Devices and Data and Agents, Oh My: How Smart Home Abstractions Prime End-User Mental Models

MEGHAN CLARK, University of Michigan, USA
MARK W. NEWMAN, University of Michigan, USA
PRABAL DUTTA, University of Michigan, USA

With the advent of DIY smart homes and the Internet of Things comes the emergence of user interfaces for domestic human-building interaction. However, the design trade-offs between the different representations of a smart home's capabilities are still not well-understood. In this work, we examine how four different smart home abstractions affect end users' mental models of a hypothetical system. We develop four questionnaires, each of which describes the same hypothetical smart home using a different abstraction, and then we collect responses depicting desired smart home applications from over 1,500 Mechanical Turk workers. We find that the choice of abstraction strongly primes end users' responses. In particular, the purely device-oriented abstraction results in the most limited scenarios, suggesting that if we want users to associate smart home technologies with valuable high-level applications we should shift the UI paradigm for the Internet of Things from device-oriented control to other abstractions that inspire a greater diversity of interactions.

CCS Concepts: • **Human-centered computing → User studies**; **Ubiquitous and mobile computing design and evaluation methods**;

Additional Key Words and Phrases: smart homes, IoT, mental models, priming, interface design, natural language processing

ACM Reference Format:
Meghan Clark, Mark W. Newman, and Prabal Dutta. 2017. Devices and Data and Agents, Oh My: How Smart Home Abstractions Prime End-User Mental Models. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 44 (September 2017), 26 pages.
DOI: http://doi.org/10.1145/3132031

1 INTRODUCTION

While the notion of smart spaces has existed for decades, a confluence of forces is at last bringing physical computation to the home market. Advances in energy storage and computer architecture have enabled a proliferation of low-power embedded sensors and actuators with networking capabilities, individually called "smart devices" and collectively dubbed "The Internet of Things" (IoT). Now the IoT community is exploring ways to orchestrate these smart devices and expose their resources to user-facing applications.

While domestic IoT architecture may take many forms, every potential solution will need to abstract the system to provide interfaces for end-user interaction. The abstractions we use to present the system collectively reflect some conceptual model or metaphor for how the user is expected to interact with the smart home. For example, a smart home app that abstracts a smart lighting system by providing virtualized interactive representations of the individual bulbs in the app conveys a device-oriented model of interactions with the system, whereas an app that exposes the lighting system's state as a datastream that emits and receives messages provides a data-oriented model. These abstractions can be explicitly chosen by designers to guide the users in interacting

2474-9567/2017/9-ART44
DOI: http://doi.org/10.1145/3132031

with and understanding the system, or they can reflect implicit assumptions on the part of the designers about how the system works.

There are many different smart home abstractions for IoT system designers to choose, but little work has been done so far on understanding the impact of different system abstractions on end users' mental models and expectations of these systems. Previous attempts to better support end-user interactions have attempted to "get at" end users' mental models for smart homes as though these models are an independently existing property of users, without considering the effects of priming. Studies have made claims about mental models based on user responses to single scenarios without acknowledging that the users could have been primed by the conceptual models that the researchers used in the prompts, potentially without realizing it. These studies generally attempt to reflect the way users "naturally" think about the smart home, but there is no "natural" way that people think about technology. To develop interfaces that support end users, instead we need an intentional and conscious comparison between several different system representations. To that end we propose a framework that would allow comparisons between different conceptual models, which breaks conceptual models down into several independent dimensions whose possible values we call *abstractions*. This abstractions framework would allow us to make comparisons between conceptual models and understand how the different ways we can present the same underlying system will affect users' expectations and behavior.

In fact, there is evidence to suggest that the abstractions that we are using are preventing consumers from thinking of valuable integrated applications that utilize the full range of a smart spaces's capabilities. According to Affinova's 2014 consumer report on IoT adoption, users think the current smart devices on the market are "gimmicky" [4]. While some devices may be gimmicky single-purpose products, many others can be integrated into applications by various emerging platforms. It is possible that due to the way these systems are presented, users do not even think certain useful applications are possible or within the scope of the system.

Additionally, the abstractions we use can create a gap between expectations and reality that prevent adoption. An understanding of how abstractions create expectations can allow us to manipulate and shrink this gap, both on the user side and on the system side.

In this work we make several contributions. First, we demonstrate that the abstractions chosen to represent a smart home system have significant priming effects on end users. We created a set of questionnaires depicting the same hypothetical smart home in four different ways and deployed the questionnaires via Mechanical Turk. Each worker was presented with only one of the four descriptions, and then was asked to describe the applications they wanted. We collected responses from over 1,500 participants. Though mental models cannot be directly observed, they can be thought to emit signals via various user behaviors. To gain insight into the mental models at play, we analyzed the following signals: 1) the qualitative differences in responses, 2) the characteristic words for each set of responses as determined by χ^2 feature selection, and 3) the differences in the *operation profile* of the responses, which we define as the distribution of tasks that a group of users performs on the hypothetical system – for example, the relative proportion of immediate actions, questions, conditional actions, and notifications. We find that users' mental models and the resulting operations that they attempt are heavily affected by the abstractions used to present the system. This finding highlights how critical it is to consciously choose abstractions, both during system design and when creating scenarios to study users' mental models for a particular domain.

Second, we describe what the specific priming effects (and therefore design trade-offs) are for some common abstractions. We examine the effects of two different dimensions for abstractions: How the system's capabilities are represented ("devices" vs. "data"), and whether or not the system is personified ("agent-mediated" vs. "unmediated"). Combining these two abstraction dimensions results in four distinct conceptual models: Unmediated Devices, Unmediated Data, Agent-mediated Devices, and Agent-mediated Data. While these four archetypal abstractions are far from the only possible system interfaces, and in practice are not mutually exclusive, they are a good starting point for understanding the effects of commonly-used abstractions on end-user thinking. Understanding the different workloads and expectations engendered by common abstractions for the domestic

IoT will help system designers better understand the trade-offs involved when choosing the abstractions for their system.

Finally, we also release our complete dataset and code to the community for additional analyses, with IRB approval. As far as we know, this is the first substantial public corpus of end-user descriptions of desirable IoT applications, as well as natural language commands and queries directed to smart home agents. Researchers can use the data to improve the relationship between the Internet of Things and its users, for example by building better natural language interfaces, creating desirable applications, and gaining additional insights into end-user needs and concerns in the smart home domain.

2 RELATED WORK

In his seminal work "The Design of Everyday Things," Norman defines mental models as conceptual models that users draw from to explain and predict the way that devices will behave in different scenarios [15]. People can use multiple models, and even potentially conflicting models, to explain various aspects of a system. Later work found that the initial mental models formed by users early on before they even interact with a system can have a subtle ongoing influence over how users update their understanding of it based on newly-acquired knowledge [3]. One proposed mechanism for this influence is that different presentations of the same system might result in mental models where different objects are the primary conceptual entities.

In psychology, this effect is called *priming*. Priming happens when exposure to an initial stimulus affects the way that a person processes a subsequent stimulus. There is evidence to suggest that mental models are "compiled" on-demand, so some psychologists have used priming to influence the mental models that individuals generate when making decisions about risky processes, from indoor radon to the occupational use of hazardous chemicals [5, 13, 14]. Our insight is that by a similar mechanism, system designers might influence (either intentionally or accidentally) users' mental models of smart home processes.

A number of HCI studies have tried to understand the mental models of smart home users, particularly to support end-user programming in the home. For example, iCAP performed a study on mental models so that the authors could build a programming tool to support the way users thought about context-aware applications [7]. However, the study presented end users with a scenario that used a distinctly device-oriented abstraction to describe the system. The prompt described a generic smart home as a house with "sensors" that could sense the environment and user activities and "execute services on behalf of users." The responses contained an unexpectedly high number of references to objects, which becomes much less surprising when seen through the lens of priming. Additionally, the study found that some users perceived the home as a tool while others viewed it as an assistant. This hinted at the possibility that the interpretation of the phrase "execute services on behalf of users" could potentially result in different mental models, but a followup study to explore the effects of different prompts was not performed.

Similarly, a study of IFTTT programs examined whether trigger-action programming "captures smart home behaviors that users actually desire" [21]. The study presented 318 Mechanical Turk workers with a hypothetical scenario where they had "a home with devices that are Internet-connected and can therefore be given instructions on how to behave," suggesting a strongly device-oriented abstraction of the system, and asked participants for five things they would want their home to do. Results showed that the majority of responses were programs that could be expressed with trigger-action, and most of the remaining responses were remote control interactions. However, the study also found that priming respondents with trigger-action examples had a significant impact on the proportion of trigger-action responses they got back, hinting that priming can have a major effect on the way people think about interacting with systems. While the study explored the effect of priming on the way users *express* programs, it did not reflect upon how the presentation of the smart home in the initial prompt might similarly affect the kinds of operations that users might attempt in the first place. One notable observation is that the responses did not feature any queries. In this work, we find that a lack of queries is characteristic of responses to device-based abstractions, but queries appear in force when using a data-based abstraction of the same system. This casts doubt on the assumption that a single-abstraction study could capture the entirety of behaviors that users desire in a smart home.

To some degree it could be said that the previous studies got out the mental models that they put in. However, the CAMP magentic poetry paper took a different approach [20]. Like earlier work, the CAMP study assumed that users held "natural conceptualizations of ubicomp technologies," but differed in that the researchers were explicitly aware of the potential to bias users' mental models with the study scenario. To avoid priming users, the study provided comics with pictures and dialog and asked users to describe what they thought was happening in scenes where the parents were "programming" or interacting with the system. The comics were still somewhat biasing because they showed devices and they showed the parents at the computer while programming. However, more concerning is the fundamental assumption that biasing should be reduced as much as possible. The real system *will* be priming, and the system will need to support users' primed mental models. Instead of trying to avoid priming in mental model studies, it is critical to determine the effects of different kinds of priming so that in designing the system we can select the abstractions that will prime in ways that we can predict and support.

