Discovering Smart Home Internet of Things Privacy Norms Using Contextual Integrity

NOAH APTHORPE, Princeton University, USA
YAN SHVARTZSHNAIDER, New York University, USA and Princeton University, USA
ARUNESH MATHUR, Princeton University, USA
DILLON REISMAN, Princeton University, USA
NICK FEAMSTER, Princeton University, USA

The proliferation of Internet of Things (IoT) devices for consumer "smart" homes raises concerns about user privacy. We present a survey method based on the Contextual Integrity (CI) privacy framework that can quickly and efficiently discover privacy norms at scale. We apply the method to discover privacy norms in the smart home context, surveying 1,731 American adults on Amazon Mechanical Turk. For $2,800 and in less than six hours, we measured the acceptability of 3,840 information flows representing a combinatorial space of smart home devices sending consumer information to first and third-party recipients under various conditions. Our results provide actionable recommendations for IoT device manufacturers, including design best practices and instructions for adopting our method for further research.

CCS Concepts: • **Security and privacy → Human and societal aspects of security and privacy**; **Privacy protections**; • **Human-centered computing → Empirical studies in ubiquitous and mobile computing**;

Additional Key Words and Phrases: Contextual Integrity, Internet of Things, Privacy

ACM Reference Format:
Noah Apthorpe, Yan Shvartzshnaider, Arunesh Mathur, Dillon Reisman, and Nick Feamster. 2018. Discovering Smart Home Internet of Things Privacy Norms Using Contextual Integrity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 2, Article 59 (June 2018), 23 pages. https://doi.org/10.1145/3214262

1 INTRODUCTION

Internet of Things (IoT) devices for consumer "smart" homes can observe sensitive details about users' in-home activities and transmit this information on the Internet [14, 23, 31, 32]. The sensitivity of smart home data and the nascency of the IoT calls for effective scalable methods to discover and understand people's privacy norms regarding these devices. Understanding these privacy norms will allow manufacturers to design devices that consumers are comfortable incorporating into their homes and help government regulators and consumer advocates identify and contextualize privacy violations. Several previous studies have attempted to measure privacy norms pertaining to IoT devices (Section 7) but none discover privacy norms at scale in a manner that is based on a formal theory of privacy.

Authors' addresses: Noah Apthorpe, Princeton University, Computer Science Department, 35 Olden St, Princeton, NJ, 08540, USA, apthorpe@cs.princeton.edu; Yan Shvartzshnaider, New York University, Computer Science Department, 251 Mercer St, Room 305, New York, NY, 10012, USA, Princeton University, Computer Science Department, 35 Olden St, Princeton, NJ, 08540, USA, yansh@nyu.edu; Arunesh Mathur, Princeton University, Computer Science Department, 35 Olden St, Princeton, NJ, 08540, USA, amathur@cs.princeton.edu; Dillon Reisman, Princeton University, Computer Science Department, 35 Olden St, Princeton, NJ, 08540, USA, dreisman@princeton.edu; Nick Feamster, Princeton University, Computer Science Department, 35 Olden St, Princeton, NJ, 08540, USA, feamster@cs.princeton.edu.

2474-9567/2018/6-ART59
https://doi.org/10.1145/3214262

In this paper, we present a general, scalable survey method for discovering consumer privacy norms based on the *Contextual Integrity (CI)* privacy framework [44] (Section 3). CI is a well-established theory that defines privacy norms as the generally accepted appropriateness of specific information exchanges, or "information flows," in specific contexts. Information flows and associated contexts can be described using five parameters: sender, recipient, subject, attribute, and transmission principle. This precise formulation makes it possible to thoroughly investigate the combinatorial space of contextual information flows and associated privacy norms with an automated, large-scale survey on a crowdsourcing platform. Our use of CI also ensures that the method is repeatable, both for the same types of devices over time, as well as for entirely new classes of devices.

The method we develop is effective for discovering privacy norms in general. In this paper, we focus on applying the method to discover smart home privacy norms. We conducted a survey with a population of 1,731 adults from the United States on the Amazon Mechanical Turk (MTurk) platform. The survey cost $2,800 and allowed us to query the acceptability of 3,840 information flows involving smart home devices in less than six hours and identify associated privacy norms (Section 4). Our results provide insightful observations and actionable recommendations for IoT device manufacturers, regulators, and consumer advocates (Section 5).

Device manufacturers can use our survey method to perform their own research on how consumers might view the use of data that their products collect. We designed the method to make it easy to customize with new information flows and contexts, allowing manufacturers to discover privacy norms relevant to specific products, including ones we have not studied in this paper. The results will indicate whether existing or proposed devices may violate established privacy norms, providing an opportunity to preempt negative user feedback, public relations debacles, or regulatory scrutiny. The method is also relatively inexpensive, allowing manufacturers to investigate a broad set of privacy norms for a fraction of product development costs or conventional surveys.

Prevailing trends and patterns in our results suggest several best practices for smart home device manufacturers. Manufacturers should note that information flows not specifically related to core device features are universally viewed as egregious violations of privacy norms. For example, the results quantifiably demonstrate that a fitness tracker sending recorded audio is considerably less acceptable than the same device sending exercise data. Information flows to Internet service providers (ISPs), for advertising, or for indefinite storage were also especially unacceptable across all devices. In general, many types of information flows were viewed unfavorably by consumers, suggesting that setting default limits on the types of data collected and where that data is sent is a reasonable starting point for meeting consumer privacy expectations.

Regulators with rulemaking authority (e.g., the United States Federal Communication Commission, which has previously written rules on privacy [19]) could use our method to discover existing privacy norms and to better inform rulemaking. For example, we find that despite increased cultural awareness of the "smart home," the average survey respondent still views information flows from smart home devices to recipients outside of the home as generally unacceptable unless the device owner has specifically granted consent. In light of known issues with current consent mechanisms (privacy policies and EULAs) [41], this result further motivates the implementation of rules requiring granular consent options within clearly specified contexts [39].

Consumer advocacy groups could use our method to determine what device behaviors are of particular concern. This could help direct efforts to pressure manufacturers into changing practices. For example, our results found that social media accounts were one of the most unacceptable recipients of user information from smart home devices. This finding might motivate consumer advocacy groups to create guidelines for how smart home devices might better integrate data sharing capabilities with social media platforms.

In summary, this work makes the following contributions:

(1) We present a survey method for privacy norm discovery that integrates a formal theory of privacy with combinatorial testing at scale.

(2) We provide insights into existing privacy norms in the smart home context, contributing to the ongoing discussion about the evolving privacy impact of IoT technology [4, 10, 13].
(3) We provide actionable recommendations based on discovered privacy norms, including development best practices for device manufacturers and instructions for further research using our survey method.

We encourage others—including manufacturers, regulators, consumer advocates, and academics—to use our CI survey method for their own research. Follow-up studies could focus on different aspects of IoT devices, as well as technical innovation in other contexts (e.g., mobile application development). Repeated surveys using our method could also provide data for longitudinal analyses of how cultural privacy norms change in response to rapidly evolving technology.

2 CONTEXTUAL INTEGRITY BACKGROUND

The rapid pace of technological innovation introduces new ways to share and acquire information, forcing us to rethink established definitions of privacy. Conventional privacy definitions, such as Role-base Access Control (RBAC) [21, 22], Attributed Base Access Control (ABAC) [45], and Enterprise Privacy Authentication Language (EPAL) [3, 17], focus on classifying information types and establishing sophisticated access control mechanisms to prevent information from reaching unintended recipients. They do not incorporate the many contexts and norms that govern information exchange. In our daily activities, we traverse many environmental settings and take on many different roles. For example, many people work from home or during transit, often communicating personal information from work and work information from home.

CI defines privacy in terms of conformance of a given informational exchange to contextual information norms [44]. Norms come as a part of a context, or a specific setting established by law, policy, common practice, social pressures, and beliefs. An information flow that is misaligned with established norms potentially violates privacy expectations in the given context. CI is becoming more relevant as the lines between professional, personal, and public spaces are shifted and blurred by the increasingly rapid pace of information exchange.

CI describes norms in terms of five information flow parameters: (1) the *sender* of the information, (2) the *recipient* of the information, (3) the *attribute* or type of information, (4) the *subject* of the information, and (5) a *transmission principle* that states the condition under which the information flow is permitted. A change to any parameter may cause a privacy norm violation. For example, you (the sender/subject) might feel that it is acceptable to provide your doctor (recipient) with medical records (attribute) with the requirement of confidentiality (transmission principle). However, if the recipient is your boss, you might reconsider.

CI has been used in many legal and computer science studies as an alternative framework to define and reason about privacy [24, 51, 52]. The computer science community has focused on formal expression of contextual privacy norms to detect infraction using formal logic and to propose methods for accountability and enforcement [6, 9, 12]. Legal privacy scholars have used CI as a new lens to re-evaluate existing regulations, such as the meaning of "reasonable expectation of privacy" in the Fourth Amendment [49], and to find gaps in the privacy settings of existing systems, including Facebook [29] and Google [62]. As a common framework, CI has inadvertently fostered synergy between these communities, helping them advance the common goal of preserving technology users' privacy.

In addition to providing a way to capture and evaluate information flows against established privacy norms, CI can also capture new and evolving norms, such as when new technology is introduced. In this case, CI argues for reassessing information flow parameters with respect to "their merits as a function of their meaning and significance in relation to the aims, purposes, and values of the context" [44]. When toasters begin transmitting information to the cloud, they become senders in an entirely new set of information flows and associated privacy norms.

