Understanding the Long-Term Use of Smart Speaker Assistants

FRANK BENTLEY, Oath, USA
CHRIS LUVOGT, Oath, USA
MAX SILVERMAN, Oath, USA
RUSHANI WIRASINGHE, Oath, USA
BROOKE WHITE, Oath, USA
DANIELLE LOTTRIDGE, Oath, USA

Over the past two years the Ubicomp vision of ambient voice assistants, in the form of smart speakers such as the Amazon Echo and Google Home, has been integrated into tens of millions of homes. However, the use of these systems over time in the home has not been studied in depth. We set out to understand exactly what users are doing with these devices over time through analyzing voice history logs of 65,499 interactions with existing Google Home devices from 88 diverse homes over an average of 110 days. We found that specific types of commands were made more often at particular times of day and that commands in some domains increased in length over time as participants tried out new ways to interact with their devices, yet exploration of new topics was low. Four distinct user groups also emerged based on using the device more or less during the day vs. in the evening or using particular categories. We conclude by comparing smart speaker use to a similar study of smartphone use and offer implications for the design of new smart speaker assistants and skills, highlighting specific areas where both manufacturers and skill providers can focus in this domain.

CCS Concepts: • **Human-centered computing** → Empirical studies in ubiquitous and mobile computing; **Human-centered computing** → Field studies; **Information systems** → Speech / audio search

Additional Key Words and Phrases: Smart Speaker; Voice Assistants; Google Home; Voice I/O; Mid-Scale Data Collection

ACM Reference Format:

Frank Bentley, Chris LuVogt, Max Silverman, Rushani Wirasinghe, Brooke White, Danielle Lottridge. 2018. Understanding the Long-Term Use of Smart Speaker Assistants. *Proc ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3 Article 91 (September 2018), 24 pages. https://doi.org/10.1145/3264901

1 INTRODUCTION

Smart speakers, such as the Amazon Echo or Google Home, have become popular additions to homes throughout the world in the past two years. These devices provide a new way for people to interact with computing systems in their home, by voice, without the need to touch a device. These devices are becoming pervasive in the United Sates and throughout the world. Amazon has sold over 30 million Echo devices, and

Author's addresses: Oath, Inc. 701 First Ave. Sunnyvale, CA, USA fbentley@yahoo-inc.com

https://doi.org/10.1145/3264901

Forrester Research expects that half of American homes will have a smart speaker device by 2022.[1] Google claims to be selling a Google Home device every second (a pace of 31.5M/year).[2]

These devices enable users to issue a broad set of commands on a wide range of topics. In addition to playing music, users can ask in natural language about the weather, stock market, travel plans, store hours, online shopping orders, and other general information topics. These devices also provide smart home integration to control lights, heating systems, and set timers and alarms. Both Amazon and Google add new functionality periodically as well as support add-on "skills" from third party developers. These skills often involve music playback, tasks such as ordering rideshare cars, or specific entertainment experiences. As of January 2018, the top skills in the Amazon ecosystem were Jeopardy, Box of Cats, Jurassic Bark, Twenty Questions, and Thunderstorm Sounds.[3] Other skills from more well-known companies include ordering pizzas from Dominos, getting a ride from Lyft or Uber, and weather forecasts from Accuweather and Weather Bug.

Smart speaker devices represent a vision of pervasive computing initially laid out in research agendas around Calm Computing [24] and demonstrated in non-functional concept videos such as the Knowledge Navigator [1] from Apple in 1987. These visions imagined computing retreating into the background and involved people using more natural interactions such as voice to engage with smart computing services instead of using a keyboard or touchscreen to interact with a graphical interface. Similar interfaces were actually built in the late 1990s and early 2000s in prototype systems in several research labs. [10, 17] While voice assistants have been a dream of the HCI, Ubicomp, and AI communities for decades, this new era of smart voice speakers represents the first-time ambient voice interfaces are readily available in home environments, for millions of people.

While these devices are becoming increasingly popular, especially as their prices are driven down (the Google Home Mini was available for $19 over the 2017 holidays), little has been studied on how they are used longitudinally. As this is the first time these types of voice assistants have been integrated into real home environments for months at a time at a large scale, understanding how and when they are used is critical to understand both their usefulness and opportunities to create new experiences for future voice assistants. This moment in time for smart speakers is quite similar to the early years of smartphone availability, where the Ubicomp and Mobile HCI research communities focused on understanding how real-world users would engage with technology that they had been studying in the lab and in small deployments for years. Specifically, we had the following research questions:

1. What types of commands do people make to their smart speakers (e.g. weather, music playback, smart home control, etc.) and in what percentages do they use these different features?
2. How are these devices used at different times of day or days of the week? Are there differences in the categories of commands?
3. How do these types of commands change over time as users become more familiar with these devices? Do the topics change? Does the length of commands change per topic over time?
4. Are there any differences in use of these devices in different age groups or household sizes?

To answer these questions, we gathered full usage logs of Google Home use from 88 diverse households located throughout the United States. Importantly, these were all existing Google Home owners who shared their past usage history with us, so that usage was not affected by being part of a study. These logs provided timestamps and command strings for a household's entire history of interaction with their device since the date of purchase. We analyzed these logs to answer the questions above and will conclude with several implications for the design of new conversational agents and smart speaker devices.

[1] https://www.forrester.com/report/Forrester+Data+Smart+Home+Devices+Forecast+2017+To+2022+US/-/E-RES140374
[2] https://www.blog.google/products/assistant/how-google-home-and-google-assistant-helped-you-get-more-done-in-2017/
[3] https://www.amazon.com/b?node=13727921011

2 RELATED WORK

The history of voice interaction in the home did not start with this current generation of smart speakers. Science fiction shows, such as Star Trek or 2001: A Space Odyssey, popularized the idea of giving voice commands to a "computer" long before creating such systems was technically feasible. As technology progressed, a variety of systems have been built over the past twenty years that have enabled a variety of types of voice interaction and control in home or office environments.

The Intelligent Room project at MIT [7, 10], built in the late 1990s, was an intelligent environment controlled via voice and gesture. Voice commands could be issued to control lights, play music, move blinds, or ask some general-purpose questions such as weather. Similar to today's smart speakers, it used a wake word (in this case "computer") at the beginning of each command. This system was fully deployed in the lab, however required a room full of computers and hundreds of meters of cabling to operate and was never studied in real home environments over time, serving as more of a technical proving ground than a system integrated into daily life.

The ComHOME project from the Interactive Institute [16] created an apartment with a variety of voice controls. In this system, voice was mainly used to initiate video calls between apartments. In a similar manner, the Aware Home project at Georgia Tech created a voice paging system in their smart home environment. Inhabitants of the home were able to initiate paging within the home by calling out the name of the person that they wanted to talk to [17]. However, for both of these systems, there was not long-term usage data from a wide variety of participants given the difficulty of custom-built deployments, so it is unclear how these types of systems would have been used over months of interaction in a diverse set of homes.

Early criticisms of smart homes and voice assistants (e.g. [4]) stated that they were motivated by what was technically possible rather than what was desirable. And indeed the interactions with these environments were quite clunky. Voice recognition technology at the time was relatively poor compared to current levels of accuracy. As an example, Google has lowered word error rates by 30% in the past few years using tensor-based systems [11]. And use cases around turning on lights or lowering shades were not seen as useful enough reasons to add a room or closet full of additional hardware to support always-on listening in a home environment. We were curious if use of these new smart speakers, given expanded feature sets, minimal cost, and easy setup, might show higher retention than these earlier deployments.

Before voice interaction came to the homes of consumers, some smart home technology became more readily available. Many of the features of the Intelligent Room and other research projects from the 1990s were widely available in DIY kits or from professional installers in the early 2010's. At this time, control was mainly via wireless remotes or smartphone GUI applications. Brush et al. [5] studied smart home interactions in 2011 and included photos of some of the physical controls for interacting with smart homes that were used in the wild at the time. Nowhere in their paper is voice interaction mentioned, as this technology was still not mature enough just seven years ago. Mennicken and Huang [20] studied the use of smart homes as well, and the differences in use between the person who set up the technology and the other members of the household, noting potential conflicts, often around using these clunky remote controls and understanding complex scripts of controls that were mapped to single buttons.

Voice interactions came into everyday use with mobile applications such as Siri, Google Assistant, and Cortana. Luger and Sellen [19] explored the use of phone-based smart assistants and found them to be "like a really bad PA [personal assistant]." They found that users expect a greater amount of functionality than the agents could provide and were frustrated when assistants could not answer questions that they thought to be simple. The use of these assistants continued to spread. However, while 98% of iPhone users had tried Siri, 70% said that they did not use it regularly.[4]

When these assistants, such as Alexa, moved into the home, a very different interaction could be created. The placement of devices in a kitchen or living room, combined with the hands-free nature of interaction, harkens back to earlier work on intelligent rooms and smart homes. Beyond understanding how these devices are used in

[4] http://www.businessinsider.com/98-of-iphone-users-have-tried-siri-but-most-dont-use-it-regularly-2016-6

a home setting, we were curious to see if their use resembled the patterns seen with Siri where frequent use was quite rare for most iPhone owners.

Given the recent release of Google Home and the Amazon Echo, there is little published work exploring how people are using these devices. Purington et al. [21] analyzed online reviews of the Amazon Echo, finding that people made an emotional connection to Alexa and that personification of the voice led to increased satisfaction with the device as shown by ratings given along with the review. Robinson et al. [22] recently explored the use of smart speakers in a single-time probe in an economically disadvantaged area of India, finding that participants asked combinations of fact-based, contextual, and opinion-related questions. However they only collected 83 questions throughout the trial of the probe.

Yet, no one has studied the use of these devices over time. Particularly, our research questions that were laid out in the introduction have not been answered by any of this related work. Understanding use over time, and the categories of use that users engage with on their smart speakers has large implications for the field of Ubiquitous Computing. These devices are similar to the introduction of the smartphone in creating mass market adoption of a concept that originated within our community. Similar to the study of smartphone use, we wanted to understand how these devices were used throughout the context of a day and over months of deployment in real home settings. Thus, we set out to conduct our own study to gather usage logs of interactions with a smart speaker device.

For this study, we were inspired by the work of Böhmer et al. [6] in their large-scale study of mobile application use that was conducted not long after the use of mobile phone applications became a mainstream phenomenon. Many of their research questions were similar to ours in understanding use over time and at various times of day, but in the domain of smartphone apps. Thus we set out to collect as many logs of Google Home commands as we could from a broad audience of participants so that we could conduct a similar type of analysis of use in the domain of smart speaker assistants.

