Data Abstraction Elephants: The Initial Diversity of Data Representations and Mental Models

Katy Williams
kawilliams@arizona.edu
University of Arizona
Tucson, Arizona, USA

Alex Bigelow
alex.bigelow@stardog.com
Stardog
Arlington, Virginia, USA

Katherine Isaacs
kisaacs@sci.utah.edu
SCI Institute and Kahlert School of Computing
Salt Lake City, Utah, USA

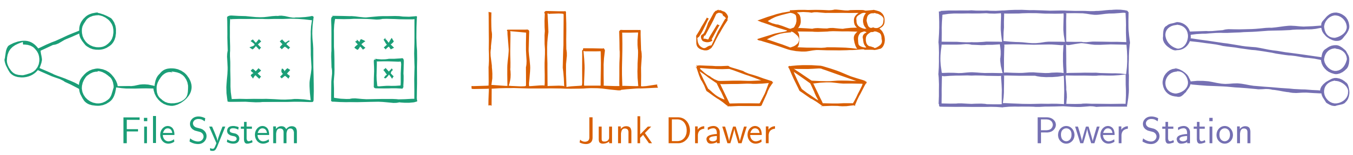


Figure 1: We created three datasets that permit a variety of different data abstractions. The file system dataset sketches include hierarchies and nested sets. The junk drawer dataset sketches include bar charts and drawings of the physical objects. The power station dataset sketches include tables and node-link graphs. We found a variety of data abstractions across each dataset.

ABSTRACT

Two people looking at the same dataset will create different mental models, prioritize different attributes, and connect with different visualizations. We seek to understand the space of data abstractions associated with mental models and how well people communicate their mental models when sketching. Data abstractions have a profound influence on the visualization design, yet it's unclear how universal they may be when not initially influenced by a representation. We conducted a study about how people create their mental models from a dataset. Rather than presenting tabular data, we presented each participant with one of three datasets in paragraph form, to avoid biasing the data abstraction and mental model. We observed various mental models, data abstractions, and depictions from the same dataset, and how these concepts are influenced by communication and purpose-seeking. Our results have implications for visualization design, especially during the discovery and data collection phase.

CCS CONCEPTS

• **Human-centered computing** → **Visualization theory, concepts and paradigms**.

KEYWORDS

Human-centered computing, visualization theory, data abstractions

ACM Reference Format:
Katy Williams, Alex Bigelow, and Katherine Isaacs. 2023. Data Abstraction Elephants: The Initial Diversity of Data Representations and Mental Models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3544548.3580669

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3580669

1 INTRODUCTION

A viewer of a data visualization brings their wisdom, experiences, biases, and interests to their viewing. This internal knowledge and their understanding of the data visualization comprise their mental model of the visualization [17]. A mental model is a personal understanding of a topic that may consist of representations of objects, background knowledge about the topic, and connections to related topics. In the field of data visualization, research has been done on mental models arising from dashboards of political data [36], trees and hierarchies [46], social networks [38], and scientific visualizations [39]. But what about the mental model that exists *before* the visualization is made, when the visualization designer and the domain expert are discussing the dataset? As visualization designers, what steps should we take to elicit and understand our viewer's mental model and how should we design following that mental model to maximize understanding and utility?

For visualization designers, our usual starting point with a new dataset is to connect it with an existing data abstraction, like a table or a network. A data abstraction is a mapping of domain-specific data to an abstract data type [30]. By selecting a data abstraction that has been repeatedly used and refined, we narrow the scope of possible visualizations to create and increase the likelihood of success by building on others' prior work in visualization. A data abstraction provides an intermediary for the designer and viewer, providing a structure for the viewer's intangible mental model and guiding the designer toward visualization design choices that will resonate with the viewer. However, often there is more than one data abstraction that may work for a given dataset. The same dataset may be initially matched with a hierarchical data abstraction, but a set data abstraction could work instead. While sometimes there is a *better* data abstraction choice for a dataset, more likely there is simply an alternative data abstraction that provides different insights.

The variety of abstractions and their implications for visualization reminded us of two stories involving elephants: the parable of the five people in the dark, reaching different conclusions about the nature of the elephant; and John von Neumann's overfitting of data to an elephant, whether or not the data truly represents one.

The parable, which has appeared in Hinduism, Jainism, Sufism, and Buddhism [41], describes five people who are unable to see encountering an elephant for the first time. Each person touches a different part of the elephant and comes to a different conclusion. For example, the person who feels the tusk says the elephant is hard and smooth like a spear, and the one who feels its side thinks it is like a wall.

Von Neumann's purported observation was, "With four parameters I can fit an elephant, and with five I can make him wiggle his trunk;" in other words, the data can be made to fit what we want to see.