Table 1. CI parameter values chosen to generate smart home information flows. The *subject* parameter is not listed and was set to "its owner", referring to the owner and primary user of a device. We included a "null" transmission principle to generate unconditional information flows.

Sender	Recipient	Attribute	Transmission Principle
a sleep monitor	the local police	{subject}'s location	if {subject} has given consent
a security camera	government intelligence agencies	{subject}'s eating habits	if {subject} is notified
a door lock	{subject}'s doctor	the times {subject} is home	if the information is kept confidential
a thermostat	an Internet service provider	{subject}'s exercise routine	if the information is anonymous
a fitness tracker	its manufacturer	{subject}'s sleeping habits	if the information is used to perform maintenance on the device
a refrigerator	other devices in the home	audio of {subject}	
a power meter	{subject}'s immediate family	video of {subject}	if the information is used to provide a price discount
a personal assistant (e.g. Amazon Echo)	{subject}'s social media accounts	{subject}'s heart rate	if the information is used for advertising
		the times it is used	if the information is used to develop new features for the device
			if the information is not stored
			if the information is stored indefinitely
			if its privacy policy permits it
			in an emergency situation
			null (no transmission principle)

3 SURVEY METHOD

This section describes our method of discovering privacy norms by surveying information flow acceptability. We chose information flows relevant to smart home devices, but the method could be used to discover privacy norms in any context.

3.1 Selecting CI Information Flow Parameters

The first step in our survey method is to select CI information flow parameters (senders, recipients, attributes, subjects, and transmission principles) relevant to the context of interest. We chose parameters relevant to smart homes by surveying the academic literature and popular press coverage about consumer IoT devices (Table 1). These lists are not exhaustive, but they cover a range of smart home information flows and demonstrate the generality of our method. Device manufacturers, regulators, consumer advocates, or academic researchers could use our survey method with different parameters in order to discover privacy norms about specific devices or information collection practices.

3.1.1 Sender. We chose the list of senders as a variety of commercially available smart home IoT devices. The devices represent a range of potential privacy concerns, including physical presence (security camera and door lock), various behaviors (sleep monitor, refrigerator, and fitness tracker), and energy usage (power meter and thermostat). This list also includes device types that many consumers would have heard about in the popular press, including "a [smart] thermostat" (e.g., a Nest thermostat).

We chose not to provide specific device names, such as "Sense Sleep Monitor," to limit the effect of participants' opinions about specific companies on survey responses. "A personal assistant (e.g., Amazon Echo)" is an exception, because "a personal assistant" requires additional explanation to ensure participants envision the correct type of device.

3.1.2 Recipient. We chose the list of recipients to capture the range of first and third-parties that can obtain user information from existing smart home devices.

We included device manufacturers because they receive the most consumer information from IoT devices. Many devices communicate sensor recordings and user interactions directly to manufacturer-operated cloud servers. Device manufacturers usually provide privacy policies and enter end-user license agreements with consumers; however, these policies many not reflect generally accepted privacy norms.

We included the local police and government intelligence agencies in consideration of recent court cases involving data obtained from IoT and mobile devices [42].

We included ISPs because of recent scrutiny over ISP access to potentially private consumer data. The Federal Communications Commission's (FCC) Broadband Privacy Rules, passed in October 2016 and overturned by Congress in April 2017 [18], were intended to limit the ability of ISPs to collect and sell consumer information obtained by network monitoring. Existing academic literature also demonstrates that ISPs can infer user behaviors from IoT network traffic even when the traffic is encrypted [2].

We included other devices in the home because several large manufacturers have IoT ecosystems that cooperate to control multiple devices in a smart home. These ecosystems include Apple HomeKit, Samsung SmartThings, and the many devices that can be controlled via an Amazon Echo or Google Home. Communications between IoT devices raise additional privacy concerns, because devices may be made by different manufacturers (with different privacy policies) or may allow sensitive information to be inferred through the analysis of data from multiple devices.

We included the immediate family of a device owner in order to investigate the privacy norms of IoT devices acting as intermediaries in information flows otherwise acceptable in non-IoT contexts. A user's immediate family typically knows the user's sleeping habits and when the user is home. We were curious whether being able to learn this information about an immediate family member via an IoT device violates privacy norms.

We included social media accounts because some available IoT devices have the ability to post directly to user profiles (after user setup and password entry) [54]. We expect that more manufacturers will connect IoT devices to social networks in the future.

3.1.3 Attribute. We chose the list of attributes to incorporate a variety of information types that a recipient could obtain about a IoT device owner from device communications. Some of these attributes are raw sensor recordings (e.g., audio, video, location) with obvious privacy implications. Others are user behaviors that can be inferred from sensor recordings by the IoT device itself or through post-hoc analysis by the recipient (e.g., eating habits, sleeping habits, when the user is home, etc.).

3.1.4 Subject. We chose the IoT device owner as the only subject in our survey. This helped limit the number of questions and reduced cognitive fatigue for participants. "Owner" is less jargon than "user," and an IoT device's owner is the must likely subject of sensor and interaction information recorded by the device.

We also considered several additional subjects that could be the focus of follow-up studies. For example, the device owner's child, spouse, roommate, or guest are all individuals who could interact with IoT devices in a smart home and be the subject of information flows with different privacy norms.

3.1.5 Transmission Principle. The transmission principle parameter provides a condition under which an information flow occurs. Our list of transmission principles cover a variety of use cases for information from IoT devices. Some of these cases are specifically mentioned in many device privacy policies, such as "if the information is used to develop new features for the device" or "if the information is kept anonymous." Other transmission principles involve the storage of the information ("indefinitely" and "not stored") or previous actions by the subject ("if {subject} has given consent" and "if its privacy policy permits it"). We also included a *null* entry in our list of transmission principles as a control to measure the acceptability of unconditional information flows.

Variations in survey responses with varying transmission principles indicate that IoT privacy norms depend on how information will be used and under what circumstances it will be transmitted. Information about these

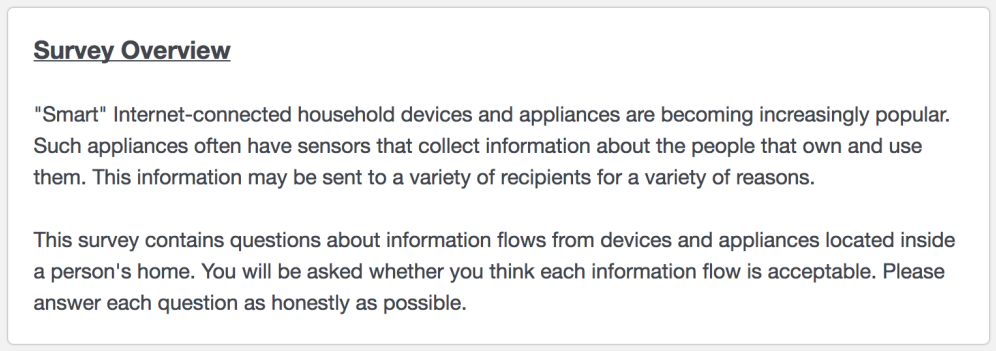

Fig. 1. Overview presented to participants at the beginning of the survey. The overview provides context and instructions without priming participants to view information flows as inherently acceptable or unacceptable.

conditional norm variations is important for IoT device manufacturers to avoid violating consumer trust and for policymakers to regulate the IoT in ways that align with consumer expectations.

3.2 Survey Design

The next step in our method is to create a survey that queries the acceptability of information flows generated from selected CI parameters. The following sections describe each part of our survey in detail.

3.2.1 Consent & Survey Overview. The first page of the survey contained a consent form. Participants who did not consent were prevented from taking the survey. The following page provided a brief background on the Internet of Things and an introduction to the CI questions (Figure 1).

3.2.2 Contextual Integrity Questions. The main section of the survey contained questions about the acceptability of information flows generated from our lists of CI parameters (Table 1). We considered all possible combinations of the parameter lists and eliminated combinations that did not make sense given the current generation of IoT devices. For example, we ruled out all combinations with a refrigerator sending its owner's heart rate to any receiver under any transmission principle. While it is conceivable that an IoT refrigerator could be designed to measure a user's heart rate, we are unaware of any existing refrigerators with heart rate monitors. Filtering such less intuitive combinations allowed us to incorporate more parameter values into the survey while limiting the total number of questions. We ended up with 3,840 five-parameter information flows after filtering.

We next generated questions querying the acceptability of all 3,840 five-parameter information flows. 3,840 questions is obviously too many to ask an individual participant, so we divided the flows into 48 sets, where all information flows in a set had the same sender and attribute. Each participant was randomly assigned a set and asked questions about all flows in that set. For example, an individual participant might have been asked questions about a sleep monitor sending audio of its owner to various recipients under various transmission principles.

In addition to reducing the number of questions asked to each participant, grouping the information flows into sets by sender and attribute assured the independence of responses across receivers and transmission principles. By preventing participants from taking the survey more than once, we are certain that all participants answered questions about all transmission principles and all recipients. This made it possible for us to conduct significance testing across recipients and transmission principles.