3 BACKGROUND

In this section we review the relationship between conceptual models, mental models, and user operations. By understanding how these elements interact with each other, we can lay the foundations for a methodological approach that will allow us to compare the impact of conceptual models on end users' expectations and behavior.

As described by Norman, designers have some implicit or explicit notion for how users should think about system, particularly with regards to what the different parts are and how they work together. This notion is called the *conceptual model*, or sometimes the *interaction metaphor*. The designer conveys the conceptual model to the end user through the *system image*. The system image is everything about the system that is visible to the user, which consists of two main parts: 1) the concrete products and interfaces created by the designer, which collectively convey the designer's conceptual model or interaction metaphor, and 2) external factors that the designer does not control. The first part includes artifacts like the physical system interface and documentation, and the second part includes the user's prior experiences, news reports, anecdotes from friends, and so on.

As users perceive the system image, they generate a mental model for how their interactions affect the system and how the system affects them. Prior work has decribed mental models as a "yoked state space" where users form a mapping between a device (or system) state space to a goal state space [17]. The conceptual model, as a manifestated by the system image, is therefore critical in influencing what states the user perceives the system to have and what goals the user thinks the system can satisfy.

While a user's mental model cannot be directly perceived, the types of operations that the user assumes are available and would like to employ can give us some insights into their mental model. In the yoked state space theory, users perform operations on a device in order to navigate the device state space to reach states that, according to their model, map to desired goal states [17]. This means that by observing the kinds of operations users expect to be able to perform, we can get a sense for how they are internally modeling the system states and what goals they perceive the system as being able to satisfy.

Understanding what kinds of operations users will assume are available when provided with a particular conceptual model would help system designers minimize the "gulf of execution" that separates user goals from the actions they need to take to make the system satisfy those goals. By knowing what operations users will expect to able able to do, system designers can design their system to support those operations as first-order primitives. In the opposite direction, designers can also discover which conceptual models will prompt users to invoke operations that the system can best support and which will deliver the desired user experience.

Given the relationship between conceptual models, mental models, and operations, it should be possible to present different conceptual models to users and observe how the types of operations they perform change due to the different mental models they generate from the system image. However, how can we control for those aforementioned external influences that are also a part of the system image, like prior experiences? In this paper we overcome that challenge by looking at population-level distributions of operation types. When comparing what populations as a whole attempt to do in response to a given conceptual model, the individual variations in external influences essentially come out in the wash. A population-level analysis is also useful because smart home system designers often design a single system that will be used by a large number of people, in which case it can be beneficial for designers to know in advance what the system's overall workload will look if it is presented in a certain way.

3.1 Abstractions Framework for Conceptual Models

To understand the trade-offs between different conceptual models, we need to articulate a set of dimensions that we can use to describe and compare them. We propose breaking conceptual models down into several independent dimensions that each have a set of possible values called *abstractions*. Under this approach, conceptual models can be constructed or described by selecting an abstraction along each dimension. While many such dimensions may exist, in this paper we focus on two: *capabilities* and *personification*.

Capabilities. Conceptual models can use many different abstractions to represent a smart home's capabilities. A common approach is to represent capabilities in terms of the available devices, but it is also possible to describe available functionality using other fundamental conceptual entities, such as actions, data streams, facts, or events. In this work, we compare two abstractions used to represent capabilities, which we call "devices" and "data."

Because Internet of Things systems are frequently conceptualized as interconnected devices or physical artifacts at the architectural level, it is common for system designers and even HCI researchers to assume that the higher layers presented to users will share a similar device-oriented abstraction. In particular, there is a great deal of IoT literature that assumes device-oriented interfaces will bubble up to users from the lower layers [1, 10, 16, 23]. However, there is no reason that the higher level of abstraction presented to users needs to resemble the ways that system designers think about the system. In fact, in this work we show that depending on the designer's goals and intended audience, using a device-oriented abstraction may not always be the most beneficial way to model and present the system.

Personification. We also examine two abstractions that can be used to represent different degrees of system personification: "unmediated" and "agent-mediated." An unmediated system is one with no personification layer, and an agent-mediated system is one with an agent that acts as an intermediary between the user and the rest of the system. Prior work has shown that users will interact with functionally-equivalent systems differently depending on whether the system is personified as human-like agent or not. For example, while both Google Now and Siri are phone-based voice interfaces, research has found that users treat Google Now like "voice-activated search," whereas they treat Siri like a social actor due to the latter's presentation as a human-like agent [11].

There are other dimensions that we do not examine in this work, but which would nevertheless be useful in an abstraction-based framework. For example, the dimension of *initiative* is independent from personification. Systems can behave with initiative without being presented as agents. Prior work has compared the user experience of systems with different degrees of initiative in a variety of domains [8, 9, 18]. While these systems act with various levels of agency, they are not presented in an anthropomorphic fashion.

There is also the dimension of *input modality*, which can be visual, written language, voice, gesture, tactile, and more. Some related works have compared the user appeal of different modalities in the smart home domain, such as voice, voice and gesture, touch screen GUIs, and touch screen text interfaces [6]. We do not explicitly compare the effects of these different modalities on user thinking, though we do look at what modalities users assume given abstractions along other dimensions. We leave exploring the effects of additional dimensions beyond the two we consider to future work.

While conceptual models are used implicitly thoughout smart home research, we are the first to attempt to explicitly characterize and compare them systematically. Abstraction dimensions such as the ones we propose here could form the basis of an evaluation framework that allows designers to compare the impact of various conceptual models and make claims about why one particular abstraction would be a better choice for an interface than another.

3.2 Classification of User Operations

In order to determine how conceptual models affect the kinds of operations that users would like to perform, we must next be able to classify user operations. We separate smart home interactions broadly into *immediate interactions*, which take place and complete right away, and *conditional interactions*, which may result in interactions at a later time.

Immediate interactions. Remote control commands like "turn on the lights" are one form of immediate interaction. However, since there are other immediate requests for action that may not fall under a remote control mindset, such as "wake up my children," we use the more general term *immediate actions* to refer to this subcategory of operations. In addition to actuation there are also queries, an often-overlooked form of smart home interaction. Immediate queries can take two forms: *direct questions* and *indirect questions*. Direct questions are any query that would properly end with a question mark. Indirect questions are requests for information that are formulated as a command, such as "tell me how much I weigh." We make the distinction between the two forms to highlight that in our operation taxonomy we consider indirect questions to be requests for information rather than actions.

Conditional interactions. Trigger-action statements, best exemplified by "when I come home, turn on the lights," are one type of ongoing interaction. However, since there can be other ways to express conditions on actions besides event-based triggers (using words like "until," "unless," "before," and "while"), we use the more general umbrella term *conditional actions*. Just as it is possible to have conditional actions, it is also possible to have conditional queries, which we call *notifications*. The difference between notifications and indirect questions can sometimes be subtle. A good rule of thumb is that indirect questions are usually requests for facts, and notifications are usually requests to be informed of events. For example, "tell me when my husband gets home" is most likely a notification, whereas "tell me when my husband will get home" is an indirect question.

This is not meant to be a definitive operation taxonomy for the smart home. However, we believe that including both actions and queries is an important step for developing a more comprehensive way to understand mental models. By disinguishing between requests for actuation and requests for information, we hope to be able to tell a more complex story than an analysis of just actions or just queries would allow.

4 METHODOLOGY

In this work we look at abstractions along two dimensions: whether or not the system is personified by an agent who acts as an intermediary ("unmediated" or "agent-mediated"), and whether capabilities are represented by the available "devices" or the available "data." The possible combinations along these two dimensions result in four different conceptual models: Unmediated Devices, Unmediated Data, Agent-mediated Devices, and Agent-mediated Data.

To examine the impact that each of these four conceptual models has on users' mental models, we devised questionnaires describing a hypothetical smart home using each of the conceptual models. We presented Mechanical Turk workers with one of the four questionnaire prompts and asked them to write either about what applications they would want in their smart home or what they would want their agent to do. We then analyzed the entities and operations present in the written responses.

Given only a few minutes with a questionnaire prompt, users cannot explore a conceptual model to the same depth as they can given a more fully-realized system over a longer period of time. However, prior work has shown that initial mental models formed by users before they even interact with a system can have an ongoing influence over how users update their understanding of the system based on newly-acquired knowledge [3]. These persistent effects, combined with the fact that even our brief descriptions resulted in observable differences, indicates that presentation matters, even if more sustained engagement might shape the users' mental models even further.

Table 1. Overview of questionnaires. We administered questionnaires with two scenarios on Mechanical Turk, each of which had two possible treatments, which resulted in four unique prompts describing smart home conceptual models. The unmediated scenario asked what applications end users wanted in their hypothetical smart home, while the agent-mediated scenario asked end users to tell a hypothetical smart home AI what they wanted it to do. For each scenario, the smart home's capabilities were described either by a list of devices or by a list of data streams. Participants were only presented with one of the four conceptual models.

Personification	**Capabilities**	**Conceptual Model**	**Responses**
Unmediated	Devices	Unmediated Devices	313
	Data	Unmediated Data	302
Agent-mediated	Devices	Agent-mediated Devices	442
	Data	Agent-mediated Data	478

4.1 Questionnaire Overview

We devised two scenarios to collect data about the way end users describe smart home applications, which can be read in their entirety in the online appendix. The unmediated scenario asked respondents to prepare to imagine IoT applications, and then gave one of two treatments with equal likelihood: 1) a list of smart devices that users had at their disposal in their smart home broken down by sensor, actuator, or online service, or 2) a list of data streams that could reasonably be synthesized from the devices listed in the other prompt, broken down by read-only data (sensor readings), read-write data (actuator statuses), or online service data.

