Transformational Provocations for Usable Privacy and Security

Designing Beyond Compliance and Expert Norms

Yuxi Wu
Northeastern University
Boston, Massachusetts, USA
yux.wu@northeastern.edu

Alexandra To
Northeastern University
Boston, Massachusetts, USA
a.to@northeastern.edu

Emilee Rader
University of Wisconsin-Madison
Madison, Wisconsin, USA
ejrader2@wisc.edu

Keith Edwards
Georgia Institute of Technology
Atlanta, Georgia, USA
keith@cc.gatech.edu

Sauvik Das
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
sauvik@cmu.edu

Abstract

A long standing goal of usable privacy and security (UPS) research is to align end-user behaviors with expert suggestions, such as through developing tools to increase user control and awareness of security and privacy (S&P) threats. These approaches, while necessary, are not sufficient to combat the ubiquitous slow violence of S&P harms that impede on people's motivation to change and erode their trust in the institutions in which experts are often entrenched. We propose that we can only effect true change and rebuild trust with users if we first deconstruct our norms of telling users what's best and expecting them to comply. We draw from concepts in critical computing, design, and games research, to propose a research agenda—"Transformational Provocation"—that involves provoking transformation in not just users' S&P knowledge, skills, or behaviors, but also their senses of truth, self, relationships, and society. Moreover, we demonstrate the value and application of our framework through illustrative case studies. With this framework, we call upon the UPS community to pursue new design opportunities for engendering organic and enduring intrinsic motivation for people to act on their S&P, both on their own and together.

CCS Concepts

• **Security and privacy → Human and societal aspects of security and privacy**; • **Human-centered computing → HCI theory, concepts and models**.

Keywords

transformational provocation, usable privacy and security, reflective play, provocative design, critical consciousness

ACM Reference Format:
Yuxi Wu, Alexandra To, Emilee Rader, Keith Edwards, and Sauvik Das. 2025. Transformational Provocations for Usable Privacy and Security: Designing Beyond Compliance and Expert Norms. In *New Security Paradigms Workshop (NSPW '25), August 24–27, 2025, Aerzen, Germany*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3774761.3774765

NSPW '25, Aerzen, Germany

ACM ISBN 979-8-4007-1875-5/25/08
https://doi.org/10.1145/3774761.3774765

1 Introduction

People feel unmotivated to engage in expert-recommended security and privacy (S&P) suggestions. Canonically, the S&P community has explained this phenomenon by characterizing end-users as "the weakest link"—i.e., lazy, reckless, and/or uneducated [2, 62, 130]. Exacerbating this adversarial observation is the so-called "privacy paradox" [21, 95, 115], or the idea that while users might *say* they're concerned about privacy, they don't *act* accordingly. As a result of this characterization, much of usable privacy and security (UPS) research has been oriented toward encouraging users to behave more in line with best practices recommended by S&P experts [1, 35]. Underlying this objective is the assumption that what experts recommend is the gold standard for user behavior [78].

Yet, experts are typically entrenched in the same large institutions that create and reinforce the environment of privacy and security harms that users are forced to navigate [136]. In this technofeudalist world [44, 122], we are employed by, receive research funding from, and make our careers in institutions that perpetuate digital slow violence—through ubiquitous surveillance and algorithmic marginalization—and erode people's senses of autonomy and self [135]. By definition, then, expert incentives are poorly aligned with user motivations, interests, and concerns [ibid]. In a period where the public's mistrust in institutions is deepening more than ever, [38], and users' tendencies for fatalist "hyperbolic scaling" about S&P concerns make the *death* of S&P control seem inevitable [112], it is unlikely users would view learning expert-recommended S&P behaviors and skills as a top priority. In other words, despite our best intentions, experts cannot alone dismantle structures of systemic inequity that lead to widespread S&P concern and resignation [8, 54, 139].

We propose that we can only effect true change and rebuild trust with users if we first deconstruct our norms of telling users what's best and expecting them to comply. The UPS community needs to center a more radical research agenda of *Transformational Provocation*. Transformation is a concept borrowed from the game design research field: "transformational play" refers to games that encourage people to reflect, act, and change beyond learning new things [11, 32]. Simultaneously, provocativeness—a key concept from critical computing and design literature—can help us intentionally cause these changes. The field of game design offers several valuable lessons for the UPS community to pursue Transformational Provocation. Games themselves contain intentional

"mechanics", or rules and elements of gameplay, for encouraging critical reflection and change. Game design research also offers both conceptual anchors for how to think about transformation and provocation, and process-oriented guides for design practice.

These mechanics and guides inform the development of our own vision of Transformational Provocation in UPS, which is governed by four prescriptive standards. **First**, to address the motivation barrier that inhibits end-user S&P behaviors [35], we must go beyond educational and awareness interventions and also create change in users' senses of self, relationships, truth, and society as they relate to S&P. **Second**, to make these changes persistent, we should look toward creating designs that are provocative and trigger personal dilemmas about S&P. **Third**, to be provocative includes challenging community norms centered around user protection, and augmenting them with notions of S&P promotion. And **fourth**, to sustain this promotion and grow more proactive as a community, we should cultivate conditions and systems that are ready to help people transform their S&P. By examining prior work in UPS and synthesizing related literature in critical computing, design, and games research, we also provide paths for implementing these standards in future work. With these standards in mind, we can re-orient ourselves toward better addressing the user motivation challenge.

To execute this vision, we present a toolkit for designing ways to provoke S&P transformation in users. Drawing from provocative design [13, 98], designing for flourishing [120], transformational game design [32], reflective play [71, 89], and critical consciousness [49], we propose three dimensional heuristics for UPS researchers to consider when provoking transformation in users:

Velocity The speed at which a user is expected to transform after being exposed to a provocation.

Does the provocation shock and disrupt the user, or does it poke and prod at the user over time?

Interiority/Exteriority The degree to which a provocation might induce transformation within an individual user versus a broader community of actors.

Could the provocation lead to a user seeking external social support, or does it generate further introspection?

Contextual Potential The levels of critical reflection and action that we expect a provocative design to spark in people.

On a scale from stoking curiosity about an app's data use to catalyzing incitement against a privacy-violating institution, what is the scope of the provocation's impact?

We incorporate these heuristics into a design process of questions and methods for researchers to envision specific user transformations and construct appropriate evaluations. To illustrate how our toolkit can be practically considered by researchers and designers, we also present two fictional scenarios of worked examples.

In summary, we offer **a provocation for the UPS research community to, in turn, provoke transformation in users.** Our framework encourages the UPS community to consider alternative methods of design and user participation in S&P beyond individually protective behaviors; emphasizes the strength of social ties in how people make decisions and form opinions about S&P; and examines the broader environmental and societal conditions that influence how effective researchers' interventions can be. Through engaging with Transformational Provocation, we hope UPS researchers and designers can better align themselves with user interests and motivations and create enduring, organic change.

2 Background

We first provide a brief overview of the concepts of transformation and reflection as they appear in the field of game design. Then, we motivate why provocation is an appropriate mechanism for inducing transformation. Finally, we describe how these concepts can inform exciting new priorities in the UPS community.

2.1 Transformation and Provocation

Not unlike the digital slow violence [50, 94, 135] that people experience from S&P intrusions associated with mass surveillance and algorithmic stigmatization [5], the virtual harms that marginalized people experience in games and gaming culture bind to their "real life" experiences [56]. Beyond these parallels, however, the field of game design also offers many relevant valuable insights for UPS. Games themselves contain a wealth of intentional interaction patterns and rules, or "mechanics", for encouraging players to reflect, act, and change. Simultaneously, past games *literature* offers both conceptual anchors for thinking about player transformation and process-oriented guides for design in practice.

Barab et al. [11] first theorized *transformational play* in 2010, arguing that the "opportunity to have a personal, agentic, and consequential role in resolving a dilemma is a significant component of both content learning and potentially more enduring outcomes." Specifically, they propose that transformational play positions the elements of *person, content, and context* in three ways: person with intentionality, content with legitimacy, and context with consequentiality. In other words, transformational play involves a person—taking on the role of a character—developing and applying academic understandings to transform both the virtual world of the character and the person's own real world.

Upon this foundation, game designer Sabrina Culyba introduced the Transformational Framework, a set of eight exploratory questions to guide game developers in creating transformational games (i.e., games that transform the player on purpose) [32]. Within Question #3—*"How should players be different after playing your game?"*—Culyba posits that "learning" insufficiently encapsulates the changes that people undergo during transformational play. Instead, she introduces eight types beyond knowledge and skill where transformation might take place during play: physical, disposition, experience, behavior, belief, relationships, identity, and society. The UPS field has focused historically on changing user S&P knowledge, skill, and behavior to better match S&P experts' conception of best practices, and less so on Culyba's transformation types, as we explain in the following subsection.

One key component of transformation is *critical reflection.* As games researcher Rilla Khaled describes in her 2018 agenda for *reflective game design* [71], "reflection involves revisiting and reassessing previous beliefs intentionally, consciously and carefully."

Doing this effectively in games requires mechanics that run counter to the design of many mainstream entertainment games, which prioritize providing clear cut solutions, disguising learning within the game, and making people comfortable and fully immersed in the game environment. Instead, Khaled argues, reflection requires triggering players to ask questions that might not have clear answers, so that they can have ownership over their view of the world; explicitly asking players to engage and reflect, so that game-derived transformations can be integrated with real-life knowledge; and disrupt players, inviting them to actively interpret their experiences and remember them.

While critical reflection is a necessary ingredient of transformation, it also requires a willingness to suspend personal judgment, which may be difficult for an individual user or researcher. Reflection is triggered when we are uncomfortable; it occurs when we encounter situations that our prior experiences and solutions cannot effectively address [71]. To intentionally induce people to reflect—and consequently, transform—we need to foreground unexpected or uncomfortable beliefs and experiences. In other words, we need to *provoke*. As Bardzell et al. [13], quoting Dunne and Raby [43], wrote, it is fairly easy to understand provocativeness conceptually: something that is not so novel as to be dismissed as weird, but not too mundane to be ignored.

Embodied symbols can be one modality for provocation [98]. Also known as provocative prototypes—i.e., "provotypes"—embodied symbols involve choosing or creating objects that represent user concerns, and ideating about how to modify those objects in a way that embodies those concerns, especially if those concerns may cause tensions between different stakeholders [19]. Recent work in human-computer interaction has created provotypes to help BIPOC (Black, Indigenous, People of Color) communities coping with racist interactions, including a "Racism Alarm" that *"is constantly listening for racist speech and sounds an alarm when racism is detected"* [119]. While the Racism Alarm was intended to be a neutral agent that could speak out on behalf of targets of racist behaviors, the authors found that design workshop participants strongly disliked how disruptive, sensitive, and socially awkward it was. As another provotype example, Bruun et al. [22] introduced Pup-Lock, an animated puppy that appears progressively sadder as more people in a family household choose to unlock their mobile devices; in this case, the authors wanted to provoke families to think about engaging in more meaningful interactions with each other instead of using their mobile devices.