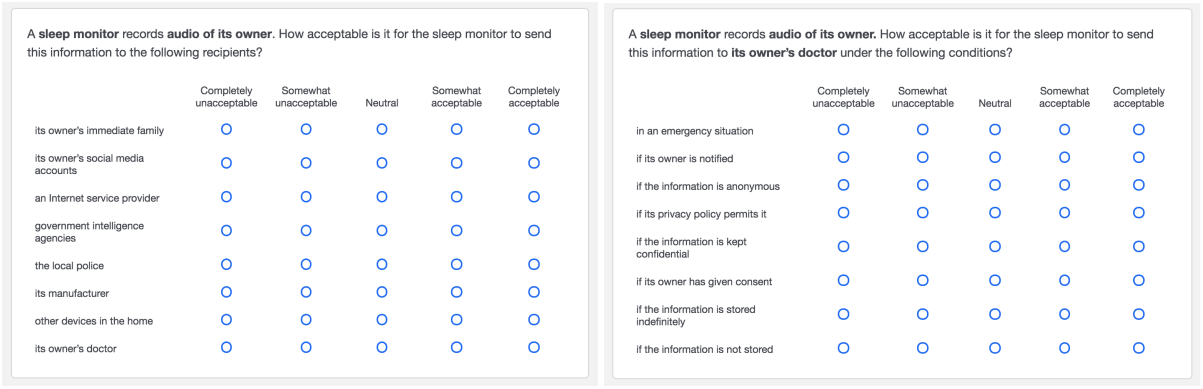

Fig. 2. Example matrix questions presented to survey respondents. Each respondent first saw one matrix question about unconditional information flows (*null* transmission principle) with varying recipients (*left*). The respondent then saw a series of questions for each recipient about information flows with varying transmission principles (*right*).

To further reduce participant cognitive load, we combined sets of multiple information flows into a question matrix format (Figure 2). The columns of each matrix were a five-point acceptability scale: "Completely Unacceptable", "Somewhat Unacceptable", "Neutral", "Somewhat Acceptable", and "Completely Acceptable." The rows of each matrix contained the parameter values that varied across the flows in the matrix. We randomized the order of the rows across matrices and participants. The non-varying parameters in each matrix were set in bold font. Each participant rated 82 total information flows over 9 question matrices.

The first question matrix presented to each participant corresponded to information flows with varying receivers and the *null* transmission principle (i.e., always transmit). We always initiated the survey with this question to prevent priming effects of the other transmission principles. The remaining question matrices presented to each participant corresponded to information flows with a fixed recipient and varying transmission principles. The order of these matrices was randomized for each participant. Figure 2 shows an example of both question matrix formats. We also inserted an attention check question ("Select 'Somewhat Acceptable'") as one row of one question matrix on each participant's survey.

3.2.3 Demographics and Technical Background Questions. The final section of the survey had a series of demographics and technical background questions presented to every participant, followed by a field for optional open-ended comments. The Appendix details the self-reported demographics of survey participants. The participants were split equally between males and females, on terms with the general US Internet user population [61]. A large fraction of the participants had a Bachelor's degree or higher and were also younger compared to the US population [61]. These differences are largely due to the MTurk platform, which tends to attract more technology-savvy users [27]. Further, 36% of survey respondents report owning at least one smart home device ("a 'smart' (Internet-connected) device or appliance in [their] home besides a smartphone, tablet, laptop, or desktop computer"). 78% of the respondents who reported owning at least one smart home device set up the device(s) themselves.

3.3 Survey Deployment

We created and hosted our survey using the Qualtrics [47] platform. The survey content and participant recruitment process were approved by our institution's Institutional Review Board (IRB). We tested the survey using UserBob [57], a usability-testing service that recruits MTurk workers to record their screen while interacting with

a website and providing audio commentary about their experience. We collected five 8–15 minute interactions with our survey from UserBob. All five UserBob workers completed the survey within 15 minutes (even while recording audio feedback). We used their feedback and misunderstandings to subsequently refine question wordings. We did not include responses from the UserBob workers in the final results. This practice, called "cognitive interviews" [56], is common in survey design and development.

We administered the final version of the survey using TurkPrime [35], an online service for researchers to easily create and manage MTurk Human Intelligence Tasks (HITs). We recruited 2,000 total participants and paid each participant $1.00 for completing the survey. Running the entire survey cost $2,800 and took less than six hours to complete. The survey was limited to workers in the US with a 90–100% HIT approval rating. We did not impose any limit on the prior number of approved HITs. Workers were automatically paid upon entering a completion code displayed at the end of the survey. Collecting all 2,000 responses took less than six hours. To avoid priming workers during task selection, we advertised the survey as a "Home Technology Survey," not as a survey specifically about privacy.

3.4 Response Analysis

The results from our survey method are amenable to a variety of analysis techniques. We employed the following straightforward aggregation approach to discover privacy norms pertaining to specific CI parameters. Adopters of our method could apply other analysis techniques, including multivariable modeling, to explore more complex relationships in the survey response data.

3.4.1 Filtering & Scaling Responses. We first filtered the responses to remove those from 269 participants who misanswered the attention check question. This yielded a total of 1731 responses with an average of 37 responses per CI matrix question (standard deviation 1.3). We used a Likert scale for the responses with the following Likert items: Completely Acceptable (2), Somewhat Acceptable (1), Neutral (0), Somewhat Unacceptable (−1), Completely Unacceptable (−2).

3.4.2 Average Acceptability Scores. We discover privacy norms by analyzing how different combinations of CI parameters affect survey responses. We grouped all information flows with the same sender and attribute and all information flows with the same transmission principle and recipient. These sender/attribute and recipient/transmission principle groupings coordinated with how the questions were presented to participants (Section 3.2.2). Each recipient/transmission principle group contains flows with all tested pairs of senders and attributes, while each sender/attribute group contains flows with all tested pairs of recipients and transmission principles. Every participant scored exactly one information flow in every recipient/transmission principle group, while every participant scored all flows in exactly one sender/attribute group.

We then computed the average[1] Likert acceptability score of all flows in each group. This statistic allows us to compare the pairwise effects of CI parameters on information flow acceptability. If a parameter has a notably high or low average acceptability score across all pairwise groups (e.g., a sender with a high average acceptability score regardless of attribute), this implies that the parameter is individually important to participants' privacy norms. If a pair of parameters has a notably high or low average acceptability score, this implies that there is likely a privacy norm involving the interplay between the two parameters.

3.4.3 Significance Test. The division of CI survey questions into sets as described above ensures that responses across pairs of senders and attributes are independent. This allows us to perform non-parametric Wilcoxon signed-rank tests [36] to measure the effect of recipients and transmission principles.

To study the effect of transmission principle, we perform the Wilcoxon test to compare acceptability scores between pairs of informations flows with the same sender, attribute, and recipient, but with *null* and non-*null*

[1]There was little difference between average and median acceptability scores due to the limited dynamic range of the Likert scale.

transmission principles. For example, we compare scores of the flow "A fitness tracker records its owner's exercise routine and sends this information to its owner's doctor" against scores of the flow "A fitness tracker records its owner's exercise routine and sends this information to its owner's doctor *if its owner has given consent*."

To study the effect of recipient on information flow acceptability, we chose the recipient "its owner's immediate family" as a baseline. We perform the Wilcoxon test to compare acceptability scores of flow pairs with baseline and non-baseline recipients and the same other parameters. For example, we compare scores of the flow "A refrigerator records its owner's eating habits and sends this information to its owner's immediate family if its owner is notified" against scores of the flow "A refrigerator records its owner's eating habits and sends this information to *its manufacturer* if its owner is notified."

Since we conducted several such tests—one for each recipient and transmission principle—we accounted for multiple comparisons using the Bonferroni correction method [50]. We took the standard 0.05 cutoff threshold and divided it by the number of tests to generate the new threshold. For transmission principle, we use a threshold of 0.00008 (0.05 / 576). For recipient, we use a threshold of 0.0001 (0.05/336).

4 RESULTS

Our analysis of the survey responses provides insight into privacy norms in the smart home context. In this section, we describe the results of the analysis procedures described in Section 3.4.

4.1 Average Acceptability Scores

For each sender/attribute and recipient/transmission principle pair, we calculate the average acceptability score of all information flows containing that pair of parameters (Figure 3). The only parameters with positive average acceptability scores in any pair are the transmission principles "if its owner has given consent" (scores range from 0.58 to 1.38), "in an emergency situation" (−0.33 to 0.86), and "if its owner is notified" (−0.82 to 0.40) and the recipients "its owner's immediate family" (0.49 to 1.38) and "other devices in the home" (−1.24 to 1.35). The transmission principle "if its owner has given consent" is notable as the only parameter with positive pairwise acceptability scores regardless of recipient.

The parameters with the lowest pairwise average acceptability scores are the transmission principles "if the information is used for advertising" (−1.51 to −1.24) and "if the information is stored indefinitely" (−1.53 to −0.87) and the recipients "government intelligence agencies" (−1.52 to 0.58), "its owners social media accounts" (−1.51 to 0.84) and "an Internet service provider" (−1.53 to 0.83). Additionally, "if the information is used for advertising" and "if the information is stored indefinitely" are the only transmission principles with average acceptability scores across all flows containing those parameters below that of the same flows with the *null* transmission principle.

There is less variation in average acceptability score across sender/attribute pairs (−0.23 to −0.87) than across recipient/transmission principle pairs (−1.53 to 1.38). The senders "a power meter" (−0.51 to −0.23) and "a fitness tracker" (−0.53 to −0.27) are the most acceptable across attributes, while "a security camera" (−0.87 to −0.26) and "a refrigerator" (−0.87 to −0.26) are the least.