3 METHOD

We used Amazon Mechanical Turk to collect full device usage logs from 88 diverse Google Home owners, in a manner similar to the phonebook data collection study by Bentley et al. [3]. Participants were provided with detailed instructions on how to access their Google account history on the web, and how to filter this history to include data from the Google Home device. Participants were given the opportunity to remove any entries that they did not feel comfortable sharing with the research team.

In addition to uploading their logs, participants were asked a few questions about their device use, including how long they had owned the device and where the device was located in the home. The survey concluded by capturing basic demographic information including the composition of their household. Participants were paid \$5 for uploading their logs, and the entire survey took an average of six minutes to complete. Data was collected in the summer of 2017, less than a year after the launch of the Google Home device, so these participants are likely early fairly adopters. All data was collected in accordance with our institution's policies for research with human subjects and data retention.

Of the participants, 47% were female and ages ranged from 18 to 64. They lived in 27 different states in all regions of America. All owned a Google Home, and 17% also owned an Amazon Echo device. Previous research [2] has shown that samples from MTurk using our same screening criteria can be quite reliable in understanding technology use when compared to large-scale professional market research surveys or usage logs from large corporations. Dozens of studies over the past three years that have compared these fast survey platforms to various ground truths (e.g. usage logs at scale) have typically observed findings within 5-7% of the ground truth. Given the time and expense of collecting thousands of logs, we believe this method gives us a strong dataset to understand the use of these devices in the wild.

Overall, 31% of the households were comprised of just a single individual, comparing well with national statistics (27% of Americans live alone according to data from the Census Bureau[5]). An additional 32% of

[5] https://www.census.gov/programs-surveys/cps.html

participants' households consisted of two people. Fifteen percent had three people, and another 15% had four. Only 6% had more than four inhabitants. We will return to household size in the Findings section to explore how households of different sizes use these devices. Of the participants, 40% were single, 21% in a relationship, and 32% were married.

The full command history was extracted from the survey responses. In total, we collected 65,499 commands (an average of 744 commands and 110 days of data per participant). We created a fixed set of nine domains (e.g. information, music, home automation, etc.), with 66 subdomains (e.g. finance, light control, television) based on the capabilities of the Google Home device. To create the domains, we took a random sample of 1,000 commands and performed a team-based grounded theory analysis to identify themes in the commands. We created a two-level hierarchy, which represent the domains and subdomains discussed in the remainder of this paper. We iterated on the domains until there was agreement from all researchers involved.

Next, we used a group of editors to manually label 20,000 queries in these domains and subdomains. A subset were labeled by multiple editors with high inter-rater reliability. We then created a feature set of n-grams from these 20,000 human-labeled commands to train a support vector machine using cross validation. This achieved an 88% accuracy at the sub-domain level.

In addition, for each user, relative date stamps were calculated to examine behaviors on the n^{th} day of owning the device regardless of purchase date so that we could compare use over time. We also used a MANOVA to investigate relationships between household sizes and usage patterns and well as found correlations and behavioral clusters between several domains of use and frequency of use of the device.

4 FINDINGS

We will now explore the data and describe how it answers our research questions in understanding how smart speaker technology is being used over time in real world homes who have purchased these devices. Starting by examining use by hour of day and day of week, we will explore the features that are gaining widespread traction in use in real home environments. We will then explore the types of commands that households are making over time, how use changes as people use these devices for months, and differences in use based on age/household composition. Finally, we will explore between-user trends and behavioral clusters of use. Section 5 will then contextualize these results in comparison to studies of smart phone use and historical functions of voice assistants as well as offer implications for the design of new smart speaker services.

4.1 Daily Use

In order to understand how these devices are being integrated into people's homes and lives, it is useful to begin by looking at overall daily usage. Figure 1 shows that the median household issued 4.1 commands to their Google Home device per day, over an average of 110 days of data that we collected per household. This usage was much higher than we expected, given the lower regular use of smart assistants on mobile phones such as Siri[6]. This high usage shows how these devices are being integrated into a wide variety of tasks throughout the day and evening. Looking at the broader distribution, the 25^{th} percentile household issued 2.5 commands per day and the 75^{th} percentile user issued 17.7. Interestingly, this shows that the vast majority of our Google Home users were quite frequent users, which is quite different from the rare use of smartphone-based assistants as discussed above.

Commands were often given to the device in sequence. We applied a standard 10-minute idle time session boundary [13-15, 23] to segment the data to examine how many commands were used in each interaction with the device, as described in [18]. While this might segment longer timers or extended audio sessions such as podcasts into multiple sessions, we find the use of these devices to be more similar to mobile phones than desktop computers, where shorter session boundaries are more commonly used. Similar to mobile phones, users are not sitting in front of the device for hours on end to perform a set of tasks, but rather pick up and put down

[6] http://www.businessinsider.com/98-of-iphone-users-have-tried-siri-but-most-dont-use-it-regularly-2016-6

(or in this case, speak to) the device as needed in shorter bursts throughout the day. Figure 2 shows a plot of commands per session. Across all of the households, 39% of all sessions consisted of only a single command and 60% of all sessions had only one or two commands. Ten percent of sessions were comprised of more than ten commands, frequently around music playback (skipping songs, adjusting volume, etc.).

Turning to the content of the sessions, almost half (48%) of all sessions only involved a single domain (as shown in Figure 3). The vast majority of sessions (77%) involved only one or two domains. Of the two-domain sessions, 35% contained one command about home automation and 34% contained one command about music (typically starting or stopping).

Next, we were interested in how use of the device within the household varied by hour of day. Figure 4 shows the volume of commands that our participants made at different hours of the day. All hours in the graph are in the local time for the device. The basic shape of the graph matches the times when most people are awake. Use sharply increases from 6-7am, has another increase before lunchtime between 12-1pm, and then has a daily high spike around 5-6pm when many people return from work. In addition to this high-level pattern, we can see some clear differences in domains used throughout the day.

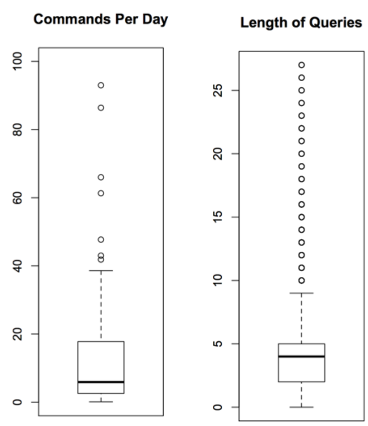

Fig. 1: Commands issued per day to Google Home (left) and average length of commands in words (right).

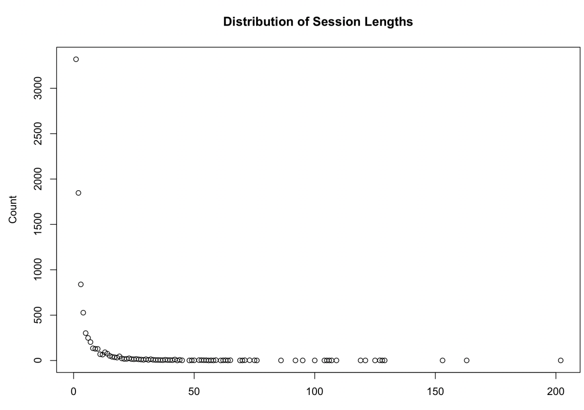

Fig. 2: A distribution of commands per session across all interactions with the device. 39% of all sessions contained only a single command.

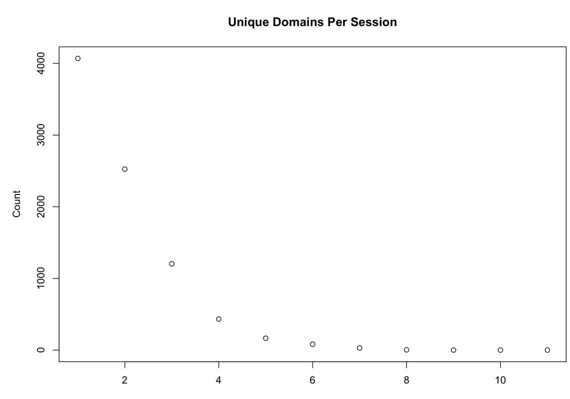

Fig. 3: Number of unique domains used within a session. 48% of all sessions only involved a single domain.

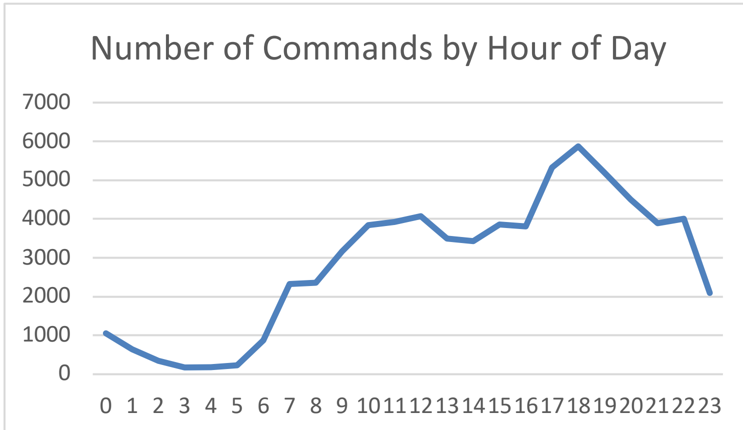

Fig. 4: Overall Google Home use by hour of day.

Table 1: Google Home use for each high-level domain as a percentage of use for each hour of the day. The conditional formatting colors indicates overall high vs. low usage for a given domain in a given hour so that temporal variations are easier to see at a glance.

