Six Ways to enact Privacy by Design: Cognitive Heuristics that predict Users’ Online Information Disclosure

S. Shyam Sundar  
Media Effects Research Laboratory  
College of Communications  
Pennsylvania State University  
University Park, PA 16802, USA  
sss12@psu.edu

Jinyoung Kim  
Media Effects Research Laboratory  
College of Communications  
Pennsylvania State University  
University Park, PA 16802, USA  
juk315@psu.edu

Andrew Gambino  
Media Effects Research Laboratory  
College of Communications  
Pennsylvania State University  
University Park, PA 16802, USA  
aug268@psu.edu

Mary Beth Rosson  
Center for Human–Computer Interaction  
College of Information Sciences and Technology  
Pennsylvania State University, University Park, PA 16802, USA  
mrosson@ist.psu.edu

Abstract  
The primary barrier preventing the implementation of Privacy by Design is a lack of understanding of user psychology pertaining to privacy-related decision-making in everyday online and mobile contexts. Through focus groups and a representative survey of US adults, we discovered six rules of thumb (or “cognitive heuristics”) employed by online users while making decisions about disclosing or withholding information. We describe these cognitive heuristics and propose that they be leveraged by the design community, by brainstorming design heuristics, in order to promote secure and trustworthy computing.

Author Keywords  
Privacy by design; privacy heuristics; information privacy; personal information disclosure

ACM Classification Keywords  
H.5.2. [Information interfaces and presentation]: User Interfaces; J.4. Social and Behavioral Sciences

Introduction  
Approaching privacy from a “design thinking” perspective [1] holds the promise of a preventive
approach to addressing privacy concerns, by avoiding data breaches and minimizing privacy risks through combined technical, design, and governance efforts [9]. According to Ann Cavoukian [2], the Ontario information and privacy commissioner who first coined the term “privacy by design” (PbD), the strongest protection of privacy might be achieved by following the seven foundational design principles, such as building data protection system into all interfaces as a default and being transparent in presenting privacy-related policies and procedures [1].

As Spiekermann [9] notes, PbD has been regarded as a panacea for alleviating privacy-related concerns, with designers seeking to simply add a few more security features to information systems. However, developing a “fault-proof” system that converts the PbD principles into real-world design elements has proven quite challenging for engineers as well as designers [13]. In fact, the majority of system developers and designers who attended the latest workshop (sponsored by The Computing Community Consortium, 2015) were of the consensus that it is difficult to put PbD into practice. They acknowledged that they lacked adequate design-related heuristics to follow when developing various online/mobile products. Although there exist many heuristics for evaluating the usability of interfaces, including Niels [6]’s heuristic-based evaluation check-list, we do not yet have a stable set of heuristics that can operationalize the seven principles of PbD.

Our position is that, in order to arrive at reliable design heuristics, we need a better understanding of user psychology as it pertains to online privacy and security. A greater focus on users’ criteria for online disclosure as well as withholding of information is necessary. This will inform us of common privacy pitfalls that can be obviated by good design. As we have argued before [11], user decision-making regarding privacy is seldom effortful, but often based on mental rules of thumb, called “cognitive heuristics.” Considering the cognitively demanding nature of many online transactions, especially in the mobile context, careful evaluations of system security are typically inhibited by users’ time constraints, lack of technological efficacy [4], and information overload [10]. Moreover, the complexity of privacy settings may hinder our adherence to the privacy protection ideals that we self-report in surveys. Therefore, our central contention is that online user behaviors do not always result from effortful thinking, but may occur due to expedient decision-making in the heat of the moment. However, this decision-making is not random, but based on predictable responses to specific cues on the interface [12].