Interactions between CI parameters also have noticeable effects on acceptability scores. For example, the recipient "other devices in the home" has positive average acceptability scores when paired with six transmission principles but negative scores when paired with five other transmission principles. The transmission principle "in an emergency situation" has positive average acceptability scores when paired with all recipients except "government intelligence agencies." The attribute "its owner's eating habits" has a less negative score when paired with the sender "a refrigerator" (−0.35) than its scores with all other senders (−0.73 to −0.53). These cases highlight that CI parameters are not independent, but combine to to create meaningful contexts relevant to privacy norms.

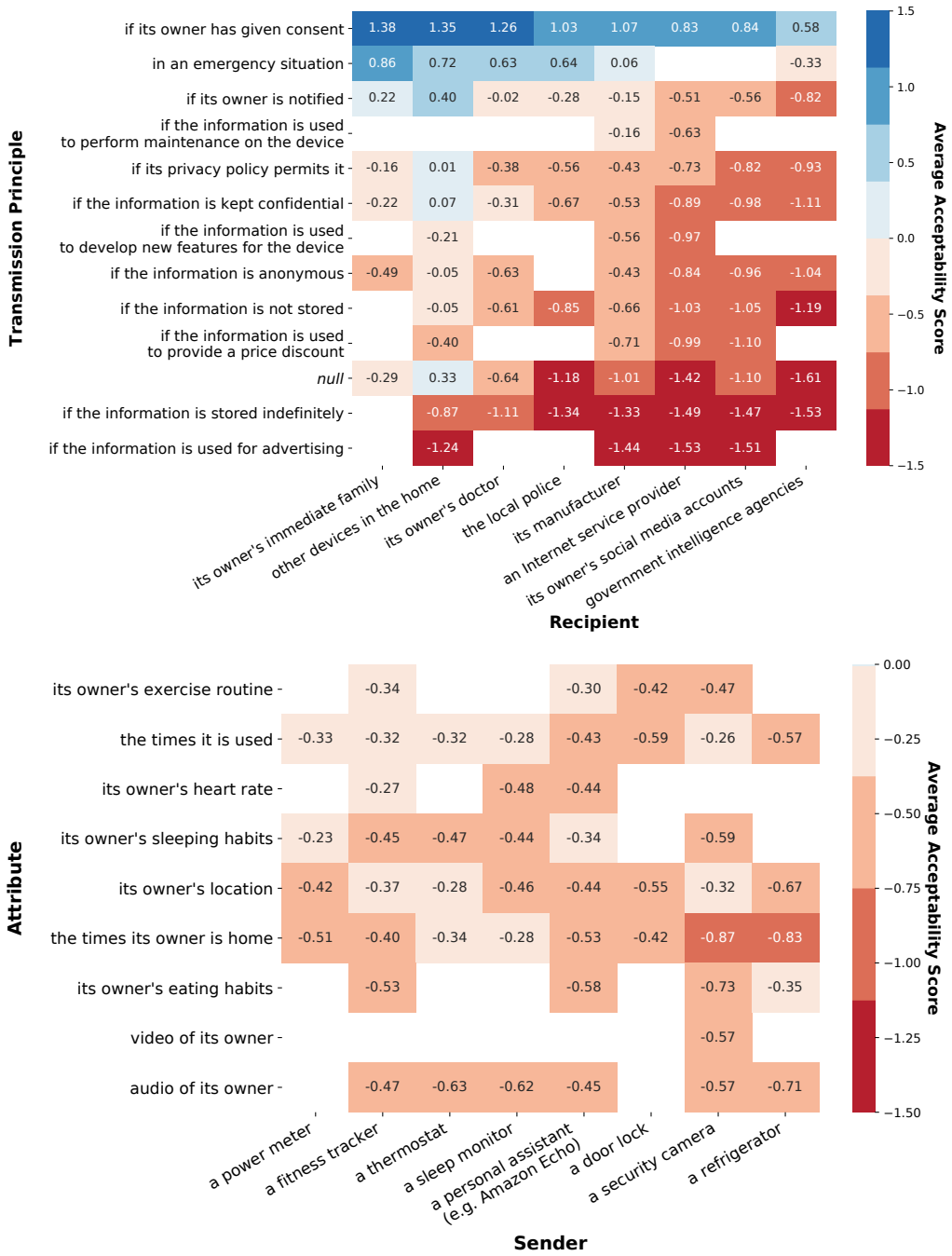

Fig. 3. Average acceptability scores of information flows with given recipient/transmission principle pairs (*top*) and given sender/attribute pairs (*bottom*). Empty locations correspond to less intuitive information flows that were excluded in the survey. Parameters are sorted from top to bottom and left to right by descending average acceptability score for *all* information flows containing that parameter.

4.2 Effect of Transmission Principle and Recipient on Information Flow Acceptability

Using the significance test described in Section 3.4.3, we evaluate the effect of adding transmission principles to unconditional information flows. We calculate the percentage of instances for which the inclusion of a particular transmission principle (as opposed to the *null* transmission principle) results in a statistically significant difference in average acceptability score (Figure 4). We also perform the same procedure for recipients, comparing average acceptability scores to information flows with the baseline recipient "its owner's immediate family" (Figure 4).

The transmission principles "if its owner has given consent," "in an emergency situation," and "if its owner is notified" result in significantly different scores in $\geq 85\%$ of instances. This indicates that these parameters heavily influenced participants reactions to information flows. In comparison, the transmission principles "if the information is not stored," "if the information is used to provide a price discount," and "if the information is anonymous," resulted in significantly different scores in $< 10\%$ of instances.

The recipients "government intelligence agencies," "an Internet service provider," and "its owner's social media accounts" result in significantly different scores in $\geq 98\%$ of instances. In comparison, the recipient "other devices in the home" resulted in significantly different scores in only 12% of instances.

4.3 Effect of Smart Home Device Ownership

Figure 5 compares the average acceptability scores of information flows with each transmission principle between the 36% of respondents who self-reported that they own at least 1 smart home device and the 62% of respondents who own 0 devices. The remaining 2% of respondents who answered "I don't know" to the device ownership question were omitted from this comparison.

Owning a smart home device generally increased participants' average acceptability scores by a small amount. This effect is most pronounced on information flows with the transmission principles "if the information is used to develop new features for the device" ($\Delta 0.17$) and "if the information is used to perform maintenance on the device" ($\Delta 0.17$). The strongest exception is the transmission principle "if the information is stored indefinitely" which received higher acceptability scores from participants who do not claim to own any smart home devices ($\Delta 0.14$). However, there are no transmission principles for which owning smart home devices changed the average acceptability score from negative to positive or vice versa.

5 OBSERVATIONS AND RECOMMENDATIONS

Analysis of our survey responses yields rich insights into IoT privacy norms in the home context. The following discussion synthesizes our results into observations and recommendations for IoT device manufacturers, policy-makers and regulators. These sections, like our choice of information flow parameters, favor breadth over depth in order to demonstrate the range of issues that can be better understood using our survey method. We hope that this potential will inspire others to adopt our survey technique and perform their own investigations of contextual privacy norms.

5.1 Device Manufacturers Should Survey Contextual Privacy Norms

Suppose a smart home device manufacturer is designing a product that will collect data about its users and communicate with first and third parties to support a variety of features and revenue sources. The manufacturer can use its complete knowledge of the device's intended behavior to generate very specific five-parameter information flows involving the device. The manufacturer can then use our survey method to determine if any of these information flows disrupt established contextual norms. If so, the manufacturer can modify device behavior during the design process before releasing a device that might be subject to consumer backlash. Even if a manufacturer is unsure what information flows are relevant for their devices, our survey method allows

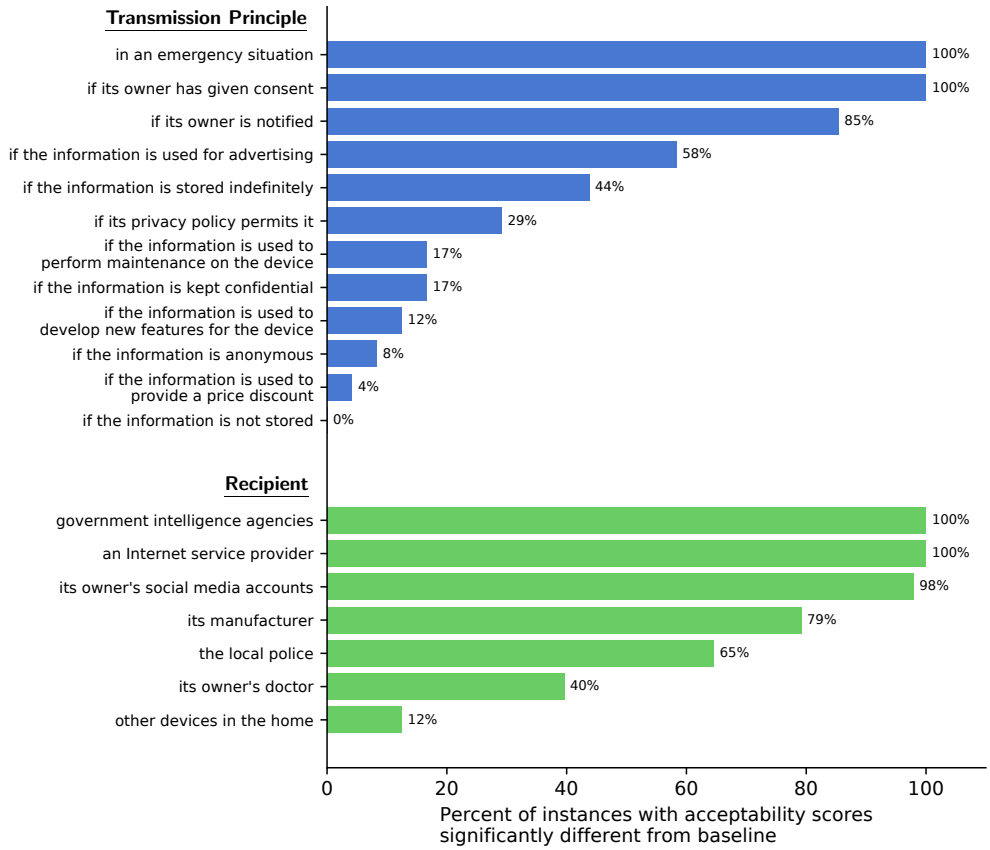

Fig. 4. The percentage of instances where the inclusion of the specified transmission principle or recipient resulted in a statistically significant difference in acceptability score compared to the same information flow with the *null* transmission principle or the "its owner's immediate family" recipient.

investigation of a broad set of potentially relevant privacy norms for relatively little cost. We were able to survey the acceptability of 3,840 information flows for only $2,800, a fraction of typical product development budgets.