To create the two sets of capabilities, we first generated a list of common smart home sensors, actuators, and online services, which became our list of devices. Then we generated the list of data streams by converting each device from the first list. For example, while the device list has "Controllable RGB Lights," the data list has "Whether the lighting is on or off," "What color the lighting is," and "How bright the lighting is." The full list of devices and data streams that we provided can be found in the online appendix.

In order to give respondents time to look over the list of capabilities, they were not allowed to continue until a one-minute timer expired. On the next page they were asked to write for five minutes in a text box about what kinds of applications they wanted in their smart home, with the list of devices or data streams displayed above the text entry box for inspiration. Afterwards, we asked the participants basic demographic questions and assessed their general familiarity with IoT and technology.

The agent-mediated scenario introduced an artificial intelligence agent (an "AI") as the intelligence behind the smart home controls. In this scenario, respondents were first asked to select a gender for their AI's voice from a randomly ordered list of male, female, and androgynous, then respondents were asked to name their AI. Our narrative highlighted that the agent was trustworthy and wanted to help the participant. We then gave respondents one of two prompts with equal likelihood, just as in the unmediated scenario: 1) a list of IoT devices that the agent had access to in their smart home broken down by sensor, actuator, or online service, or 2) a list of data streams that could reasonably be synthesized from the devices listed in the other prompt, broken down by read-only data (sensor readings), read-write data (actuator statuses), or online service data. The respondents were given a minute to look over the list. Afterwards, we told respondents that their AI wanted to help them by automating their home on their behalf. We asked them to tell the AI what it should do, and began the writing prompt with, "OK, <AI name>..." to encourage respondents to communicate directly with the smart home AI. In the agent-mediated scenario, we did not enforce a time minimum on the response page, and we did not provide the device or data list for reference, in order to discourage respondents from pasting parts of the list into the textbox. As in the unmediated scenario, we finished with demographic data collection and questions to assess technological familiarity.

4.2 Subjects

We submitted these two questionnaires to Amazon Mechanical Turk [2]. Since we planned to analyze the linguistic characteristics of the responses, we limited respondents to those located in five countries with large populations of native English speakers (United States, United Kingdom, Canada, Australia, and New Zealand) in order to increase the proportion of native English speakers in the eligible respondent pool.

Subjects were recruited to participate in an academic study about either Internet of Things applications or smart homes. The unmediated questionnaire was advertised with, "We are conducting an academic survey about Internet of Things applications. We want to understand how you think about and describe Internet of Things applications." The agent-mediated questionnaire was advertised with, "We are conducting an academic survey about smart homes. We need to understand the way that you would talk to a smart home artificial intelligence." Workers who completed the unmediated questionnaire were compensated $0.80, and those who completed the agent-mediated questionnaire were compensated $0.40. The difference was due to the difference in the enforced time limit (the agent-mediated questionnaire did not have a time minimum, so users spent less time on the task).

As shown in Table 1, we received 1,534 responses once we filtered out 53 bad actors who pasted parts of the prompt, URLs, or jibberish into the submission form. There could potentially be overlap between the participants in the unmediated questionnaire and the agent-mediated questionnaire, but the two questionnaires were administered weeks apart and most workers had very few prior HITs, so we do not expect many duplicate participants.

5 FINDINGS

Despite Mechanical Turk's reputation for skewed demographics, we found that our respondents were fairly representative of our ideal study population. Over 99% of respondents were from the United States. Respondents were slightly more male, with 58% male and 41% female. Most respondents were young but older respondents were still present, with 20% age 18-24, 43% age 25-34, 20% age 35-44, 10% age 45-54, and 6% age 55+. The distribution of educational attainment was skewed slightly higher than that of the US population [22]. Most respondents had earned at least a bachelors, with 34% having earned a high school degree, 50% having earned a bachelors (compared to 32% in the 2015 U.S. census), 12% having earned a masters, and 2% having earned a PhD (compared to 12% advanced degree holders nationally). However, despite higher levels of education, most respondents were not particularly knowledgeable about computer science. Only 9% of respondents categorized their occupation as computer worker, and 76% of respondents reported their exposure to CS concepts as either "low" or "none." 66% had never heard of the Internet of Things before.

5.1 Overview of Analysis

Mental models are difficult to observe directly, so we approach the analysis of the questionnaire responses from multiple angles. We compare three different attributes of the prompts: qualitative differences, characteristic words, and the operation profiles.

Table 2. Example responses for each abstraction. This table shows several responses for each abstraction, with individual responses separated by blank lines. These examples provide an intuitive feel for how the responses differ between prompts.

Unmediated Devices	Unmediated Data	Agent-mediated Devices	Agent-mediated Data
I would definitely want the smart watch to control the majority of the devices and controls in the house. The controllable lights would be a nice feature to have. I would definitely look for the smart door lock and smart thermostat. Those are things that make life easier for the convenience factor. The smart car would also be a good investment. Not only is it easier to use, it is energy efficient. Motion sensors in each room would save on electricity. [...] First and foremost, I would love to have a controllable TV that is hooked up to my video library. This would be not only an incredible time saver, but also a space saver as well. A temperature sensor on my smart phone that is hooked up to my thermostat would also be beneficial, as a traditional thermostat only takes into account the temperature near that thermostat. Along with the controllable TV, “smart" speakers that are hooked up to my music library would be nice. In fact, an app that turns my smart phone into a universal remote control might be my favorite “Internet of Things” application. It would offer tremendous convenience, as my smart phone is always within arms-reach. [...] I think it’d be cool to be able to put a sensor at the end of the driveway so that when certain cars drive up it will open up the garage doors [...]	I would want interface between my security system, smoke alarms, CO alarms, and cell phone. I would also want to be able to control the climate control systems (A/C and heat) from my cell phone, and monitor the temperature. It would also be nice if I could see my electricity usage in realtime, and customizable alerts sent to my phone would be quite helpful. I would like an app that tracks my heart rate along with the calories burned, sleep cycles, activity, whether I was calm or not. Having one in regards to my home, I would like to have one where I can see my child and watch the rooms where my caregiver is. Knowing where they are, or what they did. I wouldn’t be using this all the time as I trust my caregiver, but at times I would like to see how much time is spent in specific room, such as the living room. Or to track what TV apps (like Hulu or Amazon prime) was run on my TV and for how long. Maybe what was watched and the length of time. [...] I would like the ability to know how much water, electricity, and gas I use, with a running ticker of how much it is costing me. I would also like a breakdown of which rooms/objects are using the most. I would also like to know what lights are on, and if there is a window open in a room that is running heating or air conditioning. [...]	Set the temperature to 70 degrees. Lock the door. Close the blinds. Fetch and read my email. Please wake me up at 9:00 am with some pleasant music. Please make coffee at exactly 10:00am. Please set the alarm before I leave the house. Please water the yard at noon for 20 minutes. Please turn on the porch light before 7pm. Lower the lights, lock the doors, begin to play Madonna’s last two albums on shuffle from every speaker. Also synchronize the house RGB lights, my clothing lights and vibration patterns to the rhythm and tone of the music. Start my car and turn on the heat and radio. Open the garage. Lock the house doors. Adjust the temperature on the thermostat. Check all appliances and make sure they are off. Close garage after I drive away. Please be sure the door is locked and the temperature is set at 75 degrees. Next, be sure the volume is set on #15 as I want to listen to some music. Turn on my air conditioner 30 minutes before I get home at 4:00 pm. Turn TV on. Lock all doors. Adjust temperature to 70 degrees Celsius.	Please turn the red lights on dimly in my bedroom and start playing Marvin Gaye music. Dim all the other lights in the house. I would like you to make sure that when I leave the house, all lights, AC, and electronics are turned off and the door is locked. While I am gone, I would like you to monitor the house, and call my phone if anything strange happens (anyone enters the house, any objects are moved, etc.). In addition, I would like to use your knowledge of transportation to plan when to leave the house to catch the bus to work. You can also alert me to any poor weather conditions before they arrive. Thanks! Tell me my electricity consumption and gas consumption. How has my sleep been lately? When do I wake up? Am I exercising enough? Who is in the home with me? Where is my car and how fast is it going? How is my heart rate? I want you to turn on the lights when I walk into every room and play music whenever I am in the mood for it. Turn on the dishwasher when I get home from work so I can serve dinner, and make sure to alert me via a text message a half hour prior to my wife getting home so I have time to put the finishing touches on the meal.

Qualitative differences. We read all 1,534 of the responses, which was necessary to remove responses from obvious bad actors as described in Section 4.2. To guide our understanding of the structure of the responses, we also performed the exercise of sketching out a natural language grammar for a subset of the responses that included the top 100 most common nouns and verbs. We share some representative responses for each prompt in Table 2 to help the reader gain an intuition for the qualitative differences that underly the quantitative results.

Characteristic words and phrases. We used a χ^2 test to determine the keyness of each word used in each of the four response sets. Words with higher keyness are more suggestive of statistical differences between the response sets. This is a standard feature selection technique used in NLP to discover which words are particularly distinctive in a particular corpus compared to other corpora. For each word we calculated the rate of its occurrence per 1,000 words in the responses to each of the four prompts. Applying a χ^2 test to these rates produced the keyness values, which are simply the χ^2 statistic for each word. We display the rates and keyness values for the top 25 most key nouns, verbs, and other words in Figure 1.