Over 30 years of research in social cognition has shown that humans are “cognitive misers” [14] who make decisions that maximize efficiency at the cost of thoroughness. They do so based on cues that trigger cognitive heuristics (i.e., mental rules of thumb) to desired outcomes by reducing the need for careful analysis of each transaction. So, what are the key cognitive heuristics that drive privacy-related behaviors? The Modality-Agency-Interactivity-Navigability (MAIN) Model [12] formally identifies several different heuristics (i.e., rules of thumb) triggered by technological affordances, which lead to judgments of source and content credibility. These affordances manifest themselves on the interface as visually salient cues. Extended to the domain of online information disclosure, a variety of privacy and security heuristics can be triggered by various ‘cues’ on the
interface, which will affect users’ perception of the interface as well as their information disclosure behavior. For example, when a website solicits personal information, users are likely to be more forthcoming if the site belongs to a trusted authority, such as one’s own bank, a government agency (irs.gov), or well-known brand (Amazon.com). They will reveal information without carefully reading the content of the site or pausing to think about the consequences of their revelation of personal information. This is because they are implicitly applying the “authority” heuristic (an authority can be trusted with personal information).

Methods
Based on this theoretical premise, we report a battery of data-driven research findings that reveal such privacy-related rules of thumbs (i.e., cognitive heuristics) which may be affecting online and mobile users’ privacy-related decision-making. For this position paper, we have chosen six such rules or cognitive heuristics that we hope will lend themselves to design heuristics for practical use in PbD. These are derived from two studies—a large-scale national survey ($n = 786$) and focus-group interviews ($n = 41$). We conducted a survey of a national representative sample ($n = 786$) in the United States, measuring respondents’ level of agreement with six different heuristics pertaining to privacy, stated as general rules of thumb. Users’ agreement with each heuristic was measured after users had provided us their disclosure intentions in a series of scenarios representing common online and mobile contexts, including social networks, online shopping, online community, and mobile chatting. Each of these scenarios had a cue capable of triggering a heuristic. For the interview study, eight focus group interviews were conducted, with 41 participants (group sizes ranged from 2 to 9) taking part in the interviews. Three groups consisted of university students; five consisted of diverse groups of individuals recruited through an ad in a local newspaper.

If indeed heuristics predict disclosure decisions, then the respondents’ level of agreement with a given cognitive heuristic would predict their disclosure intention in a scenario featuring a cue that is hypothesized to trigger that heuristic. Support for this hypothesis would not only serve to capture the operation of heuristics in users’ privacy decision-making, but also guide designers to build more secure and trustworthy computing systems by focusing on interface cues that trigger the desired heuristic under a given situation. It will also identify the most prominent heuristics that underlie online information disclosure.

Findings
In both studies, the most dominant cues are those pertaining to social status or authority. They are important for assessing credibility of interfaces. For example, the presence of an authoritative source in the vicinity (e.g., logo of FDIC, or a trusted brand name) may assure users about the safety of the site, because users instinctively apply the authority heuristic (popular name, brand, or organization can guarantee the security of a website). In the survey study, a screenshot of a security seal (i.e., TRUSTe) on an online banking website was shown [see Figure 1], which implied online transactions within the site was secured by an authority. Users’ intention to provide their personal information (e.g., phone number, mailing address) to the site was measured by a single item (i.e., “How likely are you to provide your personal information to use online banking services?”) in each
scenario on a 7-point scale measure (1 = very unlikely; 7 = very likely; the same 7-point scale was used throughout the survey). Interestingly, the survey results showed that the greater the belief in the authority heuristic (stated above), the greater the likelihood of personal information revelation on the banking site, $\beta = .10$, $p < .001$. Likewise, participants in the interview sessions echoed the sentiment of reliance on authority cues, with one stating, “I look for that better business bureau seal, or those lock icons.” Additionally, authority cues were particularly helpful when users judged less known websites, as one participated stated “The unknown website linked back to Trip Adviser, so therefore we felt we could trust it.”