	0 1.5%	1 0.9%	2 0.5%	3 0.3%	4 0.3%	5 1.3%	6 3.4%	7 3.4%	8 4.6%	9 5.6%	10 5.7%	11 5.9%	12 5.1%	13 5.0%	14 5.6%	15 5.6%	16 7.8%	17 8.6%	18 7.6%	19 6.6%	20 6.6%	21 5.7%	22 5.8%	23 3.0%	% of Total
Music	43%	38%	37%	44%	43%	44%	31%	35%	42%	42%	40%	46%	43%	47%	45%	41%	41%	38%	37%	36%	41%	32%	34%	32%	40%
Information	10%	14%	19%	11%	10%	10%	15%	22%	22%	21%	21%	14%	18%	17%	15%	18%	18%	16%	18%	20%	13%	17%	18%	12%	17%
Automation	18%	20%	9%	10%	10%	13%	9%	4%	5%	3%	5%	5%	5%	4%	5%	6%	6%	9%	9%	11%	11%	15%	15%	23%	9%
Smalltalk	9%	6%	9%	5%	7%	4%	10%	9%	5%	6%	6%	7%	8%	7%	8%	8%	8%	9%	7%	9%	8%	13%	12%	10%	8%
Alarm	4%	9%	5%	11%	9%	9%	5%	3%	4%	9%	4%	9%	6%	5%	9%	8%	6%	10%	8%	7%	6%	4%	3%	5%	6%
Weather	5%	4%	13%	4%	5%	6%	14%	9%	12%	9%	8%	7%	6%	6%	5%	6%	6%	6%	6%	4%	4%	5%	4%	5%	6%
Video	4%	5%	1%	5%	0%	1%	0%	1%	2%	2%	5%	4%	4%	5%	4%	4%	3%	4%	4%	4%	6%	4%	3%	5%	4%
Time	1%	0%	2%	7%	7%	11%	11%	14%	7%	4%	3%	4%	4%	3%	3%	3%	4%	4%	3%	2%	2%	2%	2%	2%	4%
Lists	1%	1%	0%	2%	0%	1%	1%	1%	1%	1%	4%	2%	2%	3%	3%	3%	3%	2%	3%	2%	3%	1%	1%	2%	2%
Other	6%	3%	4%	2%	8%	2%	4%	3%	1%	2%	4%	2%	5%	4%	3%	4%	4%	3%	4%	4%	6%	5%	8%	5%	4%

Table 1 further breaks down the hourly usage pattern by the domain of the command. Perhaps the most interesting observation is how the relative percentages of domains stay fairly consistent throughout the day (as seen by the relative uniformity of colors in each row). As we will discuss in the following section, this is quite different from patterns observed in smartphone use, where different app categories have larger hourly differences. However, there are some interesting changes in the percentage of use of some domains throughout the day. Music playback dominates throughout the day, remaining the most used category. From 9pm to 2am, home automation rises in use. Turning off lights, adjusting the thermostat, and other nightly routines dominate this usage. Weather jumps from 6% to 14% at the 6am hour and remains high through the 9am hour, before falling back below 10% for the rest of the day. Requests for the time are much higher from 3am-9am, likely as users lay in bed wondering how many hours they have left before they need to get up.

This use matches the rhythm of a person's day and shows how different domains are more or less relevant at different hours. When building assistants of any kind, it is important to consider the user's context and what is important to them at particular times of the day. This becomes increasingly important as smart speakers add screens (such as the Echo Show) and can proactively show relevant information, a point we will return to in the Discussion.

Turning to Figure 5, we can see the use of these devices by day of the week. Weekend use is significantly higher than use during the week (t=2.02, p=0.04), with Mondays being the day of least use. This matches with when people are likely to be at home, and thus able to use the device. Interestingly, the categories of use remain quite stable throughout the week and weekend as a percent of the total use, showing that there are not specific use cases that are being more heavily utilized on specific days.

We will return to these findings on temporal use in the Discussion, where we highlight differences between the use of voice assistants and the use of mobile phone applications in a direct comparison to data from Böhmer et al.'s [5] exploration of mobile phone app use. Overall, the patterns described above are quite different from how people are using their mobile phones, and we find many of the differences in sessions lengths, domains per session, and hourly use to be quite interesting for the design of future voice assistants.

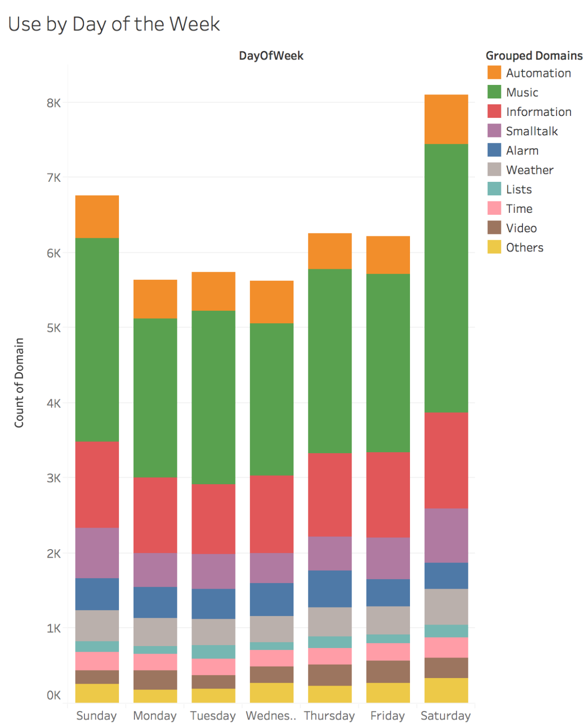

Fig. 5: Google Home use by day of the week, broken out by category. Use on the weekends is significantly higher than use during the week.

4.2 Changing Use over Time

Now, we will turn to look at the use of these devices over time, as a user integrates them into their lives over weeks and months. As a reminder, we collected an average of 110 days of use from our participants, enabling us to discover how usage changes as the devices become a fixture in the home and less of a novelty. Figure 6 shows the types of commands that users make over time, and the percentage of total commands from each of these categories.

Use by category is fairly stable with changes over time being relatively small. We find this interesting as users seem to quickly settle on what they will use the device for on the first few days and rarely change this use. In the first 7 days, Music represents 35% of use, staying fairly steady but declining to 26% of use in the 145-150th days. We can see that Automation use cases follow the opposite trend, starting at only 3% of commands on day one, but then growing to 16% of all commands by the 145-150th days as users connected additional smart devices. Other domains such as weather and setting alarms stay relatively constant over time, so that overall the increasing trend of automation and decrease in music account for most of the changes over these five months.

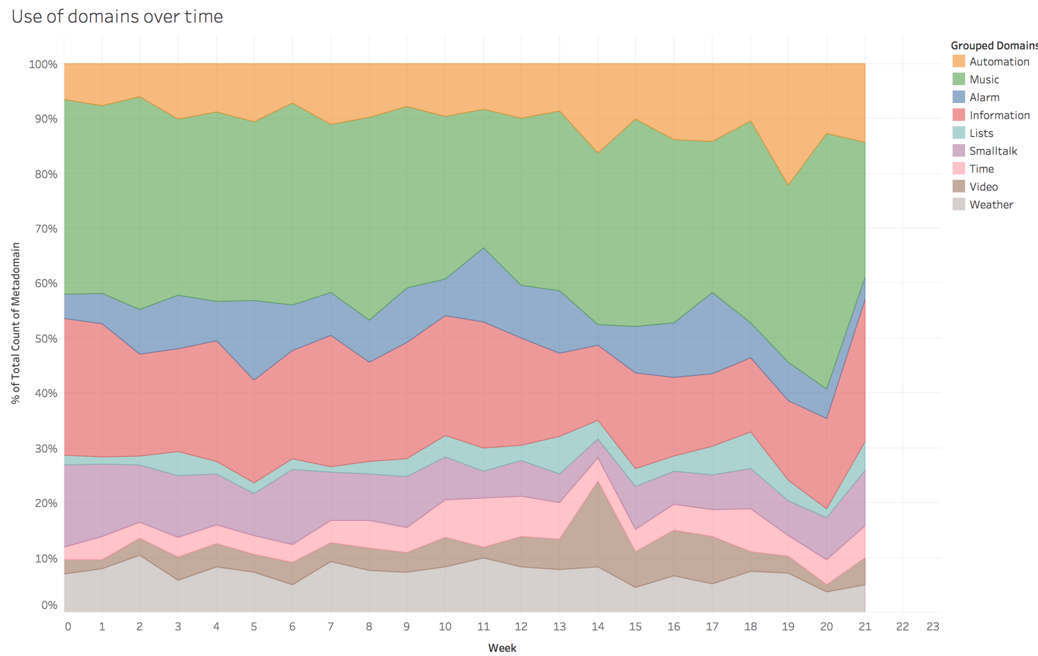

Fig. 6: The types of commands that users make over months of device ownership.

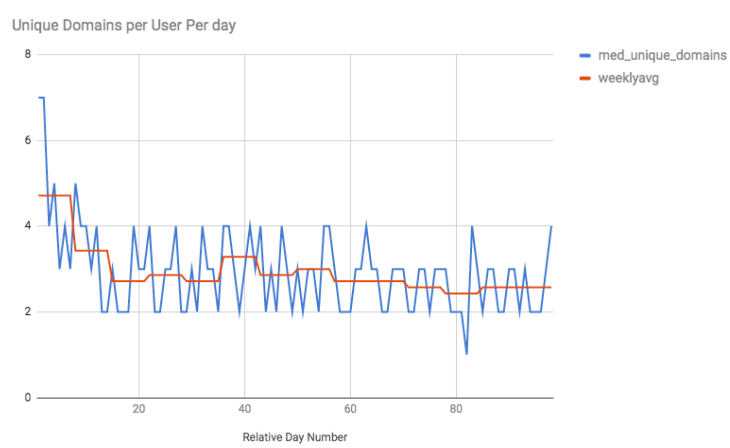

Fig. 7: Average of unique domains that users queried per day over time.

While the overall domains are fairly static over time, we were interested in the behavior of specific users. Do they try previously unexplored domains, or are they fairly set with the domains that they begin with? Figure 7 shows that by the third week users are fairly set in the domains that they use, only using the device for approximately three different domains, and that increased exploration does not occur. What is most interesting to us is that these domains are quite different across users with each user settling in on the domains that are most useful for them. In total, 22 different subdomains were still being used by participants on week 14, yet in that week only 28% of users had tried a new subdomain and only 4% tried a new top-level domain.

4.3 Specific Commands

We wanted to explore the content of the commands themselves in order to better understand what users were actually asking about in particular domains as well as how the length of commands changed (or did not change) over time. Participants uttered a total of 19,376 unique commands out of the total 65,499 commands in the data set. Table 2 shows the top 20 commands that were made across all usage. In total, these top 20 commands make up 22.3% of all commands that were uttered.

Table 2: Top 20 commands made to the Google Home across all users and the count of how many times these exact commands were made.

Command	**Percent**
stop	7.3%
what time is it	2.7%
pause	1.1%
how much time is left	1.1%
pause TV	0.8%
play	0.8%
skip this song	0.8%
volume up	0.8%
tell me a joke	0.8%
what's the temperature	0.8%
resume TV	0.7%
volume down	0.7%
next song	0.6%
next	0.5%
turn on kitchen	0.5%
what's the temperature outside	0.5%
stop TV	0.5%
turn on table	0.5%
turn off kitchen	0.5%
turn off living room	0.4%

We were also interested in understanding the different words that were used for commands in each domain. We performed TF-IDF on the commands for each domain, with the most distinct words highlighted in Table 3. Music requests are most uniquely about “play”ing “songs” or “music,” “skip”ping, and looking for music “by” an artist. Automation is about “turn”ing on and off devices in the “kitchen” on the “table” or in a “room.” Lists are most frequently about “shopping” and “add”ing to “lists” for items like “cheese” and “milk.” What is most interesting about this analysis is the wide diversity of different tasks that these devices are supporting in daily life – everything from asking for stock prices, performing math such as division, playing video on the TV, querying the weather forecast, controlling lights, keeping a grocery list, and playing music. Unlike phone-based assistants, our participants have found many diverse uses for their smart speakers and use them regularly.

