our design is independent of any specific smartwatch platform. In practice, any smartwatch could directly communicate with the microcontroller that actuates the wristband.

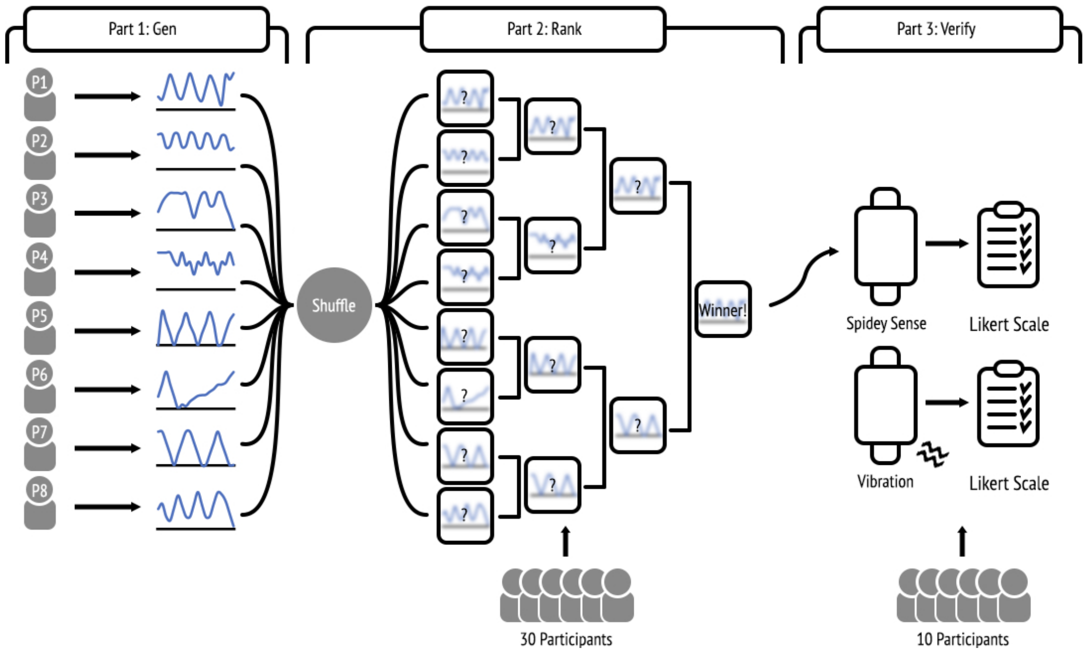

Fig. 3. We used the "Gen-Rank-Verify" user study structure: (Part 1) Gen—Participants create their best squeezing pattern; (Part 2) Rank—Participants rank patterns created by the participants in 'Find'; (Part 3) Verify—Participants evaluate the appropriateness of the top-ranked squeezing pattern of Spidey Sense and a smartwatch vibrotactile feedback for security warnings.

6 EVALUATION

As Baumann et al. have studied that different emotions can be elicited by varying patterns of squeeze haptics [4], our evaluation was designed to examine if such squeeze affective haptics can be applicable to alert users situated in our specific target scenarios — i.e., for urgent cybersecurity warnings. Thus, we consider the broader examination of how psychological events are affected by squeeze haptics to be out-of-scope.

6.1 Study Design: "Gen-Rank-Verify"

The high-level objective of our study was to systematically explore the design space of squeeze haptics in order to find a squeeze notification that improves people's perception of urgent cybersecurity warnings. To do so, we employed a three-part study we call "Gen-Rank-Verify." (Figure 3). Broadly, the goal of this three-part study was to enlist three separate sets of participants to converge on an affective wrist-squeeze pattern that was well-suited to a target scenario.

We started by selecting two "grounding" videos that exemplified the target scenarios for which we designed Spidey Sense. Then, we primed participants to think about our target scenarios by having them watch those videos. In the "Gen" phase, one set of eight participants independently generated squeeze notifications that they believed would capture their attention if they found themselves in the scenarios depicted in the grounding videos. In the "Rank" phase, a second set of 30 participants rank-ordered these independently generated notifications through a series of pairwise comparisons. Finally, in the "Verify" phase, a third set of 10 participants compared the appropriateness of the winning squeeze notification from the previous phases against a vibrotactile baseline.

We drew inspiration of this study pattern from Bernstein et al.'s 'Find-Fix-Verify' methodology that allow crowd worker collectives to collaboratively edit word documents [5], and Obrist et al.'s 'Find-Fix-Verify' methodology to design novel affective haptic notifications [40]. Both methodologies were designed to systematically explore an open-ended design space with a clear fitness function by leveraging the collective wisdom of many individuals. The "Gen-Rank-Verify" methodology shares many similarities with "Find-Fix-Verify", but we made a few adaptations to tailor the methodology to the task of designing affective haptics for cybersecurity warnings.

First, since lab studies poorly approximate situations in which users make cybersecurity decisions [8], we introduced a "priming" step in which participants were shown emotionally visceral target scenarios using grounding videos. Prior work suggests that video (e.g., film excerpts, short vignettes) is an effective medium to help viewers engage in scenes that approximate real-life situations [17, 27].

Second, whereas Obrist et al. had their "Rank" participants directly provide a total rank-order of all the haptic patterns their "Find" participants generated [40], we asked participants to make pairwise comparisons in a tournament with randomly-ordered brackets (shown in Figure 3). We made this modification because while it is challenging for people to meaningfully compare eight squeeze notifications simultaneously, it is relatively easy for them to compare two options at a time and assess which of the two is better suited to the target scenario. The tournament set-up with randomly-ordered brackets afforded a total-order ranking without requiring participants to assess all eight squeeze notifications simultaneously.