To better relieve users’ privacy concerns, several sites and apps have become proactive about explicitly displaying the details of their data management practices, which are often stated in their privacy policy. Urban et al. [7]'s consumer privacy survey reported that 62% percent of Internet users expressed little concern about their personal data when they could see the detailed conditions and terms of privacy policy. By explaining what and how users’ information is used, statements about privacy policy and explicit permission demonstration can imbue credibility [5]. We call this tendency to trust data holders due to their full disclosure of policy the **transparency heuristic** (if a website makes its privacy policy transparent, then my information is safe). It is a rule of thumb because it does not involve actual perusal of the content of the policy, and is based on the appearance of propriety. In the survey, a screenshot of privacy policy statement on a mobile application installation page was shown [Figure 2], which explained why and what user information would be collected and retained. Users’ intention to provide their personal information was then assessed using a statement (i.e., “How likely are you to provide your personal information to use this application?”). Similar to the authority heuristic, individuals who had greater belief in the transparency heuristic showed a higher tendency to allow access to their personal information (e.g., contact list, photos) when the mobile application explicitly explained the details of their data usage, $\beta = .26$, $p < .001$. Therefore, proactive statement of data handling policy can by itself serve as a credibility-enhancing cue, thereby encouraging greater disclosure of personal information.

Within the interview sessions, an emergent theme was users placing more trust when they felt their information disappeared. One participant’s description of this tendency showed their reliance on quick decision making, “I do feel that Snapchat’s servers are secure, because the information [disappears]. That gives me a sense of security, but I do know that someone could take a screen capture, which makes it less secure but I don’t think about that in the moment.” The rule of thumb that “if the information is not stored, but disappears, it is safe to share,” which we termed the **ephemerality heuristic**. This heuristic is coupled with the rise in popularity of messaging applications with ephemeral functions, such as Snapchat, Line and WeChat.

On the other hand, when information systems transfer user information to third parties, it can lower perceived security and lead to a reduction in disclosure tendencies. For example, the presence of unexpected personalization features on an interface (e.g., pop-up offers for products browsed on a different site) might
trigger the **fuzzy boundary heuristic** *(users’ online information is vulnerable)* [3, 11]. In contrast to the previous three heuristics (i.e., authority, transparency, and ephemerality), the fuzzy boundary heuristic and the next two rules of thumbs are negative heuristics because they are likely to discourage, rather than encourage, users’ information disclosure online. The fuzzy boundary heuristic was tested in the survey, using the context of backing up private information to an online storage service, such as iCloud and Microsoft OneDrive. Users’ intention to regularly synchronize their personal data to the online storage was assessed, using a statement (i.e., “How likely are you to sync your information to an online storage?”). Consistent with our expectation, the more people believed in the fuzzy boundary heuristic, the less likely they were to synchronize their personal information to the drive, $\beta = -.17, p < .001$. Participants in the interview sessions expounded on their experience with the fuzzy boundary phenomenon, citing feelings of “creepiness” and “uncertainty” when faced with targeted ads that were clearly based on covert tracking of their prior browsing.

Another extension of this tendency is the distrust of wireless networks, especially in public places, such as a café, library, or airport. Klasnja and his colleagues [8] found that a majority of Internet users expressed greater privacy concerns when they were on a wireless network. It appears that users see wireless transactions as being public, leaving them feeling vulnerable. The rule of thumb behind this **public-ness heuristic** *(It is not secure to manage my personal business in public)* is an old dictum: You don’t discuss private matters out in public. In the survey, respondents were asked to imagine that they were browsing an online shopping website on their smartphone while using free public WiFi at an airport and state their likelihood of making an online purchase in this situation (i.e., “How likely are you to make an online purchase using free public WiFi at an airport?). Similar to the pattern of results with the fuzzy boundary heuristic, the degree of belief in the public-ness heuristic was negatively associated with their intention to disclose personal information to make an online purchase using the public WiFi at the airport, $\beta = -.22, p < .001$. In the interview sessions, participants specifically identified behaviors, such as using public computer terminals, and even using LTE/4G networks as inherently less safe than their private machines and networks. One interview participant noted their concern with public networks stating “I feel unsafe in public. I don’t know... Someone could be sitting next to you, and have some kind of app on their phone and they can actually intercept your data.” Here, the participant identifies their fear of individuals and unknown systems that may be present within the public network.