We were interested in exploring if users made longer commands over time, demonstrating an understanding of more complex requests that they could make to the device, or if commands got simpler over time with users optimizing what they could say. The right side of Figure 1 shows a boxplot of command lengths. The median command to the Google Home device was four words long, with 25% of all commands being just two words or less and 25% of commands being longer than five words in length.

Table 3: TF-IDF analysis of words used in each domain. The most unique ten words in each domain are shown.

Music	**Automation**	**Smalltalk**	**Alarm**	**Video**	**Time**	**Weather**	**Information**	**Lists**
play (0.041)	turn (0.062)	you (0.011)	timer (0.073)	tv (0.148)	time (0.111)	weather (0.039)	divided (0.008)	shopping (0.308)
song (0.017)	kitchen (0.047)	why (0.008)	set (0.026)	pause (0.068)	what (0.065)	what's (0.037)	stock (0.008)	list (0.087)
music (0.016)	table (0.034)	do (0.007)	alarm (0.026)	resume (0.056)	is (0.030)	temperature (0.032)	price (0.008)	add (0.078)
skip (0.010)	room (0.031)	your (0.007)	time (0.021)	rewind (0.040)	it (0.030)	the (0.020)	what's (0.007)	cheese (0.007)
by (0.092)	off (0.028)	birthday (0.007)	for (0.013)	episode (0.016)	in (0.001)	forecast (0.015)	weather (0.006)	grocery (0.007)
kitchen (0.008)	living (0.025)	good (0.007)	an (0.012)	stop (0.015)	california (0.001)	today (0.010)	what (0.005)	paper (0.006)
pause (0.008)	lights (0.022)	thank (0.006)	14 (0.011)	play (0.014)	what's (0.001)	outside (0.009)	score (0.004)	milk (0.005)
the (0.007)	lamp (0.020)	have (0.005)	cancel (0.009)	video (0.009)	zone (0.001)	like (0.008)	spell (0.004)	creamer (0.005)
table (0.006)	small (0.018)	happy (0.005)	longer (0.007)	next (0.006)	the (0.001)	tomorrow (0.008)	who (0.004)	wipes (0.005)
next (0.005)	basement (0.016)	what (0.005)	how (0.007)	family (0.005)	times (0.001)	what (0.006)	who (0.004)	towels (0.005)

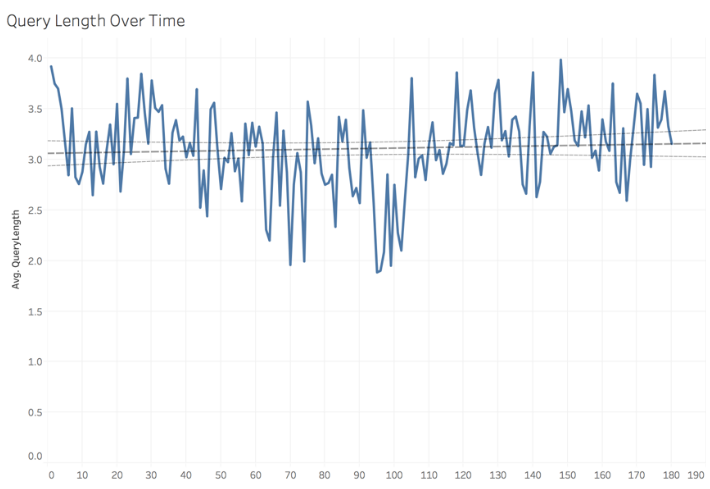

Fig. 8: Length of commands made to the Google Home device over time.
The x-axis represents days since the device was first used.

Figure 8 shows command lengths over time. We find it interesting that the average command length does not change over time. Users are not significantly changing how they interact with the device after the first few days, which can also be seen above as the categories of use do not change much over time as well. However, when focusing in on particular command types, we found that commands for Time/Task management (r^2=0.1, p=0.0001), Info requests (r^2=0.05 p=0.0037), and Entertainment (r^2=0.085, p=0.0001) get longer over time, while commands for Small Talk decrease in length (r^2=0.05, p=0.003). When looking at the actual utterances, we can see that for weather (part of Info requests), users start to ask for specific features of the weather, such as the high temperature for the day, or for weather in other cities as they interact with the device over time. Entertainment/Music requests start to include asking for specific songs, increasing command length.

When reading individual commands, it is interesting that users frequently change the way that they ask for specific information. For example, users who have been using the device for months still ask in different ways for the weather and seem not to settle on a particular “command.” For example, one user asked “What’s the weather forecast?” “What’s the high temperature for today?” and “What’s the high going to be for today?” all within a few days of each other. It seems extremely important that strong language processing is able to extract meaning from widely diverging commands, as users have a hard time remembering specific syntax that has worked in the past and speak to assistants with frequent variations in commands, even within the same sessions. As mentioned above, we saw 19,376 unique commands made to control a much more modest number of device capabilities.

4.4 Usage Patterns

In addition to the usage across all users discussed above, we wanted to further investigate differences between users. We began by running bivariate correlations using the proportional usage variables (commands per day, sessions per day, percent use by time of day, percent use by each domain, percent weekday/weekend use) to investigate usage trends across our 88 participants. We observed nine significant large correlations (with absolute values >.350), all significant with p<.001, 2-tailed, with n=88.

We observed three correlations around amount of use. We found strong positive relationships between:

- Commands per day and Number of domains used (r=.590)
- Commands per day and Sessions per day (r=.545)
- Commands per day and Commands per session (r=.378)

The correlations with commands per day reflect patterns of greater overall usage of the tools: the more commands per day, the greater the number of sessions per day, commands per session and domains involved. Users who engaged more heavily with the devices each day also explored more domains and had more complex sessions of use.

We observed two correlations around midday use:

- Percent Midday and Percent Small Talk (r = -.390)
- Percent Midday and Percent Night (r = -.367)

These correlations suggest that users who use the device more in the middle of the day use it less at night and vice versa. Further, those that use it midday do not engage in as much small talk. We found a smaller correlation between percent small talk and percent night (r = .254, p =. 017), suggesting that users who tend to use the device at night engage in small talk compared to those who use it midday.

In addition, we observed another strong timing-related relationship:

- Percent Evening and Percent Weekend (r = .398)

Thus, usage during evenings and weekends were correlated. Those who heavily used their devices in the evenings also did so on the weekends, compared to mid-day users who had a more consistent usage level across different days of the week. These likely represent those who work outside the home, and households where no one is home during the day. This relates to the correlation in the previous section showing that people who use the device more mid-day use it less at night.

Lastly, we found positive and inverse relationships related to usage of devices for information.

- Percent Information and Average Command Length (r = .371)
- Percent Information and Percent Music (r = -.478)
- Percent Information and Percent Small Talk (r = .387)

Usage of the assistant for information retrieval was associated to longer commands and more small talk. Information and music were inversely related with a quite strong correlation, suggesting that users who tend to use the tools for information don't tend to use it as much for music. This is interesting, as it points to users having their own particular dominant usage of the device – whether that's music or information.

4.5 Demographic Differences

We will now explore some demographic differences in use. As our dataset was quite diverse across ages, income levels, household composition, and gender of the device owner, we were able to break down our analysis and focus on age or family composition-based differences in the use of smart speaker systems.

Table 4: Age-related differences in Google Home use for the top 20 sub-categories. Values are the average use of that domain for a participant per day in each age range.

	18-24	25-34	35-44	45-54	55-64
music	0.74	1.67	1.24	0.06	0.21
device	0.38	0.62	0.78	0.02	0.16
volume	0.12	0.51	0.47	0.03	0.11
lightcontrol	0.22	0.54	0.02	0.00	0.02
smalltalk	0.59	0.50	0.48	0.05	0.11
alarm	0.34	0.4	0.42	0.01	0.22
weather	0.20	0.38	0.40	0.01	0.15
video	0.04	0.24	0.02	0.01	0.00
time	0.22	0.17	0.13	0.00	0.01
math	0.37	0.15	0.09	0.01	0.02
define	0.08	0.09	0.23	0.00	0.04
games	0.14	0.10	0.20	0.02	0.01
shoppinglist	0.00	0.14	0.08	0.00	0.00
sports	0.06	0.09	0.14	0.01	0.01
local	0.08	0.1	0.09	0.01	0.04
joke	0.02	0.07	0.12	0.00	0.00
animals	0.04	0.12	0.09	0.00	0.01
finance	0.00	0.09	0.01	0.00	0.00
otherautomation	0.19	0.05	0.01	0.00	0.01
notableperson	0.01	0.06	0.12	0.00	0.00

Table 4 shows a breakdown of usage by age of the account holder that set up the device. Here, we are showing the lower-level categories that were labeled before grouping to the metadomains to better understand the specific topics that are queried. There are some clear patterns. Adults 25-44 are more likely to change the volume on the device ($t=3.64$, $p=0.0006$), however those 18-24 are much more likely to ask the device for the time, especially compared to 45-54 year olds ($t=2.41$, $p=0.02$). In addition, 18-24 year olds are less likely to listen to music on the device compared with 25-44 year olds ($t=2.34$, $p=0.04$), perhaps because it requires a paid subscription or they prefer to use other devices for music playback.

Figure 9 and Table 5 show use by household size. Larger households did not use the device significantly more often than smaller households ($r^2=0.04$, $p=0.74$), which was surprising to us as we assumed that additional family members would all try to use it as well. The largest homes, those with 4+ inhabitants, were the least likely to use a breadth of features from the devices. Specifically, Home Automation features were used much less frequently with this group, yet this group used basic functionality such as music playback and information requests at a similar rate to other groups.

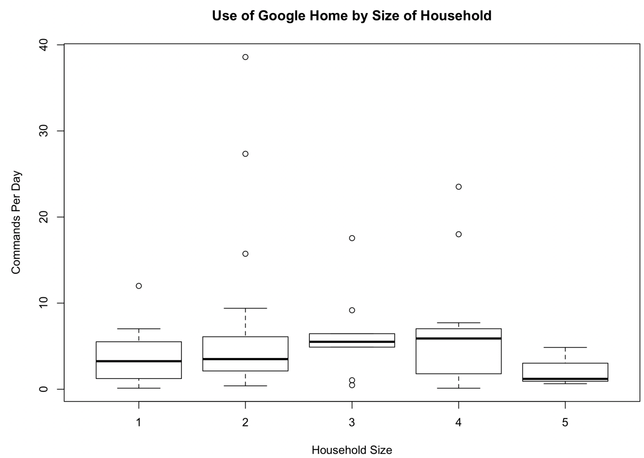

Fig. 9: Use of Google Home by size of household. Note that usage does not increase significantly for larger households.