Third, while Obrist et al. had their "Verify" participants rate, on an absolute scale, how well the top-ranked haptic notification approximated an emotional sensation, we converted this phase into a within-subjects experiment. In random order, our participants rated the appropriateness of not only the top-ranked Spidey Sense squeeze notification, but also a "control" vibrotactile notification from a commodity smartwatch (Mobvoi Ticwatch E). This adaptation allowed us to comparatively evaluate how our best performing Spidey Sense squeeze notification performed relative to a realistic baseline.

Finally, we renamed the "Find" phase to "Gen" to better capture the idea that the first phase of the methodology involves participants engaging in a generation task.

6.2 Study Procedure and Results

Details about each study phase will be described in the following subsections. Our study protocol was reviewed and approved by an Institutional Review Board.

6.2.1 Priming Target Scenarios With Grounding Videos. For each of the three study phases — "Gen", "Rank", and "Verify" — an important first step was to have participants clearly understand an example of the target scenarios for which they were designing or assessing Spidey Sense squeeze notifications. To do so, we showed participants one or two "grounding" videos that exemplified target scenarios for which we would expect Spidey Sense to help with in practice.

For the laptop use-case, we created an abridged version of the Black Mirror [2] episode "Shut Up and Dance". In short, the video illustrates the perils of installing untrusted software on one's laptop; the main character installs malware disguised as an anti-virus, which affords remote attackers full access to his system that they then use to blackmail him.

For the smartphone use-case, we found a news report from the Canadian Broadcasting Corporation detailing the perils of blanket permission-granting for Android apps. In the video, journalists ask unwitting pedestrians to install a custom app they had developed that could track users' locations, read message histories and access their phone cameras.

[2]Black Mirror is a Netflix television show in which each episode depicts a human-interest story in a technology-fueled dystopia.

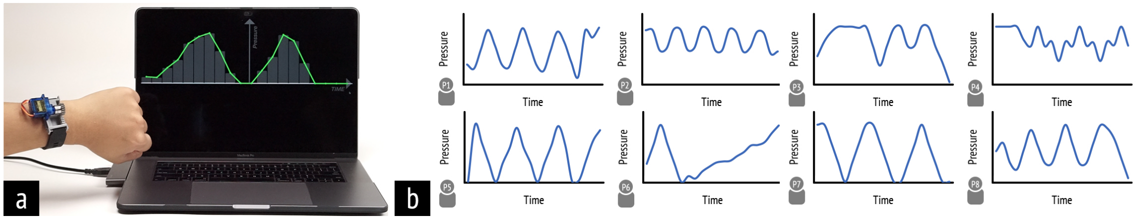

Fig. 4. (a) Design GUI used in "Gen"—a user creates their own best squeezing pattern by adjusting pressure over time, periodicity, and infinite / finite repetition; (b) Squeezing patterns (over one second) designed by the "Gen" phase participants—all the participants chose to have their squeeze notifications repeat indefinitely.

The journalists met these pedestrians a few days later for a follow-up, showing them the extent of the data the app was capable of collecting. The intention of the news report was to alert the general public of the importance of being vigilant about the permissions one grants to smartphone apps. Both of the grounding videos we showed participants are provided in the supplementary materials.

6.2.2 Gen - Generating Task-Appropriate Squeeze Notifications. To find a squeeze notification suitable for improving people's awareness of urgent cybersecurity warnings, it was first necessary to systematically explore the design space of squeeze notifications that Spidey Sense could produce. In the "Gen" study, we recruited participants to independently design novel squeeze notifications that they believed would capture their attention if they found themselves in a scenario depicted in the grounding videos. Below, we outline the high-level procedure participants ran through.

Design GUI: To facilitate participants' generation of these squeeze notifications, we created a simple GUI to manipulate the Spidey Sense wristband along the three configurable parameters — pressure over time, periodicity and infinite / finite repetition (Figure 4). By default, the parameters were set as minimal pressure over time and infinite repetition.

The GUI displayed a 2D line graph where the x and y axes represented time and squeezing pressure, respectively. This line graph represented the "pressure over time" parameter. Participants could manipulate this graph to vary how strongly the wristband squeezed their wrist through a repeatable one-second sequence. The one-second sequence was split into 20 contiguous buckets representing 50 ms each. Participants could alter the position of the graph in the y-axis for any of these 50 ms buckets; raising or lowering the line would increase or decrease the pressure applied at that interval, respectively. Participants could have the full one-second sequence repeat infinitely or could manually specify the number of repetitions. To facilitate manipulation of the interface, participants used a touch screen to directly draw a pressure-over-time graph into the interface. We included a video figure in the supplementary materials to demonstrate how the Design GUI works, in practice.

Tutorial and Interface Exploration: We created a tutorial video explaining all of the design interface affordances. The tutorial video is provided in the supplementary materials. After watching this tutorial video, participants were given 5 minutes to familiarize themselves with the interface by trying to adjust the three parameters before we proceeded to the next step.

Grounding Videos: We next showed participants the grounding videos described above.

Designing Custom Squeeze Notifications: Participants were then asked to design a squeeze notification that would best capture their attention if they were in a situation analogous to the ones depicted in our grounding videos. Once they finalized their custom squeeze notification, we asked them to save the pattern. If they later wanted to change this pattern, they were able to create and save a new one.

While participants were designing their notifications, we had them think out-loud to gain insight into their design process.

Recruitment and Compensation: We recruited participants by placing flyers advertising our study around our institution. Participants who completed the study were offered $5 in the form of an Amazon gift card.

6.2.3 Gen - Results. We recruited eight participants for the "Gen" phase (mean age: 25.6; four female; mean smartwatch experience years: 2).

The pressure-over-time graphs of the squeeze notifications that our participants generated are shown in Figure 4 (b). Each participant created their own pattern from scratch without being primed by any reference patterns. In analyzing participants in-process thoughts, we found that three factors influenced their designs: large pressure differentials, rapid pressure changes, and irregular patterning.