Operation profiles. To paint a more rigorous picture of what users want to do in response to each prompt, we developed a set of labels to categorize sentences based on the operations defined in Section 3.2. These labels were "immediate action," "conditional action," "direct question," "indirect question," and "notification." After reading the responses to the unmediated prompts (see Table 2), we also added the labels "wants device," "wants remote control," "wants automation," and "wants to know." We also included a "none of the above" option for other kinds of sentences. Once we determined the set of labels that we were interested in, we took the first sentence of each response (for a total of 1,534 sentences) and asked three trained annotators to label the operations in each sentence, permuting the list of available labels that we displayed with each sentence to prevent biasing. The resulting Cohen's kappa statistics between the pairs of annotators were 0.76, 0.76, and 0.78, indicating a good level of inter-rater agreement. From the three sets of labels, we calculated the mean and standard deviation of the percentage of the responses to each prompt that contain a particular operation. Since a sentence could have multiple independent phrases in it with different kinds of operations (e.g., "Turn on the light and tell me my weight"), the percentages of responses that contain particular operations are independent from each other and do not sum to one hundred. The resulting operation profiles are shown in Figure 2. These profiles give us a sense for what respondents to a particular prompt wanted to do in their hypothetical smart home.

5.2 Finding 1: Priming has a powerful effect on users' mental models

As illustrated in Table 2, the answers we received to the four different prompts exhibited distinct qualitative differences, particularly between the unmediated and agent-mediated prompts. Responses to the unmediated prompt are expressed as hypothetical situations about what the respondent *would like* in their home ("I would like" and "I would want" were the most common 3-grams), whereas the responses to the agent-based prompt are expressed as executable directives ("turn on the" was the most common 3-gram for both the Device and Data prompts). The Unmediated Devices responses emphasize the devices that the respondent would like and why, whereas the Unmediated Data responses specify high-level "apps" that the user would like. While both agent-mediated prompts include immediate actions like "turn on the lights" and conditional actions like "turn on the lights when I get home," the Agent-mediated Data responses contain more questions and requests for monitoring and notification.

These observations are also reflected in the striking differences between operation profiles clearly visible in Figure 2. The Unmediated Devices responses are dominated by users expressing desires for devices, whereas the Unmediated Data responses show more requests for automation and information. The agent-based prompts both have immediate actions and conditional actions, but the Agent-mediated Data prompt also shows a large proportion of questions and notifications. Presenting the same smart home system in four different ways resulted in four distinct operation profiles.

NOUNS	app	sensors	phone	home	meter	devices	sensor	application	house	motion	things	system	water	alerts	internet	power	voice	ability	electricity	thermostat	think	food	device	time	security
Keyness	343	329	237	224	208	193	175	174	165	154	152	148	132	128	128	123	114	111	111	108	101	98	89	86	85
Rate in Unmediated Devices	1.2	3.9	5.0	7.4	2.1	3.0	2.2	0.6	7.8	2.2	3.0	3.0	3.0	0.5	2.0	2.6	2.1	1.0	0.7	3.1	1.5	1.0	1.4	2.5	2.0
Rate in Unmediated Data	4.8	0.2	2.2	9.2	0.0	0.5	0.4	2.4	7.3	0.2	2.7	2.5	4.0	2.3	0.4	1.2	2.5	1.9	2.9	0.7	1.5	1.9	0.7	2.9	1.7
Rate in Agent-mediated Devices	0.0	0.9	1.4	6.8	0.2	1.1	0.0	0.0	8.1	1.0	1.0	1.0	1.8	0.2	0.7	1.0	0.6	0.1	0.5	6.4	0.0	0.2	0.0	1.2	0.6
Rate in Agent-mediated Data	0.0	0.0	1.2	7.1	0.0	0.2	0.0	0.0	6.7	0.0	0.4	0.3	2.2	0.9	0.0	0.3	0.9	0.0	4.4	1.2	0.0	0.3	0.0	2.6	0.7

VERBS	be	like	have	want	control	think	know	love	use	turn	tell	cool	google	help	sleep	see	do	track	determine	monitor	dry	open	get	need	save
Keyness	1021	436	363	226	186	101	97	86	82	80	74	55	51	49	48	42	40	36	33	32	31	31	30	30	30
Rate in Unmediated Devices	20.4	11.0	10.7	9.2	4.7	1.5	1.6	1.5	2.6	4.1	1.1	1.2	0.9	1.3	0.3	1.2	2.1	0.9	0.0	2.3	0.4	1.2	2.3	2.0	1.0
Rate in Unmediated Data	17.9	12.4	7.5	8.8	3.4	1.5	4.0	0.9	2.6	4.2	3.2	0.4	0.2	1.5	1.5	1.6	2.0	2.3	0.4	1.7	0.0	2.0	1.6	1.6	0.9
Rate in Agent-mediated Devices	4.2	6.5	4.0	8.1	2.9	0.0	1.9	0.0	1.7	23.0	4.1	0.5	0.2	0.7	0.5	1.1	1.4	2.9	0.0	3.5	0.0	3.9	2.5	2.4	0.6
Rate in Agent-mediated Data	5.1	6.3	6.7	7.4	0.7	0.0	5.1	0.2	1.6	20.5	11.3	0.4	0.1	0.9	2.3	1.0	3.4	2.5	0.2	2.0	0.1	1.4	2.8	2.0	0.9

MISC.	would	smart	that	it	could	with	able	my	how	controllable	is	also	please	if	much	degrees	your	whether	what	connected	hooked	nice	great	this	when
Keyness	1613	1276	1005	605	428	316	304	288	240	231	204	201	189	175	157	144	144	137	127	124	118	110	104	101	101
Rate in Unmediated Devices	32.7	18.8	22	13.2	7	7.4	4.8	23.2	1.4	2.2	6	7.4	0	6	1.3	0.1	1.9	0.2	2.6	1.4	1.2	2	1.8	2.5	7
Rate in Unmediated Data	24.7	4.6	16.8	15.5	6.9	4	5.4	22.5	6.7	0	11.9	6.1	0	8.9	4.5	0.1	2.6	2.2	5.7	0.1	0	1.6	1.1	2.9	9.3
Rate in Agent-mediated Devices	6.6	2	5.6	5.3	0.8	3.1	0.8	34.2	1.9	0	7.6	5.8	8.1	7.9	0.7	7	0.2	0.2	2.5	0.3	0	0.4	0.1	1.2	12
Rate in Agent-mediated Data	6.6	1	6.1	7.5	0.4	1.6	0.4	28.1	8.9	0	19.1	4.7	7.6	7.3	5.9	5.9	0.3	1.4	9	0	0	0.1	0.1	1.8	11.6

Fig. 1. Characteristic nouns, verbs, and miscellaneous words for the different abstractions. This figure shows the top 25 nouns, verbs, and remaining words sorted by keyness as determined by their χ^2 statistic. Words with higher keyness are more suggestive of statistical differences between the response sets. The rates show how often each word occurs per 1,000 words within each set of responses. Differences between prompts can be identified when the rates vary within a column. Notable outcomes here are that the Unmediated Devices responses are characterized by the words like "sensor[s]", "device[s]", "meter", and "things," with an emphasis on "control" and "controllable," through a "phone." This suggests users were focused mainly on remote control of devices through a phone. The Unmediated Data responses, on the other hand, emphasize "apps" and "applications," as well as "alerts" and "know[ing]," though they also still score highly on "control" and words associated with conditional actions and notifications like "if" and "when." This supports the notion that this conceptual model will encourage users to express various kinds of requests for information and automation in addition to remote control commands. The two agent-based conceptual models showed mostly similar word rates, suggesting that placing an intermediary between the system's devices or data capabilities smoothes out some of the differences seen in the two unmediated prompts. Both showed high rates for "please" and "turn." However, for the agent-mediated data prompt, there was more of an emphasis on "tell" and less on "control", a mental shift reflected in other query-related words like "know," "how," "much" and "what."

5.3 Finding 2: Different abstractions produce distinct mental models

While the observation that abstractions have a priming effect on mental models has implications for research, we think that characterizing the specific effects of each abstraction would be useful as well, particularly for system designers hoping to understand the implications and trade-offs involved in choosing a particular abstraction over another. Below we take a deep dive and characterize the trends associated with the responses to each prompt.

Islands. The operation profile for responses to the Unmediated Devices prompt, as shown in Figure 2, is dominated by "wants device," with 53% of sentences expressing a desire for a particular device or devices. The example responses provided in Table 2 give some intuition for how users conveyed this, with sentences like, "I would definitely want the smart watch[...]" and "I would love to have a controllable TV." In Figure 1 you can see that particularly characteristic of responses to this prompt are words like "sensor[s]", "device[s]", "meter", and "things," with an emphasis on "control" and "controllable." The word "phone" is particularly associated with this conceptual model as well. The overall picture that emerges is that users responded to the Unmediated Devices prompt by expressing a desire to have controllable smart devices that the user can control manually, primarily with their phone, with surprisingly few automatic applications running on the smart home system.