Table 5: Differences in use by domain for different household sizes

	1	2	3	4+
Music	42%	35%	40%	47%
Information	15%	16%	21%	18%
Automation	12%	11%	9%	3%
Alarm	7%	7%	2%	8%
Weather	9%	5%	9%	4%
Smalltalk	7%	9%	7%	9%
Time	5%	2%	7%	3%
Video	1%	7%	1%	1%
Lists	0%	3%	2%	2%
Other	3%	5%	2%	5%

We conducted a MANOVA with age and household size as between-subjects factors with proportional use as dependent variables. Age was not significant and household size was significant with Hotelling's trace (p=.05). We observed a relationship between household size and the proportion of commands that were given on weekdays versus weekends. Larger households (of 3 or more) use the device proportionally more on weekends compared to smaller households ($M_{householdsize1}$= .238, $M_{householdsize2}$=.265, $M_{househouldsize3+}$ = .380; $F[2,58] = 4.651$, p=.014). A post-hoc analysis using the LSD method found that the single and pair households differed significantly from households of 3 or more (p=.03 and p=.02 respectively), but did not differ significantly from each other.

These demographic differences highlight that these devices are being used different by households of different compositions. While younger adults are more likely to ask for time, and those 25-44 are most likely to use the device for music, larger households engaged more on weekends.

4.6 Behavioral Clusters

Finally, we conducted a K-means cluster analysis with all participants in order to identify behavioral clusters between participants. We chose a 4-cluster solution in order to highlight maximal differences between users that remained significant. Clusters converged in 3 iterations. We included 20 outcome variables in the cluster analysis (As shown in Appendix 1). The cluster centers were significantly differentiated for 7 variables with a strong statistical trend for an 8th variable. These are included in the table below.

Table 6: Four behavioral clusters emerged across our n=88 participants.
The features used in our final clustering are included in this table.

	Super Users	**Explorers**	**Medium Users**	**Light Users**	**p**
Commands Per Day	30.44	7.48	5.55	2.61	<.001
Sessions Per Day	5.93	2.38	2.36	2.30	.001
Commands Per Session	5.40	3.98	2.74	2.42	.005
Percent of Use at Night	0.20	0.18	0.13	0.29	.007
Percent of Use on Weekends	0.30	0.35	0.33	0.23	.04
Percent of Use on Music	0.18	0.17	0.24	0.29	.068
Percent of Use on Video	0.05	0.01	0.01	0.01	.003
Number of Subdomains Used	51	44	29	12	<.001
	n=8	n=21	n=34	n=25	

Super users have the largest total number of commands and sessions per day as well as use the highest number of domains. Beyond overall higher usage in terms of commands per day and sessions as well as number of subdomains explored, their most distinguishing thematic feature is that they proportionally issue more commands for video.

We have two medium usage clusters with a similar number of sessions per day. They differ in their style of use: Explorers tend to have more commands per session compared to the other medium cluster. In addition, they explore significantly more subdomains of use, at 44 compared to 29 for the non-explorer medium group. The Medium group was our largest cluster, with 34 of our 88 participants falling into this cluster, suggesting that this may be more illustrative of common behavior.

The Light user group use the fewest commands and domains but have other distinguishing characteristics. Proportionally, they utter more commands at night and fewer over the weekend, compared to all other groups. We also observed a trend where they use more music related commands. Note that. there were no significant demographic differences between clusters (p>.1).

Groups differed in expected ways of heavier to lighter use, but also in unexpected ways in terms of domains of use. The heaviest users used proportionally more video and less music. All groups had higher weekend use while the lightest group concentrated on night time use. Our medium groups differed in their style of use, where both groups averaged about the same number of sessions per day but the Explorer group uttered more commands per session. This was interesting to us as sessions and commands did not increase linearly. The Medium group was our largest cluster, suggesting that uttering fewer commands per session is a more common behavior.

5 DISCUSSION AND LIMITATIONS

This study has explored how people are interacting with smart speaker assistants in their daily lives over months of ownership. We analyzed the voice history logs of 65,499 interactions with existing Google Home devices from 88 homes for a period of over 3 months.

We began by quantifying what usage of these devices looks like overall. Participants made 4.1 commands per day to their devices on average, with 40% of all queries being for Music followed by Information (17%) and Automation (9%). Weekend use was higher than weekday use and the median user settled on three domains that they engaged with per week, not exploring new domains as time went on.

We found that specific types of commands were made more often at particular times of day. Over time, we found that commands in some domains increased in length but that exploration of new topics was low. We found relatively few demographic differences though we did observe that larger households use the devices proportionally more on the weekends.

Correlation analysis found a relationship for overall use of the device: the more commands, the more sessions and the more domains. Different times of day were correlated: evening use was related to weekend use, but midday use was inversely related to nighttime use. More midday use was related to less small talk and more information use. The use of information commands increased with longer commands but was associated to less music use, suggesting those that use the device for information don't use it for music and vice versa. This pattern was reflected in the cluster analysis. The heaviest users used proportionally less music compared to the lighter groups. General use was the strongest variable that shaped the clusters, though some other patterns emerged. We found four user groups: one heavy group, the "super users", two medium-use groups that differed in the number of commands per session, and one lighter group. The super users and those who explored more domains were similar in that they used on average over 40 subdomains. The lightest users and those who used shorter commands were similar in that they used less than 30 subdomains but listened to proportionally more music.

We have learned much more than what we could capture in a single lab setting by looking at the actual commands that users made in the contexts of their own homes and lives. While this form of data collection doesn't replace the need for interviews and more in-depth design research, it gives us insight into longer-term natural rhythms and patterns that can be used for fitting new experiences into people's lives.

We also find this method of data collection interesting as a way for any researcher to gain access to usage data of systems that they do not control. As more companies provide user-accessible databases of stored interaction history, as is becoming law in some places [12], researchers can start to understand the use of systems in the wild and at a large enough scale to find statistically significant patterns of use. Anyone in the world with a few hundred dollars to compensate participants could have executed this study, regardless of their institution. Most importantly with this method, each user gives their consent for their data to be used for this research, similar to already prevalent survey or in-person methods of data collection. We believe that methods like this one can help address some of the ethical issues of data scraping brought up in recent ethics discussions such as the recent town hall at CSCW [8].

5.1 Comparison to Self-Reported Data

Around the same time our study was conducted, two nationally representative surveys were conducted, one by VoiceLab [7] and one by NPR[8]. These studies did not collect logs of use, and instead focused on asking respondents which features they used and at what time of day. We will briefly compare our findings to their results.

As we saw in Table 1, music accounted for 40% of the use of the device for our participants. This further quantifies existing research, such as the VoiceLab study that found that "Playing Music and Books" was the most "liked" feature of voice assistants. The Smart Audio Report from NPR also explored the use of voice assistants through a survey asking users to self-report which domains they used at particular times of day. They found Traffic, Weather, and News all overindexed in the 5-9am hours. We also saw Weather as a more popular domain in these hours compared to the rest of the day. Traffic and News both fit under our Information domain, which also sees a spike from 7-10am. Similar to their report, we saw Automation peaking later at night, however we did not see a rise in alarms at these later hours. Given the self-reported nature of the NPR data, participants may have forgotten times when they snoozed or set extra alarms in the morning, or due to social desirability did not want to admit to this morning laziness.

5.2 Comparison to Smartphone Application Use

We now turn to directly comparing our data with findings from Böhmer et al.'s [6] analysis of mobile phone use. This paper is what inspired our current analysis, as we wanted to understand the smart speaker ecosystem at a similar time in its growth to where mobile phone apps were at the time of Böhmer's study. Smart speaker devices are currently solving many use cases that people have turned to their phones for in the past decade. These devices can give information about the weather, set alarms and timers, give news updates, answer general information questions, and provide access to sports scores and finance tickets. Despite the similarities in functionality provided, there are some striking differences in use in early adopters of these two different types of devices and how they fit into the routine of a person's day.

The first key difference is in count of domain invocations in a typical session. When looking at phone use, Böhmer et al. found that the vast majority of sessions only involved a single app, and that only 5.7% had more than 3 apps launched. Our analysis points to a longer tail of interactions, with only 39% of interactions consisting of a single command and 10% having more than ten commands. When looking at specific domains we observed similar differences. Böhmer et al. found that 68% of mobile phone sessions only involved one app, where we found that only 48% of sessions with the Google Home were only in a single domain.

Another key difference lies in the use by time of day. While the overall arc of use looks nearly identical, the variation in types of domains used throughout the day is much less pronounced on the smart speaker systems. Most categories are fairly stable over time (as seen by the colors staying stable horizontally in Table 1). However, Böhmer et al. found much more variation in times of use for specific applications, so much so that it was the title of their paper ("Falling Asleep with Angry Birds, Facebook, and Kindle"). Other than asking for time and weather in the morning, and doing more with home automation at night, most use of Google Home stayed within a few percentage points throughout the day.

Similar to Böhmer et al.'s analysis, we found that smart speakers have their "hit" applications. Whereas the domain accounting for half of all mobile phone use is communication/messaging applications, music (40%) for smart speakers accounts for the most frequent use of the device.

Our data set is important to the Ubiquitous computing field in that it explores the use of smart speaker systems over time, in the wild, with no researcher intervention. Similar to the work of Böhmer early on in the smartphone era, we hope that this can guide future researchers in understanding the rhythms of use of this newly mass-market technology. Understanding these patterns can help us to build more contextual services in the future that best adapt to a user's context.

---

[7] https://s3-us-west-1.amazonaws.com/voicelabs/report/vl-voice-report-exec-summary_final.pdf
[8] https://www.nationalpublicmedia.com/wp-content/uploads/2018/01/The-Smart-Audio-Report-from-NPR-and-Edison-Research-Fall-Winter-2017.pdf

5.3 Historical Context

We found that many of the most common uses, including music playback, device control, and commands about weather, mimic original use cases developed in the Intelligent Room [10] and other early home AI environments from the 1990s. While systems at the time were not practical for broad deployment (requiring a closet full of computers to run), now the processing capability has been moved to the cloud, allowing for a small and simple in-home device that millions of people can use. These devices represent the first time that millions of users are able to experience many of these technologies that were developed in research labs 15 years ago, and the opportunities to study their adoption and for what purposes they are used is an exciting area of research that has only recently opened up in the past year as these devices reach consumer price points and ease of setup and use.