Following this recommendation could have prevented several recent privacy-related public relations snafus by IoT device manufacturers. For example, knowing that sending information about users' eating habits and home occupancy is especially likely to be deemed unacceptable could have prevented notorious design choices made by manufacturers of smart TVs and headphones [26, 40].

As a detailed example of how using our survey method could have protected device manufacturers, consider the high-profile cases of IoT toys "spying" on children [23]. Some of the backlash to these devices is due to their insecure implementations, allowing hackers to obtain sensitive user information. However, many consumers were justifiably upset that these devices collected and indefinitely stored audio recordings of their children regardless of security considerations. The manufacturers of these toys could have used our survey method to determine how to better fit device features to user privacy norms. In this case, surveyed information flows would all have the same sender, attribute, recipient, and subject parameters. The sender would be the toy itself. The recipient

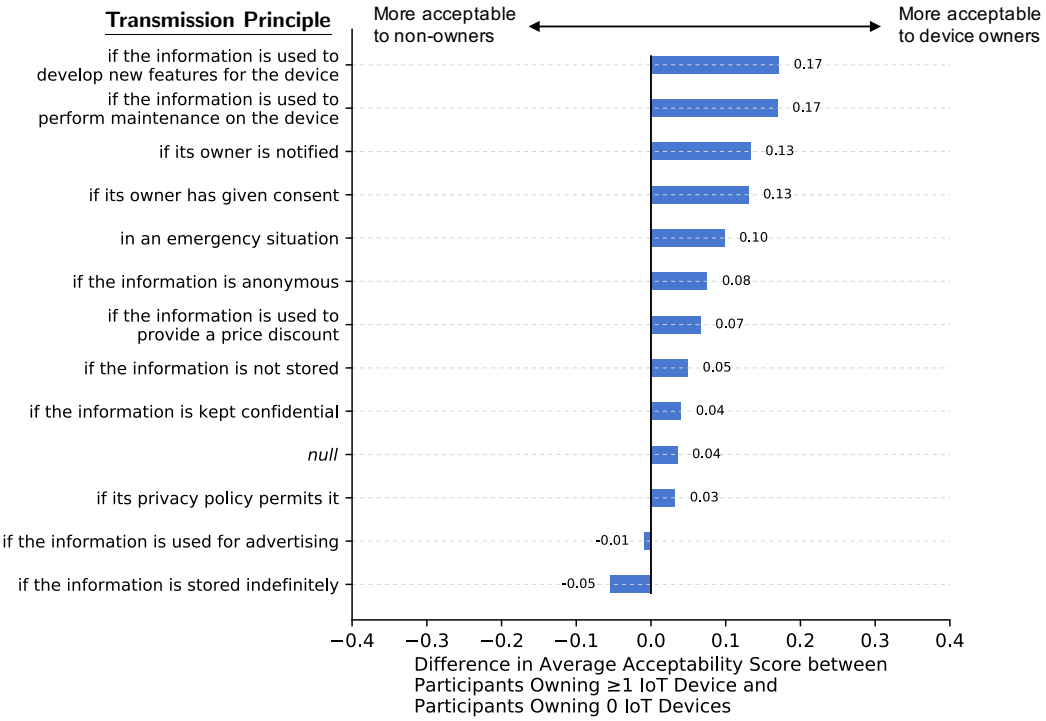

Fig. 5. Difference in average acceptability scores of information flows with specified transmission principles between respondents who do and do not own any home IoT devices. Positive values indicate that respondents owning IoT devices rated the associated information flows as more acceptable on average.

would be the device manufacturer. The attribute would be recorded audio, and the subject would be the child playing with the toy. The transmission principles would then vary to reflect different implementation options, such as "if the information is automatically deleted within 24 hours," "if the information can be manually deleted by an adult user through a web interface," etc. Because A/B testing of privacy-related designs is difficult to do ethically, such a survey would allow for a data-driven decision about which potential implementation to choose.

5.2 Privacy Norms Support Restricting IoT Device Communications

A recent report on IoT security and privacy by the Broadband Internet Technical Advisory Group (BITAG) recommends that "IoT devices should be restrictive rather than permissive in communicating" [25]. The negative acceptability scores of most information flows in our survey support this recommendation, indicating that restricting IoT device communications is in keeping with consumer privacy norms.

This reinforces the importance of critically evaluating device communications during the software development process in order to avoid communications not strictly necessary for core device behavior. Manufacturers should be especially careful of third-party services and libraries used by their devices, which often invoke their own information flows to potentially unknown recipients.

5.3 Privacy Policies Should Clearly State Transmission Principles

Our results complement another recommendation from the BITAG report: "IoT devices should ship with a privacy policy that is easy to find & understand" [25]. The use and utility of privacy policies has been examined by many prior research efforts [11, 15, 16, 30, 38]. Previous studies have shown that "consumers believe the term 'privacy policy' on a website means that the site protects their privacy" [55]. In our results, information flows with the transmission principle "if its privacy policy permits it" are more acceptable than unconditional flows but still had negative average acceptability scores across eight out of nine recipients. Simply abiding by a privacy policy does not seem to make an IoT device automatically adhere to consumer privacy norms, regardless of the policy's clarity. However, easily accessible and interpretable privacy policies could allow manufacturers to benefit from other positive effects on information flow acceptability observed by our survey. For example, information flows that occur only in emergency situations were generally deemed as acceptable, while their unconditional counterparts were generally unacceptable. If a device communicates certain information only in emergency situations, its manufacturer should use the privacy policy to clearly highlight this criteria to reduce consumer privacy concerns. This is particularly relevant for IoT security systems, such as smart smoke detectors, carbon monoxide detectors, and glass break detectors, as well as for certain health monitoring devices, such as fall detectors for the elderly.

5.4 User Consent is Broadly Important for Information Flow Acceptability

The transmission principle, "if the owner has given consent" has the highest average acceptability scores across all recipients and is the only information flow parameter with positive average acceptability scores across all conditions. In other words, survey respondents tended to accept information transfers with user consent irrespective of the rest of the context. This result aligns with notice-and-consent initiatives advocating for meaningful and timely data usage notifications and for requiring user consent to collect information [20].

However, our results also provide a clear illustration of the idea argued by Nissenbaum in [44] that consent-based definitions of privacy are just a part of a bigger picture. For example, information flows taking place "in an emergency situation" had positive acceptability scores with five of six recipients without any specification of consent. Such insights further support calls to adopt new a privacy paradigm for IoT that will "respect the context in which personally identifiable information is collected" [60].

5.5 Local Data Sharing Should Consider Secondary Information Flows

The recipient "other devices in the home" was the second most acceptable recipient in the survey and the only recipient with a positive average acceptability score in information flows with the *null* transmission principle. This may indicate there is a broader privacy norm distinguishing recipients inside the home versus outside the home. This is promising for manufacturers of interconnected smart home IoT systems that share information among a network of devices.

However, device manufacturers should also consider secondary information flows from their devices through an intermediary. At first glance, the increased acceptability of information flows to other devices in the home might indicate that devices that only communicate with an IoT hub or smartphone by Bluetooth or local WiFi can be less concerned about sending user information. However, a device that sends its owner's location to a nearby smartphone and then the smartphone sends the location to a manufacturer's cloud server would have participated in two information flows: a flow with the recipient "other devices in the home" and a flow with the recipient "its manufacturer." While the first flow may be acceptable given user privacy norms, the second may not. This is particularly relevant for manufacturers of single-purpose sensors and actuators designed to connect to an IoT hub made by a different company. Such manufacturers should avoid having their devices send potentially sensitive information with greater frequency or precision than necessary. Obfuscation techniques could also be employed to prevent privacy violating information flows from an intermediary device.

5.6 User Familiarity and Market Penetration Impact Privacy Norms

The power meter, fitness tracker, and thermostat have the least negative acceptability scores averaged across all tested attributes, while the refrigerator, security camera, and door lock have the most negative. We may speculate that these results reflect exposure to smart devices of particular types. Fitness trackers and Nest-like smart thermostats, widely available and well-advertised, have adjusted users' expectations with regards to the connectivity and information sharing behavior of these devices. Smart refrigerators and door locks have seen comparatively less market penetration. This interpretation suggests that greater exposure to specific types of smart home devices has modified or created privacy norms, making information flows involving these devices more acceptable. However, it would require further longitudinal surveys to indicate whether privacy norms are indeed being modified by exposure to IoT devices over time. Such surveys, especially focusing on a single sender with a wider variety of other information flow parameters, would be valuable for IoT device manufacturers to see how their products fit into evolving cultural norms. This could give an indication of when or how gradually to release new products or new features.