These stories reflect possible pitfalls when applying data abstractions to a dataset. The first reminds us that each person may have a different mental model of the data, shaped by their expectations, prior knowledge, and inferences; this mental model impacts their preferred choice of data abstraction. The second reminds us that an individual, perhaps a data visualization designer, can force-fit a data abstraction where it may be unhelpful or misleading. With these pitfalls in mind, we set out to better understand the breadth and form of data abstractions arising from people's mental models and how they communicate their mental models before they are influenced by abstractions or representations chosen by other parties. We seek to build a foundational understanding to drive more concrete guidelines and methodologies for eliciting mental models and exploring data abstractions during the design phase of visualization design [37].

As visualization researchers, we recognize that our users may have mental models of the dataset that could prove to be valuable resources to leverage during the design process. Often in the case of design studies, we are creating a visualization tool where none has previously existed. The only mental model the user has is one of the data space and their interactions with the dataset. The user's mental model may include aspects of an insufficient visualization that serves some but not all of the user's needs. In situations like these, the data and tasks are often still fluid and need to be stabilized.

This instability in the data and tasks during the initial stages of the design process can be beneficial, as it provides more options to explore and does not impose a bias on the design. As we show in this paper, the creativity and lateral thinking shown during our interviews about mental models of datasets suggest that domain experts can offer creativity coupled with domain knowledge that could lead to more productive brainstorming and collaborating early in the visualization design process. However, the best practices for eliciting mental models of data and incorporating their related data abstractions in visualization design are unclear.

As a visualization community, we would like to develop more concrete guidelines and methodologies for eliciting mental models to help steer our data abstraction choices. However, we need to understand fundamentally how internal representations of data are translated to external representations and how difficult pinning down that mapping can be. To begin, we consult existing literature on mental models and their elicitation in areas like education, natural resource management, artificial intelligence, cognitive science, and psychology.

Mental models are notoriously difficult cognitive phenomena to elicit [8]. Klein and Hoffman describe the multitude of reasons why we should not study mental models, yet argue that *because* of their slippery, elusive behavior, we should continue to strive to find best practices for eliciting, describing, and analyzing mental models [23]. We continue the conversation by asking these questions to further understand how understanding mental models can help with visualization design: How do we avoid choosing a non-fruitful data abstraction during visualization design? How many abstractions should we include if multiple abstractions provide insight into the data [2]? There is an inclination toward selecting a single "good" abstraction, but by doing so, how much do we compress the space of reasonable abstractions? Is there a breadth in how people think about these data abstractions in their existing mental model of a dataset? How big is this breadth? We do not attempt to answer all of these questions but provide this paper as a starting point for the community to investigate mental models at the start of the design study process, before the existence of a visualization, to strategically explore suitable data abstractions.

Specifically, we begin by asking the following research questions:

- What factors influence people's initial mental models of data?
- What encodings and visualizations do people commonly use to communicate their mental model?
- How do people describe how they think about the data? How do people describe their sketches?
- How difficult is it for people to sketch and/or describe their mental model? How difficult is it for us to understand?

With the answers to these questions, we can have a better understanding of how users attempt to convey their mental models of datasets, which allows us to incorporate aspects of their mental model in our choice of data abstraction and visualization design. While no one use case is a perfect representative of a "typical" design study, we conduct this experiment using small, incomplete datasets in paragraph form to represent design studies where the data are evolving and both the designer's and the user's mental models of the data are shifting throughout the design process. Further research is needed into techniques for improving the elicitation process in a visualization design context, but our study shows how semi-structured interviews and eliciting representations in the form of sketches can be effective means of clarifying how a person thinks about a dataset. These results have implications for how designers approach the initial stages of the design study methodology, and how effectively and efficiently they can execute the design study.

Recognizing the open-ended nature of these research questions, we conducted a study into the mental models and data abstractions people create from a dataset. Rather than presenting participants with tabular data, which has been shown to influence design choices [1], we presented participants with one of three datasets in paragraph form. We observed a wide variety of mental models, data abstractions, and depictions from the same dataset, as well as how these concepts are influenced by communication and purpose-seeking. We present our collection of core concepts and their implications for visualization design.

In summary, our contributions are:

(1) A set of themes, supported by codes, that describe the diversity of initial mental models and data abstractions, their depictions and influences leading to them, and how they are communicated (section 4),
(2) Implications of these themes and codes for visualization and data design (subsection 5.3), and
(3) An open database of the sketches and transcripts resulting from the study. [1]

We discuss background in mental models, data abstractions, and sketching (section 2). Next, we detail our study methodology, the motivations behind our three synthetic datasets, and our analysis process (section 3). We explain how our interviews and sketches support our codes, which in turn motivate our themes (section 4). We discuss our research questions (subsection 5.1), and the limitations of our study (subsection 5.2), and we provide implications for the visualization community (subsection 5.3).