However, without reliable heuristics for evaluating provocativeness, it can be quite difficult to create designs that are actually provocative for users. In particular, as we argue in the following section, UPS researchers may find it difficult to suspend normative community beliefs about what is "best" for the user, or what the "most" secure and private behaviors are.

2.2 Situating Transformation and Provocation in Usable Privacy and Security

Within usable security and privacy, transformational games have gradually been used to address the user motivation challenge. For example, Alotaibi et al. introduced two mobile games for helping people become aware of malware attacks and the importance of using strong passwords [4], and Dabrowski et al. explored competition and gamification in a security course at a university [34]. Other works have aimed to improve user *self-efficacy*—an individual's belief that they have the ability to achieve their goals and complete tasks successfully [10]—in adopting S&P behaviors [11, 27]. Most recently, Krsek et al. [76] developed C.A.L.Y.P.S.O., a transformational role-playing game (RPG) that leads the user through a series of S&P choices with demonstrated in-game consequences on not just the player, but other characters in the game. Key to gameplay is the creation of a "brave space", where players can explore the consequences of their cybersecurity decisions without ever putting themselves in real danger.

However, even outside of fully-featured game design, transformation and provocation form an exciting lens for envisioning a future agenda in the UPS community. One persistent problem in UPS has been the mismatch between what S&P experts believe to be best practices for protecting people's S&P, and what users actually do. The UPS field, after all, originally stemmed from a curiosity from S&P researchers about why users make errors when using S&P tools, characterizing users in early years as the "weakest link" [131]. Whereas S&P experts might recommend that users prioritize installing software updates, using two-factor authentication, and using a password manager to protect their S&P, users tend to prefer using antivirus software, visiting only known websites, and changing passwords frequently [24, 66]. Simultaneously, through their lived experiences with S&P problems, users can bring unique expertise and capabilities that experts may lack [125].

These differences can lead to an awkward tension between experts and non-experts. For example, Poole et al. [103] found that tech experts who may be initially enthusiastic about helping non-experts with their computer problems can quickly become bored of doing so; also, to maintain an aura of expertise, they become uncomfortable when they cannot solve non-experts' problems. And, as Adams and Sasse wrote in their now-canonical 1999 work, "Users Are Not the Enemy" [2], security experts within organizations tend to characterize non-expert users as, "*at best...a security risk that needs to be controlled and managed, at worst...the enemy within.*" But this attitude is not simply a relic of the past: work from as recently as 2024 found that S&P experts felt that security measures should simply be "pushed on the workforce" if the experts felt it was right to do so [78]. However, these experts are often deeply entrenched in the institutions that inflict privacy harms upon people en masse, and ultimately may be disincentivized to advocate for non-reformist changes [54] as their expert status is predicated on their deep knowledge of existing institutions and systems.

The adversarial relationship between experts and users is exacerbated by the so-called "privacy paradox". As Barry Brown [21] first noted in 2001, while users express concerns over their privacy, they seem to take very little action to protect their privacy. Myriad studies [14, 73] have uncovered cases where users have stated they were very interested in protecting their personal data, but were also not willing to pay small amounts of money to do so (or, vice versa, were willing to give up their personal information for small discounts). Historically, this paradox has been explained one of two ways: either users truly do not value privacy as much as they say they do, or their privacy decisions are distorted by information asymmetry and thus do not correctly represent their privacy values. However,

providing better user controls or promoting individual privacy self-management, for which UPS work has historically striven, is an illusory solution to broader issues of mass institutional surveillance. As Solove [115] argues, privacy itself has intrinsic value beyond whether or not people choose to trade their personal data.

Despite these persistent themes, it is infeasible to teach millions of non-expert users the "right" way to think about complex legal processes or algorithmic delivery systems, so that a grassroots movement may collectively rise up and critique surveillance in formats that experts find workable. This is due to two key reasons: (1) expert concerns are poorly aligned with user interests and incentives, and (2) experts alone cannot dismantle institutional systems of privacy harm. **Now, more than ever, as public mistrust in academic and scientific institutions escalates [38], it is imperative for the UPS community to recalibrate its goals and motivations for creating user interventions and communicating with users.** As we argue in the following section, researchers and users can effect true change together if and only if we, as researchers, challenge our own norms of telling users what's best and expecting them to comply. Through our own reflection, we can more effectively provoke transformation in *users*.

3 Prescriptive Standards for Transformational Provocation in UPS Research

In this section, we provide a roadmap for adopting a Transformational Provocation approach to usable security and privacy (hereon referred to as TP-UPS). First, in Section 3.1, we outline expectations for what TP-UPS is and is not. Then, in the remaining subsections, we discuss in depth the four prescriptive standards of TP-UPS. For each standard, we define its notable properties and overall scope. We then provide suggestions for implementation of the standard in future work, drawn from additional examples in game design, pedagogy, geography, law, and of course UPS itself.

3.1 What Is TP-UPS, and What Is It *Not*?

What does it mean to take a TP-UPS approach to research? Briefly, TP-UPS should involve adhering to the following four standards, which we discuss in further detail beyond this subsection: (1) going beyond S&P knowledge and skill, (2) triggering personal S&P dilemmas, (3) enhancing user protection with S&P promotion, and (4) cultivating a readiness to transform.

However, adopting these standards does not mean designing more interventions or overloading users with new information and choices; TP-UPS is not a solutionist [33] call for simply triggering personal dilemmas in every interaction we have with users. Nor does TP-UPS strive solely for targeted behavioral change in users—be it through paternalistic nudges à la Thaler and Sunstein [118], or via the COM-B model of behavior introduced by Michie et al. [88] and implemented by Sasse et al. [111] within UPS. In other words, TP-UPS does not aim to guide users into doing what experts actually want through creative learning principles or define criteria for what kinds of responses and behaviors are "right".

At the same time, our standards should not be used to evaluate or deflate the value of prior work, but used as aspirations for *future* work: TP-UPS means creating opportunities for pro-social alignment or solidarity between experts and users. We pick up a thread of arguments from Herley [63] and continued by Coles-Kemp et al. [30] and Das Chowdhury et al. [36]: that UPS researchers ought to critically and continually examine their prevailing notions of what S&P is and how best to achieve it. We should allow ourselves, as researchers and designers, to sit with discomfort and uncertainty in user responses and behaviors. **Through the lens of TP-UPS, we can invite this critique and reflection from users, too, enriching relationships between the two parties.**

3.2 Going Beyond S&P Knowledge and Skill

3.2.1 Standard Overview. Cybersecurity user interventions, be they in games or otherwise, are typically grounded in learning theories, which ends up providing people with the knowledge and awareness to potentially protect themselves. However, these interventions don't necessarily incorporate behavioral change theories, which can nudge people to "act upon this knowledge and protect themselves" [27]. On this basis, we offer our first standard:

> ***Standard #1***
>
> TP-UPS research should aim for change that purposefully goes beyond increasing S&P knowledge and skill in users, and aims toward changes in users' dispositions, experiences, beliefs, relationships, and identities.

Past work has found that Americans overwhelmingly feel powerless over how the government and companies handle their personal data. Even as people become more knowledgeable about methods to protect their data, they also grow more skeptical that the methods will work; simultaneously, users being more concerned about S&P can go hand-in-hand with feeling more overwhelmed by the actions they must take to protect it [86]. As Seberger et al. [113] argue, these protection-oriented "solutions" may simplify S&P problems into bite-size, solvable pieces, but they fail to address the larger problem of the normalization of affective discomfort. Echoing Herley [62] on why users choose not to take certain security advice, we contend the same for privacy: privacy dashboards, private browsing, and cookie blocking will not fully assuage people's privacy concerns, since they only address issues at the interaction level [113].

Education, awareness, and behavior changes don't fully address the core issues of people's aversion to S&P; it is unlikely that better educating users about data use and S&P will lead to people feeling more positively or confidently about their own security and privacy. *What does it matter that people are acting more securely and privately if they don't believe in what they are doing, or if they don't even understand why they are making those choices? From a self-efficacy lens, how can the UPS community help people feel that they have power and control over S&P?*

3.2.2 Implementations. Users are often confronted with a barrage of privacy invasions and an overwhelming set of S&P decisions that they have to make, without breathing room to think about how they feel about those decisions. **UPS researchers should design and support spaces or opportunities for users to slow down and engage in mindfulness and more deliberate reflection** [57], as a direct counter to this constant barrage.

Game designers have created specific mechanics for players to slow down, purposefully giving players spaces to think about their actions and reflect on the journeys they and their characters have. One such mechanic type is "lingering defeat", or giving people the opportunity to return to challenges in a game that they have lost previously and assess (and reassess) how they might make their choices differently the next time [89]; similarly, "kill-cams" can also encourage players to revisit their experiences. Relatedly, giving players a quiet space to calm down and silently reflect on the environment around them—e.g., a picturesque landscape or a longer cutscene transition—can allow them to mentally recharge.

Slowing down can also have significant broader impacts beyond the immediate benefits of silence and calm. One crucial reason people may feel so powerless and resigned over their S&P is that although the number of violations they experience may be overwhelming, the individual effects of the violations are difficult to observe, so people cannot devote the time, effort, and resources to addressing them all. The term "slow violence" refers to incremental and accretive events that may be near invisible to people when they are initially experienced, but inflict significant harms when aggregated over long periods of time and across populations [94]. While the concept has long been applied in the context of industrial pollution and environmental degradation, only in recent years has slow violence been referenced in UPS-adjacent contexts—specifically, in domains of state surveillance and policing [75], digital technology abuse [23], and online behavioral advertising [50, 135].

Robert Nixon, the originator of the term, suggested one disarming agent against slow violence: *"to devise arresting stories, images, and symbols adequate to the pervasive but elusive violence of delayed effects"* [94]. In other words, gathering specific evidence of this violence and publicly reporting on it in an emotionally-appealing way may be an effective vector for *transforming* the public and inspiring resistance against slow violence. Over time, repeated user reflection over their observations [37] of S&P slow violence can also encourage users to construct a personal narrative related to S&P. As such, Wu et al. [134] also recommend integrating the collection and collation of people's observations and narratives of privacy harm into their day-to-day lives. In these ways, we can more formally scaffold how people already go about building their personal theories of security, privacy, harm, and mitigations.

In particular, past work [7, 66, 69, 124] has invariably found that people's mental models tend not to align with actual technical implementation. However, this work has primarily been point-in-time studies: they capture people's mental models once, but not how they might evolve. *How do people's mental models and folk theories of security and privacy develop and change over time? Do they seek evidence for, revisit, or update these theories?* Incorporating fictional inquiry techniques, such as storyboarding or other vignette-based methods, can create "brave spaces" for people to interact with these models and inhabit environments of potential discomfort. Fiction can be a "purposeful, deliberate, direct participant in the practices of science fact" that allows us to understand, explore, and question alternate futures [17]. Design fiction can be particularly useful for understanding how users think about figures of authority in S&P, such as government regulatory bodies, the news media, technology companies, and the UPS research community itself.