Beyond these original use cases, the existence of smartphones and smart TVs have enabled new types of interactions, such as managing grocery lists that are mirrored onto the phone and controlling video playback on internet-connected televisions. One interesting difference between today's smart speaker devices and the original Ubicomp systems is that commands given to today's devices are much more about creating action or getting an answer right now, instead of the more agent-based approach that was taken in much of the earlier research. Users are not asking the devices to set up routines or take if-then style interactions but are more often using these devices as a "voice-based switch" to turn on a light, start or stop a timer, or get weather in a very transactional way. Instead of asking "always let me know if it will rain in the morning" users simply ask "will it rain today?" each day when they think it might rain. This shows that there is still much work to do before these devices truly match the ambitions of the earlier research and much promise in making these devices even more useful and proactive.

5.4 Limitations

While we are quite happy with the sample of households that we were able to reach in this data collection study, they are indicative of early adopter users of the device in America. Use in other countries may differ, and use for the next generation of users coming in with the heavily discounted prices ($19 for Google Home Mini) of the 2017 holidays may also be different.

We were also not able to ascertain who was using the device for each command, so our analysis was conducted at the household level. We do know the household makeup, but with our log access we did not receive the device's best estimate as to who was speaking. Note that our data was collected before Google enabled the multi-user mode of the device. We believe it is still quite valuable to understand adoption at the household level, and it might be interesting future research to also download the audio files of each interaction so that speaker identification can be performed.

Finally, while the process of saving and uploading logs was relatively straightforward, we received many empty or invalid submissions to the survey. Some of these users might have been trying to scam us given the high reward for completing the MTurk HIT, while others might not have been able to follow the instructions that we provided due to lower technical literacy (we did describe every step and included screenshots, but there were several steps to bring up, filter, save, and upload the results). Therefore, our sample may tend a little more towards the more technically proficient users of these devices.

6 IMPLICATIONS FOR DESIGN

In exploring the findings, we have uncovered a number of implications for the design of new smart speaker assistants. These implications leverage the patterns that we observed in use over time as well as the types of commands that participants were making.

6.1 Introducing the User to New Domains

One of the most striking findings to us is that use does not change much over time, and that users in general settle on only three domains that they use on a weekly basis and do not frequently try new ones. With the rapidly growing set of third party skills available, there is a large opportunity in teaching users about new capabilities that might help them. It seems that the periodic emails from Google about new capabilities are not helping users to successfully discover new commands that they can make. The assistant itself could help users to find new domains, based on interactions that they make as well as by exploring interactions on other platforms such as mobile phone apps that they have installed or email receipts. For example, the device could see that a user rides Lyft, and then mention that a skill from that company is available the next time they ask for directions or local venue information.

6.2 Opportunities for Timed Multi-Modal Interactions

As new devices come to market that have ambient display screens as well as voice input mechanisms, such as the Amazon Echo Show, devices will have the ability to anticipate people's needs and provide relevant information on an always-visible touchscreen. In our data, we saw that people often turn on music or set timers of specific lengths at dinner time, or ask for the weather in the morning. We can create systems that show these options to the user, or even have the information waiting for them at a glance at the right times, without the need to talk or interrupt others in the household with a response. Users could then choose to refer specifically to the information on the screen, asking for more information about "that" or other quick access words based on what is on the screen such as "star the second email." Much of the previous Ubicomp research on ambient displays will be useful here in creating these multi-modal interactions.

6.3 Suggesting Related or Shortened Commands

We can also help the user with repetitive tasks by engaging in dialog to name specific actions and to create shortcuts. For example, a 12-minute timer at dinner could be renamed "pasta timer" allowing for easy setting later. Lighting combinations can be remembered, such that commands can be simple if the "usual" action is to occur at that time. For example, "TV time" could replace "Turn on the living room lights, turn off the kitchen lights, and turn on the television." This simplification can make interacting with a voice assistant easier, yet also runs into issues in shared homes of remembering commands that other have made such as those identified by Mennicken et al. [20]

We find it interesting that there are relatively few commands for topics of structured information. For example, 1.2% of requests were for sports, 6% were for weather, and 0.9% were for finance. More research is needed to explore this further and understand if there is a lack of demand for these types of information, if it is not clear what types of structured data are available to be queried in these domains, or if users just do not know how to ask the device for these types of information in a way that it can understand.

Finally, it is interesting that users are not making longer commands over time. Conversational assistants can provide ways for users to learn how to make commands more complex. A sports assistant might offer up tips on querying specific players or stat types. An email assistant could help users to understand that they could make longer commands about package tracking, flight status, or personal finance (e.g. "How much have I spent on Lyft this week?") instead of sticking to the basic commands that users are currently making. Interaction with these devices is very one-way right now, with users making specific one-utterance commands. There is a great opportunity in making interactions more of a dialog, which can also introduce users to more advanced features of the device in particular domains.

6.4 Supporting Deeper Agent-Based Interactions

A final implication is in supporting deeper agent-based interactions. As mentioned in the discussion, a stark difference between today's smart speakers and earlier research projects on voice assistance is the single command nature of today's devices. Whereas previous research focused on creating "smart" assistants that could learn over time and remember large amounts of context, supporting rule chaining and other complex logic, today's devices are typically made to answer one specific question or perform one action at a time, by direct command.

Future systems can learn and respond to commands such as "tell me if it's going to rain" or "wake me up early if traffic is bad" or other types of commands that involve using state or context to act on the user's behalf. While this was the vision of many early Ubicomp systems with voice interfaces, these types of agent-style commands are not currently supported as seen in our data as people ask for specific pieces of information daily (e.g. "will it rain today?).

7 FUTURE WORK

This work sheds light on the need for future work in this domain. There are several areas that we see as ripe for next steps. Some are simple extensions to this type of study using other devices and services, while others build on clear findings from this research to build new systems.

The first, and simplest form of future work, would be to replicate this study in other countries or using other devices. The use of these devices might differ in varied cultural contexts or in environments with multiple generations living in a home. Differing expectations of how children use technology in the home may also affect use.

Beyond studying Google Home devices, it would also be helpful to study the use of other products such as Amazon's Alexa devices and compare the different categories of use between devices as well as use over time. Perhaps one platform leads to better exploration of new categories over time, or perhaps people use different language to converse with "Alexa" over "Google" due to the personification of the device. These would be interesting hypotheses for future research to address. Amazon has a similar page with command history that could be scraped using a similar method to this study.

Another way to study existing use would be to examine use within a household and how different family members use the device in different ways. Our logs only captured the use at a household level, and it would be interesting to see if specific family members used the device in different ways. The raw audio files are available on the command history pages, so this should be possible to identify different speakers in an analysis that also collects these files.

Going past simply studying existing behaviors, new systems can be created to encourage users to explore broader ranges of topics or to try new types of commands within a domain. Multiple strategies can be deployed here including exploring conversational approaches on the device, different types of emails or notifications on mobile devices, or ways to share commands that were useful with friends or family. Since we observed a fairly static set of domains over time, helping users to discover features that might be useful for them is increasingly important. Since voice interfaces do not have visible affordances, increasing user's mental models of what they can do is a critical area of future research.

Finally, we have seen that these devices are currently mostly used for music playback. While this was a feature of many of the early voice assistants in research such as the Intelligent Room [10], these systems had many more advanced functionalities that current voice assistants lack. Exploring ways that voice assistants can be more agent-based and proactive is a promising area of future research with these devices. Currently, the vast majority of commands to the device are single sentence commands, that do not take advantage of existing context or past actions. Checking the status of timers or stopping music is about as far as current devices go. Learning music tastes over time, as well as proactively notifying users of specific events, traffic anomalies, weather conditions, etc. could provide many added benefits for users of these devices.

We hope that this work opens up many future questions for continued research in this exciting domain as voice assistants become a part of hundreds of millions of households throughout the world and this area of previous Ubicomp research makes its way into the living rooms, kitchens, and bedrooms with high daily usage.

8 CONCLUSION

Through collecting usage logs from a diverse sample of users, we were able to see how the use of a smart voice assistant fits into a household's daily life and how the use of these devices changes over an average of 110 days of use. Through this investigation, we have developed a deeper understanding of how these assistants fit into the rhythm of a day, the types of commands that people issue as they use the device over weeks and months, and how these commands change in terms of topic or length as familiarity with the device increases. Specifically, we have found that users are heavily using these devices compared to phone-based assistants, issuing a median of 4.1 commands per day that have a median of 4 words, that entertainment and home automation commands peak in the evening while weather and time requests peak in the early morning hours, that only 4% of users tried a new domain in week 14, and that a variety of age and household composition differences occur in use of these devices. Developing this understanding is important as these devices become more common in our lives and is also important in providing insights for creating new conversational assistants or skills.

We hope that this work can spark future studies of smart speaker assistants as well as other devices that have a downloadable interaction history posted online. Specifically, it would be interesting to compare these findings to the use of Amazon's Alexa or Apple's newly-released HomePod or to run a similar study in another year or two as these devices become more common in people's homes.

Finally, more qualitative studies of device use can complement this quantitative examination of command logs to better understand the lived experience of having one of these devices in the home and aspects of use that are most enjoyable or most frustrating. The quantitative approach that we have used opens many questions for further research including a deeper investigation of the "smalltalk" commands where users are attempting to talk to their assistants as if they were people.

ACKNOWLEDGMENTS

We would like to thank our participants who provided valuable usage data as well as Yahoo and the UX Research team for providing the time and space to conduct this research.