The last security consideration has to do with the portability of mobile devices. Chin et al. [4] showed greater user concerns over saving personal information on mobile devices due to higher risks of being stolen, compared to larger, fixed machines. Users may be applying the **mobility heuristic** *(Mobile devices are inherently unsafe in protecting my information)* whenever they are reminded that they are on a mobile device, and may therefore avoid storing private information on portable devices. This heuristic speaks to the security of the physical device, which when invoked may encourage a preference for non-mobile devices or even nonphysical entities, such as a cloud storage. This heuristic was tested by asking the survey respondents to imagine using a tablet or mobile phone.
Their likelihood of making an online purchase with it was measured (i.e., "How likely are you to make an online purchase using a table PC or mobile phone?"). In line with our supposition, the more they agreed with the mobility heuristic, the less likely the were to reveal their private information to make an online purchase using mobile devices, $\beta = -.20$, $p < .001$. This sentiment was echoed by interview participants, though when probed further, the basis for these feelings were found to be quite thin. One participant bluntly shared “I don’t know why I would feel safer on my desktop, I just do,” which suggests that mobility played a role, even though the user lacked an explanation for it.

Discussion

Our data show that cognitive heuristics indeed play a significant role in predicting users’ tendency to instinctively trust (or distrust) certain online systems and divulge (or withhold) their private information. This provides support to the theory that privacy-related decision-making is driven by mental rules of thumb about online privacy and security.

The question is how do these heuristics get triggered in the interaction context. Our survey provided specific cues on the interface pertaining to the heuristic under investigation. Support for our hypotheses suggests that interface cues play a critical role in triggering mental rules of thumb among users, which in turn dictate their disclosure behaviors.

In our analysis we have uncovered six cognitive heuristics that may inform PbD practices. While the cognitive heuristics refer to user behavior, for designers, these rules of thumb can be leveraged for ethical design practices to build safer systems. However, analysis of the individual heuristics reveals tensions and dilemmas faced when trying to employ PbD.

Designers may utilize the authority heuristic by merely asking “does this interface feature any appeals to authority?” Because the authority heuristic leads to more information sharing, the placement of authority cues needs to be handled with care. Coupling a seemingly innocuous symbol of authority with an information entry page may lead to users exhibiting reckless behaviors.

The transparency heuristic brings to light a dilemma for system design. While inherently we want to design more digestible privacy statements and user agreements, it appears as though full disclosure engenders a feeling of trust in the user. This causes a tension between user experience and ethical design practice.

Design features that utilize any sense of information disappearance may give their users a false sense of security. The ephemerality heuristic can be implemented as a design principle, and a sound one, if it actually entails the practice of data minimization. However, if data are actually stored and the function is used merely to increase user enjoyment, then the practice is deceptive, and should be reconsidered.

Heuristics that inhibit information sharing, such as the fuzzy boundary heuristic, may be implemented as safeguard measures. Reminders on the interface, that a user’s information is being shared, may force users to rethink their risky habits.
However, the public-ness and mobility heuristics manifest partly because of where a user is situated. Though it may seem difficult to trigger, the underlying mechanism behind these heuristics may be a distrust of public networks (and as such, outside or unknown parties). Therefore, designers may seek to design systems that properly alert users to the public nature of their communication when they truly are at risk.

In sum, these six heuristics could be used to design appropriate nudges that serve to incentivize users to protect their privacy. We look forward to brainstorming ideas for doing so at the workshop.

Acknowledgement
This research is supported by the U. S. National Science Foundation (NSF) via Standard Grant No. CNS-1450500.

References