3.3 Triggering Personal S&P Dilemmas

3.3.1 Standard Overview. Provocation is key to transformation, as we noted in Section 2.1. Bardzell et al. [13] remark that for a design to be provocative, it should possess a "slight strangeness" that is neither too novel nor too mundane, and, consequently, challenge a person's existing feelings and assumptions about the state of the world. Relatedly, designs that trigger personal dilemmas—i.e., those that involve an individual's values and goals—can be particularly effective in provoking self-reflection [98]. Following this logic, we present our second standard:

Standard #2

TP-UPS designs should trigger personal dilemmas and lead to critical reflection about security and privacy in people, while maintaining agency.

Past literature has given us a few strategies for triggering personal dilemmas—putting up barriers to default behaviors, asking users to make forced choices, and creating embodied symbols or provotypes [98]—which may be helpful to consider for S&P. Creating behavior barriers to the most habitual or automatic choices that a user might make is likely the most familiar strategy for the UPS community. For example, as the National Institute of Standards and Technology (NIST) puts it, making authentication more "usable" means that it's "hard to do the wrong thing" [47]. (Indeed, there has been extensive past work to shift users from making their default "wrong" or insecure choices, to more easily adopting "correct" behaviors.) However, this canonical interpretation of usability is inherently at odds with Transformational Provocation, since it elides not only the opportunity for user reflection but also users' agency over their decisions. As we expound in the following subsection, being forced to make S&P decisions on behalf of another party can be particularly uncomfortable, as the decision may raise questions of paternalism versus autonomy. While we design ways for users to deliberate such tensions, we should also search for opportunities for *UPS researchers* to confront these dilemmas too.

3.3.2 Implementations. One straightforward way to trigger personal S&P dilemmas in users is to shock users enough that they abruptly stop what they are doing and consider new perspectives and feelings. Nthala and Rader [96] first laid groundwork for such triggers in their conceptual model for provoking privacy *speculation*. They proposed that showing users surprising inferences about themselves might encourage users to inspect how and why their data was being collected and used, and think more critically about the parties collecting and using the data. In other words, these inferences can create a sense of internal conflict within users and cause them to immediately seek understanding of the conflict.

While past work [127, 129] has explored how people react to seeing the inferences that ad platforms and social media sites make about them, the primary goal of these efforts was transparency and awareness, with a secondary hope of "improved" user behavior. However, it is unclear that awareness, as a broader concept, is persistently transformational. Over a decade ago, it was revealed that Facebook conducted an experiment about "emotional contagion" to manipulate people's moods via their home page feeds [74]. Public

outrage was widespread, and researchers warned about the dangers of spending too much of our lives online. Today, users of TikTok actively engage in "training" their own "For You" pages [114], so that they can only view content algorithmically curated for them or even use the platform for psychiatric self-diagnosis [52].

UPS researchers should thus explore new conceptions of what people might find surprising or disruptive. However, to do so effectively, we ourselves should also be targets of provocation. As we referenced in Section 3.4, it may be hard for us to figure out what is actually surprising or disruptive to users, due to the challenges of suspending our normative community beliefs. To counter this constrained perspective, UPS researchers should be more actively engaged in the methods of design fiction and fictional inquiry, as aforementioned. Within UPS reseach, especially in challenging the hegemony of large tech institutions, there has already been some design fiction work. For example, Wong et al. [133] employed design workbooks—collections of conceptual designs—to explore questions about user privacy stemming from a science fiction novel. More recent works have also used a fictitious job placement app that invasively collects data to explore unemployed individuals' forced consent [64]; pondered intimate data collection practices through a marketing brochure for a near-future menstrual tracking app [48]; and imagined a future society where people can report their privacy harms to the government [134]. To increase provocativeness, we can ask users to articulate the threats and outcomes they might find most *undesirable* and purposefully confronting them with scenarios in which those negative outcomes become realized.

Games provide additional mechanics of disruptive provocation. A game might purposefully instill sharp discomfort in a player, such as through taking on the role of a nurse making care decisions in their daily work with life-or-death consequences [89]. As an S&P example, we might ask people to play the character of a person with a family member who is unfamiliar with technology and S&P. This roleplay can lead to dilemma-inducing questions such as, "*What if someone has sensitive information about a vulnerable family member of yours? How would you act in trying to protect that information? How would you take that family member's own free will into consideration when making these protective actions?*" Through these forced choices [98], people would be pushed to balance paternalism and stewardship [93] when considering the technology use of older adults, who are especially at risk of attacks like phishing and scams.

Additionally, as we suggested above, directly asking people to manipulate or interact with embodied symbols [98] can be provocative. What might such symbols look like for UPS? One striking recent example is Eyecam, an anthropomorphic webcam stylized to look like a realistic human eye [117]. The creators of Eyecam hoped to provoke reflections on our relationships to ubiquitous sensing devices, including the S&P vulnerabilities of webcams. The visual of a human eyeball can make a user become more aware of how they are being surveilled. The "fleshiness" and responsiveness of Eyecam, in particular, also evoke an almost human bond with the webcam, and call into question how we assign agency to sensors collecting data about us, especially as they grow more autonomous. **UPS researchers should explore additional modalities and domains of interaction with embodied symbols to encourage user reflection about S&P.**

Be it due to resignation over S&P woes or a relative lack of technical education, non-expert users are not commonly tasked with contributing ideations about the future of these institutions. As past work has suggested, users are enthusiastic about simply being asked their opinion at all [136]. *If we asked users to write a story about a day-in-the-life of their government spy or Big Tech employee, what tales would they spin?* As one cultural example, the popular meme of the "Government Agent Watching Me" has been referenced by users on social media since at least late 2017 [72]; more recently, the "My Chinese Spy" meme emerged out of the January 2025 ban of TikTok in the United States [45]. Both memes paint an almost appreciative picture of a state-level employee surveilling and personally curating what users see. The proliferation of these characters suggests that people *do* keep S&P in their minds, even if their inability to adhere to expert-recommended S&P behaviors doesn't necessarily demonstrate so. *What roles can these memes play in forming people's mental models of online S&P? What other embodied symbols can we create for users to inspect or inhabit?*

3.4 Enhancing Protection with S&P Promotion

3.4.1 Standard Overview. One challenge that UPS researchers might face in generating provocative personal S&P dilemmas can be the ability to suspend normative community beliefs and judgments around best practices and protections for users. This could prove especially difficult if there seems to be what we feel is a clearly "right" choice for users to make. For example, experts tend to value software updates, while users remain skeptical or even "betrayed" by them [66, 121]. While existing approaches in UPS—which have focused on the risks, threats, and harms that people face, along with the negative S&P behaviors (or lack of positive ones) that people may exhibit—have been important and valuable, they can also lead to exacerbating tensions between experts and "regular" users. Our paternalistic propensity for correcting individual user behavior can, as Daniel Solove writes [115], lead users to see best-practices S&P behavior as an insurmountable pile of homework assigned by S&P experts, which they struggle to prioritize [108]. Drawing upon Coles-Kemp et al.'s critiques of what it means to be secure and how to achieve that security [30], we propose the following standard:

Standard #3

TP-UPS interventions should guide people toward taking S&P-*promoting* actions, which don't necessarily have to be S&P-*protecting*.

The field of positive design, as one example, aims to incorporate pleasure, personal significance, and virtue into design, to support and increase people's subjective well-being [39]. Building on this concept, "positive *computing*" [25] and "designing for flourishing" [120] have also emerged as frameworks that center joy and wellbeing in technology use and design, especially for BIPOC communities. These paradigms propose several design tenets that the UPS community may find helpful for moving beyond a focus on protection and problem-solving, including but not limited to seeking opportunities for self-actualization, refraining from damage-centered reductionism, and evaluating for self-sustainability.

This is not to say we should begin ignoring the real harms that people may face and abandon user protection efforts altogether. Rather, incorporating positive design principles into UPS can be another way to help align expert and user motivations. For example, as we suggest in the next section, incorporating humor into the conversations that people have about S&P may help groups of people cope with the traumatic effects of S&P violations. In particular, introducing S&P topics in a less "serious" manner can ease the shame that people may feel when seeking help from others for issues like identity theft and data breaches. Additionally, S&P-related memes can reveal people's folk theories and mental models and provide accessible touchpoints for co-design.

3.4.2 Implementations. Guiding users to generate transformation *externally* can introduce new incentives for people to engage with S&P. The effects of S&P violations, such as identity and resource theft, data breaches, creepy personalization, and surveillance without consent, can be deeply distressing and traumatic to an individual. Peer support can be key to coping with this trauma, especially in cases where people feel deep shame for not "protecting themselves enough" after falling victim to S&P threats [26, 109]. At a basic level, **UPS researchers should continue exploring new ways for people to have conversations about S&P with others in their lives.** As Rader et al. found in 2012 for undergraduate students [105], and as Pfeffer et al. replicated ten years later with a more diverse sample [101], sharing anecdotes about S&P threats and experiences are highly influential in (1) people's formation of mental models and (2) increasing sense of camaraderie and social safety. Watson et al. [126], in interviews with small social groups, found that the group format of the interviews specifically served as a forcing function for participants to share helpful S&P advice and tools with each other that they wouldn't have otherwise.

Sharing stories about S&P can be a way for groups to cope with trauma from those events, but not all conversations about S&P need to be "serious" or high-stakes to help with that coping. We can look for opportunities to incorporate joy and humor into S&P conversations. As one example, "PowerPoint parties", or social gatherings where friends present comedic slideshows about niche interests, have increased in virality in recent years [6]. These gatherings could serve as a low-stakes testing environment for publicly discussing S&P. To develop conversations further, we could encourage people to visualize the ways that people's S&P actions negatively or positively affect the people around them.

Transformational games have been effective in creating a safe or brave space for exploring these consequences; however, engaging social groups in S&P roleplaying in real-life can be a low-cost alternative to full-fledged game development. For example, a player may chat with an S&P expert character within a game about SIM swapping, and then be given an opportunity to enable alternative two-factor authentication methods to protect themselves. If the player chooses not to secure their accounts, they might learn that others around them have fallen prey to scams, due to an attacker nefariously impersonating the player online [76]. Through these roleplaying activities, users can exchange and discuss each others' threat models and S&P strategies. Extending the model of privacy speculation [96], researchers could also explore how surprising or disruptive inferences influence the S&P of social groups; we could also ask users to speculate about *each others'* data relationships with technology companies. This can confront users about how differences in personal characteristics—some more "vulnerable" than others—change their experiences of online safety and harm.

These conversations might generate cognitive dissonance or introspection. As we referenced in Section 3.3.2, asking people to roleplay as a character who is responsible for others' S&P can bring up conflicting feelings of paternalism and stewardship. Similarly, we should directly ask people to reflect on how their S&P and digital behaviors affect the S&P of other people online. For example, Hasan et al. [60] found that people who used aggressive and self-deprecating humor were more likely to violate other people's privacy by sharing photos of them online. These studies can reveal new explanatory personal traits, beyond S&P skill or knowledge, that influence people's S&P behaviors. Zooming out, we should continue exploring large-scale measurements and collations of sentiments about S&P; this could help us understand how to better tailor provocations to different groups of people. For example, Hasan et al. [61] created a psychometric scale to measure how much an individual values the privacy of *other people*. We might imagine that individuals who value other people's privacy more might be better candidates for starting these conversations or confrontations about S&P in their own social groups.