First, several participants found that large pressure differentials helped make the squeeze notification more obvious (P1, P5, P7). Intuitively, this makes sense: adjacent sensations with high contrast tend to be more noticeable (e.g., a hot shower on a cold day). In addition, several participants articulated that rapidly changing the amount of pressure that the wristband applied would be more noticeable. P1 and P2, for example, opted for short-period designs that repeated as frequently as possible. P3, P4, and P6 created front-loaded patterns with rapid pressure changes at the beginning to attract a potential user's attention quickly. Third, several participants designed irregular patterns to mitigate habituation. However, each participant integrated irregularity in their own way. For example, P1 added a slight variation towards the end of an otherwise regular pattern. P4 considered irregularity the most important design factor and integrated it throughout the entire pattern. P8 designed for irregularity by adjusting the pressures at each peak of an otherwise regular pattern.

Additionally, some participants drew external inspiration in designing their squeeze notifications. For instance, P1 tried to approximate an ambulance's siren sound. P6 designed a a slow ramp-up in pressure to imply a heightening risk the longer one delays action. We note that all participants chose to have their squeeze notifications repeat infinitely, rather than selecting a finite number of repetitions.

6.2.4 Rank - Finding the best squeeze notification. With eight distinct squeeze notifications in hand, we next needed to assess which of these would be best suited to the task of alerting people to urgent cybersecurity warnings. In the "Rank" study, we approached this problem through a series of pairwise comparisons. While it is easier for an individual to compare two haptic notifications at a time than eight, we also could not have each participant do every possible pairwise comparison — that would require each participant to make $\binom{8}{2} = 28$ comparisons. Instead, we arranged the eight squeeze notifications in random order and had participants make *seven* pairwise comparisons in a tournament. Below, we outline the high-level procedure in sequence.

Tournament Structure and Scoring: We randomly arranged our eight squeeze notifications into a tournament structure with three rounds. The random arrangement ensured that ordering effects did not influence the outcome. There were four brackets in the first round, two brackets in the second round, and one bracket in the third. Thus, each participant made only seven pairwise comparisons in total. Each bracket pit two of the squeeze notifications against each other in a pairwise competition, i.e., participants had to evaluate which of the two was more appropriate to our target scenarios as primed by the grounding videos. The winner of a bracket advanced to the next round, where it would be placed against the winner of a different bracket from the previous round. The tournament concluded after the third round.

Each participant experienced a randomly-bracketed tournament: i.e., the squeeze notifications that were pit against each in the first round were randomly selected without replacement from the eight squeeze notifications generated

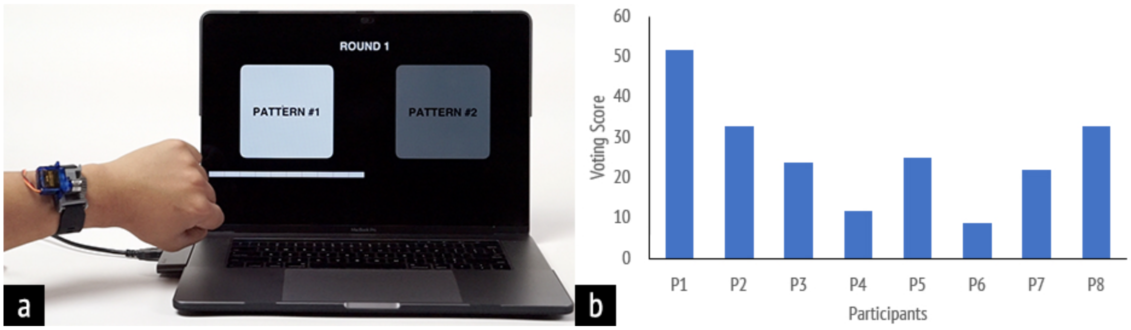

Fig. 5. (a) Tournament GUI used in "Rank"—participants were presented with a squeeze notification pair sequence; (b) A group of participants in "Rank" stage ranked the notifications created by the first group of participants in "Gen" stage from best to worst. We show the squeeze notification pattern voting result in the 'Rank' session

in the "Gen" study. Participants were not told that the pairwise comparisons they were making were ordered in a tournament structure.

We tallied the total "wins" each pattern had across all pairwise competitions, and ranked the notifications from best to worst.

Tournament GUI: We designed a GUI tool that led participants through the tournament bracket (Figure 5 (a)). The GUI presented participants with a sequence of squeeze notification pairs, each of which represented a bracket in the tournament. For each pair, a participant was shown a screen with two buttons. Each button represented one of the two haptic notifications in the bracket. Participants could click on a button to "experience" its assigned haptic effect. In each round, participants were asked to select which of the two notifications was better for the target scenario presented in the grounding video. We included a video figure to demonstrate the Tournament GUI in the supplementary materials.

To ensure that participants experienced both haptic effects in a bracket before selecting a winner, the system required participants to try both squeeze notifications before making a selection and imposed a 30-second delay before allowing participants to select a winner. The delay was meant to provide participants with enough time to think about which of the two squeeze notifications was better suited to the target scenario.

Tutorial Video: We prepared a short tutorial video explaining the tournament GUI. Participants first watched this video to familiarize themselves with the interface. This video is provided in the supplementary materials.

Grounding Video: We next showed participants the "grounding video." In contrast to the 'Gen' session, we showed participants only the Black Mirror grounding video. We made this decision partially because showing both grounding videos would take 10 minutes, which was too long for on-the-spot recruitment. Moreover, the Black Mirror grounding video was framed as a personal narrative, which better captured our target scenarios.

Tournament: After watching the grounding video, participants then used our Tournament GUI to make the pairwise comparisons in their randomly bracketed tournament. Their goal for each of the seven comparisons was to pick the squeeze notification that they believed would better capture their attention if they found themselves in a scenario similar to the one in the grounding video.