REFERENCES

[1] Apple Computer. 1987. Knowledge Navigator. Available at: https://www.youtube.com/watch?v=-jiBLQyUi38 Accessed 9/5/2017.
[2] Frank R. Bentley, Nediyana Daskalova, and Brooke White. 2017. Comparing the Reliability of Amazon Mechanical Turk and Survey Monkey to Traditional Market Research Surveys. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17). ACM, New York, NY, USA, 1092-1099. DOI: https://doi.org/10.1145/3027063.3053335
[3] Frank R. Bentley and Ying-Yu Chen. 2015. The Composition and Use of Modern Mobile Phonebooks. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, 2749-2758. DOI: https://doi.org/10.1145/2702123.2702182
[4] Anne-Jorunn. Berg. 1994. "A Gendered Socio-technical Construction: The Smart House", Cynthia Cockburn and Ruža First-Dilić (eds.), Bringing Technology Home: Gender and Technology in a Changing Europe, Buckingham: Open University Press.
[5] A.J. Bernheim Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, Stefan Saroiu, and Colin Dixon. 2011. Home automation in the wild: challenges and opportunities. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). ACM, New York, NY, USA, 2115-2124. DOI: https://doi.org/10.1145/1978942.1979249
[6] Matthias Böhmer, Brent Hecht, Johannes Schöning, Antonio Krüger, and Gernot Bauer. 2011. Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage. In Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11). ACM, New York, NY, USA, 47-56. DOI: https://doi.org/10.1145/2037373.2037383
[7] Rodney A. Brooks. 1997. "The intelligent room project." In Cognitive Technology, 1997. Humanizing the Information Age. Proceedings., Second International Conference on, pp. 271-278. IEEE.
[8] Amy S. Bruckman, Casey Fiesler, Jeff Hancock, and Cosmin Munteanu. "CSCW Research Ethics Town Hall: Working Towards Community Norms." In Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 113-115. ACM, 2017.
[9] Frank R. Bentley and Ying-Yu Chen. 2015. The Composition and Use of Modern Mobile Phonebooks. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, 2749-2758. DOI: https://doi.org/10.1145/2702123.2702182

[10] Michael Coen, Luke Weisman, Kavita Thomas, and Marion Groh. 2000. "A context sensitive natural language modality for an intelligent room." In Managing interactions in smart environments, pp. 68-79. Springer, London.
[11] Jeffrey Dean. July 7, 2016 Large-Scale Deep Learning with TensorFlow for Building Intelligent Systems, ACM Webinar
[12] European Union General Data Protection Regulation. http://www.eugdpr.org/ Accessed 9/4/17
[13] Michael Grace-Martin and Geri Gay. Web browsing, mobile computing and academic performance. Educational Technology & Society 4, no. 3 (2001): 95-107.
[14] Kirstie Hawkey and Kori Inkpen. 2005. Web browsing today: the impact of changing contexts on user activity. In CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05). ACM, New York, NY, USA, 1443-1446. DOI=http://dx.doi.org/10.1145/1056808.1056937
[15] Daqing He and Ayse Göker. "Detecting session boundaries from web user logs." In *Proceedings of the BCS-IRSG 22nd annual colloquium on information retrieval research*, pp. 57-66. 2000.
[16] Stefan Junestrand, Konrad Tollmar, Sören Lenman, and Björn Thuresson. 2000. Private and public spaces: the use of video mediated communication in a future home environment. In CHI '00 Extended Abstracts on Human Factors in Computing Systems (CHI EA '00). ACM, New York, NY, USA, 16-17. DOI: http://dx.doi.org/10.1145/633292.633304
[17] Cory Kidd, Robert Orr, Gregory Abowd, Christopher Atkeson, Irfan Essa, Blair MacIntyre, Elizabeth Mynatt, Thad Starner, and Wendy Newstetter. 1999. "The aware home: A living laboratory for ubiquitous computing research." Cooperative buildings. Integrating information, organizations, and architecture (1999): 191-198.
[18] Mounia Lalmas, Heather O'Brien, and Elad Yom-Tov. "Measuring user engagement." Synthesis Lectures on Information Concepts, Retrieval, and Services 6.4 (2014): 1-132.
[19] 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 (CHI '16). ACM, New York, NY, USA, 5286-5297. DOI: https://doi.org/10.1145/2858036.2858288
[20] Sarah Mennicken and Elaine Huang. 2012. "Hacking the natural habitat: an in-the-wild study of smart homes, their development, and the people who live in them." Pervasive computing (2012): 143-160.
[21] Amanda Purington, Jessie G. Taft, Shruti Sannon, Natalya N. Bazarova, and Samuel Hardman Taylor. 2017. "Alexa is my new BFF": Social Roles, User Satisfaction, and Personification of the Amazon Echo. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17). ACM, New York, NY, USA, 2853-2859. DOI: https://doi.org/10.1145/3027063.3053246
[22] Simon Robinson, Jennifer Pearson, Shashank Ahire, Rini Ahirwar, Bhakti Bhikne, Nimish Maravi, and Matt Jones. 2018. Revisiting "Hole in the Wall" Computing: Private Smart Speakers and Public Slum Settings. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems(CHI '18). ACM, New York, NY, USA, Paper 498, 11 pages. DOI: https://doi.org/10.1145/3173574.3174072
[23] Myra Spiliopoulou, Myra, Bamshad Mobasher, Bettina Berendt, and Miki Nakagawa. "A framework for the evaluation of session reconstruction heuristics in web-usage analysis." *Informs journal on computing* 15, no. 2 (2003): 171-190.
[24] Mark Weiser and John Seely Brown. "The coming age of calm technology." Beyond calculation. Springer, New York, NY, 1997. 75-85.

Appendix 1: Usage by Household

User ID	CommandsPerDay	SessionsPerDay	CommandsPerSession	PercentMorning	PercentMidDay	PercentEvening	PercentNight	PercentWeekend	AvgCommandLength	PercentMusic	PercentInformation	PercentAutomation	PercentSmallTalk	PercentAlarm	PercentWeather	PercentVideo	PercentTime	PercentLists	NumberDomains
1	24.28	3.93	6.18	0.20	0.23	0.42	0.15	0.24	4.57	0.06	0.11	0.28	0.17	0.06	0.04	0.01	0.01	0.00	61
2	1.20	0.31	3.88	0.05	0.05	0.13	0.77	0.05	4.34	0.08	0.05	0.00	0.12	0.02	0.07	0.00	0.01	0.00	14
3	2.00	0.94	2.13	0.05	0.51	0.39	0.05	0.25	4.01	0.25	0.06	0.09	0.00	0.41	0.02	0.00	0.00	0.00	14
4	3.30	0.67	4.91	0.13	0.13	0.55	0.20	0.41	4.90	0.42	0.23	0.02	0.04	0.00	0.05	0.00	0.01	0.00	41
5	4.50	1.64	2.74	0.13	0.65	0.11	0.11	0.00	3.63	0.02	0.11	0.30	0.16	0.03	0.10	0.00	0.00	0.02	18
6	12.24	3.80	3.22	0.09	0.29	0.51	0.11	0.28	3.53	0.22	0.15	0.06	0.04	0.12	0.01	0.06	0.00	0.03	46
7	2.14	0.55	3.89	0.07	0.49	0.32	0.12	0.62	4.55	0.28	0.27	0.00	0.04	0.03	0.04	0.00	0.02	0.00	35
8	34.21	8.53	4.01	0.15	0.31	0.19	0.34	0.33	3.87	0.24	0.14	0.02	0.08	0.07	0.04	0.02	0.05	0.00	51
9	18.08	4.81	3.76	0.11	0.20	0.40	0.28	0.33	4.10	0.05	0.05	0.61	0.08	0.04	0.02	0.01	0.01	0.00	44
10	4.00	1.66	2.41	0.05	0.19	0.65	0.10	0.46	4.78	0.39	0.13	0.03	0.02	0.25	0.01	0.02	0.00	0.01	31
11	6.46	0.61	10.63	0.04	0.34	0.51	0.11	0.60	4.39	0.06	0.38	0.00	0.18	0.00	0.02	0.00	0.00	0.00	43
12	23.83	5.42	4.40	0.06	0.21	0.42	0.30	0.29	5.23	0.07	0.13	0.31	0.03	0.03	0.01	0.10	0.01	0.00	44
13	37.67	4.08	9.23	0.15	0.39	0.32	0.14	0.39	3.79	0.29	0.10	0.01	0.07	0.04	0.11	0.01	0.02	0.00	54
14	10.08	4.31	2.34	0.14	0.14	0.56	0.13	0.38	4.12	0.05	0.06	0.47	0.05	0.18	0.00	0.00	0.00	0.06	22
15	1.05	0.86	1.21	0.00	0.43	0.46	0.11	0.07	3.28	0.83	0.07	0.00	0.07	0.00	0.04	0.00	0.00	0.00	5
16	14.33	4.92	2.92	0.06	0.41	0.40	0.13	0.17	3.49	0.26	0.08	0.44	0.01	0.01	0.01	0.00	0.00	0.00	29
17	2.72	0.51	5.37	0.09	0.30	0.56	0.05	0.32	3.52	0.41	0.06	0.03	0.08	0.00	0.01	0.07	0.03	0.00	32
18	1.35	0.91	1.49	0.14	0.36	0.41	0.09	0.30	4.07	0.14	0.17	0.00	0.03	0.13	0.02	0.00	0.20	0.00	21
19	19.89	4.22	4.71	0.17	0.03	0.32	0.48	0.30	4.12	0.02	0.14	0.01	0.15	0.16	0.35	0.00	0.01	0.00	19
20	1.08	0.85	1.27	0.08	0.19	0.71	0.00	0.23	6.47	0.13	0.31	0.00	0.08	0.02	0.00	0.00	0.00	0.31	17
21	5.67	2.78	2.04	0.12	0.45	0.33	0.10	0.20	4.12	0.02	0.22	0.00	0.04	0.10	0.06	0.06	0.04	0.00	20
22	6.06	0.57	10.69	0.42	0.26	0.26	0.05	0.33	3.66	0.06	0.11	0.00	0.05	0.13	0.02	0.01	0.03	0.09	39
23	5.16	2.24	2.31	0.13	0.42	0.29	0.16	0.28	4.55	0.03	0.23	0.14	0.08	0.09	0.08	0.00	0.02	0.00	52
24	14.21	3.89	3.66	0.20	0.23	0.31	0.26	0.22	4.16	0.05	0.11	0.38	0.10	0.01	0.05	0.08	0.01	0.00	56
25	9.84	2.64	3.72	0.28	0.21	0.46	0.05	0.39	4.63	0.21	0.20	0.01	0.17	0.04	0.09	0.01	0.01	0.05	46
26	3.17	1.89	1.68	0.58	0.25	0.12	0.05	0.02	4.39	0.00	0.30	0.00	0.09	0.04	0.09	0.00	0.02	0.25	19
27	1.92	1.07	1.80	0.33	0.25	0.26	0.16	0.37	3.59	0.30	0.08	0.00	0.09	0.05	0.07	0.02	0.01	0.04	30
28	34.08	6.87	4.96	0.14	0.40	0.37	0.09	0.37	3.71	0.23	0.04	0.28	0.04	0.08	0.05	0.04	0.01	0.01	44
29	6.01	2.01	2.99	0.05	0.15	0.37	0.43	0.36	4.86	0.03	0.27	0.00	0.16	0.02	0.22	0.01	0.04	0.00	46
30	1.04	0.48	2.18	0.00	0.08	0.63	0.29	0.13	2.71	0.71	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.00	5
31	17.55	5.33	3.29	0.14	0.42	0.34	0.10	0.29	4.72	0.22	0.28	0.00	0.02	0.00	0.12	0.00	0.16	0.00	44
32	9.40	3.20	2.94	0.02	0.30	0.35	0.33	0.33	4.85	0.06	0.29	0.17	0.18	0.00	0.04	0.01	0.00	0.04	22
33	0.47	0.73	0.64	0.00	0.17	0.72	0.11	0.26	4.06	0.16	0.23	0.00	0.17	0.00	0.06	0.02	0.01	0.00	21
34	5.79	1.29	4.50	0.00	0.05	0.83	0.12	0.65	4.12	0.19	0.25	0.00	0.41	0.00	0.04	0.00	0.00	0.00	18
35	2.40	0.61	3.91	0.03	0.24	0.50	0.23	0.52	4.85	0.27	0.23	0.02	0.25	0.03	0.08	0.00	0.01	0.01	37
36	2.00	0.76	2.63	0.00	0.07	0.00	0.93	0.24	4.55	0.00	0.50	0.00	0.50	0.00	0.00	0.00	0.00	0.00	6
37	6.45	1.79	3.61	0.27	0.42	0.21	0.10	0.58	5.31	0.18	0.39	0.04	0.12	0.00	0.08	0.00	0.00	0.00	34
38	1.20	0.61	1.97	0.29	0.60	0.11	0.00	0.36	4.68	0.18	0.13	0.02	0.09	0.18	0.12	0.00	0.01	0.00	25
39	38.58	6.48	5.96	0.18	0.43	0.25	0.15	0.24	3.85	0.31	0.15	0.00	0.04	0.12	0.02	0.01	0.01	0.05	51
40	6.03	1.92	3.14	0.04	0.52	0.35	0.09	0.35	3.64	0.35	0.09	0.06	0.06	0.04	0.08	0.00	0.06	0.01	25
41	27.34	6.88	3.97	0.03	0.31	0.40	0.26	0.26	3.27	0.05	0.08	0.37	0.04	0.01	0.03	0.23	0.00	0.04	52
42	1.11	0.40	2.78	0.16	0.25	0.26	0.34	0.35	4.08	0.29	0.09	0.00	0.11	0.09	0.10	0.00	0.04	0.01	23
43	6.19	2.33	2.65	0.34	0.11	0.35	0.20	0.33	4.17	0.12	0.12	0.17	0.09	0.08	0.05	0.01	0.17	0.00	38
44	1.26	1.17	1.08	0.18	0.39	0.30	0.13	0.27	4.61	0.74	0.01	0.05	0.01	0.01	0.01	0.00	0.00	0.00	13
45	5.01	3.39	1.48	0.01	0.41	0.49	0.10	0.26	4.78	0.42	0.13	0.01	0.08	0.06	0.04	0.00	0.00	0.16	40
46	3.35	1.70	1.96	0.13	0.20	0.58	0.09	0.51	4.14	0.37	0.12	0.01	0.04	0.12	0.16	0.00	0.01	0.00	29
47	4.85	3.51	1.38	0.13	0.21	0.64	0.02	0.15	4.31	0.52	0.15	0.04	0.05	0.00	0.04	0.01	0.00	0.02	32
48	2.91	1.01	2.89	0.09	0.22	0.65	0.03	0.52	4.19	0.42	0.10	0.06	0.08	0.01	0.02	0.00	0.01	0.00	29
49	4.82	1.42	3.39	0.17	0.07	0.07	0.69	0.28	4.05	0.32	0.20	0.00	0.08	0.04	0.01	0.02	0.04	0.00	19
50	5.58	3.08	1.81	0.36	0.45	0.17	0.02	0.17	5.43	0.07	0.41	0.01	0.21	0.01	0.13	0.00	0.00	0.00	46
51	2.59	1.89	1.37	0.12	0.22	0.61	0.06	0.40	4.24	0.15	0.09	0.39	0.09	0.02	0.07	0.01	0.01	0.00	31
52	5.23	1.54	3.40	0.05	0.56	0.31	0.08	0.27	3.64	0.48	0.09	0.07	0.07	0.05	0.05	0.00	0.00	0.00	35
53	6.25	1.35	4.64	0.03	0.45	0.41	0.11	0.46	4.77	0.30	0.20	0.01	0.05	0.03	0.05	0.00	0.01	0.06	44
54	5.69	0.99	5.74	0.00	0.22	0.60	0.18	0.73	5.00	0.15	0.35	0.02	0.10	0.02	0.04	0.05	0.01	0.00	44
55	7.72	2.30	3.36	0.16	0.34	0.42	0.08	0.34	3.63	0.15	0.13	0.00	0.04	0.23	0.03	0.00	0.00	0.01	28
56	18.00	3.77	4.77	0.12	0.47	0.29	0.12	0.44	3.62	0.33	0.05	0.01	0.05	0.11	0.06	0.00	0.01	0.00	34
57	6.33	0.40	15.83	0.00	0.05	0.43	0.52	0.00	3.78	0.36	0.05	0.01	0.16	0.00	0.02	0.00	0.00	0.00	14
58	7.11	4.44	1.60	0.17	0.16	0.44	0.23	0.39	2.77	0.45	0.02	0.20	0.02	0.00	0.09	0.00	0.03	0.00	11
59	2.55	7.78	0.33	0.38	0.25	0.26	0.10	0.28	5.40	0.09	0.25	0.00	0.19	0.03	0.07	0.00	0.06	0.00	18
60	15.73	4.47	3.52	0.31	0.46	0.13	0.11	0.07	3.63	0.20	0.16	0.00	0.11	0.07	0.06	0.00	0.08	0.00	34