However, as a community, **UPS researchers should be prepared to accept that the conversations that people have may not align with our norms or expectations.** Wei et al. [128], for example, investigated how users on TikTok spread *anti*-privacy and security advice in the form of tutorials on how to surveil and control *other* people's activity online. They found that the social norms of "TikTok culture"—e.g., framing intimate partner surveillance as cheating detection, and child surveillance as a "parenting hack"—made it more acceptable for users to spread S&P advice that S&P experts might find normatively incorrect, wrong, and/or harmful. *Beyond paternalistic platform monitoring and algorithmic curation, is there a way to align users' strong personal and social motivations for being interested in such content with UPS norms? Are there norms that the UPS community should reconsider?*

Finally, as we noted in Section 3.2.2, "slowing down" can encourage users to engage in critical reflection about S&P. Past design research has proposed journaling as a vehicle for carrying out these "slow" mechanics against a broader context of personal wellbeing. While past work in UPS has frequently employed diary study methods as a way to collect empirical data about user S&P behaviors and feelings [107], how the methods themselves might affect those users not been extensively studied. As one example, Barbosa et al. [12] designed a technology probe, *Who Am I*, that provided users a way to visualize the cumulative inferences made about them over an extended period of time. Similar to how smartwatches and fitness trackers enable users to be actively engaged in their understandings of their physical health, while granting them the ability to communicate with their doctors with high-level concrete data, we might imagine that taking regular measurements of S&P experiences can empower people to speak more definitively about their experiences with professional advocates and policy experts.

3.5 Cultivating a Readiness to Transform

3.5.1 Standard Overview. The rich body of UPS research has often left addressing large-scale user S&P issues as an exercise for regulators, legislators, and tech companies, rather than something in which users themselves can hold a personal stake. However, user responses to large-scale S&P events can give us important pointers for provoking organic transformation in people, particularly in terms of user coordination and environmental preparation. As a hopeful, collaborative example, after the 2022 overturning of Roe v. Wade via Dobbs v. Jackson in the United States, Song et al. analyzed how people made sense of privacy issues and risk mitigation strategies on Reddit [116]. They found that the collective nature of the sensemaking about privacy fostered a sense of mutual support and reduced biases and blind spots in an individual's privacy assessment. In this vein, we propose our last standard:

Standard #4

TP-UPS researchers should cultivate conditions and systems that are ready to help people transform their S&P.

In times when users continue to feel that S&P issues are secondary concerns, or that the consequences of S&P violations are not concrete enough to take immediate action, we should set up infrastructure ready to absorb the swell of action triggered by highly provocative external events. *What does a collective action platform for security and privacy issues look like [136]? How can technical systems, supplementing legal ones, support people's access to technological due process [29, 31]? Where can they direct demands for redress for the harms and wrongs that they have endured [58]?*

As we consider these questions, however, we can't simply wait around for large flashy S&P events that disrupt multiple stakeholders, inflicting collective S&P trauma in the hopes of motivating users. (These events might not even be particularly reliable for firing up users: past work in dealing with people's responses to data breaches has repeatedly found evidence of user resignation after the initial shock of finding out about the breach [84, 85].) *Is it possible to create conditions outside of such events to* proactively *increase transformation potential?*

3.5.2 Implementations. In 2019, Fox et al. created a "product catalog" for Vivewell, a fictional menstrual tracking app, to interrogate the emotional, social, and political implications of growth in intimate biosensing technologies [48]. During that time, the consequences of intimate health data surveillance, though significant, still felt closer to science *fiction* than fact. After the overturning of Roe v. Wade three years later, the risks are more much tangible to American users: in the state of Texas, for example, laws now incentivize people to report illegal abortions to the government with a cash "bounty" [20]. **UPS researchers should be sensitive to the environmental conditions that can lower or raise a provocation's potential to transform.**

We can also encourage large groups of users and experts alike to document and respond to S&P events collectively. Bennett and Segerberg [16] termed *connective action* to refer to the way that laypeople's personal stories involving popular ideological or political motivations spread across weak ties [55] on social media; they differentiate connective action from *collective* action, which may involve strong formal organization coordination to manage a larger base of individuals. Activism infographics on social media are one way to bridge connective and collective action: they pair credible information from experts with calls for protests, financial action, and organizing in a format that is easy to disseminate [70].

Concurrently, **UPS researchers should lead people to critique how S&P manifests in their lives; they should also model critical behavior themselves.** One straightforward way, building on existing awareness campaigns by S&P advocacy organizations, can be to communicate directly with users about S&P. Engaging with public forums, such as Privacy Guides[1] and the r/privacy community on Reddit, through "Ask Me Anything (AMA)" sessions can give people the opportunity to ask questions about, e.g., what privacy rights they can exercise and why regulations are the way they are. However, experts must be willing to accept user critique, dispute, and refusal, not just push best practices.

In the same vein, we should provide spaces for people to ask questions about specific devices or technologies they encounter. Even though people harbor folk theories about how a technology functions, they often don't have a way to test these theories in a satisfying or helpful way [104]. For example, users are often concerned about smart speakers eavesdropping on their conversations; typically, the only way they can confirm that the devices are *not* listening is through unplugging the device from power [41]. However, this binary action does not enhance the user's understanding of how their data is being stored, analyzed, and used when it *is* plugged in, and reduces the user's self-efficacy to whether or not they remember to power on the speaker. We might imagine building a smart speaker simulator as a technology probe [65], with virtual representations of data storage and analysis, that can allow the user to poke around and figure out how the device works in a judgment free zone and enhance their mental models.

Additionally, as previously mentioned, transformational S&P games can also create "brave spaces" for users to test out different S&P choices and visualize the consequences of their actions on the people around them. While the consequences inside games are not necessarily "real", giving users a space to make these mistakes and see their effects can ingrain self-efficacy in users. Also, engaging social groups in roleplaying can be a lower cost way of mimicking these effects without developing a fully-featured game. However, researchers should also consider devising additional modalities for critique. *Beyond creating self-contained environments, how can we support the creation of "brave" and "safe" spaces for people to make mistakes and test hypotheses about S&P?*

UPS researchers should continue to explore ways to embed the questions and conversations people already have about S&P directly into technical systems, through a lens of social governance and oversight [28, 81]. Akter et al. [3] developed CO-oPS (Community Oversight for Privacy and Security), a mobile app to help communities seek S&P advice on the apps installed on their phones. In the app, users within small communities can review the apps and permissions that other community members have installed and granted on their phones; through this mechanism, members can give feedback on each others' behaviors and share knowledge.

[1] discuss.privacyguides.net

Systems like CO-oPS could be extended to a broader space of social cybersecurity [137] systems that allow people to collaboratively, formally construct community-wide resources things like threat models, accountability measures, and shared narratives [90].

Games continue to offer helpful mechanics for UPS to borrow. Janik [67] and Dooghan [42] both reference the use of the "bloodstain" mechanic in the *Dark Souls* series as a form of asynchronous communication between players. When player characters "die" in the game, they leave behind visible bloodstains in the game environment that other players can interact with; through these interactions, players can see how other players died and receive advice on how to proceed. These mechanics remind the player that a world of other real people exist outside of the game and create a sense of player camaraderie, especially when players leave behind humorous messages for each other. An analogue for S&P, then, could be providing in-situ notifications or persuasion-based messages related to S&P, perhaps through a browser extension. When a user encounters a potentially unsafe situation online, they might see the consequences of those who made insecure choices ahead of them and volunteered their experiences to warn others. For example, if a user is prompted to change a password after a data breach, a pop-up character could tell a story about how a criminal used their old password to make fraudulent purchases, recalling applications of the Protection Motivation Theory to passwords [83, 110, 138].

Such messages need not be focused on damages to the user; rather, they could raise curiosity or speculation about how other stakeholders, such as regulatory bodies or technology companies themselves, might operate. González Cabañas et al. presented Facebook Data Valuation Tool (FDVT), a tool for Facebook users that provides "personalized and real-time estimation of the revenue they generate for Facebook" [53]. A notification with FDVT data could read, "*50 users in your neighborhood did not log into Facebook today, costing Meta $50 in personal data this month.*" This might encourage the user to weigh the value of their own data, both as an individual and aggregated with other users. Relatedly, as the Federal Trade Commission has written in the past, "an injury may be sufficiently substantial...if it does a small harm to a large number of people" [46]. Along these lines, provoking smaller reflections on a collective level can add up to a large potential for transformation. As Vincent et al. have argued through their concept of *data leverage* [123], if members of the public, in aggregate, can reduce, stop, redirect, or otherwise manipulate their data contributions to technology companies, they can create a meaningful change in the balance of power between themselves and the companies.

4 A Design Toolkit for Transformational Provocation

In this section, we present a toolkit for UPS researchers to design for Transformational Provocation. We first define three dimensional heuristics for the UPS community to explore and consider: **Velocity**, **Interiority/Exteriority**, and **Contextual Potential**. We then offer a generalized design process for researchers and designers to meet the TP-UPS standards in future work. (In Section 5, we present two fictional scenarios as worked examples of the process).

4.1 Dimensional Heuristics of Provocation

As Bardzell et al. wrote [13], the provocativeness of a design can be difficult to gauge. As such, we developed three dimensional heuristics of provocation in S&P to guide UPS researchers throughout the design life cycle, from ideation to evaluation. In particular, we hope these heuristics can help UPS researchers delimit specific transformations that they hope to provoke in people and construct appropriate measurements or expectations for those transformations. We frame these heuristics as *dimensional* since varying levels of each heuristic can result in very different designs. The four pillars of Khaled's agenda for reflective game design [71]—prioritizing questions over answers, clarity over stealth, disruption over comfort, and reflection over immersion—were highly informative in our nascent conception of these dimensions.

Our heuristics are not perfectly continuous, mutually exclusive, or all-encompassing. We also anticipate more heuristics to be added as the nature of security and privacy evolves. For example, the rapid proliferation of artificial intelligence and machine learning—through both its unique capabilities and data requirements—can both amplify existing S&P harms and expose users to new S&P risks [79]. These unknowns, however, only reinforce the necessity for the UPS community to reckon with transformation.

4.1.1 Velocity. **Velocity** refers to the speed at which a user is expected to transform after being exposed to a provocation. For example, presenting a surprising personalized inference to a user one time may be quite shocking and cause them to immediately ask questions about data use. The velocity of a provocation can range from low to high: while a provocation should always lead to a user to stop and think, it might do so gently over time (low velocity) or very abruptly (high velocity).

4.1.2 Interiority/Exteriority. Many user S&P behaviors are inherently social: people share authentication methods with friends and family to access streaming services, pose questions about S&P when unsure of what to do, and curate their self-presentation on social media through sharing and withholding personal information [137]. **Interiority and exteriority** deal with the degree to which a provocation might induce transformations within an individual versus a broader community of actors. However, interiority and exteriority are not necessarily oppositional: a provocation focusing on *interiority* should ask the user to inspect their own choices and beliefs, which may include beliefs about other people. One focusing on *exteriority* should involve encouraging the user to engage in S&P behaviors with friends, family, acquaintances, or the public.