Recruitment and Compensation: We recruited participants in two ways: (1) through flyers promoting our study around our institution, and (2) by setting up a booth in a well trafficked building in our institution and canvassing. Participants recruited via canvassing were able to complete the study directly at the booth. Participants could enter a raffle to win 20 USD in the form of a Amazon gift card as compensation, though study participation was not mandatory to enter the raffle.

6.2.5 Rank - Result. We recruited 30 participants for the "Rank" study (mean age: 28.8; 13 female with one participant not providing gender; mean smartwatch experience years: 0.68).

We refer to each of the eight "Gen" squeeze notifications by the participant, in the "Gen" study, who created it. Figure 5 (b) shows the total scores. We awarded one point for each bracket win: i.e., 3 points to the top ranked pattern, 2 points to the first runner-up, and 1 point to the rest of semi-finalists. P1's squeeze design obtained the most points (51 points) by far, followed distantly by squeeze patterns designed by P2 and P8 (both gained 33 points). The winning pattern, P1's squeeze pattern, is an infinitely repeating pattern that has 4 large pressure peaks followed by rapid decreases in pressure, and a more irregular, smaller pressure peak at the end. In other words, the winning pattern pulsed intensely and quickly.

The results also show that squeeze patterns created by P2, P8, and P5, which are the tied for second and fourth, respectively, in overall rank were similar. P2's squeeze pattern had five repetitions of small pressure pulses. Both of the squeeze patterns designed by P5 and P8 consisted of four large pressure pulses. The less successful patterns had fewer than four repetitions. It seems, therefore, that squeezing patterns with rapid pressure pulses are better suited for alerting people to cybersecurity warnings.

6.2.6 Verify - Comparatively evaluating the best squeeze notification vs. a vibrotactile baseline. While the tournament setup helped us find the best squeezing pattern among the eight patterns with which we initially started, we next wanted to quantitatively evaluate if this "best" pattern was actually "good" compared to an established vibrotactile baseline. We define a "good" pattern as one that is deemed, by end-users, to be appropriate for our target scenarios.

Setup: We asked participants to evaluate two conditions: (i) a *control* condition in which participants experienced vibrotactile feedback; and, (ii) a *treatment* condition in which participants experienced the top-ranked squeeze notification from the "Rank" study. For the control condition, we used a Mobvoi Ticwatch E smartwatch that provided a vibrotactile notification (500-ms vibration). We sent the notification via Firebase Cloud Messaging. We only chose the vibrotactile as our baseline because Spidey Sense is designed as a secondary communication channel to communicate cybersecurity warnings to cover various situations including a case where a user may not engage with visual information from a screen (e.g., walking) and/or with auditory feedback (e.g., in a meeting). Thus, evaluating Spidey Sense only for specific cases where a user can engage with visual or auditory feedback where we could potentially choose visual and auditory notifications as the evaluation baselines for the evaluation is out of our scope in this paper. For the treatment condition, since the top-ranked squeeze notification is an infinitely repeating pattern, the repeating squeeze pattern was provided until a participant started to rate.

Grounding Video: Participants watched the Black Mirror "Shut Up and Dance" abridged episode that we spliced.

Evaluation - Spidey Sense and Vibrotactile Feedback: We ran a within-subjects experiment with two conditions. We had participants experience both the "best" Spidey Sense pattern from the "Verify" study and a vibrotactile notification, in counterbalanced order: i.e., five participants were randomly selected to evaluate Spidey Sense first, and the other five evaluated the Tic Watch's default vibrotactile notification first. Then, we had participants rate, on a Likert scale, the appropriateness of both the squeeze notification and the vibrotactile feedback for the scenario in the video (1 - really inappropriate, 3 - neutral, 5 - really appropriate).

Recruitment and Compensation: We advertised our study by placing flyers around our institution to recruit participants. We offered participants $5 USD in the form of an Amazon gift card if they completed the study.

Ratings	**1**	**2**	**3**	**4**	**5**
Vibration	4	4	1	1	0
Spidey Sense	0	1	0	6	3

Table 1. The number of ratings for smartwatch vibration and Spidey Sense's squeezing sensation for the evaluation of the appropriateness of each haptic pattern for security warning communication

	Coefficient	**p-value**	
Spidey Sense vs. Vibrotactile	2.20	0.001	**

Table 2. Coefficients for a Poisson regression model comparing the Spidey Sense and a commodity Smartwatch. The result of this formative study shows the potential that Spidey Sense can be an appropriate communication channel for urgent security warnings.

We hypothesized that Spidey Sense would be rated as more appropriate for alerting people to urgent security warnings than the vibrotactile baseline.

6.2.7 Verify - Result. We recruited 10 participants (mean age: 25.2, three female, mean smartwatch experience years: 0.4). We analyzed the results with a random-intercepts Poisson regression. The dependent variable was participants' Likert scale evaluation of "appropriateness", i.e., an integer value between 1 to 5. The independent variable was the condition that participants rated: *Spidey Sense* or *vibrotactile*. We used a Poisson regression because the Poisson distribution better approximates count data and integer-dependent variables [20]. We incorporated a random-intercepts term to account for repeated observations (each user provided two ratings, one for Spidey Sense and one for vibrotactile, in our within-subjects design).

The results of our analysis are shown in Table 2. Coefficients represent the estimated difference in how participants rated the "appropriateness" of Spidey Sense to our target scenario, compared to the vibrotactile baseline. The random intercept accounts for the fact that each individual might have a different baseline rating. The results show that participants found Spidey Sense to be significantly more appropriate for urgent cybersecurity warnings than the vibrotactile control ($b = 2.20$, $p < 0.001$). More specifically, the model estimates that Spidey Sense is, on average, rated 2.20 points higher on the 5-point appropriateness Likert scale.

