IoT Interfaces for Everyday People

Meghan Clark
Intel/NSF CPS-Security Final PI Meeting
7/12/2018

mclarkk@berkeley.edu
Consumer IoT is coming.

Over 40 billion IoT devices by 2021

Consumer IoT is coming.

But how will we interact with it?
Little has been done to explore the trade-offs between these representations.
What if certain representations discourage users from thinking of classes of interactions?
What are mental models and how are they formed?

What are mental models and how are they formed?
What are mental models and how are they formed?
What are mental models and how are they formed?
What are mental models and how are they formed?

In this talk, like Norman, I use conceptual model and mental model to refer to separate ideas.
What are mental models and how are they formed?

Initial exposure to conceptual models may prime user’ mental models.
We should incorporate priming into the design process

Works that assume “natural” mental models for IoT technologies:

- iCAP: Interactive Prototyping of Context-aware Applications [1]
- Practical Trigger-Action Programming in the Smart Home [2]
- CAMP Magnetic Poetry [3]
  - Tried to avoid biasing users by providing scenarios in comic form

Unlike prior work, we assume that:

- All scenario descriptions and interfaces express conceptual models that will prime users
- Priming is not inherently bad
- We should incorporate priming in our system design to learn how to intelligently shape user interactions and align them with system capabilities


In this work, we examine the following three questions:
In this work, we examine the following three questions:

• To what degree do different abstractions prime end-user mental models?
In this work, we examine the following three questions:

• To what degree do different abstractions prime end-user mental models?
• What are the specific effects and trade-offs of common conceptual models?
In this work, we examine the following three questions:

• To what degree do different abstractions prime end-user mental models?
• What are the specific effects and trade-offs of common conceptual models?
• Do different populations respond differently to conceptual models?
How can we compare conceptual models and their effects?
How do we control for individual differences in prior experiences?
How do we control for individual differences in prior experiences?
How do we compare conceptual models?
We examine four *abstractions* along two *dimensions*, for a total of four conceptual models.
Now we can compare conceptual models.
How do we compare user responses to different abstractions?

THE STUDY
We deployed four questionnaires to Mechanical Turk.
Write for five minutes about what applications you want

Pick gender and name for “smart home AI”

“Okay, <AI name>...”
We had 1,535 respondents in total

<table>
<thead>
<tr>
<th>Conceptual model</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unmediated Devices</td>
<td>313</td>
</tr>
<tr>
<td>Unmediated Data</td>
<td>302</td>
</tr>
<tr>
<td>Agent-mediated Devices</td>
<td>442</td>
</tr>
<tr>
<td>Agent-mediated Data</td>
<td>478</td>
</tr>
</tbody>
</table>
Our subjects are representative of who we want to study

- Age, gender, and education similar to U.S. pop
- Overall non-technical, an improvement over past studies

<table>
<thead>
<tr>
<th></th>
<th>Our study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not a computer worker</td>
<td>91%</td>
</tr>
<tr>
<td>CS exposure low/none</td>
<td>76%</td>
</tr>
<tr>
<td>Never heard of IoT</td>
<td>66%</td>
</tr>
</tbody>
</table>
Are our subjects representative of who we want to study?

Our study population is slightly more male than the US population
Are our subjects representative of who we want to study?

Our study population skews younger than the US population
Are our subjects representative of who we want to study?

Our study population skews more educated than the US population

[Bar graph showing education levels: High school, Bachelors, Advanced degree for 'Our study' and 'US population']
Are our subjects representative of who we want to study?

<table>
<thead>
<tr>
<th></th>
<th>Our study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not a computer worker</td>
<td>91%</td>
</tr>
<tr>
<td>CS exposure low/none</td>
<td>76%</td>
</tr>
<tr>
<td>Never heard of IoT</td>
<td>66%</td>
</tr>
</tbody>
</table>

Our study population is overall non-technical
What do we do with this dataset?

ANALYSIS
How do we analyze the dataset?

- Qualitative differences
How do we analyze the dataset?

- Qualitative differences
- Characteristic words
How do we analyze the dataset?

- Qualitative differences
- Characteristic words
- User operations profile

“Turn on the lights and tell me how much I weigh”

Immediate Action  Indirect Question
User operation schema

• Immediate interactions
  – Immediate actions ("Turn on the lights")
  – Direct questions ("What is my weight?")
  – Indirect questions ("Tell me my weight")

• Conditional interactions
  – Conditional actions ("When I come home turn on the light")
  – Notifications ("Let me know when my children get home")
How did the four prompts affect subjects?