5.7 Device Communications Should Directly Support Primary Functionality

Examining acceptability scores across pairs of senders and attributes indicates that information flows may be more acceptable when the attribute is more closely related to the primary function of the device. For example, a fitness tracker is likely to need heart rate data and exercise routine information in order to provide fitness recommendations. These information flows are correspondingly more acceptable than a fitness tracker sending less fitness-relevant information, such as recorded audio or home occupancy. Similarly, a smart refrigerator sending users' eating habits (associated with the food-related purpose of the device) is deemed more appropriate than a refrigerator sending users' locations. However, the effect size is small and not consistent across all sender/attribute pairs. These results motivate more extensive follow-up surveys to disentangle pairwise sender/attribute effects on privacy norms.

5.8 ISP Data Collection, Advertising, and Indefinite Data Storage Generally Violate Privacy Norms

ISPs are among the least acceptable recipients across all transmission principles. A comparison of the median value of the average acceptability score for the ISP recipient against the the median value of the average score for the baseline recipient "its owner's immediate family" (-2 vs. 0) further indicates that the ISP recipient has a negative effect on information flow acceptability. This result is relevant given the recent nullification of the FCC broadband consumer privacy resolution [19], which effectively allows ISPs to mine and sell customer information. Our results provide a strong indication that consumers do not favor the idea of devices transmitting information to ISPs despite this regulatory action.

The transmission principles "if the information is used for advertising" and "if the information is stored indefinitely" had the lowest acceptability ratings averaged across all recipients and were the only transmission principles to have lower average scores than unconditional flows with the *null* transmission principle. User dislike of sharing data for advertising aligns with previous studies [46], and a mistrust of persistent cloud storage reflects recent high-profile data breaches. However, advertising and indefinite storage (for a variety of purposes) are still standard for many companies collecting data from non-IoT provenances. Our survey indicates that these privacy concerns, originally developed in non-IoT contexts, are still very relevant to the nascent home IoT. This recommends that IoT manufacturers remain cautious with their data storage and advertisement practices or risk breaching privacy norms and alienating users.

5.9 IoT Device Owners Are More Accepting of Smart Home Information Flows

Responses to the technical background questions allow us to compare privacy expectations between participants who own one or more IoT devices in their homes compared to those who do not. We chose to compare responses across transmission principles, because all participants answered questions about information flows with all transmission principles.

Acceptability scores are higher for participants owning smart home devices across all transmission principles, but the differences between respondent groups are small. These differences motivate follow-up studies to see whether this effect is consistent across larger populations. These studies could be combined with standardized privacy concern scales, such as IUIPC [37], to see how people with different underlying privacy concerns and different exposures to novel technologies react to information flows. More fundamentally, this line of research would follow the original formulation of CI, which calls for examining established norms in terms of their "merits as a function of their meaning and significance in relation to the aims, purposes, and values of the context" [44].

5.10 Smart Home Privacy Norms Exhibit Similarities to Smartphone Privacy

Our results also warrant a comparison with existing research that has explored various dimensions of mobile application permissions [1, 33, 34, 58]. Previous studies [33, 34] have shown that smartphone users are generally uncomfortable sharing their data—including their location, contacts, and account data—with advertisers and social networks services. Users also cite concerns about personal data as reasons to avoid downloading applications [1]. The vast majority of smartphone participants have been shown to deny at least one information flow—as defined by CI—in current smartphone systems [58].

While only some of the information flows that we considered directly map onto current flows in smartphone systems, we noted some similarities. For example, our participants found flows with the transmission principles "if the information is used for advertising" and "its owner's social media accounts" largely unacceptable [33, 34]. Nevertheless, our participants were generally comfortable with flows that were conditioned on consent. Future studies could use our survey methodology with "smartphone" as a sender to enable a more direct comparisons between smartphone privacy norms and those we discovered for IoT devices. Combining such surveys with field experiments and interventions [58] could also indicate how and when smartphone users will deny consent to specific flows.

6 LIMITATIONS AND FUTURE WORK

We acknowledge a number of limitations of our approach, which we believe can be addressed with future studies and refinement of CI-based survey methods.

6.1 Parameter Ambiguity

Our goal was to demonstrate that the CI survey method is effective in broad contexts, even if some interpretation of information flow parameters is left to survey participants. Such broad contexts are likely of interest to regulators or consumer advocates seeking holistic overviews of privacy norms.

Nevertheless, we recognize that participant interpretations may have caused uncontrolled variations in information flow acceptability scores. For example, "its owner's immediate family" could be interpreted as various nuclear to extended family groups. The generic device types (e.g., "thermostat" rather than "Nest Learning Thermostat") may have caused participants to envision different on-market devices with different features or to imagine entirely non-existent devices. However, the acceptability scores of information flows that did include a specific name-brand device ("a personal assistant (e.g. Amazon Echo)") had a standard deviation within 0.02 points of the standard deviations of flows with all other senders. This suggests that using generic device types

Table 2. Selected representative open-ended comments relevant to participant rationale categorized by general value of concern. Participant ID numbers in parentheses.

Value	Participant Comments
Trust	• I'm starting not to trust smart technology because of the capability that "big brother" may be listening. (P513) • I don't trust my ISP or my government or companies. I used to trust my last doctor but not my current one. (P863)
Security	• I do not and will not own smart appliances due to the danger of hacking. Sending info in any situation is absolutely completely unacceptable to me. (P302) • I am personally leery about connected devices that keep information, especially considering their lack of security. (P1623)
Privacy	• I personally value my privacy more than convenience given the option. (P407) • For me it's totally unacceptable, for any reason, to have audio on my thermostat. A person should be speak in his/her own home without being recorded by anything or anybody. (P1167) • ISPs should never have access to anyone's physical location. It is idiotic to alert social media where you are, especially if you are traveling. (P1494) • I understand that "privacy" is really an oxymoron but I'm sad (and a little concerned) that we're not even safe in our own homes from information gathering devices. (P1253)
Transparency	• Sharing personal information is only ever completely acceptable when the owner gives consent. (P1749) • Even if consent is given, I still have privacy concerns about some of those situations. (P983) • In general I think transmitting that information is unacceptable, although owner consent & emergencies may be okay if that info is transmitted to police, doctors and immediate family. (P1665) • Just a thought I had: a customer can give consent by checking that they agree to privacy policy, while not having read the policy. On reflecting on this, I think I may be more careful to agree that it's okay if a person gives consent this way. (P1282) • Storing and sharing any of the information in any of the situations is unacceptable unless the individual consumer for some reason wants them to, is clearly informed what they're agreeing to, and gives clearly expressed permission. (P79) • I believe many people will volunteer to provide their data and that is perfectly okay and the way we should move towards. (P917)

did not significantly increase acceptability score variability, but follow-up surveys with overlapping device types and specific devices would be needed to confirm.

Device manufacturers or regulatory investigators could use our CI survey method with more clearly defined information flow parameters to discover expectations about product behavior in specific scenarios. We look forward to seeing such applications of our survey method in future research.

6.2 Participant Rationale

Our CI survey method offers a way to discover privacy norms regarding information flows. An even deeper understanding would require evaluating participant rationale and the tradeoffs between values (e.g., trust, national security, safety and/or security) and aims/purposes (e.g., making a system more efficient or a user more productive) that contribute to the emergence of the norms.

We examined participants' optional open-ended comments for indications of the values motivating their responses. About 30 participants gave comments related to the IoT content of the survey. Several of these express general concerns about privacy (e.g., "Worried about this type of technology in the future"). Those that do provide further insight are collected in Table 2. Many say that information sharing from IoT devices is either never acceptable or acceptable only with owner consent or in emergencies. This matches the aggregated acceptability scores from the full survey results. Concerns about IoT device security and lack of trust in various entities (ISP, government, companies, smart technology, doctor) are cited as reasons for disapproving information

flows. Ultimately however, the limited number and length of these responses restricts their ability to explain nuanced variations in acceptability scores across information flows. Researchers who use our CI survey method to systematically appraise a wide space of privacy norms may elect to perform follow-up studies to identify the values underlying particularly interesting or surprising discovered norms.

6.3 Complex Modeling

Users' privacy expectations likely depend on additional factors beyond the five CI information flow parameters (e.g., demographics information). The regular format of responses from our survey methodology makes it possible construct statistical models to test the interactions of these factors. However, the complexity of such models increases substantially with each new factor. With sufficient data and more complex models (e.g., mixed-effect models [8]), future research could identify how other factors influence privacy norms.

6.4 Limitations of the Platform

Because we conducted our survey on Amazon Mechanical Turk, our results may be less generalizable to the broader US population. However, while the MTurk population is limited in its diversity, research [7, 53] has shown that it is similar to participants from university campuses and other online participant pools. Future research could validate and extend our findings to more diverse participants.

6.5 Online Survey Tool

Our next goal is to create an online tool which will allow anyone to easily create and conduct surveys using our CI method. Users will be able to input CI parameters relevant to their particular domain, choose from provided default parameters, and specify parameter combinations to exclude. The tool will automatically generate an online survey using our format and deploy the survey on MTurk. We will provide an interface that plots results, highlights notably acceptable or unacceptable parameters or information flows, and allows the user to re-deploy the survey at a later date to collect longitudinal data. We hope that the results of this research, combined with a user-friendly online tool, will incite further studies using our CI survey method.