7 DISCUSSION

We implemented and evaluated a novel haptic notification system, Spidey Sense, designed to be better suited to the task of alerting users of urgent cybersecurity warnings. Through our "Gen-Rank-Verify" study, we systematically explored the design space of squeeze notifications for cybersecurity warnings, and found empirical evidence confirming that participants found Spidey Sense to be more appropriate for alerting people to urgent cybersecurity warnings than a vibrotactile baseline.

Spidey Sense illustrates the potential of exploring wearable, affective haptics to improve people's awareness of urgent cybersecurity warnings. Advances in tangible interaction and haptic technologies have historically been considered impractical for applications in usable security but, with the increasing market penetration of on-body wearable devices, this assumption of impracticality may no longer hold. With Spidey Sense, we show that it should be possible to tap into the rich design space of tangible interaction and haptic notifications to create end-user cybersecurity systems that are more aligned with people's corporeal threat perception.

To date, there has been relatively limited exploration of affective haptics in the context of improving cybersecurity warnings. Our work can serve as a bridge to encourage future collaborations between usable security, affective haptics, and tangible computing researchers.

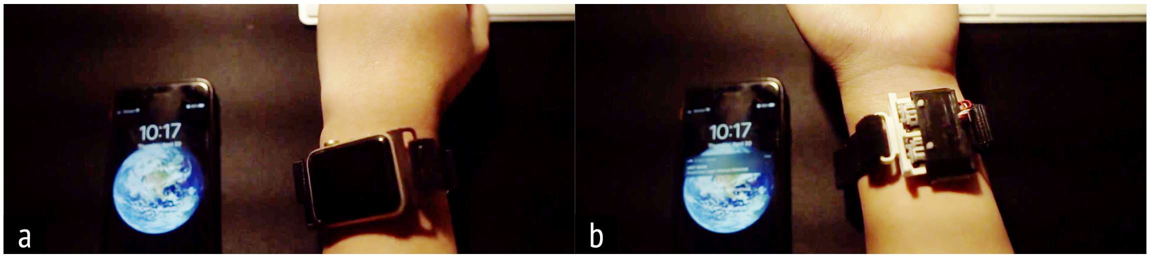

Fig. 6. Spidey Sense wristband can be integrated with an Apple Watch: (a) Spidey Sense wristband integrated with an Apple Watch 2 keeps still if there is no important notification; (b) The Spidey Sense wristband provides squeeze haptics when an important notification arrives.

7.1 Design Implications, Limitations, and Future Work

We will discuss the design implications, limitations, and considerations of our work that should shape future work.

7.1.1 Deployability. We took some liberties in creating the Spidey Sense proof-of-concept, but have refined our design to work with existing Apple Watches. Figure 6 shows a Spidey Sense prototype that we integrated with an Apple Watch Series 2. In our refined design, we use two tiny linear servo motors (GS-1502), which reduced the overall height of the actuation module to 9.8mm (Figure 1). Additionally, we integrated a customized board design based on nRF52832 (Bluetooth Low Energy and Arm Cortex-M4 CPU) that communicates with iOS and has a rechargeable lithium polymer (150mAh) battery, which allows Spidey Sense to be fully untethered and deployable. In the future, if there comes to exist a stronger thread actuator than ModiFiber, we could reduce the size even further [19].

Power consumption is another important consideration for deployability. Spidey Sense needs a stable power source to produce a squeeze effect. This could be solved by having one's smartwatch directly power the actuator, though the smartwatch battery would drain more quickly. In practice, however, the power requirements should be minimal. For example, the GS-1502 servo motor used for the Spidey Sense deployable module consumes idle current 32mA at a 4.2V supply, which is marginal relative to the battery capacity of modern smartwatches (e.g., Apple Watch Series 2 42mm's battery capacity is 334mAh). Moreover, integrating an additional tiny Li-Po battery could extend battery lifespan.

In cases where a separate device is impossible or impractical for users to procure, it may be possible to use the existing hardware in commodity smartwatches to produce a similar effect. For instance, Apple's expressive Taptic engine has been used to produce a granular vibrotactile guide that helps people take calming deep breaths through their "Breathe" app. It should be possible, then, to implement a vibrotactile equivalent of our best performing squeeze notification (where instead of squeezes, we deliver vibrations) using the Taptic engine. In future work, we intend to comparatively evaluate Spidey Sense vis-a-vis the Taptic engine in a lab evaluation.

7.1.2 Gen-Rank-Verify: A Methodology for Exploring and Evaluating Haptic Feedback for Cybersecurity Warnings. In the process of designing our evaluation methodology, we found a gap in the existing literature: there was little guidance on how to systematically explore a new haptic design space for cybersecurity warnings. Rather, we had to assemble together best practices from the broader HCI, tangible interaction and usable privacy and security literature. This synthesis is a minor contribution unto itself.

For our study, we followed the general "Find-Fix-Verify" study structure that Bernstein et al. [5] introduced and that Obrist et al. [40] utilized to explore similarly open-ended design spaces in word processing and haptic expression of emotions, respectively. We made three key adaptations that we believe make our adapted methodology more suited to the cybersecurity warning context.

First, cybersecurity threats are abstract, adding a layer of complexity to the task of getting lay-users to design affective notifications that are meant to capture high-arousal emotions [40]. To address this complexity, we showed participants emotionally visceral grounding videos.

Second, to rank the haptic designs that were generated in the "Gen" phase, we utilized a series of pairwise comparisons in a randomly-bracketed tournament. Obrist et al., in contrast, asked participants to provide a total rank order simultaneously [40]. A key benefit of our approach is that people can make more reliable assessments of haptic sensations when comparing one sensation vs. another. In contrast, it is much more difficult to comparatively evaluate more than two distinct haptic sensations simultaneously. A challenge with this approach, however, is scalability — every additional design introduced in the "Gen" phase requires many more pairwise comparisons.