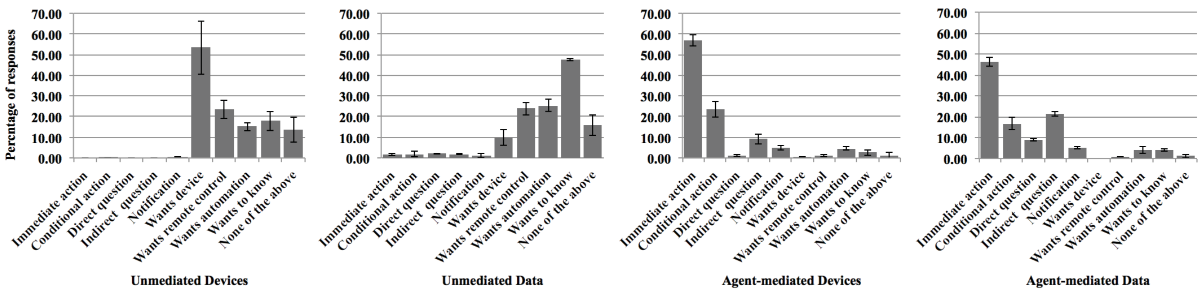

Fig. 2. Operation profiles found in the four sets of responses. These charts show the proportion of different kinds of operations in the user responses to the questionnaires. We took the first sentence of each response (for a total of 1,534 sentences) and asked three trained annotators to label each sentence from a set of labels that we provided based on a qualitative exploration of the dataset. They could provide multiple labels if there were multiple phrases, so the percentages are independent and do not add up to 100. One notable observation is that the operation profiles are all different, demonstrating that abstractions had a major priming effect on our users. When combined with qualitative analysis and the words and phrases associated with each prompt, we can get a picture of the mental models behind each of these distributions.

When presented with the Unmediated Devices abstraction, users seem not to see "smart devices" as connected devices, instead tending to treat them as isolated islands of functionality. In response to a prompt that said, "[describe] different *applications* that you would want in your smart home," the majority of users expressed a desire to have a *device*. This suggests that when presented with the Unmediated Devices abstraction, users tend to think of the device itself and the application behavior as one and the same. The relative lack of higher-level applications that run across multiple devices suggests that respondents to this prompt did not tend to perceive a global computational scope or a sense of general programmability in their smart home.

Watchdog. Users presented with the Unmediated Data abstraction showed a propensity for higher-level applications compared to the Unmediated Devices abstraction. Figure 1 shows that the words "app" and "application" were particularly associated with responses to this prompt. The operations profile in Figure 2 shows that users expressed much more desire for automation when responding to this conceptual model compared to the Unmediated Devices model.

However, the majority of sentences were labeled as "wants to know." Many of the applications described in the Unmediated Data responses focus on monitoring things and providing alerts to the users. You can see this in Figure 1 where the words "alerts," "know," "see," and "track" were found to be strongly associated with this prompt. Also ranked highly were measurable phenomena, like "water," "electricity," and "sleep". Users also wanted to be able to solicit information from the system on demand, as evidenced by the relative keyness of "what," "is," "how," and "much." Another significant word was "if," which is associated both with *doing* (in the form of trigger-action automation) as well as with *knowing* (notifications and alerts). We called this trend in mental models *watchdog* to convey that users tended to think of the system as an unintelligent global observer whose primary job is to monitor and inform users, and whose global scope gives it the ability to run integrated automation applications across the system.

Delegate. The operation profile of the responses to the Agent-mediated Devices prompt shown in Figure 2 reveals that most responses (57%) were immediate actions like "turn on the lights," and nearly a quarter of responses were conditional actions. This means that respondents were primarily tasking the agent with undertaking actions on their behalf, either immediately or in an automated way, hence the name *delegate*. However, the users were also very polite, as "please" was the 20th most common word, and "I would like you" and "would like you to" were the second and third most common 4-grams respectively. This politeness suggests that the users perceived the system to have social agency.

An interesting pattern that occurs not just here, but in both agent-based prompts, is providing a reason or goal after specifying actions to take. One example from the Agent-mediated Devices prompt is, "adjust the blinds every morning around 5, so that I wake up easier when my alarm goes off." Other reasons that users expressed to explain actions were things like, "so I can save energy," "so I can watch Netflix," "so I can go to sleep," and "so I can let [someone] know [something]." Providing goals as a part of instructions suggests that users potentially expect agents to understand basic goals. If the system were able to support this, it would provide the agent with a great deal of flexibility in achieving the specified goals. For example, if in the first example the system could detect that the user would like aid in waking up based on what they said, the system could potentially suggest (or attempt) supplemental strategies like turning on the lights to full brightness, or playing additional sounds. The ability to comprehend even basic goals would give a smart home a lot of power to improvise and make suggestions to aid its occupants.

Assistant. Like the Agent-mediated Devices prompt, the Agent-mediated Data prompt produced mental models that supported primarily immediate actions (46%) with some conditional actions (17%), but the operation profile shows that unlike in the responses to any other prompt, 39% of the operations were labeled as questions (both direct and indirect) or other requests for information. While requests for information have seldom appeared in previous research on smart home mental models, it is clear that this underrepresentation in the literature is not out of a lack of user interest in queries as a form of interaction. We called this mental model *assistant* to reflect the fact that in addition to performing all the same functions as the delegate (specifically, remote control-oriented immediate actions and automation-oriented conditional actions), the assistant also provides a substantial amount of informing and notifying of users.

One interesting observation is that despite the strong presence of questions in the Agent-mediated Data response set, the word "why" does not appear a single time in any of the Agent-based responses (but does appear in responses to the Unmediated prompts). This suggests that users do not think that agents are capable of either introspection ("why did you turn on that light?") or explanation ("why is my energy bill so high?"). It is possible that this is due to the limitations of our method of data collection, which did not involve repeated interactions with the agent. These questions might arise if users lived in a smart home and interacted with an agent over a long period of time. Nevertheless, it is interesting that not a single user thought to ask an agent "why."

5.4 Finding 3: Subpopulations behave differently in response to some, but not all, abstractions

Prior work has shown that occupants interact differently with smart homes depending on their technical expertise and age. Technical occupants often assume the role of programming and maintaining the system, whereas non-technical occupants communicate their wants and issues to the technical member for translation into programmatic instructions [12]. Ethnographies have also shown that older occupants and younger occupants may have different needs and expectations from smart home interfaces [19].

We analyzed demographic subpopulations to better understand differences in the way these populations respond to conceptual models. We determined technical expertise by whether the respondent self-reported their exposure to computer science concepts as None ("No exposure to ideas of computer science"), Low ("Some exposure to computer science concepts"), Medium ("Undergraduate computer science student"), or High ("Computer science graduate student or professional"). When determining age, we define older occupants as those whose age is 55 or older, and younger occupants as those who are 34 or younger. We used a χ^2 test on the observed frequency of operations in the populations' responses to each prompt to determine whether any of the operations significantly differed between the two populations we compared.

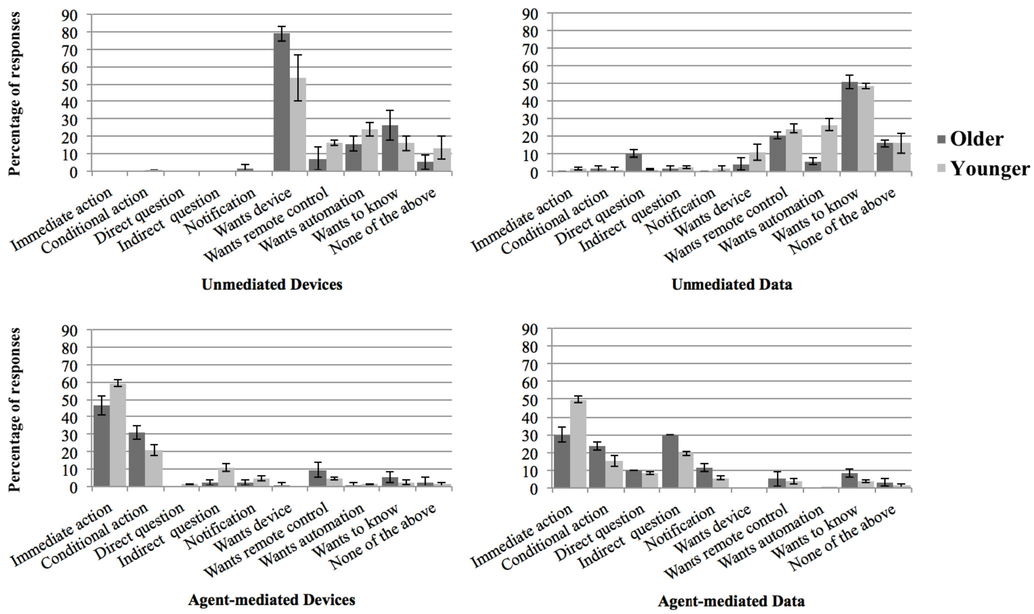

Fig. 3. Comparison of operation profiles for older users vs. younger users.

Figure 3 shows that older respondents respond more strongly to priming than younger respondents, whereas younger respondents tend to adhere less to the primed model. Given the Unmediated Devices conceptual model, older respondents were significantly more likely to want a device ($p = 0.025$), and in the Unmediated Data responses, older respondents were more likely to ask direct questions ($p = 0.007$). Conversely, when presented with an Unmediated Data conceptual model younger respondents were more likely than older respondents to want remote control capabilities (0.0003). Younger respondents were also more likely to ask indirect questions in the Agent-mediated Devices responses despite the device abstraction ($p = 0.014$), and express immediate actions in the Agent-mediated Data responses despite the data abstraction ($p = 0.026$).

Figure 4 tells a somewhat similar story. While there are no significant differences between respondents with high CS exposure and no CS exposure when presented with the familiar Unmediated Devices conceptual model, the populations diverge when presented with the remaining three less-familiar models. Given the Unmediated Data abstraction, those with high CS exposure are more likely to recognize the ability to perform automation ($p = 0.037$), whereas those with no CS exposure are much more likely to stick with expressing a general desire to know things ($p = 0.029$). Presented with the Agent-mediated Devices prompt, respondents with high CS exposure were more likely to say indirect questions, whereas those without the training would still say they wanted a device ($p = 0.0463$) or automation ($p = 0.001$). Finally, when presented with the Agent-mediated Data abstraction, those with high CS exposure were significantly more likely to request that the system perform immediate actions ($p = 0.0234$). As with older vs. younger respondents, those with no CS exposure tended to conform closely to the primed conceptual model, whereas those with high CS exposure were able to invoke functionality that did not conform as closely to the model. The similarity between these results is not due to correlation between younger age and more computing expertise, as those who self-reported higher CS exposure tended to be older.