4.1.3 Contextual Potential. As educator Paulo Freire first defined in his landmark text, *Pedagogy of the Oppressed*, *critical consciousness* is a person's ability to recognize, analyze, and transform their inequitable social conditions, individually or collectively [49]. As a descendant concept in education and social work, *transformative potential* refers to the variations in levels of critical consciousness between different people [68]. To avoid confusion with our own agenda of transformation, we use the term **contextual potential** to describe the levels of critical consciousness elicited or sparked by a provocative *design* (in contrast with the transformative potential of an individual *person*). We present contextual potential based on the two pillars of critical consciousness: *reflection* and *action* [40].

These pillars include both the nature of decisions or actions that we hope to provoke people into taking regarding S&P, and the expected duration of the provocation's effects. In particular, we note the importance of understanding a user's broader environment—be it personal, social, or political—when anticipating how that user might respond to a provocation. For example, in times of political turmoil, it may be appropriate to provoke collective user action; in calmer times, we might encourage people to continually reflect. In other words, on a scale from stoking curiosity to catalyzing incitement, what is the intended reach of transformation?

4.2 Design Process

Before we begin designing provocative interventions, we should take time to **define desired transformations**. When doing so, we should ask ourselves the following questions:

- In addition to education and awareness, how might you expect users to transform?
- How will users' motivation to adopt S&P behaviors and actions change after transformation?
- How can this transformation benefit users beyond protection against a single threat?

As noted in Section 2.1, Culyba [32] offers several alternative objectives for individual transformation—beyond knowledge, skill, and behavior—that we believe the UPS community should strive to provoke in end-users. Below, we define these transformations for S&P and provide examples of potential desired transformations:

Disposition An individual's feelings about S&P are changed. (*Example:* A user feels more negatively about tech companies and is upset about data use practices.)

Experience An individual's personal anecdotes about S&P are changed. (*Example:* A user experiences the challenges of having their identity stolen.)

Belief An individual's sense of truth about S&P is changed. (*Example:* A user now believes that S&P should be a primary concern in their lives.)

Relationships The nature of an individual's social relationships are changed by S&P. (*Example:* An individual feels more comfortable sharing stories about S&P with their friends and understands that they are not alone.)

Identity An individual's sense of self is changed. (*Example*: An individual might initially believe that only "paranoid" [100] people care about privacy. Afterwards, they might now think, *"I can see myself as a privacy advocate."*)

After defining the desired transformation, we can move on to the provocation heuristics proposed in the previous subsection. Below, we provide suggestions for appropriate methods and research questions along each heuristic of provocation. When creating provocations, we should consider the heuristics in the following ways:

4.2.1 *Velocity.*

- *What threats and outcomes do people might find most undesirable or alarming?* These responses can seed high velocity design fictions, which purposefully confront users with scenarios in which those undesirable outcomes are realized.
- *What S&P harms do people find too small to devote direct attention to, even if it continues to bother them?* These instances of slow violence [94] can become the basis for low velocity provocations. These may require more longitudinal or ethnographic methods (as opposed to high velocity ones, where the effect sizes may be big enough to measure one time).
- *How can velocity-adjacent techniques already familiar to the UPS community, e.g., developing browser extensions and apps to infer from browsing activity and alert users, be adapted for transformation?* The content of these alerts could focus on encouraging critical reflection and action, rather than persuading users to do a "correct" behavior.

4.2.2 *Interiority/Exteriority.*

- *How can we can engage people in thinking about potential futures?* Researchers could consider exploring design fiction methods such as story completion and comicboarding, which can encourage users to think outside of immediate impacts on their own lives and generate new personas and scenarios.
- *How might the interaction design methods we use differ based on whether a provocation is focused on interiority or exteriority?* Provocations focused on increased interiority might require individual cognitive walkthroughs or in-depth surveys and interviews to understand transformational effects in users; those focused on exteriority could require group evaluation techniques like focus groups or group interviews.

4.2.3 *Contextual Potential.*

- *How can we anticipate people's openness to reflection and action?* Researchers should engage in broader contextual inquiry by evaluating the current social and political conditions. Relatedly, we should considering adapting measures like the Critical Consciousness Scale [40] for UPS.
- *How might we strengthen ties with people who are already motivated or curious about S&P, so that they might shift from reflection to action?* Co-design and participatory design methods should be of particular focus for the UPS community.
- *What expertise can we gain from people in improving our own recommendations of best practices?* As Rahwan argued in his "society-in-the-loop" agenda, public opinion shapes the values by which we should evaluate experts[106].

5 Vignettes of Transformational Provocation

We present two fictional scenarios to illustrate how our Transformational Provocation design process can work across contexts of life. The first scenario (Section 5.1) explores the challenges of digitally untangling from an intimate partner when the relationship ends. The second scenario (Section 5.2) examines how parents may be confronted with choices about their children's presence online. We chose these two cases as poignant illustrations of how traditional approaches and tools in UPS may not be compatible with the personal, social, and political complications of users' lives, echoing prior work [128]. Specifically, we argue that even if users do the traditional "right" thing that experts might recommend in these

scenarios, they might not feel fully engaged with or satisfied with their S&P outcomes.

For each of our scenarios, we provide a background sentence that hints at the possible space of S&P issues for a given character. We first provide a short example of the actions that character may take in a traditional UPS world. Then, we propose an example desired transformation that a researcher might work toward for the character in that scenario. We then use a series of short vignettes to illustrate what provocations might look like. We highlight the parts of these vignettes that refer to a specific provocation heuristic: green for Velocity, pink for Interiority/Exteriority, and yellow for Contextual Potential.

5.1 Scenario: Ending Intimate Relationships

Over the course of an intimate relationship, partners may share a plethora of digital resources, ranging from lower stakes media and entertainment subscriptions, to the more sensitive calendars, smart devices, and financial accounts [80, 99]. However, prior work has suggested that existing models of resource sharing and access control can result in social friction [97]. These frictions are particularly evident when relationships end (to say nothing of those escaping intimate partner violence situations, who face even greater stakes for making clean separations from their abusers). For example, users may feel uncomfortable directly conversing with their ex-partner about ending sharing practices, or they may be uncertain about whether sharing has actually ended despite changing passwords [97]. More recently, social media[2] and entertainment[3] platforms have also introduced "algorithmic blending" tools that serve content targeted at a user *and* a partner *together*, leading users to be more digitally intertwined with their partners than ever.

5.1.1 Scenario Overview. Based on this prior work, let's imagine a scenario with a character named Jordan, who recently broke up with their partner of three years, Reese. Jordan would like to move on from this relationship, particularly digitally. Traditional approaches from UPS might lead Jordan to normative practices of changing passwords to shared accounts:

Traditional UPS Outcomes

Jordan takes inventory of all the accounts and services that they shared with Reese. They do a pretty thorough job, but after the third day of changing passwords, Jordan gets tired and abandons the endeavor. A short while later, Jordan finds out Reese has locked Jordan out of the Netflix account that Jordan pays for.

Even though it may be the "correct" S&P action to take in this scenario, changing passwords alone cannot cover the additional social anxieties or uncertainties that Jordan might continue to have about Reese. Are there other changes that Jordan could go through to feel more certainty or agency over their S&P?

[2]https://www.theverge.com/news/650333/instagram-blend-reels-feed-friends-dm-group-chat

[3]https://newsroom.spotify.com/2021-08-31/how-spotifys-newest-personalized-experience-blend-creates-a-playlist-for-you-and-your-bestie/

5.1.2 Transformational Provocation Approach. Taking a TP approach, we should first define a desired transformation. We use the following as a working example:

Desired Transformation

Jordan feels more observant of how their life becomes digitally intertwined with others in an intimate relationship.

Based on this, we might envision a technical intervention that helps Jordan achieve this goal of increased observance. In the vignettes below, we write about a fictional app named "NoContact", designed to help Jordan critically reflect on their digital footprint, especially in relation to Reese. To do so, it may send notifications from time to time that Jordan could find provocative. When considering velocity in the design of this intervention, what kinds of events would Jordan find undesirable or alarming? In this first vignette, NoContact sends alarming notifications to remind Jordan of S&P-protecting measures necessary for untangling their life from Reese. By including a report of progress, the app encourages Jordan to completely update all of their accounts and settings, rather than letting Jordan give up due to tedium.

Vignette 1: Velocity

Jordan signs up for NoContact, a tool that claims to help them untangle their digital life from Reese. The tool scans Jordan's device and periodically sends some notifications: *"You have 15 apps that use Reese's email to log in!" "You've shared 245 photos with Reese!" "Reese can see you're out at a bar!"* At the end of one month, the tool prompts Jordan to review the "progress" they've been made in moving on. Jordan wonders if Reese has noticed Jordan's data- and account-sharing practices change over time.

As Muise et al. [92] suggested, social media may create jealousy and suspicion in intimate relationships, due to the ambiguity of how and what information is presented to users. In the midst of a breakup, in particular, people may use social media to stalk or potentially harass ex-romantic partners [82]. In this next vignette, we consider interiority/exteriority by enabling NoContact to gently shame Jordan out of pursuing cyberstalking behaviors:

Vignette 2: Interiority/Exteriority

Jordan is feeling a bit lonely and thinks about Reese. Jordan pulls up Reese's Twitter account; after not seeing any new posts, Jordan heads to Reese's Instagram account, where Reese has posted a new photo with a stranger. Jordan then opens up Venmo [15] to see if any charges have occurred between Reese and the stranger. A notification from NoContact pops up: *"You've looked at three of Reese's social media profiles in the last five minutes. Please type* 'I want to keep infringing on Reese's privacy' *to continue."* Jordan feels embarrassed about stalking Reese's accounts.

Finally, a single provocative design like NoContact obviously cannot exhaustively cover all aspects of transformation in users; relatedly, not all provocations need to come from researchers. As Jordan begins to move on, they may already start organically noticing how the inferences that algorithmic content feeds make about Jordan and Reese can be at odds with their real life experiences. Uncovering these small opportunities for personal agency is key to encouraging users to commit to caring about S&P:

Vignette 3: Contextual Potential

Jordan browses on TikTok and sees a video about one of Reese's favorite brands of kitchen cookware, reminding Jordan of Reese. This video causes Jordan to wonder whether TikTok's algorithm "thinks" they are still together. Jordan becomes frustrated: they want to move on, but the algorithm keeps butting in. They become concerned about what else the algorithm thinks about their personal life, and start coming up with ways to "trick" the algorithm.

5.2 Scenario: Parent-Child Relationships

Children recognize S&P behaviors, such as identifying information as sensitive and understanding what is appropriate to share with others, at a young age; conversely, parents tend to defer teaching kids about S&P until they are older and more socially engaged online [77]. However, parents also violate their children's privacy by participating in "sharenting", or posting images of children on social media [18]. All the while, tools for parents to surveil their children have become more normalized in recent years, resulting in what Wei et al. describe as a "tension between parents' desire for information and control to ensure safety with teens' desires for autonomy and privacy" [128].