THE FINDINGS
Unmediated Devices
“I would definitely want the smart watch to control the majority of the devices and controls in the house. I would definitely look for the smart door lock and smart thermostat.”

Unmediated Data
“I would want an 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.”

Agent-mediated Devices
“Set the temperature to 70 degrees. Lock the door. Close the blinds. Fetch and read my email. Please wake me up at 9 AM with some pleasant music.”

Agent-mediated Data
“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, monitor the house, and call my phone if anything strange happens (anyone enters the house, any objects are moved, etc.). Tell me my electricity consumption and gas consumption. How has my sleep been lately?”
What mental models did people form when presented with each conceptual model?
Unmediated Devices $\rightarrow$ “Islands”

“I would definitely look for the smart door lock and smart thermostat.”
Unmediated Devices ➔ “Islands”

- Wanted devices instead of application
- Manual remote-control
- One-on-one interactions
- “sensor,” “phone,” “device”
- Lack of higher-level applications

“I would definitely look for the smart door lock and smart thermostat.”
Unmediated Data ➔ “Watchdog”
Unmediated Data ➔ “Watchdog”

• Majority of sentences were “wants to know”
• Also wanted apps and automation
• “app” and “application”
• “alerts,” “know,” “see,” and “track”

“It would be nice if I could see my electricity usage in real time, and customizable alerts sent to my phone would be quite helpful.”
Agent-mediated Devices $\rightarrow$ “Delegate”

“Turn on the lights.”
Agent-mediated Devices ➔ “Delegate”

- 57% of sentences were immediate actions
- Nearly a quarter were conditional actions
- “Please” was characteristic of both agent-mediated response sets

“Turn on the lights.”
Agent-mediated Data ➔ “Assistant”
Did subpopulations show differences?

Populations from the literature:
• Older vs. younger
• Technical vs. non-technical

Our populations:
• 55 and older vs. 34 and younger
• High CS exposure vs. no CS exposure
Subpopulations responded differently

Findings:

• Older subjects responded more strongly to priming than younger subjects
• Subjects with no CS exposure responded no differently than subjects with high CS exposure on Unmediated Devices prompt
• Subjects with no CS exposure responded more strongly to priming on remaining three prompts

Implications:

• Smart home work motivated by elder care needs to evaluate with older subjects
• Evaluations for general population must involve non-technical users, especially since the primary users of home technology are often non-technical “passive users.”*

Older subjects responded more strongly to priming than younger subjects.
Older subjects responded more strongly to priming than younger subjects.
Older subjects responded more strongly to priming than younger subjects
Older subjects responded more strongly to priming than younger subjects
How about technical vs. non-technical populations?
There was no significant difference between experts and non-experts for the most common conceptual model.
However, there were differences for the remaining three conceptual models.
Subjects without CS exposure responded more strongly to priming than those with high CS exposure.
Subjects without CS exposure responded more strongly to priming than those with high CS exposure.
Bonus Finding: User-friendly IoT systems require native AI support
Bonus: Systems require native AI support

• Prediction ("Have coffee ready before I wake up")
Bonus: Systems require native AI support

- Prediction
- Goals ("Turn on the lights at 8 AM so that I wake up")
Bonus: Systems require native AI support

• Prediction
• Goals
• Explanation
  • Not a single agent-based response included “why”!
So what are the takeaways?

THE IMPLICATIONS
Takeaways

• Mental models researchers should do comparative studies
Takeaways

• Mental models researchers should do comparative studies
• System designers should be aware that the interface choices will influence system workload
Takeaways

• Mental models researchers should do comparative studies
• System designers should be aware that the interface choices will influence system workload
• Choose your test subjects well
Takeaways

• Mental models researchers should do comparative studies
• System designers should be aware that the interface choices will influence system workload
• Choose your test subjects well
• Both questions and commands are important
Takeaways

• Mental models researchers should do comparative studies
• System designers should be aware that the interface choices will influence system workload
• Choose your test subjects well
• Both questions and commands are important
• AI should be built into low levels of the IoT stack
Takeaways

• Mental models researchers should do comparative studies
• System designers should be aware that the interface choices will influence system workload
• Choose your test subjects well
• Both questions and commands are important
• AI should be built into low levels of the IoT stack
• And finally...
Unmediated Devices considered harmful?
IoT Interfaces for Everyday People

Meghan Clark
Intel/NSF CPS-Security Final PI Meeting
7/12/2018

mclarkk@berkeley.edu