Third, for the "Verify" phase, we ran a counterbalanced, controlled experiment. This adaptation affords a more rigorous evaluation of the novel haptic sensation against an established baseline. Prior approaches used the "Verify" phase to examine if the "Fix" phase results were good on an absolute scale. Our approach affords this evaluation, as well as if the new haptic sensation is "more appropriate" than existing approaches.

7.1.3 Ecological Validity, Habituation and Future Directions. Spidey Sense is an instantiation of a broader class of affective haptics that motivates end-user responses to security warnings, which can help catalyze a new direction of cybersecurity research. However, as with any study, ours had limitations that should be considered in contextualizing our findings.

A limitation of the present work is that we have yet to run studies to examine the system's usability and effectiveness "in the wild." Doing so will be our key focus moving forward.

Prior work suggests that informing participants that they are participating in a security-related study might affect their natural behavior responsive to security warnings [7, 8, 15, 29, 30]. Accordingly, in-lab studies are difficult to run without employing complicated deception schemes. Field studies, in contrast, can afford more ecologically valid data but at the cost of control. People may encounter just one urgent cybersecurity warning every month "in the wild." Moreover, detecting which cybersecurity warnings might be "urgent" is a challenge in and of itself.

Another key limitation of the present work is that we did not test Spidey Sense's resilience to habituation effects. Prior work suggests that people become habituated to cybersecurity warnings: as they are repeatedly exposed to the same warning, they gradually learn to ignore it [46, 47, 49]. As we did not test repeated exposures to Spidey Sense, it is presently unclear how its effect varies over time. Prior work provides an evidence that varying squeeze patterns of Spidey Sense may reduce *habituation* since a variety of squeeze patterns can be designed on Spidey Sense. While polymorphic warnings have been studied as a way to reduce habituation by repeatedly varying warning appearance [3], our intention with Spidey Sense design was to explore an entirely different channel of communication that is not overloaded by non-security related alerts. The two approaches, thus, should be able to complement one another.

Lastly, in our evaluation, we evaluated Spidey Sense only compared to a vibrotactile notification, considering that Spidey Sense is designed for a variety of cases including the cases that disallow users to engage with visual (e.g., walking) or auditory notifications (e.g., in a meeting). However, in practice, the available notification types could differ depending on context, and in the controlled setup, it would be challenging to replicate such various cases where we may need to choose a different set of baselines.

As a first step towards addressing these challenges, we are working with industry partners to explore a field deployment of Spidey Sense. The industry partners, with their more comprehensive data, will handle the "urgent warning detection" problem. We will handle the "urgent warning delivery" problem with Spidey Sense. We hope to evaluate the Spidey Sense's usability and effectiveness in practice via this collaboration.

8 CONCLUSION

We introduced and evaluated Spidey Sense, a smartwatch wristband that produces programmable squeeze notifications to alert users to urgent cybersecurity warnings. We made two key contributions with this work. First, we designed and implemented a functional smartwatch wristband that can produce a novel, wrist-squeeze haptic sensation. Second, we utilized a three-part "Gen-Rank-Verify" study methodology to find an optimal wrist-squeeze notification that users deemed to be more appropriate for our target scenarios than a vibrotactile baseline. Our work provides initial evidence that Spidey Sense's squeeze haptics are more appropriate for urgent cybersecurity warnings than a standard vibrotactile baseline. These contributions, in turn, should serve as a stepping stone towards a fruitful interdisciplinary research agenda at the intersection of tangible computing, affective haptics and usable security.

ACKNOWLEDGMENTS

This research was generously supported, in part, by the National Science Foundation through grant SaTC-2029519. We thank Dr. Frank L. Hammond III, Dr. Omer T. Inan, Dr. Sang-won Leigh, Jennifer Molnar, Joshua Lee, and Samer Mabrouk for their constructive discussions for the early prototype. We are also grateful to Mehrab Bin Morshed, Vedant Das Swain, Sindhu Kiranmai Ernala, Matthew Hong, Karthik S. Baht, Jacob Logas, Hayeong Song, and members of GT Ubicomp lab and GT SPUD lab for their feedback to help improve our study design. Lastly, we would like to thank the reviewers for their valuable reviews.

REFERENCES

[1] Anne Adams and Martina Angela Sasse. 1999. Users are not the enemy. *Commun. ACM* 42, 12 (1999), 40–46.

[2] Devdatta Akhawe and Adrienne Porter Felt. 2013. Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness. In *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*. USENIX, Washington, D.C., 257–272. https://www.usenix.org/conference/usenixsecurity13/technical-sessions/presentation/akhawe

[3] Bonnie Brinton Anderson, C Brock Kirwan, Jeffrey L Jenkins, David Eargle, Seth Howard, and Anthony Vance. 2015. How polymorphic warnings reduce habituation in the brain: Insights from an fMRI study. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2883–2892.

[4] Matthew A Baumann, Karon E MacLean, Thomas W Hazelton, and Ashley McKay. 2010. Emulating human attention-getting practices with wearable haptics. In *2010 IEEE Haptics Symposium*. IEEE, 149–156.

[5] Michael S Bernstein, Greg Little, Robert C Miller, Björn Hartmann, Mark S Ackerman, David R Karger, David Crowell, and Katrina Panovich. 2010. Soylent: a word processor with a crowd inside. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. ACM, 313–322.

[6] Leonardo Bonanni and Cati Vaucelle. 2006. A framework for haptic psycho-therapy. *depression and anxiety* 12 (2006), 24.

[7] Cristian Bravo-Lillo, Lorrie Cranor, Julie Downs, Saranga Komanduri, Stuart Schechter, and Manya Sleeper. 2012. Operating system framed in case of mistaken identity: measuring the success of web-based spoofing attacks on OS password-entry dialogs. In *Proceedings of the 2012 ACM conference on Computer and communications security*. ACM, 365–377.