5.2.1 Scenario Overview. From these findings, let's imagine a scenario with Perry, who is the parent of a young child, Cameron. Perry has to make many decisions about how to approach sharing information about Cameron online. Traditional UPS approaches might see Perry delving deep into parental control device settings and "locking everything down":

Traditional UPS Outcomes

Perry learns about parental controls on Cameron's device and turns on the "maximum" security settings. Through those settings, Perry can monitor everything Cameron does online. While neither Perry nor Cameron is comfortable with this level of control and surveillance, Perry doesn't really know how else to protect Cameron.

While existing UPS literature has highlighted the importance of building more foundational S&P behaviors early on in a child's life, specific recommendations have often consisted of new kinds of educational materials [77] or updating existing parent monitoring software to be less focused on restriction and more on direct engagement [132]. For one, these recommendations for acquiring more knowledge or skills simply add to the "homework heap" of S&P management advice that Solove argues is impossible to complete [115]. For another, these forms of tech solutionism [91] do not confront the underlying conflict between paternalism and autonomy that parent and child alike face.

5.2.2 Transformational Provocation Approach. To weigh this dilemma, we can use this working example of a desired transformation:

Desired Transformation

Perry constructs a personal understanding of how their own S&P practices can influence the degree that Cameron is exposed online.

In the following vignettes, we highlight how different designs can address the same desired transformation in different ways. In the velocity vignette, we adapt Barbosa et al.'s *Who Am I?* system [12], which sends both immediate notifications and weekly digests of ad inferences. This gives a mix of high and low velocities:

Vignette 1: Velocity

Perry receives a weekly digest of their activity online, with highlights of activities related to Cameron. Perry also installs an app that alerts them about possible inferences that advertising platforms might make about them [12]. Cameron gets sick in the middle of the night, and Perry takes them to the doctor. The app warns that Perry may see ads about sleep aid supplements for children in the future.

While the desire to share good news about family members with others is natural, the path toward *overly* sharing information about children online can be a slippery slope. In particular, the benefits of social connection from sharenting may seem so enticing that notions of S&P can fall to the wayside. News media has frequently reported instances of children asking their parents to stop sharing, often framing the child's request as jarring to the parent [51, 59]. Is there a way to achieve this provocation without forcing children to take on the burden of confrontation? In this next vignette, a bot first encourages introspection in Perry, which leads to Perry seeking out a conversation with Cameron:

Vignette 2: Interiority/Exteriority

Perry shares a photo of Cameron on social media. A child safety bot comments, "If you were my parent, I'd be furious that you posted this." Perry has a knee-jerk reaction of blocking the bot, but the comment bothers them days later. Perry contemplates how else they may have violated Cameron's consent online in the past. Perry has a conversation with Cameron about S&P.

Finally, as with the intimate relationships scenario, researcher-side provocations are the not the be all and end all of Transformational Provocation. If circumstances allow—e.g., Cameron becomes

repeatedly exposed to increasingly violating school surveillance with seemingly little benefit—Perry may feel incited to take action:

> ***Vignette 3: Contextual Potential***
>
> Cameron's school has recently decided to deploy a new system that monitors students' activities online. Over the last year, the school has already asked parents for permission to install other monitoring software that hasn't seemed to work. Perry is getting fed up, and feels that these systems go too far into invading Cameron's privacy. Perry starts a group chat with other parents to discuss the monitoring system. Perry is relieved that other parents feel the same way. The parents collectively protest the system through phone calls and emails.

5.3 Additional Toolkit Considerations

While the majority of provocations we presented in the vignettes can be easily imagined as researcher-induced—i.e., through the user installing an app or subscribing to alerts—we wish to re-emphasize that not all provocations can or may come from us. As we expressed in Section 3.5 and throughout the two scenarios, external events can be highly provocative. We include such examples in our vignettes to highlight that researcher interventions are not the be-all and end-all of users' engagements with S&P: to successfully provoke user transformation, we should consider the rich context of user experiences beyond what we explicitly design. We should understand our capabilities and limits as researchers and designers, do our best to create safe and positive S&P experiences for users, and be comfortable with relinquishing control over downstream impacts.

Additionally, although each highlight within the vignettes only refers to one heuristic, the vignettes also hint at how provocations span multiple heuristics. For example, in the second vignette in the parent-child scenario, the provocation directly asks Perry to consider how their actions affect Cameron, aligning with interiority/exteriority; however, Perry's anger at seeing the comment suggests a high level of velocity, and Perry's conversation with Cameron suggests reflection and action. As we noted in the Design Toolkit outline (Section 4), the heuristics are not mutually exclusive.

6 Future Work

In Section 3, we presented suggestions for practical implementations of four aspirational standards for the UPS community, which offer a concrete basis for future UPS research. For one, we should design and support spaces or opportunities for users to slow down and engage in mindfulness and more deliberate reflection about S&P, perhaps through the games-based "mechanics" for reflective play [89]. We should also explore new conceptions of what people might find surprising or disruptive; we might achieve this by incorporating additional modalities and domains of user interaction with embodied symbols, such as through data physicalization [9]. These new interactions can help us understand new ways for people to have natural conversations about S&P [105], embed questions and complaints directly into technical systems [134], and facilitate more positive social alignment. In doing so, we should be prepared to accept that the conversations and reflections that people have may not align with our norms or expectations.

Finally, as we noted in the previous section, the standards and toolkit themselves should be targets of continual inspection and updating. Future work on *developing the standards and toolkit* could consider values such as accessibility and inclusion—such as through the lens of differential vulnerabilities [87, 102], as Das Chowdhury et al. already have begun to do in their application of the capability approach to privacy-enhancing technologies [36]. In doing so, we can model critical and reflective behavior for non-experts.

7 Conclusion

In this paper, we introduced the concept of Transformational Provocation as a novel direction of research for the usable privacy and security community. Specifically, we prescribed for striving beyond community norms of improving users' S&P skills, knowledge, or behaviors, and attempting to create meaningful changes in people's senses of truth, self, relationships, and society as they relate to S&P. Through a lens of provocative design, we proposed specific paths of future work for implementing our agenda. We also presented a toolkit for UPS researchers to practically design for Transformational Provocation. When designing future provocations, we should consider their *velocity*, or the speed at which a user is expected to transform after being provoked; their *interiority and exteriority*, or the degree to which it might induce transformations outside of an individual user; and their *contextual potential*, or the possible levels of critical consciousness sparked by that provocation. Through engaging with Transformational Provocation, we hope the UPS research community can better align itself with user interests and motivations to create enduring, organic change.

References

[1] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, et al. 2017. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)* 50, 3 (2017), 1–41.

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

[3] Mamtaj Akter, Leena Alghamdi, Dylan Gillespie, Nazmus Sakib Miazi, Jess Kropczynski, Heather Lipford, and Pamela J Wisniewski. 2022. Co-ops: A mobile app for community oversight of privacy and security. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing*. 179–183.

[4] Faisal Alotaibi, Steven Furnell, Ingo Stengel, and Maria Papadaki. 2017. Enhancing cyber security awareness with mobile games. In *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE, 129–134.

[5] Nazanin Andalibi, Cassidy Pyle, Kristen Barta, Lu Xian, Abigail Z Jacobs, and Mark S Ackerman. 2023. Conceptualizing Algorithmic Stigmatization. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[6] Scottie Andrew. 2024. Why young people are throwing PowerPoint parties. https://www.cnn.com/2024/03/25/us/tiktok-powerpoint-parties-presentation-nights-cec/index.html.

[7] Farzaneh Asgharpour, Debin Liu, and L Jean Camp. 2007. Mental models of security risks. In *Financial Cryptography and Data Security: 11th International Conference, FC 2007, and 1st International Workshop on Usable Security, USEC 2007, Scarborough, Trinidad and Tobago, February 12-16, 2007. Revised Selected Papers 11*. Springer, 367–377.

[8] Brooke Auxier, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. 2019. Americans and privacy: Concerned, confused and feeling lack of control over their personal information. (2019).

[9] S Sandra Bae, Clement Zheng, Mary Etta West, Ellen Yi-Luen Do, Samuel Huron, and Danielle Albers Szafir. 2022. Making data tangible: A cross-disciplinary

design space for data physicalization. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.

[10] A Bandura. 1977. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* (1977).

[11] Sasha A Barab, Melissa Gresalfi, and Adam Ingram-Goble. 2010. Transformational play: Using games to position person, content, and context. *Educational researcher* 39, 7 (2010), 525–536.

[12] Natã M Barbosa, Gang Wang, Blase Ur, and Yang Wang. 2021. Who am i? a design probe exploring real-time transparency about online and offline user profiling underlying targeted ads. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–32.

[13] Shaowen Bardzell, Jeffrey Bardzell, Jodi Forlizzi, John Zimmerman, and John Antanitis. 2012. Critical design and critical theory: the challenge of designing for provocation. In *Proceedings of the designing interactive systems conference*. 288–297.

[14] Susanne Barth and Menno DT De Jong. 2017. The privacy paradox–Investigating discrepancies between expressed privacy concerns and actual online behavior–A systematic literature review. *Telematics and informatics* 34, 7 (2017), 1038–1058.

[15] Rosanna Bellini, Kevin Lee, Megan A Brown, Jeremy Shaffer, Rasika Bhalerao, and Thomas Ristenpart. 2023. The {Digital-Safety} Risks of Financial Technologies for Survivors of Intimate Partner Violence. In *32nd USENIX Security Symposium (USENIX Security 23)*. 87–104.

[16] W Lance Bennett and Alexandra Segerberg. 2012. The logic of connective action: Digital media and the personalization of contentious politics. *Information, communication & society* 15, 5 (2012), 739–768.

[17] Julian Bleecker. 2022. Design fiction: A short essay on design, science, fact, and fiction. *Machine Learning and the City: Applications in Architecture and Urban Design* (2022), 561–578.

[18] Alicia Blum-Ross and Sonia Livingstone. 2020. "Sharenting," parent blogging, and the boundaries of the digital self. In *Self-(re) presentation now*. Routledge, 70–85.

[19] Laurens Boer and Jared Donovan. 2012. Provotypes for participatory innovation. In *Proceedings of the designing interactive systems conference*. 388–397.

[20] Emma Bowman. 2022. As states ban abortion, the Texas bounty law offers a way to survive legal challenges. https://www.npr.org/2022/07/11/1107741175/texas-abortion-bounty-law.

[21] Barry Brown. 2001. Studying the internet experience. *HP laboratories technical report HPL* 49 (2001).

[22] Anders Bruun, Rikke Hagensby Jensen, Jesper Kjeldskov, Jeni Paay, Camilla Mejlby Hansen, Katarína Leci Sakácová, and Mette Hyllested Larsen. 2020. Exploring the non-use of mobile devices in families through provocative design. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 813–826.

[23] Rachel Brydolf-Horwitz. 2022. Embodied and entangled: Slow violence and harm via digital technologies. *Environment and Planning C: Politics and Space* 40, 2 (2022), 391–408.

[24] Karoline Busse, Julia Schäfer, and Matthew Smith. 2019. Replication: No one can hack my mind revisiting a study on expert and {Non-Expert} security practices and advice. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 117–136.

[25] Rafael A Calvo and Dorian Peters. 2014. *Positive computing: technology for wellbeing and human potential*. MIT press.