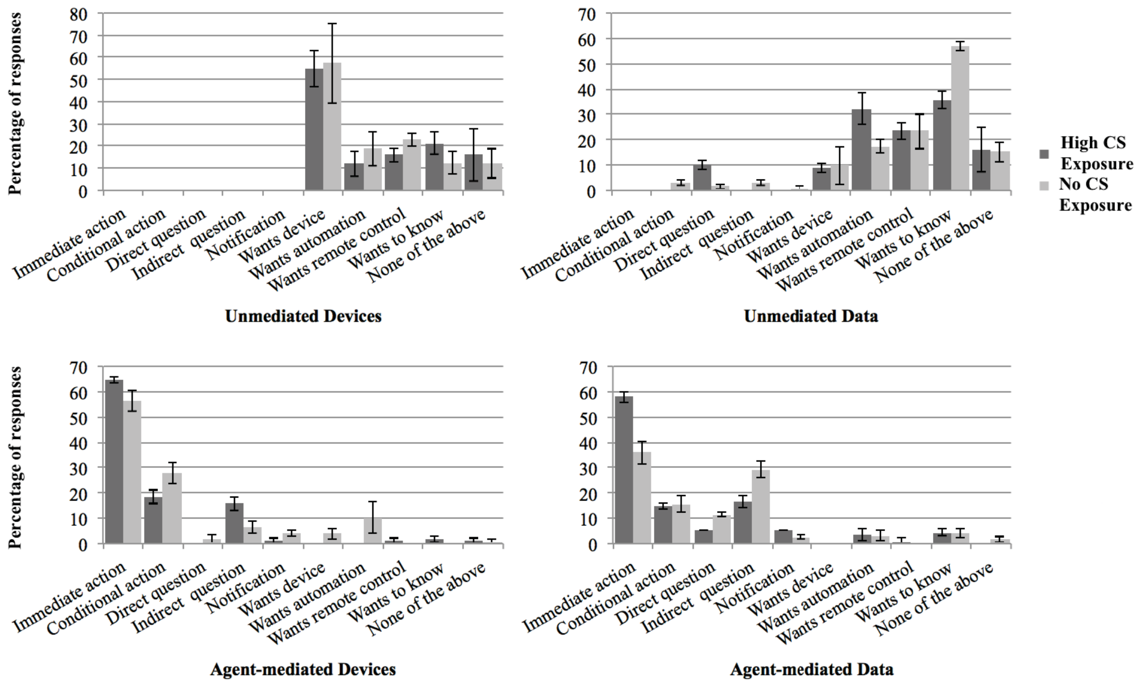

Fig. 4. Comparison of operation profiles for users with high computer science exposure vs. no computer science exposure.

6 DISCUSSION

Our findings about the effect of priming on users' mental models has significant implications for mental model research and system design. Researchers studying mental models for the purpose of building intuitive end-user interfaces need to be aware of the priming effects of system abstractions. Researchers must assume that the abstractions used in the study scenarios will prime users, and that is a good thing because so will the abstractions used to present the actual system. Researcher will need to choose abstractions consciously in a principled way, so that the mental models and operation profiles resulting from different abstractions can be compared. Building up a corpus of conceptual models and their associated mental models and operation profiles will be beneficial for system designers who want to know what abstractions they should select without having to spend too much money or time running their own user studies.

In terms of the architectural implications for the domestic Internet of Things, system designers will need to embrace priming as a part of their system design space. The abstractions used to represent the system will strongly influence the kinds of interactions that users will task the system with. This means that higher-level choices about the way the system is presented to users will have lower-level consequences for the system requirements. System designers will need to become aware of the coupling between the broad space of abstractions available to represent the system and the operational implications of each option.

System designers will also need to keep their target audience in mind. The operation profile is not just a function of the abstraction, but also the subpopulation interacting with the system. Technically-inclined users may have a deeper understanding that allows them to discover more functionality than the conceptual model conveys, whereas non-technical people will only be able to express those desires that can be clearly satisfied within the confines of the presented model. Similarly, older occupants will also be greatly influenced (and constrained) by the conceptual model.

Currently, systems exist that do not support interoperability between devices very well, or which feature single-purpose devices. For those systems, the Unmediated Devices abstraction may be a sensible choice of abstraction, as manual one-on-one interactions with individual devices are what the system can most easily support. However, the value of the Internet of Things is projected to come from the ability to run integrated applications on general-purpose physical computing systems. In transitioning from a special-purpose to a general-purpose future, and as the number of devices scales up, it will be important to keep in mind other abstractions that we could transition to away from Unmediated Devices that would encourage users to think more broadly. Since system designers may reasonably wish to avoid the implementation complexity that comes with users' expectations of agents, it is significant that our findings showed that the Unmediated Data abstraction also encouraged users to think of high-level automation applications, presumably due to the global scope that "monitoring" data implies.

Finally, we found that end users will comfortably use primitives that have traditionally fallen under the umbrella of AI rather than systems. Previous work has identified that in the smart home domain users will employ "fuzzy triggers" which specify qualitative preferences like "comfortable," "normal," and "sufficient" that must be learned by the system [21], but we identify two new concepts in user applications that would require artificial intelligence techniques to implement:

Prediction. Some commonly-occurring commands (actions relying on the condition "before," such as "have coffee ready before I wake up") and questions (often questions starting with "when will," such as "when will my husband be home") require prediction. The basic nature of these primitives suggests that machine learning for prediction tasks needs to be included in IoT systems at a fundamental level.

Goals. Many responses to the agent-based prompts describe a goal and provide example actions in order to demonstrate to the agent how to achieve the goal ("Do X so that Y"). Systems that wish to support intuitive end-user programming may therefore benefit from drawing on the extensive work done in artificial intelligence on goal-based agents, planning, and learning from demonstration.

7 LIMITATIONS

The questionnaire methodology that we used does not rule out the possibility that the priming we see is a minor initial effect that can be overruled with training and experience interacting with the system. However, even if the initial impressions do not have a lasting effect on user interactions once the user has become accustomed to the system, these primed mental models still affect purchasing and adoption decisions. A user should be able to look at the presentation of a system and envision valuable interactions before actually interacting with the system. Once interaction begins, their conceptualization of the system as a useful based on its abstraction should not be challenged by the way the system actually behaves.

The questionnaire methodology also fails to capture the repeated interactions that users would have with the home over time. Previous work has shown that there is a smart home app development lifecycle that begins with brainstorming the application, as we ask users to do, but then also includes stages where the user iteratively improves it as it runs in the house and bugs or unexpected behaviors emerge [24]. Our questionnaires do not capture the debugging phase, which means missing certain kinds of interactions like introspective queries (for example, the word "why" doesn't appear once in the agent responses, even though in practice users do want to know "why" things happen).

Finally, while the Mechanical Turk population is not entirely representative of the general populace, one area where it does excel is that we had a large percentage and a large absolute number of non-technical respondents, which is an improvement over many smart home user studies.

8 CONCLUSION

In this work we show that the abstractions used to present a smart home system to end users have a significant priming effect on users' mental models and the kinds of interactions that users expect to have with the home. We introduce a preliminary framework that identifies several dimensions along which conceptual models can be described, and then focus particularly on abstractions situated along two of those dimensions: the system capabilities (devices vs. data), and the system agency (unmediated vs. agent-mediated), the combination of which results in four archetypal abstractions. We then characterize the mental models that users form when presented with those abstractions. We use Amazon Mechanical Turk to collect over 1,500 responses to questionnaires where we describe a single hypothetical smart home's capabilities using one of the four abstractions. To gain insight into users' mental models, we analyzed the qualitative differences between the sets of responses, as well as differences in the types of tasks the users ask the system to perform, and the words characteristic of each prompt according to statistical tests. Based on our analysis of these three signals, we find that different abstractions strongly affect the mental models the users have of the system and the ways that they want to interact. We also found that older users and users without computer science expertise tend to respond more strongly to priming than younger users and users with computer science expertise.

Our findings have major implications for both HCI researchers and system designers. HCI researchers studying mental models will need to be aware that there is no "natural" way that end users think about these systems and that priming should be embraced as an omnipresent influence on mental models, requiring researchers who are trying to understand mental models to do comparative evaluations of different system abstractions. System designers should consider the choice of abstraction to be part of their design space, since it will heavily influence the kinds of operations that the system will need to support.

We also found that end users employ computational primitives that have traditionally fallen under the umbrella of artificial intelligence, such as prediction and goal-oriented planning. Consequently, HCI researchers, system designers, and AI researchers must take active measures to bridge the institutional gaps between them and collaborate closely on future platforms.

Finally, we found that the most popular abstraction used in IoT systems today, the Unmediated Devices abstraction, produced the most limited kinds of interactions, characterized by manual one-on-one interactions with individual devices. High-level applications like automation and queries often do not even occur to users presented with this abstraction, though they do occur when users are given different abstractions of the same system. This suggests that as the Internet of Things architecture continues to evolve towards interoperable general-purpose physical computing systems, we should move away from the Unmediated Devices abstraction if we want end users to think of valuable, integrated applications that operate across a network of devices.