[8] Cristian Bravo-Lillo, Saranga Komanduri, Lorrie Faith Cranor, Robert W. Reeder, Manya Sleeper, Julie Downs, and Stuart Schechter. 2013. Your Attention Please: Designing Security-decision UIs to Make Genuine Risks Harder to Ignore. In *Proceedings of the Ninth Symposium on Usable Privacy and Security* (Newcastle, United Kingdom) *(SOUPS '13)*. ACM, New York, NY, USA, Article 6, 12 pages. https://doi.org/10.1145/2501604.2501610

[9] Jeremy D Brown, Joshua N Fernandez, Sean P Cohen, and Katherine J Kuchenbecker. 2017. A wrist-squeezing force-feedback system for robotic surgery training. In *2017 IEEE World Haptics Conference (WHC)*. IEEE, 107–112.

[10] Angela Chang, Ben Resner, Brad Koerner, XingChen Wang, and Hiroshi Ishii. 2001. LumiTouch: an emotional communication device. In *CHI'01 extended abstracts on Human factors in computing systems*. 313–314.

[11] Gwénaël Changeon, Delphine Graeff, Margarita Anastassova, and José Lozada. 2012. Tactile emotions: A vibrotactile tactile gamepad for transmitting emotional messages to children with autism. In *International conference on human haptic sensing and touch enabled computer applications*. Springer, 79–90.

[12] Guillaume Dezecache, Julie Grèzes, and Christoph D. Dahl. 2017. The nature and distribution of affiliative behaviour during exposure to mild threat. *Royal Society Open Science* 4, 8 (2017), 170265. https://doi.org/10.1098/rsos.170265 arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.170265

[13] Paul Dourish, E Grinter, Jessica Delgado De La Flor, and Melissa Joseph. 2004. Security in the wild: user strategies for managing security as an everyday, practical problem. *Personal and Ubiquitous Computing* 8, 6 (2004), 391–401.

[14] Serge Egelman, Lorrie Faith Cranor, and Jason Hong. 2008. You'Ve Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. ACM, New York, NY, USA, 1065–1074. https://doi.org/10.1145/1357054.1357219

[15] Serge Egelman, Janice Y. Tsai, and Lorrie F. Cranor. 2010. Tell Me Lies: A Methodology for Scientifically Rigorous Security User Studies. In *Proceedings of the Workshop on Studying Online Behavior* (Atlanta, Georgia, USA) *(CHI '10 Workshop)*. ACM, New York, NY, USA, 4 pages.

[16] Mohamad A Eid and Hussein Al Osman. 2015. Affective haptics: Current research and future directions. *IEEE Access* 4 (2015), 26–40.

[17] Seyedeh Maryam Fakhrhosseini and Myounghoon Jeon. 2017. Affect/emotion induction methods. In *Emotions and Affect in Human Factors and Human-Computer Interaction*. Elsevier, 235–253.

[18] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. 2012. Android Permissions: User Attention, Comprehension, and Behavior. In *Proceedings of the Eighth Symposium on Usable Privacy and Security* (Washington, D.C.) *(SOUPS '12)*. ACM, New York, NY, USA, Article 3, 14 pages. https://doi.org/10.1145/2335356.2335360

[19] Jack Forman, Taylor Tabb, Youngwook Do, Meng-Han Yeh, Adrian Galvin, and Lining Yao. 2019. ModiFiber: Two-Way Morphing Soft Thread Actuators for Tangible Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 660.

[20] William Gardner, Edward P Mulvey, and Esther C Shaw. 1995. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological bulletin* 118, 3 (1995), 392.

[21] Matthew J Hertenstein, Rachel Holmes, Margaret McCullough, and Dacher Keltner. 2009. The communication of emotion via touch. *Emotion* 9, 4 (2009), 566.

[22] Matthew J Hertenstein, Dacher Keltner, Betsy App, Brittany A Bulleit, and Ariane R Jaskolka. 2006. Touch communicates distinct emotions. *Emotion* 6, 3 (2006), 528.

[23] J. Hong. 2017. The Privacy Landscape of Pervasive Computing. *IEEE Pervasive Computing* 16, 3 (2017), 40–48. https://doi.org/10.1109/MPRV.2017.2940957

[24] Da-Yuan Huang, Ruizhen Guo, Jun Gong, Jingxian Wang, John Graham, De-Nian Yang, and Xing-Dong Yang. 2017. RetroShape: Leveraging rear-surface shape displays for 2.5 D interaction on smartwatches. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 539–551.

[25] Gijs Huisman. 2017. *Social touch technology: extending the reach of social touch through haptic technology*. Ph.D. Dissertation. University of Twente. https://doi.org/10.3990/1.9789036543095

[26] Alexandra Ion, Edward Jay Wang, and Patrick Baudisch. 2015. Skin Drag Displays: Dragging a Physical Tactor Across the User's Skin Produces a Stronger Tactile Stimulus Than Vibrotactile. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 2501–2504. https://doi.org/10.1145/2702123.2702459

[27] Kinga Jurásová and Marián Špajdel. 2013. Development and assessment of film excerpts used for emotion elicitation. *Activitas Nervosa Superior Rediviva* 55 (01 2013), 135–140.

[28] Irene Kotsia, Stefanos Zafeiriou, and Spiros Fotopoulos. 2013. Affective gaming: A comprehensive survey. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 663–670.