[26] Janet X Chen, Allison McDonald, Yixin Zou, Emily Tseng, Kevin A Roundy, Acar Tamersoy, Florian Schaub, Thomas Ristenpart, and Nicola Dell. 2022. Trauma-informed computing: Towards safer technology experiences for all. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–20.

[27] Tianying Chen, Margot Stewart, Zhiyu Bai, Eileen Chen, Laura Dabbish, and Jessica Hammer. 2020. Hacked time: Design and evaluation of a self-efficacy based cybersecurity game. In *Proceedings of the 2020 acm designing interactive systems conference*. 1737–1749.

[28] Chhaya Chouhan, Christy M LaPerriere, Zaina Aljallad, Jess Kropczynski, Heather Lipford, and Pamela J Wisniewski. 2019. Co-designing for community oversight: Helping people make privacy and security decisions together. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–31.

[29] Danielle Keats Citron. 2007. Technological due process. *Wash. UL Rev.* 85 (2007), 1249.

[30] Lizzie Coles-Kemp, Rikke Bjerg Jensen, and Claude PR Heath. 2020. Too much information: Questioning security in a post-digital society. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[31] Kate Crawford and Jason Schultz. 2014. Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev.* 55 (2014), 93.

[32] Sabrina Culyba. 2018. *The Transformational Framework: A process tool for the development of Transformational games*. Carnegie Mellon University.

[33] Jay Cunningham, Gabrielle Benabdallah, Daniela Rosner, and Alex Taylor. 2023. On the grounds of solutionism: Ontologies of blackness and HCI. *ACM Transactions on Computer-Human Interaction* 30, 2 (2023), 1–17.

[34] Adrian Dabrowski, Markus Kammerstetter, Eduard Thamm, Edgar Weippl, and Wolfgang Kastner. 2015. Leveraging competitive gamification for sustainable fun and profit in security education. In *2015 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 15)*.

[35] Sauvik Das, Cori Faklaris, Jason I Hong, Laura A Dabbish, et al. 2022. The security & privacy acceptance framework (spaf). *Foundations and Trends® in Privacy and Security* 5, 1-2 (2022), 1–143.

[36] Partha Das Chowdhury, Andrés Domínguez Hernández, Kopo Marvin Ramokapane, and Awais Rashid. 2022. From utility to capability: A new paradigm to conceptualize and develop inclusive pets. In *Proceedings of the 2022 New Security Paradigms Workshop*. 60–74.

[37] Thom Davies. 2018. Toxic space and time: Slow violence, necropolitics, and petrochemical pollution. *Annals of the American Association of Geographers* 108, 6 (2018), 1537–1553.

[38] Claudia Deane. 2024. Americans' Deepening Mistrust of Institutions. *Pew Research Center* (2024).

[39] Pieter MA Desmet and Anna E Pohlmeyer. 2013. Positive design: An introduction to design for subjective well-being. *International journal of design* 7, 3 (2013).

[40] Matthew A Diemer, Luke J Rapa, Catalina J Park, and Justin C Perry. 2017. Development and validation of the critical consciousness scale. *Youth & Society* 49, 4 (2017), 461–483.

[41] Youngwook Do, Nivedita Arora, Ali Mirzazadeh, Injoo Moon, Eryue Xu, Zhihan Zhang, Gregory D Abowd, and Sauvik Das. 2023. Powering for privacy: improving user trust in smart speaker microphones with intentional powering and perceptible assurance. In *32nd USENIX Security Symposium (USENIX Security 23)*. 2473–2490.

[42] Daniel M Dooghan. 2023. Fantasies of Adequacy: Mythologies of Capital in Dark Souls. *Games and Culture* (2023), 15554120231192080.

[43] Anthony Dunne and Fiona Raby. 2021. *Design noir: The secret life of electronic objects*. Vol. 2. Bloomsbury Publishing.

[44] Cédric Durand. 2024. *How silicon valley unleashed techno-feudalism: the making of the digital economy*. Verso Books.

[45] Megan Farokhmanesh. 2025. 'My Chinese Spy' Memes Show Americans Aren't Sold on the TikTok Ban. https://www.wired.com/story/tiktok-ban-my-chinese-spy-meme/.

[46] Federal Trade Commission. 1980. FTC Policy Statement on Unfairness. https://www.ftc.gov/legal-library/browse/ftc-policy-statement-unfairness.

[47] National Institute for Standards and Technology (NIST). 2023. *Special Publication 800-63B: Usability Standards*. Technical Report. National Institute for Standards and Technology (NIST).

[48] Sarah Fox, Noura Howell, Richmond Wong, and Franchesca Spektor. 2019. Vivewell: speculating near-future menstrual tracking through current data practices. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. 541–552.

[49] Paulo Freire. 2020. Pedagogy of the oppressed. In *Toward a sociology of education*. Routledge, 374–386.

[50] Liza Gak, Seyi Olojo, and Niloufar Salehi. 2022. The distressing ads that persist: Uncovering the harms of targeted weight-loss ads among users with histories of disordered eating. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–23.

[51] Zoya Garg, Elmer Gomez, and Luciana Yael Petrzela. 2019. Opinion: If You Didn't "Sharent," Did You Even Parent? https://www.nytimes.com/2019/08/07/opinion/parents-social-media.html.

[52] R Gilmore, J Beezhold, V Selwyn, R Howard, I Bartolome, and N Henderson. 2022. Is TikTok increasing the number of self-diagnoses of ADHD in young people? *European Psychiatry* 65, S1 (2022), S571–S571.

[53] José González Cabañas, Ángel Cuevas, and Rubén Cuevas. 2017. Fdvt: Data valuation tool for facebook users. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 3799–3809.

[54] Andre Gorz. 1967. *Strategy for Labor: A Radical Proposal*. Beacon Press.

[55] Mark S Granovetter. 1973. The strength of weak ties. *American journal of sociology* 78, 6 (1973), 1360–1380.

[56] Kishonna L Gray, Bertan Buyukozturk, and Zachary G Hill. 2017. Blurring the boundaries: Using Gamergate to examine "real" and symbolic violence against women in contemporary gaming culture. *Sociology compass* 11, 3 (2017), e12458.

[57] Barbara Grosse-Hering, Jon Mason, Dzmitry Aliakseyeu, Conny Bakker, and Pieter Desmet. 2013. Slow design for meaningful interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3431–3440.

[58] Johanna Gunawan, Cristiana Santos, and Irene Kamara. 2022. Redress for dark patterns privacy harms? A case study on consent interactions. In *Proceedings of the 2022 symposium on computer science and law*. 181–194.

[59] Tara Haelle. 2016. Do Parents Invade Children's Privacy When They Post Photos Online? https://www.npr.org/sections/health-shots/2016/10/28/499595298/do-parents-invade-childrens-privacy-when-they-post-photos-online.

[60] Rakibul Hasan, Bennett I Bertenthal, Kurt Hugenberg, and Apu Kapadia. 2021. Your photo is so funny that i don't mind violating your privacy by sharing it: effects of individual humor styles on online photo-sharing behaviors. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[61] Rakibul Hasan, Rebecca Weil, Rudolf Siegel, and Katharina Krombholz. 2023. A psychometric scale to measure individuals' value of other people's privacy (VOPP). In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.

[62] Cormac Herley. 2009. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proceedings of the 2009 workshop on New security paradigms workshop*. 133–144.

[63] Cormac Herley. 2013. More is not the answer. *IEEE Security & Privacy* 12, 1 (2013), 14–19.

[64] Naja Holten Møller, Trine Rask Nielsen, and Christopher Le Dantec. 2021. Work of the Unemployed: An inquiry into individuals' experience of data usage in public services and possibilities for their agency. In *Designing Interactive Systems Conference 2021*. 438–448.

[65] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 17–24.

[66] Iulia Ion, Rob Reeder, and Sunny Consolvo. 2015. {"... No} one Can Hack My {Mind"}: Comparing Expert and {Non-Expert} Security Practices. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*. 327–346.

[67] Justyna Janik. 2019. Ghosts of the present past: Spectrality in the video game object. *Journal of the Philosophy of Games* 2, 1 (2019).

[68] Alexis Jemal. 2017. Critical consciousness: A critique and critical analysis of the literature. *The Urban Review* 49 (2017), 602–626.

[69] Ruogu Kang, Laura Dabbish, Nathaniel Fruchter, and Sara Kiesler. 2015. {"My} data just goes {Everywhere:"} user mental models of the internet and implications for privacy and security. In *Eleventh symposium on usable privacy and security (SOUPS 2015)*. 39–52.

[70] Darya Kaviani and Niloufar Salehi. 2022. Bridging action frames: Instagram infographics in US ethnic movements. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–43.

[71] Rilla Khaled. 2018. Questions over answers: Reflective game design. *Playful disruption of digital media* (2018), 3–27.

[72] Know Your Meme. [n. d.]. Government Agent Watching Me. https://knowyourmeme.com/memes/government-agent-watching-me.

[73] Spyros Kokolakis. 2017. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security* 64 (2017), 122–134.

[74] Adam DI Kramer, Jamie E Guillory, and Jeffrey T Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24 (2014), 8788–8790.

[75] Rory Kramer and Brianna Remster. 2022. The slow violence of contemporary policing. *Annual Review of Criminology* 5, 1 (2022), 43–66.

[76] Isadora Krsek, Shuyang Chen, Sauvik Das, Jason I Hong, and Laura Dabbish. 2024. CALYPSO: Designing Gameplay Features for Improving Player's Cybersecurity Self-Efficacy. In *Companion Proceedings of the 2024 Annual Symposium on Computer-Human Interaction in Play*. 372–377.

[77] Priya Kumar, Shalmali Milind Naik, Utkarsha Ramesh Devkar, Marshini Chetty, Tamara L Clegg, and Jessica Vitak. 2017. 'No Telling Passcodes Out Because They're Private' Understanding Children's Mental Models of Privacy and Security Online. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–21.

[78] Leona Lassak, Elleen Pan, Blase Ur, and Maximilian Golla. 2024. Why Aren't We Using Passkeys? Obstacles Companies Face Deploying FIDO2 Passwordless Authentication. In *33th USENIX Security Symposium (USENIX Security 24)*.

[79] Hao-Ping Lee, Yu-Ju Yang, Thomas Serban Von Davier, Jodi Forlizzi, and Sauvik Das. 2024. Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[80] Junchao Lin, Jason I Hong, and Laura Dabbish. 2021. " It's our mutual responsibility to share" The Evolution of Account Sharing in Romantic Couples. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–27.

[81] Heather Richter Lipford and Mary Ellen Zurko. 2012. Someone to watch over me. In *Proceedings of the 2012 New Security Paradigms Workshop*. 67–76.

[82] Amy Lyndon, Jennifer Bonds-Raacke, and Alyssa D Cratty. 2011. College students' Facebook stalking of ex-partners. *Cyberpsychology, behavior, and social networking* 14, 12 (2011), 711–716.

[83] James E Maddux and Ronald W Rogers. 1983. Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of experimental social psychology* 19, 5 (1983), 469–479.

[84] Peter Mayer, Yixin Zou, Byron M Lowens, Hunter A Dyer, Khue Le, Florian Schaub, and Adam J Aviv. 2023. Awareness, Intention,(In) Action: Individuals' Reactions to Data Breaches. *ACM Transactions on Computer-Human Interaction* 30, 5 (2023), 1–53.