ACKNOWLEDGMENTS

This work was supported in part by the TerraSwarm Research Center, one of six centers supported by the STARnet phase of the Focus Center Research Program (FCRP) a Semiconductor Research Corporation program sponsored by MARCO and DARPA. The work was also supported by the National Science Foundation Graduate Research Fellowship. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Special thanks to Catherine Finegan-Dollak and Laura Wendlandt for their help with annotations.

REFERENCES

[1] Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash. 2015. Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Communications Surveys and Tutorials* 17, 4 (2015), 2347–2376. https://doi.org/10.1109/COMST.2015.2444095

[2] Amazon.com, Inc. 2015. Amazon Mechanical Turk. https://www.mturk.com. (2015).

[3] Peter A. Bibby and Stephen J. Payne. 1996. Instruction and Practice in Learning to use a Device. *Cognitive Science* 20, 4 (oct 1996), 539–578. https://doi.org/10.1207/s15516709cog2004_3
[4] Mike Black, Katie Dana, Kim Gaskins, Courtney Gaynier, Al Lemieux, and Kyle McKinley. 2014. *The Internet of Things: Can it find a foothold with mainstream audiences today?* Technical Report. The Nielsen Company.
[5] Ann Bostrom, Cynthia J Atman, Baruch Fischhoff, and M Granger Morgan. 1994. Evaluating risk communications: completing and correcting mental models of hazardous processes, Part II. *Risk analysis* 14, 5 (1994), 789–798.
[6] Barry Brumitt and Jonathan J Cadiz. 2001. Let there be light" examining interfaces for homes of the future. In *Human Computer Interaction. INTERACT'01. IFIP TC. 13 International Conference on Human Computer Interaction. IOS Press, Amsterdam, Netherlands; 2001; xxvii+ 897 pp.* 375–82.
[7] Anind K. Dey, Timothy Sohn, Sara Streng, and Justin Kodama. 2006. iCAP: Interactive prototyping of context-aware applications. In *International Conference on Pervasive Computing*. https://doi.org/10.1007/11748625_16
[8] Eric Horvitz. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 159–166. https://doi.org/10.1145/302979.303030
[9] D V Keyson, M P A J De Hoogh, A Freudenthal, A P O S Vermeeren, and D Keyson. 2000. The Intelligent Thermostat: A Mixed-Initiative User Interface. *CHI* (2000).
[10] Gerd Kortuem, Fahim Kawsar, Vasughi Sundramoorthy, and Daniel Fitton. 2010. Smart Objects as Building Blocks. *IEEE Internet Computing* 14, 1 (2010), 44 – 51.
[11] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really bad PA": The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/2858036.2858288
[12] Sarah Mennicken and Elaine M Huang. 2012. Hacking the Natural Habitat: An In-the-Wild Study of Smart Homes, Their Development, and the People Who Live in Them. *LNCS* 7319 (2012), 143–160.
[13] M Granger Morgan. 2002. *Risk communication: A mental models approach*. Cambridge University Press.
[14] Jörg Niewöhner, Patrick Cox, Simon Gerrard, and Nick Pidgeon. 2004. Evaluating the efficacy of a mental models approach for improving occupational chemical risk protection. *Risk Analysis* 24, 2 (2004), 349–361.
[15] Donald A Norman. 2013 [1988]. *The Design of Everyday Things*. Vol. 16. 272 pages. https://doi.org/10.1002/hfm.20127 arXiv:arXiv:1011.1669v3
[16] Pankesh Patel and Animesh Pathak. 2011. Towards application development for the Internet of Things. In *ACM/IFIP/USENIX 12th International Middleware Conference*. 1145–711266. https://doi.org/10.1145/2093190.2093195>
[17] Stephen Payne, Helen Squibb, and Andrew Howes. 1990. The Nature of Device Models: The Yoked State Space Hypothesis and Some Experiments With Text Editors. *Human-Computer Interaction* 5, 4 (dec 1990), 415–444. https://doi.org/10.1207/s15327051hci0504_3
[18] Devika Pisharoty, U Virginia, Rayoung Yang, and Mark W Newman. 2015. ThermoCoach: Reducing Home Energy Consumption with Personalized Thermostat Recommendations. In *2nd ACM International Conference on Embedded Systems For Energy-Efficient Built Environments*. https://doi.org/10.1145/2821650.2821671
[19] François Portet, Michel Vacher, Caroline Golanski, Camille Roux, and Brigitte Meillon. 2013. Design and Evaluation of a Smart Home Voice Interface for the Elderly: Acceptability and Objection Aspects. *Personal Ubiquitous Comput.* 17, 1 (Jan. 2013), 127–144. https://doi.org/10.1007/s00779-011-0470-5
[20] Khai N Truong, Elaine M Huang, and Gregory D Abowd. 2004. CAMP: A Magnetic Poetry Interface for End-User Programming of Capture Applications for the Home. In *International Conference on Ubiquitous Computing*.
[21] Blase Ur, Elyse Mcmanus, Melwyn Pak, Yong Ho, and Michael L Littman. 2014. Practical Trigger-Action Programming in the Smart Home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 803–812. https://doi.org/10.1145/2556288.2557420
[22] U.S. Census Bureau. 2016. Educational Attainment in the United States: 2015. U.S. Department of Commerce. (March 2016).
[23] Roy Want, Bill N Schilit, and Scott Jenson. 2015. Enabling the Internet of Things THE IOT VISION. *Computer* 48, 1 (2015), 28–35.
[24] Jong-bum Woo and Youn-kyung Lim. 2015. User Experience in Do-It-Yourself-Style Smart Homes. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. https://doi.org/10.1145/2750858.2806063

A APPENDIX: QUESTIONNAIRES

In this section we provide the original questionnaire prompts that we provided to our Mechanical Turk subjects. There are four questionnaires: Unmediated Devices, Unmediated Data, Agent-mediated Devices, Agent-mediated Data.

A.1 Unmediated Devices

This prompt has three screens. The first introduces a list of devices available in the smart home, which the respondent cannot move on from until a one-minute timer expires. The second screen provides the same list of devices for the respondent to reference and asks the respondent to write for five minutes (enforced by timer) to describe different applications they would want in their smart home. The final screen consists of demographics questions.

A.2 Unmediated Data

This prompt has three screens. The first introduces a list of data streams available, which the respondent cannot move on from until a one-minute timer expires. The second screen provides the same list of data for the respondent to reference and asks the respondent to write for five minutes (enforced by timer) to describe different applications they would want in their smart home. The final screen consists of demographics questions.

A.3 Agent-mediated Devices

This prompt has four screens. The first asks the user to choose a voice gender (male, female, or androgynous) and a name for their smart home agent. The second screen presents a list of devices that the agent has access to in the home (the same list of devices as in the unmediated prompts). The third screen does *not* provide a list of the devices for reference, and asks the respondent to write (with no time limit) about what they want their agent to do. The final screen consists of demographics questions.

A.4 Agent-mediated Data

This prompt has four screens. The first is the same as the first screen in the Agent-mediated Devices questionnaire. The second screen presents a list of data that the agent has access to (the same list as in the unmediated prompts). The third screen is also the same as in the Agent-mediated Devices questionnaire. The final screen consists of demographics questions.