7 RELATED WORK

Previous work has captured and analyzed privacy norms in the IoT context. This section surveys the work that is most closely related to this study.

In 2012, Barkhuus highlighted the shortcomings of the existing privacy frameworks on self-gathered empirical data [5]. She argued for the use of CI for privacy-related user studies as a way to understand the "contextually grounded reasons for people's privacy concern or lack thereof." This observation motivates our work.

In the same year, Winter performed a small study using CI to identify specific practices in an IoT setting that could be perceived as privacy violations [59]. Our work also uses the CI framework to construct context-related questions that can help identify practices that conflict with established privacy norms. However, our study examines many more actors, attributes, and transmission principles. We also offer a more thorough statistical analysis of responses.

A 2013 study conducted by Pew Research Center, which surveyed 461 United States (US) adults and facilitated 9 online focus groups with 80 people each, found that Americans would opt to "share personal information or permit surveillance in return for getting something of perceived value" [48]. A section on "home activities, comfort and data capture" specifically explored users' perceptions towards a smart thermostat that can "learn about your temperature zone and movements around the house and potentially save you on your energy bill...in return for sharing data about some of the basic activities that take place in your house." Data analysis showed that 55% of participants found this behavior unacceptable for various reasons, while 27% deemed it acceptable. Our

survey provides a much more comprehensive analysis of such scenarios and covers a range of settings, devices, and information types to construct a more complete picture of the privacy norms surrounding IoT information flows.

In 2015, Horne et al. studied emerging privacy norms and user privacy expectations in response to new technology [28]. In three separate vignette experiments, participants were recruited on MTurk to answer a series of questions about the frequency and granularity of information collection by a smart power meter. The questions asked whether the confidentiality or sale of the collected information affected their privacy expectations. Our work also explores privacy expectations and norms, but we rely on the CI framework to rigorously scale our data collection to many more users and settings. Our method allows us to ask and examine users' privacy expectations for different devices, information recipients, and conditions under which information is shared.

A 2016 study by Martin and Nissenbaum [39] surveyed 569 respondents using a series of vignettes describing different CI information flows with varying receivers and transmission principles. This study found that specifying additional contextual information affects users' perception of what information is sensitive. They conclude that knowing "how the information is used is more important to meeting/violating privacy expectations than the type and sensitivity level of given information." While our work also uses CI to provide additional contextual information to capture users' privacy expectations, it offers a robust methodology that uses all the CI parameters.

Our survey method is based on previous work from 2016 by Shvartzshnaider et al. that used the language of CI to crowdsource discovery of users' privacy expectations in the education domain [52]. Users were asked multiple yes-or-no questions such as "Is it acceptable for the student's professor to share the student's transcript with the student's academic advisor if the student is performing poorly?" The aggregated responses were analyzed to produce a set of most approved and most disapproved norms in the surveyed community.

In 2017, Emami-Naeini et al. used a 1,007-participant vignette study to capture privacy expectations and preferences of users in a set of 380 IoT use-case scenarios [43]. MTurk workers were presented with 14 different IoT data collection scenarios, including factors such as location, biometrics, temperature, purpose, retention time, and whether the data is shared after collection. A linear model provides an indication of which factors positively or negatively affect users' comfort levels. Overall, this work is timely and shares a similar motivation to ours, but does not use a formal theory of privacy and tests many fewer conditions than the 3,840 information flows surveyed in our study. By showing that "privacy preferences are diverse and context-dependent," this work supports our use of CI.

8 CONCLUSION

This work presents a survey method for privacy norm discovery that integrates a formal theory of privacy with combinatorial testing at scale. We use the method to discover privacy norms regarding smart home IoT devices. Our survey of 1731 U.S. adults and 3,840 information flows provides actionable recommendations for device manufacturers, regulators, and consumer advocates. These results demonstrate that our CI survey method enables effective discovery of privacy norms, even in rapidly-changing technological domains. Our method is easily adaptable to arbitrary contexts with varying actors, information types, and communication conditions, paving the way for future studies informing the design of emerging technologies.

ACKNOWLEDGMENTS

We thank Helen Nissenbaum and Marshini Chetty. This work is supported by the Department of Defense through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program, a Google Faculty Research Award, the National Science Foundation through awards CNS-1535796 and CNS-1539902, and the Princeton University Center for Information Technology Policy Internet of Things Consortium.

REFERENCES

[1] Monica Anderson. 2015. Key takeaways on mobile apps and privacy. http://www.pewresearch.org/fact-tank/2015/11/10/key-takeaways-mobile-apps/

[2] Noah Apthorpe, Dillon Reisman, and Nick Feamster. 2016. A Smart Home is No Castle: Privacy Vulnerabilities of Encrypted IoT Traffic. In *Workshop on Data and Algorithmic Transparency*.

[3] Paul Ashley, Satoshi Hada, Günter Karjoth, Calvin Powers, and Matthias Schunter. 2003. Enterprise privacy authorization language (EPAL). *IBM Research* (2003).

[4] Itai Asseo, Maggie Johnson, Bob Nilsson, Neti Chalapathy, and TJ Costello. 2016. The Internet of things: Riding the wave in higher education. *Educause Review* (2016), 11–31.

[5] Louise Barkhuus. 2012. The mismeasurement of privacy: using contextual integrity to reconsider privacy in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 367–376.

[6] Adam Barth, Anupam Datta, John C Mitchell, and Helen Nissenbaum. 2006. Privacy and contextual integrity: Framework and applications. In *2006 IEEE Symposium on Security and Privacy*. IEEE, 15–pp.

[7] Christoph Bartneck, Andreas Duenser, Elena Moltchanova, and Karolina Zawieska. 2015. Comparing the similarity of responses received from studies in Amazon's Mechanical Turk to studies conducted online and with direct recruitment. *PloS one* 10, 4 (2015), e0121595.

[8] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823* (2014).

[9] Omar Chowdhury, Andreas Gampe, Jianwei Niu, Jeffery von Ronne, Jared Bennatt, Anupam Datta, Limin Jia, and William H Winsborough. 2013. Privacy promises that can be kept: A policy analysis method with application to the HIPAA privacy rule. In *Proceedings of the 18th ACM Symposium on Access Control Models and Technologies*. ACM, 3–14.

[10] Federal Communications Commission. 2017. Green Paper: Fostering the Advancement of the Internet of Things. https://www.ntia.doc.gov/other-publication/2017/green-paper-fostering-advancement-internet-things

[11] Lorrie Faith Cranor, Joseph Reagle, and Mark S Ackerman. 2000. Beyond concern: Understanding net users' attitudes about online privacy. *The Internet upheaval: raising questions, seeking answers in communications policy* (2000), 47–70.

[12] Natalia Criado and Jose M Such. 2015. Implicit Contextual Integrity in Online Social Networks. *Information Sciences* (2015).

[13] Paul Daugherty, Prith Banerjee, Walid Negm, and Allan E Alter. 2015. Driving unconventional growth through the industrial internet of things. (2015). https://www.accenture.com/us-en/_acnmedia/Accenture/next-gen/reassembling-industry/pdf/Accenture-Driving-Unconventional-Growth-through-IIoT.pdf

[14] Tom Davenport and John Lucker. 2015. Running on data: Activity trackers and the Internet of Things. https://dupress.deloitte.com/dup-us-en/deloitte-review/issue-16/internet-of-things-wearable-technology.html

[15] Julia Brande Earp, Annie I Antón, Lynda Aiman-Smith, and William H Stufflebeam. 2005. Examining Internet privacy policies within the context of user privacy values. *IEEE Transactions on Engineering Management* 52, 2 (2005), 227–237.

[16] Serge Egelman, Janice Tsai, Lorrie Faith Cranor, and Alessandro Acquisti. 2009. Timing is everything?: the effects of timing and placement of online privacy indicators. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 319–328.

[17] Enterprise Privacy Authorization Language (EPAL 1.2) 2003. https://www.w3.org/Submission/2003/SUBM-EPAL-20031110/

[18] Federal Communications Commission. 2016. FCC Adopts Broadband Consumer Privacy Rules. https://www.fcc.gov/document/fcc-adopts-broadband-consumer-privacy-rules

[19] Federal Communications Commission. 2016. FCC Releases Rules to Protect Broadband Consumer Privacy. https://www.fcc.gov/document/fcc-adopts-broadband-consumer-privacy-rules

[20] Federal Trade Commission. 2007. Fair Information Practice Principles. https://web.archive.org/web/20100309105100/http://www.ftc.gov/reports/privacy3/fairinfo.shtm#Notice/Awareness

[21] David Ferraiolo, D Richard Kuhn, and Ramaswamy Chandramouli. 2003. *Role-based access control*. Artech House.

[22] David F Ferraiolo, Ravi Sandhu, Serban Gavrila, D Richard Kuhn, and Ramaswamy Chandramouli. 2001. Proposed NIST standard for role-based access control. *ACM Transactions on Information and System Security (TISSEC)* 4, 3 (2001), 224–274.

[23] Lorenzo Franceschi-Bicchierai. 2017. Internet of Things Teddy Bear Leaked 2 Million Parent and Kids Message Recordings. https://motherboard.vice.com/en_us/article/pgwean/internet-of-things-teddy-bear-leaked-2-million-parent-and-kids-message-recordings

[24] Frances Grodzinsky and Herman T Tavani. 2010. Applying the "Contextual Integrity" Model of Privacy to Personal Blogs in the Blogoshere. *Computer Science and Information Technology Faculty Publications* (2010).