[85] Peter Mayer, Yixin Zou, Florian Schaub, and Adam J Aviv. 2021. " Now I'm a bit {angry:}" Individuals' Awareness, Perception, and Responses to Data Breaches that Affected Them. In *30th USENIX Security Symposium (USENIX Security 21)*. 393–410.

[86] Colleen McClain, Michelle Faverio, Monica Anderson, and Eugenie Park. 2023. How Americans view data privacy. *Pew Research Center* (2023).

[87] Nora McDonald and Andrea Forte. 2020. The politics of privacy theories: Moving from norms to vulnerabilities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[88] Susan Michie, Maartje M Van Stralen, and Robert West. 2011. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science* 6 (2011), 1–12.

[89] Josh Aaron Miller, Kutub Gandhi, Matthew Alexander Whitby, Mehmet Kosa, Seth Cooper, Elisa D Mekler, and Ioanna Iacovides. 2024. A Design Framework for Reflective Play. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–21.

[90] Eyitemi Moju-Igbene, Hanan Abdi, Alan Lu, and Sauvik Das. 2022. " How Do You Not Lose Friends?": Synthesizing a Design Space of Social Controls for Securing Shared Digital Resources Via Participatory Design Jams. In *31st USENIX Security Symposium (USENIX Security 22)*. 881–898.

[91] Evgeny Morozov. 2013. *To save everything, click here: The folly of technological solutionism*. PublicAffairs.

[92] Amy Muise, Emily Christofides, and Serge Desmarais. 2009. More information than you ever wanted: Does Facebook bring out the green-eyed monster of jealousy? *CyberPsychology & behavior* 12, 4 (2009), 441–444.

[93] Savanthi Murthy, Karthik S Bhat, Sauvik Das, and Neha Kumar. 2021. Individually vulnerable, collectively safe: The security and privacy practices of households with older adults. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–24.

[94] Rob Nixon. 2011. *Slow Violence and the Environmentalism of the Poor*. Harvard University Press.

[95] Patricia A Norberg, Daniel R Horne, and David A Horne. 2007. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs* 41, 1 (2007), 100–126.

[96] Norbert Nthala and Emilee Rader. 2020. Towards a conceptual model for provoking privacy speculation. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[97] Borke Obada-Obieh, Yue Huang, and Konstantin Beznosov. 2020. The burden of ending online account sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[98] Deger Ozkaramanli and Pieter Desmet. 2016. Provocative design for unprovocative designers: Strategies for triggering personal dilemmas. In *2016 Design Research Society 50 Anniversary Conference*. The Design Research Society, 2001–2016.

[99] Cheul Young Park, Cori Faklaris, Siyan Zhao, Alex Sciuto, Laura Dabbish, and Jason Hong. 2018. Share and share alike? An exploration of secure behaviors in romantic relationships. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*. 83–102.

[100] Elizabeth Paton-Simpson. 2000. Privacy and the reasonable paranoid: The protection of privacy in public places. *The University of Toronto Law Journal* 50, 3 (2000), 305–346.

[101] Katharina Pfeffer, Alexandra Mai, Edgar Weippl, Emilee Rader, and Katharina Krombholz. 2022. Replication: Stories as informal lessons about security. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. 1–18.

[102] James Pierce, Sarah Fox, Nick Merrill, and Richmond Wong. 2018. Differential vulnerabilities and a diversity of tactics: What toolkits teach us about cybersecurity. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–24.

[103] Erika Shehan Poole, Marshini Chetty, Tom Morgan, Rebecca E Grinter, and W Keith Edwards. 2009. Computer help at home: methods and motivations for informal technical support. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 739–748.

[104] Emilee Rader and Janine Slaker. 2017. The importance of visibility for folk theories of sensor data. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. 257–270.

[105] Emilee Rader, Rick Wash, and Brandon Brooks. 2012. Stories as informal lessons about security. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*. 1–17.

[106] Iyad Rahwan. 2018. Society-in-the-loop: programming the algorithmic social contract. *Ethics and information technology* 20, 1 (2018), 5–14.

[107] Elissa M Redmiles, Jessica Bodford, and Lindsay Blackwell. 2019. "I just want to feel safe": A Diary Study of Safety Perceptions on Social Media. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 405–416.

[108] Elissa M Redmiles, Noel Warford, Amritha Jayanti, Aravind Koneru, Sean Kross, Miraida Morales, Rock Stevens, and Michelle L Mazurek. 2020. A comprehensive quality evaluation of security and privacy advice on the web. In *29th USENIX Security Symposium (USENIX Security 20)*. 89–108.

[109] Karen Renaud, Rosalind Searle, and Marc Dupuis. 2021. Shame in cyber security: effective behavior modification tool or counterproductive foil?. In *Proceedings of the 2021 New Security Paradigms Workshop*. 70–87.

[110] Ronald W Rogers. 1975. A protection motivation theory of fear appeals and attitude change1. *The journal of psychology* 91, 1 (1975), 93–114.

[111] M Angela Sasse, Jonas Hielscher, Jennifer Friedauer, and Annalina Buckmann. 2022. Rebooting it security awareness–how organisations can encourage and sustain secure behaviours. In *European Symposium on Research in Computer Security*. Springer, 248–265.

[112] John S Seberger, Marissel Llavore, Nicholas Nye Wyant, Irina Shklovski, and Sameer Patil. 2021. Empowering resignation: There's an app for that. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.

[113] John S Seberger, Irina Shklovski, Emily Swiatek, and Sameer Patil. 2022. Still creepy after all these years: The normalization of affective discomfort in app use. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–19.

[114] Ignacio Siles, Luciana Valerio-Alfaro, and Ariana Meléndez-Moran. 2024. Learning to like TikTok... and not: Algorithm awareness as process. *New Media & Society* 26, 10 (2024), 5702–5718.

[115] Daniel J Solove. 2021. The myth of the privacy paradox. *Geo. Wash. L. Rev.* 89 (2021), 1.

[116] Qiurong Song, Renkai Ma, Yubo Kou, and Xinning Gui. 2024. Collective Privacy Sensemaking on Social Media about Period and Fertility Tracking post Roe v. Wade. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–35.

[117] Marc Teyssier, Marion Koelle, Paul Strohmeier, Bruno Fruchard, and Jürgen Steimle. 2021. Eyecam: Revealing relations between humans and sensing devices through an anthropomorphic webcam. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[118] Richard H Thaler and Cass R Sunstein. 2003. Libertarian paternalism. *American economic review* 93, 2 (2003), 175–179.

[119] Alexandra To, Hillary Carey, Riya Shrivastava, Jessica Hammer, and Geoff Kaufman. 2022. Interactive fiction provotypes for coping with interpersonal racism. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–14.

[120] Alexandra To, Angela DR Smith, Dilruba Showkat, Adinawa Adjagbodjou, and Christina Harrington. 2023. Flourishing in the everyday: Moving beyond damage-centered design in HCI for BIPOC communities. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 917–933.

[121] Kami E Vaniea, Emilee Rader, and Rick Wash. 2014. Betrayed by updates: how negative experiences affect future security. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2671–2674.

[122] Yanis Varoufakis. 2024. *Technofeudalism: What killed capitalism*. Melville House.

[123] Nicholas Vincent, Hanlin Li, Nicole Tilly, Stevie Chancellor, and Brent Hecht. 2021. Data leverage: A framework for empowering the public in its relationship with technology companies. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 215–227.

[124] Rick Wash. 2010. Folk models of home computer security. In *Proceedings of the Sixth Symposium on Usable Privacy and Security*. 1–16.

[125] Rick Wash, Norbert Nthala, and Emilee Rader. 2021. Knowledge and capabilities that {Non-Expert} users bring to phishing detection. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 377–396.

[126] Hue Watson, Eyitemi Moju-Igbene, Akanksha Kumari, and Sauvik Das. 2020. " We Hold Each Other Accountable": Unpacking How Social Groups Approach Cybersecurity and Privacy Together. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[127] Miranda Wei, Madison Stamos, Sophie Veys, Nathan Reitinger, Justin Goodman, Margot Herman, Dorota Filipczuk, Ben Weinshel, Michelle L Mazurek, and Blase Ur. 2020. What Twitter knows: Characterizing ad targeting practices, user perceptions, and ad explanations through users' own Twitter data. In *29th USENIX Security Symposium (USENIX Security 20)*. 145–162.

[128] Miranda Wei, Eric Zeng, Tadayoshi Kohno, and Franziska Roesner. 2022. {Anti-Privacy} and {Anti-Security} Advice on {TikTok}: Case Studies of {Technology-Enabled} Surveillance and Control in Intimate Partner and {Parent-Child} Relationships. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. 447–462.

[129] Ben Weinshel, Miranda Wei, Mainack Mondal, Euirim Choi, Shawn Shan, Claire Dolin, Michelle L Mazurek, and Blase Ur. 2019. Oh, the places you've been! User reactions to longitudinal transparency about third-party web tracking and inferencing. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. 149–166.

[130] Ryan West, Christopher Mayhorn, Jefferson Hardee, and Jeremy Mendel. 2009. The weakest link: A psychological perspective on why users make poor security decisions. In *Social and Human elements of information security: Emerging Trends and countermeasures*. IGI Global, 43–60.

[131] Alma Whitten and J Doug Tygar. 1999. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0.. In *USENIX security symposium*, Vol. 348. 169–184.

[132] Pamela Wisniewski, Haiyan Jia, Heng Xu, Mary Beth Rosson, and John M Carroll. 2015. " Preventative" vs." Reactive" How Parental Mediation Influences Teens' Social Media Privacy Behaviors. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 302–316.

[133] Richmond Y Wong, Deirdre K Mulligan, Ellen Van Wyk, James Pierce, and John Chuang. 2017. Eliciting values reflections by engaging privacy futures using design workbooks. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–26.

[134] Yuxi Wu, William Agnew, W. Keith Edwards, and Sauvik Das. 2025. Design (ing) Fictions for Collective Civic Reporting of Privacy Harms. *Proceedings of the ACM on Human-Computer Interaction* 9, CSCW2 (2025), 1–26.

[135] Yuxi Wu, Sydney Bice, W Keith Edwards, and Sauvik Das. 2023. The Slow Violence of Surveillance Capitalism: How Online Behavioral Advertising Harms People. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 1826–1837.

[136] Yuxi Wu, W Keith Edwards, and Sauvik Das. 2022. "A Reasonable Thing to Ask For"': Towards a Unified Voice in Privacy Collective Action. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[137] Yuxi Wu, W Keith Edwards, and Sauvik Das. 2022. SoK: Social Cybersecurity. In *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 1863–1879.

[138] Yixin Zou, Khue Le, Peter Mayer, Alessandro Acquisti, Adam J Aviv, and Florian Schaub. 2024. Encouraging users to change breached passwords using the protection motivation theory. *ACM Transactions on Computer-Human Interaction* 31, 5 (2024), 1–45.

[139] Shoshana Zuboff. 2023. The age of surveillance capitalism. In *Social theory re-wired*. Routledge, 203–213.