61	6.94	1.96	3.54	0.02	0.43	0.26	0.29	0.31	3.78	0.11	0.25	0.03	0.08	0.01	0.10	0.01	0.02	0.00	48
62	4.89	1.31	3.73	0.05	0.37	0.42	0.17	0.63	3.72	0.24	0.12	0.01	0.10	0.11	0.06	0.00	0.00	0.03	34
63	0.57	0.09	6.25	0.01	0.28	0.61	0.10	0.00	4.95	0.22	0.19	0.02	0.10	0.04	0.08	0.01	0.00	0.00	23
64	1.94	1.59	1.22	0.08	0.41	0.44	0.08	0.30	2.56	0.38	0.06	0.00	0.02	0.01	0.05	0.02	0.00	0.00	14
65	0.72	2.00	0.36	0.00	0.15	0.15	0.69	0.31	4.15	0.46	0.15	0.00	0.08	0.23	0.08	0.00	0.00	0.00	5
66	1.19	0.44	2.67	0.00	0.00	0.00	0.00	0.00	0.00	0.44	0.09	0.00	0.28	0.00	0.03	0.00	0.00	0.00	5
67	3.50	2.41	1.45	0.26	0.23	0.48	0.03	0.07	5.15	0.05	0.42	0.00	0.26	0.00	0.08	0.00	0.04	0.00	26
68	4.10	2.09	1.97	0.13	0.39	0.18	0.30	0.32	3.67	0.18	0.15	0.00	0.02	0.11	0.18	0.00	0.05	0.00	34
69	3.97	0.85	4.66	0.14	0.29	0.29	0.29	0.56	3.91	0.34	0.07	0.04	0.16	0.00	0.05	0.14	0.01	0.01	35
70	0.12	0.02	5.50	0.05	0.00	0.32	0.64	0.64	4.41	0.41	0.05	0.05	0.18	0.00	0.00	0.00	0.00	0.00	7
71	2.30	2.03	1.13	0.22	0.26	0.37	0.16	0.44	3.99	0.22	0.06	0.01	0.06	0.12	0.08	0.00	0.00	0.04	23
72	7.02	3.00	2.34	0.11	0.57	0.11	0.21	0.30	2.47	0.47	0.01	0.00	0.02	0.00	0.11	0.01	0.00	0.00	10
73	3.06	0.62	4.93	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.35	0.01	0.13	0.02	0.08	0.01	0.01	0.01	47
74	5.28	2.98	1.77	0.25	0.40	0.24	0.11	0.20	3.90	0.29	0.19	0.00	0.02	0.07	0.06	0.01	0.12	0.00	42
75	4.67	10.67	0.44	0.07	0.21	0.43	0.29	0.36	5.71	0.36	0.29	0.00	0.00	0.29	0.00	0.07	0.00	0.00	7
76	23.52	5.22	4.51	0.10	0.35	0.34	0.22	0.28	4.24	0.22	0.25	0.00	0.05	0.10	0.03	0.00	0.04	0.02	53
77	0.47	0.32	1.45	0.00	0.63	0.13	0.25	0.06	6.00	0.00	0.38	0.00	0.00	0.00	0.31	0.00	0.00	0.00	8
78	3.38	6.90	0.49	0.07	0.52	0.21	0.20	0.23	4.73	0.25	0.08	0.04	0.07	0.23	0.00	0.03	0.04	0.00	16
79	4.77	1.78	2.67	0.02	0.20	0.73	0.05	0.36	5.03	0.13	0.16	0.00	0.06	0.52	0.05	0.01	0.00	0.02	33
80	6.09	3.33	1.83	0.10	0.21	0.44	0.25	0.48	4.71	0.18	0.24	0.03	0.25	0.05	0.05	0.00	0.03	0.00	44
81	0.71	4.97	0.14	0.04	0.26	0.41	0.30	0.30	4.41	0.07	0.07	0.67	0.00	0.04	0.00	0.00	0.00	0.00	8
82	0.39	0.30	1.30	0.00	0.23	0.23	0.54	0.23	5.00	0.00	0.62	0.00	0.08	0.00	0.15	0.00	0.15	0.00	8
83	0.64	2.93	0.22	0.16	0.07	0.05	0.01	0.08	3.91	0.32	0.12	0.01	0.07	0.09	0.04	0.00	0.00	0.00	15
84	4.17	13.55	0.31	0.08	0.09	0.62	0.21	0.35	4.99	0.23	0.19	0.00	0.13	0.00	0.11	0.02	0.00	0.00	28
85	2.83	0.94	3.02	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.16	0.03	0.13	0.01	0.04	0.00	0.00	0.00	32
86	9.17	3.69	2.49	0.12	0.30	0.46	0.12	0.32	4.94	0.11	0.22	0.39	0.07	0.01	0.11	0.00	0.01	0.00	33
87	1.29	1.63	0.79	0.10	0.38	0.41	0.11	0.21	3.41	0.22	0.24	0.00	0.11	0.02	0.06	0.00	0.02	0.00	32
88	5.75	2.85	2.02	0.01	0.13	0.39	0.47	0.33	4.38	0.07	0.07	0.45	0.07	0.10	0.05	0.00	0.00	0.00	37

Received February 2018; revised May 2018; accepted September 2018.