Received February 2017; revised May 2017; accepted July 2017

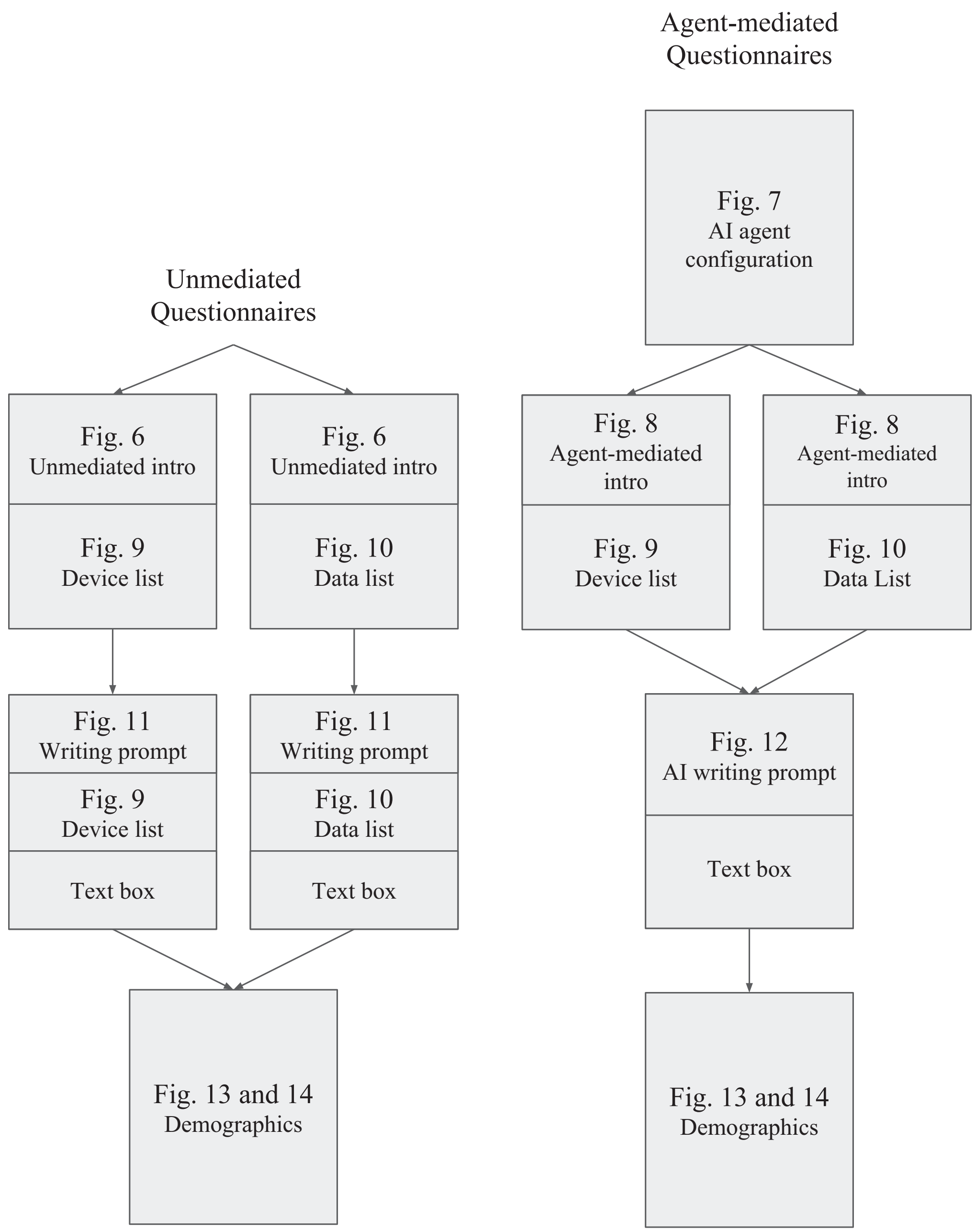

Fig. 5. An overview of the questionnaire screens and where to find their figures. Each gray block represents a single screen view. To avoid redundancy we have broken up the screens into different figures as shown above.

The blue "Next" button will appear in one minute.

Over the next few minutes, you will imagine and write about potential Internet of Things applications. For inspiration, take at least one minute to look over this list of example devices that you have at your disposal (you will be able to refer back to this list later):

Fig. 6. The introduction for the Unmediated Device questionnaire. The Unmediated Data questionnaire introduction is identical, except that it says "data" instead of "devices."

It is the future, and you have just bought your very own Artificial Intelligence (AI) agent, like Siri -- except that this AI can see and control the physical world around you as well. This AI is totally loyal to you.

Pick a gender for your AI's voice:

Pick a name for your AI:

Fig. 7. The first screen for the Agent-mediated questionnaires. On this screen, users personalize their AI by choosing a voice gender (menu order was randomized between male, female, and androgynous) and choosing a name.

The blue "Next" button will appear in one minute.

You have chosen to install your AI -- "Samantha" -- into your futuristic smart home. As your personal assistant, Samantha wants to automate your home for you. Below is a list of data that Samantha has access to in your home. Take at least one minute to look over this list and think about what you want to tell Samantha to do (you will need to write it on the next page):

Fig. 8. The introduction for the Agent-mediated Data questionnaire. The Agent-mediated Devices questionnaire is identical, except that it says "devices" instead of "data."

Controllable devices:
Controllable RGB lights
Controllable speakers hooked up to your music library and services
Controllable TV hooked up to your video library and services
Smart watch
Clothes that vibrate and/or light up
Smart door lock
Smart phone
Smart outlets
Smart thermostat
Internet-connected microphone that can speak to you and understand your voice commands, questions, etc.
Smart car (remote heating/air conditioning, locks, engine, lights, music)

Non-controllable devices:
Motion sensors in every room
Reed switches (a.k.a. close/open sensors) for doors and windows
Step counter/fitness tracker
Wireless scale
Heart rate monitor wristband (measures continuously)
Glucose monitor
EEG headset (a.k.a. brainwave reader)
iBeacons (they emit wireless signals that can be detected by nearby phones and devices)
GPS/iBeacon tags that can be attached to objects, pets
Accelerometer tags that can be attached to objects, pets
Pressure sensors
Temperature sensor
Humidity sensor
Light sensors
Smart power meter
Smart water meter
Smart gas meter

Digital "devices":
Web services

- Social networks (Facebook, Twitter, Foursquare, Instagram, etc.)
- Transportation (Public transit systems, Uber, Lyft, etc.)
- Evernote, Dropbox, Campfire
- Google Drive, Calendar, Gmail, YouTube
- RSS feeds (News sites, etc.)

Fig. 9. The list of devices for the Unmediated Prompt. The Agent-mediated device list is the same, except that the section headings include the agent, e.g. "Devices controllable by <AI name>."

Data that an application can access and change:
Whether the lighting is on or off
What color the lighting is
How bright the lighting is
What song is playing in a room
What appliances are running
Indoor temperature and humidity
Text (SMS) content/alerts, email content/alerts, social network content/alerts
Whether the door is locked
Voice announcements
Voice-recognition of commands, questions, feedback
Whether the car heating/air-conditioning is on

Data that an application can access:
Who is in what room
How much you weigh
How much you are exercising right now
How much you have exercised historically
How much you sleep and when you sleep
Whether you are asleep now
What time you will get home
What your current blood sugar level is, and your historical blood sugar trends
What your current heart rate is, and your historical heart rate trends
Whether you are currently calm, sleep deprived, focused, scatter-brained, anxious
Current weather, the weather forecast
Current locations of transportation services like buses, Uber, etc.
Where specific objects are located
Whether a surface or object is being touched
Whether an object has been moved
Whether a door or window is open or closed
What space you are currently at (home, work, specific store, specific restaurant, specific cafe, etc.)
Who you are currently with, who you have been with
How much electricity you consume, when you consume it, whether you are currently consuming it
How much electricity an appliance consumes, how much it is currently consuming
How much water you use, when you used it, whether you are currently using it
How much gas you use, when you used it, whether you are currently using it
How much electricity, water and gas cost
When you drive your car, where you drive it to
Where your car is and how fast it is traveling

Digital data available to an application:
Web service alerts and content

- Social network alerts/content (new status posts, tweets, check-ins, photos, etc.)
- Transportation alerts (approaching buses, how close the nearest Uber or Lyft is, etc.)
- Organizational alerts/content (a new Evernote memo, Dropbox file, Campfire message)
- Google app alerts/content (someone edited a doc, an upcoming calendar event, a new email)
- RSS feed alerts/content (newly posted news articles, etc.)

Fig. 10. The list of data for the Unmediated Prompt. The Agent-mediated device list is the same, except that the section headings include the agent, e.g. “Data that <AI name> can access and change.”

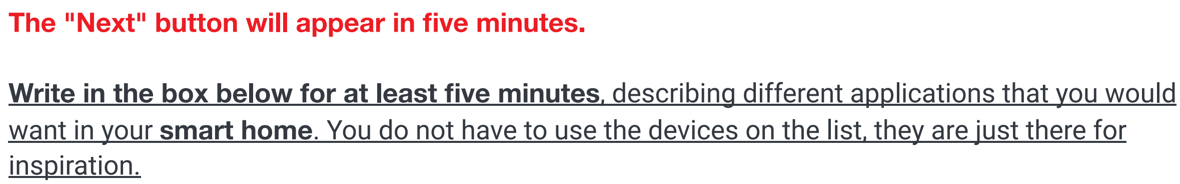

Fig. 11. The writing prompt for the Unmediated Device questionnaire. The Unmediated Data prompt is the same, except it says “data” instead of “devices.”

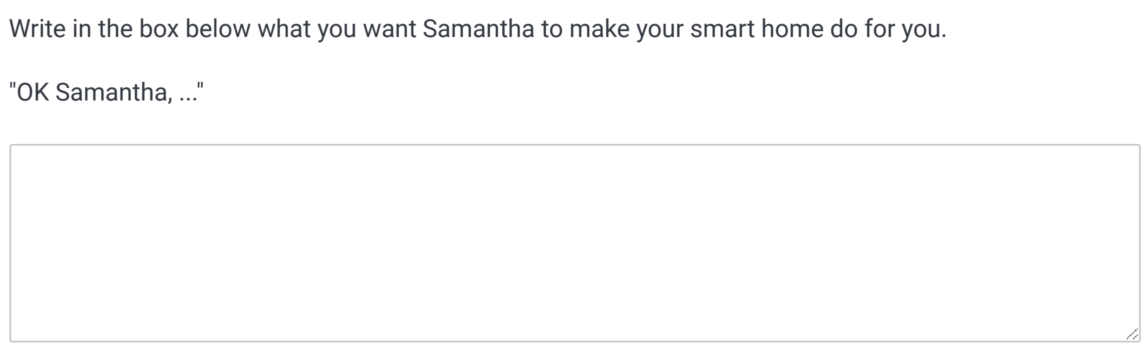

Fig. 12. The writing prompt for all Agent-mediated questionnaires.

Had you heard of the Internet of Things before this survey?

- ◯ Yes
- ◯ No

What is your current age?

- ◯ 18-24
- ◯ 25-34
- ◯ 35-44
- ◯ 45-54
- ◯ 55-64
- ◯ 65+
- ◯ Prefer not to answer

What is your gender?

- ◯ Female
- ◯ Male
- ◯ Other
- ◯ Prefer not to answer

Does your living situation include more than one person?

- ◯ Yes
- ◯ No
- ◯ Prefer not to answer

What is the highest level of education you have currently attained (or the closest equivalent):

- ◯ Less than high school
- ◯ High school
- ◯ Bachelor's
- ◯ Master's
- ◯ PhD

Fig. 13. The first part of the demographics questions presented at the end of each questionnaire.

○ Prefer not to answer

What is your level of exposure to computer science?

○ High - Computer science graduate student or professional
○ Medium - Undergraduate computer science student
○ Low - Some exposure to computer science concepts
○ None - No exposure to ideas of computer science

Which occupation group best matches your current job?

In which country do you currently reside?

Comments (optional):

Fig. 14. The second part of the demographics questions.