[25] Broadband Internet Technical Advisory Group. 2016. *Internet of Things (IoT) Security and Privacy Recommendations*. Technical Report. https://www.bitag.org/documents/BITAG_Report_-_Internet_of_Things_(IoT)_Security_and_Privacy_Recommendations.pdf

[26] Hayley Tsukayama. 2017. Bose headphones have been spying on customers, lawsuit claims. *The Washington Post* (2017). https://www.washingtonpost.com/news/the-switch/wp/2017/04/19/bose-headphones-have-been-spying-on-their-customers-lawsuit-claims/

[27] Paul Hitlin. 2016. Turkers in this canvassing: young, well-educated and frequent users. http://www.pewinternet.org/2016/07/11/turkers-in-this-canvassing-young-well-educated-and-frequent-users/

[28] Christine Horne, Brice Darras, Elyse Bean, Anurag Srivastava, and Scott Frickel. 2015. Privacy, technology, and norms: The case of Smart Meters. *Social science research* 51 (2015), 64–76.
[29] Gordon Hull, Heather Richter Lipford, and Celine Latulipe. 2011. Contextual gaps: privacy issues on Facebook. *Ethics and information technology* 13, 4 (2011), 289–302.
[30] Carlos Jensen and Colin Potts. 2004. Privacy policies as decision-making tools: an evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 471–478.
[31] David Kravets. 2016. Sex toys and the Internet of Things collide—what could go wrong? https://arstechnica.com/tech-policy/2016/09/sex-toys-and-the-internet-of-things-collide-what-could-go-wrong/
[32] Nile Lars. 2014. Connected Medical Devices, Apps: Are They Leading the IoT Revolution – or Vice Versa? https://www.wired.com/insights/2014/06/connected-medical-devices-apps-leading-iot-revolution-vice-versa/
[33] Jialiu Lin, Shahriyar Amini, Jason I. Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and Purpose: Understanding Users' Mental Models of Mobile App Privacy Through Crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. ACM, 501–510.
[34] Jialiu Lin, Bin Liu, Norman Sadeh, and Jason I. Hong. 2014. Modeling Users' Mobile App Privacy Preferences: Restoring Usability in a Sea of Permission Settings. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. USENIX Association, 199–212. https://www.usenix.org/conference/soups2014/proceedings/presentation/lin
[35] Leib Litman, Jonathan Robinson, and Tzvi Abberbock. 2017. TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior research methods* 49, 2 (2017), 433–442.
[36] Richard Lowry. 2014. Concepts and applications of inferential statistics. (2014).
[37] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.
[38] Kirsten Martin. 2015. Privacy notices as tabula rasa: An empirical investigation into how complying with a privacy notice is related to meeting privacy expectations online. *Journal of Public Policy & Marketing* 34, 2 (2015), 210–227.
[39] Kirsten Martin and Helen Nissenbaum. 2016. Measuring privacy: an empirical test using context to expose confounding variables. *Colum. Sci. & Tech. L. Rev.* 18 (2016), 176.
[40] Chris Matyszczyk. 2015. Samsung's warning: Our Smart TVs record your living room chatter. https://www.cnet.com/news/samsungs-warning-our-smart-tvs-record-your-living-room-chatter/
[41] Aleecia M McDonald and Lorrie Faith Cranor. 2008. The cost of reading privacy policies. *ISJLP* 4 (2008), 543.
[42] Eliott McLaughlin. 2017. Suspect OKs Amazon to hand over Echo recordings in murder case. https://www.cnn.com/2017/03/07/tech/amazon-echo-alexa-bentonville-arkansas-murder-case/index.html
[43] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. 2017. Privacy Expectations and Preferences in an IoT World. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association, Santa Clara, CA, 399–412.
[44] Helen Nissenbaum. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford Law Books.
[45] Bill Parducci. 2005. eXtensible Access Control Markup Language (XACML) specification. (2005).
[46] Joseph Phelps, Glen Nowak, and Elizabeth Ferrell. 2000. Privacy concerns and consumer willingness to provide personal information. *Journal of Public Policy & Marketing* 19, 1 (2000), 27–41.
[47] Qualtrics Online. 2017. http://www.qualtrics.com
[48] Lee Rainie and Maeve Duggan. 2017. Privacy and Information Sharing. http://www.pewinternet.org/2016/01/14/privacy-and-information-sharing/
[49] Andrew D Selbst. 2013. Contextual expectations of privacy. *Cardozo Law Review* (2013).
[50] Juliet Popper Shaffer. 1995. Multiple Hypothesis Testing. *Annual Review of Psychology* 46, 1 (1995), 561–584.
[51] Pan Shi, Heng Xu, and Yunan Chen. 2013. Using contextual integrity to examine interpersonal information boundary on social network sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 35–38.
[52] Yan Shvartzshnaider, Schrasing Tong, Thomas Wies, Paula Kift, Helen Nissenbaum, Lakshminarayanan Subramanian, and Prateek Mittal. 2016. Learning Privacy Expectations by Crowdsourcing Contextual Informational Norms. *The Fourth AAAI Conference on Human Computation and Crowdsourcing* (2016).
[53] Daniel J Simons and Christopher F Chabris. 2012. Common (mis) beliefs about memory: A replication and comparison of telephone and Mechanical Turk survey methods. *PloS one* 7, 12 (2012), e51876.
[54] Snap Spectacles 2017. Snap Spectacles. https://www.spectacles.com/
[55] FTC Staff. 2010. Protecting Consumer Privacy in an Era of Rapid Change–A Proposed Framework for Businesses and Policymakers. *Journal of Privacy and Confidentiality* 3, 1 (2010), 5.
[56] Seymour Sudman, Norman M Bradburn, and Norbert Schwarz. 1996. *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.
[57] UserBob - Usability Testing. 2017. https://userbob.com/

[58] Primal Wijesekera, Arjun Baokar, Ashkan Hosseini, Serge Egelman, David Wagner, and Konstantin Beznosov. 2015. Android Permissions Remystified: A Field Study on Contextual Integrity. In *24th USENIX Security Symposium (USENIX Security 15)*. USENIX Association, 499–514. https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/wijesekera

[59] Jenifer S Winter. 2012. Privacy and the emerging internet of things: using the framework of contextual integrity to inform policy. In *Pacific Telecommunications Council Conference Proceedings*.

[60] Christopher Wolf and Jules Polonetsky. 2013. An Updated Privacy Paradigm for the "Internet of Things". https://fpf.org/wp-content/uploads/Wolf-and-Polonetsky-An-Updated-Privacy-Paradigm-for-the-%E2%80%9CInternet-of-Things%E2%80%9D-11-19-2013.pdf

[61] Kathryn Zickuhr. 2013. Who's not online and why. Pew Research Center's Internet and American Life Project. http://www.pewinternet.org/files/old-media/Files/Reports/2013/PIP_Offline%20adults_092513_PDF.pdf

[62] Michael Zimmer. 2008. Privacy on planet Google: Using the theory of contextual integrity to clarify the privacy threats of Google's quest for the perfect search engine. *J. Bus. & Tech. L.* 3 (2008), 109.

Received August 2017; revised February 2018; accepted April 2018

APPENDIX

Table 3. Self-reported demographics and technical background of survey participants.

Metric	Sample	Metric	Sample
Female	51%	Live with family	61%
Male	49%	Live alone	20%
Other	<1%	Live with one or more non-family roommates	16%
Prefer not to disclose	1%	Other	2%
		Prefer not to disclose	1%
Democrat	44%		
Republican	21%	A one-family house detached from any other house	60%
Independent	33%	A building with 10 or more apartments	14%
Prefer not to disclose	3%	A building with fewer than 10 apartments	11%
		A one-family house attached to one or more houses	8%
Less than $10k	4%	A mobile home	3%
$10k - $20k	8%	A dormitory	<1%
$20k - $30k	12%	A boat, RV, van, etc.	<1%
$30k - $40k	11%	Other	1%
$40k - $50k	11%	Prefer not to disclose	1%
$50k - $60k	11%		
$60k - $70k	10%	Suburban	53%
$70k - $80k	8%	Urban	30%
$80k - $90k	5%	Rural	18%
$90k - $100k	5%		
$100k - $150k	9%	Married or domestic partnership	46%
More than $150,000	3%	Single, never married	44%
Prefer not to disclose	2%	Divorced	7%
		Separated	1%
18-24 years old	10%	Widowed	1%
25-34 years old	45%	Prefer not to disclose	1%
35-44 years old	23%		
45-54 years old	11%	No children under 16	67%
55-64 years old	6%	Children under 16	32%
65-74 years old	2%	Prefer not to disclose	1%
75 years or older	<1%		
Prefer not to disclose	1%	0-3 hours Internet use per day	15%
		4-7 hours Internet use per day	45%
Nursery school to 8th grade	<1%	8-12 hours Internet use per day	29%
Some high school, no diploma	1%	>12 hours Internet use per day	11%
High school graduate, diploma or the equivalent	9%		
Trade/technical/vocational training	3%	Own 0 IoT devices	62%
Some college credit, no degree	23%	Own ≥1 IoT device	36%
Bachelor's degree	40%	I don't know	2%
Associate degree	12%		
Master's degree	10%	I set up my IoT devices	78%
Professional degree	2%	Someone else set up my IoT devices	21%
Doctorate degree	1%	I don't remember who set up my IoT devices	1%
Prefer not to disclose	1%		