[29] K. Krol, M. Moroz, and M. A. Sasse. 2012. Don't work. Can't work? Why it's time to rethink security warnings. In *2012 7th International Conference on Risks and Security of Internet and Systems (CRiSIS)*. 1–8. https://doi.org/10.1109/CRISIS.2012.6378951

[30] Kat Krol, Jonathan M. Spring, Simon Parkin, and M. Angela Sasse. 2016. Towards Robust Experimental Design for User Studies in Security and Privacy. In *The LASER Workshop: Learning from Authoritative Security Experiment Results (LASER 2016)*. USENIX Association, San Jose, CA, 21–31. https://www.usenix.org/conference/laser2016/program/presentation/krol

[31] B. W. Lampson. 2004. Computer security in the real world. *Computer* 37, 6 (June 2004), 37–46. https://doi.org/10.1109/MC.2004.17

[32] Jaeyeon Lee and Geehyuk Lee. 2016. Designing a Non-contact Wearable Tactile Display Using Airflows. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. ACM, New York, NY, USA, 183–194. https://doi.org/10.1145/2984511.2984583

[33] Sang-won Leigh and Pattie Maes. 2016. Body Integrated Programmable Joints Interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 6053–6057. https://doi.org/10.1145/2858036.2858538

[34] Sang-won Leigh, Kush Parekh, Timothy Denton, William S. Peebles, Magnus H. Johnson, and Pattie Maes. 2017. Morphology Extension Kit: A Modular Robotic Platform for Customizable and Physically Capable Wearables. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. ACM, New York, NY, USA, 397–400. https://doi.org/10.1145/3027063.3052969

[35] Nicholas Micallef, Mike Just, Lynne Baillie, and Maher Alharby. 2017. Stop Annoying Me!: An Empirical Investigation of the Usability of App Privacy Notifications. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction* (Brisbane, Queensland, Australia) *(OZCHI '17)*.

ACM, New York, NY, USA, 371–375. https://doi.org/10.1145/3152771.3156139

[36] Tyler Moore. 2010. The economics of cybersecurity: Principles and policy options. *International Journal of Critical Infrastructure Protection* 3, 3 (2010), 103 – 117. https://doi.org/10.1016/j.ijcip.2010.10.002

[37] Steven Musil. 2020. Apple still dominates growing global smartwatch sector. https://www.cnet.com/news/apple-still-dominates-growing-global-smartwatch-sector/

[38] Daniela Napoli, Sebastian Navas Chaparro, Sonia Chiasson, and Elizabeth Stobert. [n.d.]. Something Doesn't Feel Right: Using Thermal Warnings to Improve User Security Awareness. ([n. d.]).

[39] Chris Nodder. 2005. Users and trust: A microsoft case study. *Security and Usability* (2005), 589–606.

[40] Marianna Obrist, Sriram Subramanian, Elia Gatti, Benjamin Long, and Thomas Carter. 2015. Emotions Mediated Through Mid-Air Haptics. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 2053–2062. https://doi.org/10.1145/2702123.2702361

[41] Evan Pezent, Ali Israr, Majed Samad, Shea Robinson, Priyanshu Agarwal, Hrvoje Benko, and Nick Colonnese. 2019. Tasbi: Multisensory squeeze and vibrotactile wrist haptics for augmented and virtual reality. In *2019 IEEE World Haptics Conference (WHC)*. IEEE, 1–6.

[42] Henning Pohl and Kasper Hornbæk. 2018. ElectricItch: Skin Irritation As a Feedback Modality. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. ACM, New York, NY, USA, 765–778. https://doi.org/10.1145/3242587.3242647

[43] Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. 2014. Large-scale assessment of mobile notifications. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 3055–3064.

[44] Bahador Saket, Chrisnawan Prasojo, Yongfeng Huang, and Shengdong Zhao. 2013. Designing an Effective Vibration-based Notification Interface for Mobile Phones. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. ACM, New York, NY, USA, 149–1504. https://doi.org/10.1145/2441776.2441946

[45] Florian Schaub, Rebecca Balebako, Adam L. Durity, and Lorrie Faith Cranor. 2015. A Design Space for Effective Privacy Notices. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. USENIX Association, Ottawa, 1–17. https://www.usenix.org/conference/soups2015/proceedings/presentation/schaub

[46] Stuart E Schechter, Rachna Dhamija, Andy Ozment, and Ian Fischer. 2007. The emperor's new security indicators. In *2007 IEEE Symposium on Security and Privacy (SP'07)*. IEEE, 51–65.

[47] David Sharek, Cameron Swofford, and Michael Wogalter. 2008. Failure to recognize fake internet popup warning messages. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 52. SAGE Publications Sage CA: Los Angeles, CA, 557–560.

[48] Sunghyun Song, Geeyoung Noh, Junwoo Yoo, Ian Oakley, Jundong Cho, and Andrea Bianchi. 2015. Hot & tight: exploring thermo and squeeze cues recognition on wrist wearables. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*. 39–42.

[49] Joshua Sunshine, Serge Egelman, Hazim Almuhimedi, Neha Atri, and Lorrie Faith Cranor. 2009. Crying Wolf: An Empirical Study of SSL Warning Effectiveness.. In *USENIX security symposium*. Montreal, Canada, 399–416.

[50] Dzmitry Tsetserukou, Alena Neviarouskaya, Helmut Prendinger, Naoki Kawakami, and Susumu Tachi. 2009. Affective haptics in emotional communication. In *2009 3rd international conference on affective computing and intelligent interaction and workshops*. IEEE, 1–6.

[51] Anthony Vance, David Eargle, Jeffrey L Jenkins, C Brock Kirwan, and Bonnie Brinton Anderson. 2019. The fog of warnings: how non-essential notifications blur with security warnings. In *Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019)*.

[52] Wikipedia. 2021. Public key certificate — Wikipedia, The Free Encyclopedia. http://en.wikipedia.org/w/index.php?title=Public%20key%20certificate&oldid=1019769813. [Online; accessed 26-April-2021].

[53] Graham Wilson, Harry Maxwell, and Mike Just. 2017. Everything's Cool: Extending Security Warnings with Thermal Feedback. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. ACM, New York, NY, USA, 2232–2239. https://doi.org/10.1145/3027063.3053127

[54] Anusha Withana, Daniel Groeger, and Jürgen Steimle. 2018. Tacttoo: A Thin and Feel-Through Tattoo for On-Skin Tactile Output. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. ACM, New York, NY, USA, 365–378. https://doi.org/10.1145/3242587.3242645