2 BACKGROUND AND RELATED WORK

We discuss related work in mental models, data abstractions, and sketching in visualization.

2.1 Mental Models

We draw on the abundance of research on mental models in areas like cognitive psychology [19], design [18], and HCI (e.g. [16, 38]), as well as their applications in visualization (e.g. [3, 25]), natural resource management [21], computer science education [15, 34], and engineering [17]. A *mental model* is an individual's understanding of a subject or concept that consists of their prior knowledge, understanding of the presented material, and integration of the knowledge with their worldview. Mental models are more abstract than perceptual images. They contain less detail because our brains omit details we deem irrelevant, yet contain more information than a visual image because they include our prior knowledge [35].

Research on mental models often examines how well people learn "something in the world," frequently an interactive or dynamic system [23]. Klein and Hoffman explain that the mental model is shaped by the rules, laws, and principles that govern this "something" as we observe and learn how this something exists in the world [23]. Jonassen and Henning state mental models are "representations of objects or events in systems and the structural relationships between those objects and events?' [20]. We compare our findings to those of other mental models in subsection 4.7.

Studying mental models is challenging because there is limited accuracy, they are unique to each individual, they are incomplete representations of reality, they are inconsistent and context-dependent, and are highly dynamic models [21]. Klein and Hoffman outline these issues with mental models, and why researching mental models can be controversial but worthwhile [23]. They argue that it is imperative to understand how mental models are formed and how mental models may be modified to increase the depth of understanding, with applications in education and group discussions.

[1]https://osf.io/kvnb9/

Given mental models are internal phenomena, any method to elicit a mental model can only give us a representation of the mental model. Sketching, interviews, and arranging topic cards are common due to their flexibility. Harper and Dorton created a more specific elicitation method for mental models that uses a detailed notational framework to visualize the mental models [16]. A more indirect approach is observation, such as listening to participants think aloud about their strategy in the word-guessing game Passcode, as they work with an AI to understand how the AI gives and receives clues about the word [14]. Regardless of the method, these knowledge-elicitation methods have been repeatedly tested by cognitive scientists and the strengths and weaknesses of using the methods on mental models have been discussed at length [7, 8, 21].

A popular strategy for studying mental models is to use direct elicitation. Direct elicitation requires the interviewees to represent their understanding of a given topic externally, e.g. by drawing a diagram of their mental model or by arranging a set of cards of existing concepts [21]. Interviews are also viable ways to elicit mental models. Milgram and Jodelet asked Parisians to draw a map of Paris and speak about all of the elements of the city that came to mind. From the activity and follow-up interview, they found that participants' sketches of "their" city were a combination of major city landmarks and personal touches, such as a butcher including the meat stockyards or an architect adding an avenue to connect prominent structures [27]. Like with all representations, these representations of mental models are influenced by the skill of the interviewer and the ability of the interviewee to verbalize their understanding.

2.2 Mental Models and Data Visualization

In a collaborative group setting, people share ideas and socially negotiate a community mental model that draws on collective experiences, knowledge, and wisdom from the individuals in the group [20]. This setting occurs in the use of data visualizations, such as when stakeholders are analyzing a visualization. This collaboration also occurs in the early stages of the design methodology, when domain experts and designers are negotiating the data and tasks they wish to support. Liu and Stasko argue for the inclusion of mental model research in visualization, saying that visualization can be viewed as a tool to support the formation of mental models about data and information [24]. They developed a visualization-centric definition of a mental model, stating a mental model is a "functional analog representation to an external interactive visualization system" and listing characteristics of that internal representation. They use this definition to explain how internal representations affect how people interact with external representations and vice versa. To put this theory into practice, Mayr et al. present measures and evaluation procedures to assess mental models in other domains and discuss their applicability to information visualization [25].

Visualizations are effective ways of modifying mental models to improve understanding. The addition of an effective visualization when learning new concepts can be critical to developing a viable mental model of a new subject or system, such as computer architecture [45].

2.3 Data Abstractions

A *data abstraction* is a mapping of domain-specific data to an abstract data type [30], e.g., power station supply lines can be mapped to a network, providing a more generalizable form to the data. Abstraction should happen early in the design process, during the discovery stage, and should be frequently re-examined by the domain experts to ensure correctness and cohesion with their mental model of the problem [37]. The mental model of the user might not neatly correspond to one particular data abstraction, but the discussion around the data abstraction can serve as a way for the user to make their abstract mental model more concrete to help the visualization designer. The visualization designer may need to change the abstraction based on their understanding of how the user interacts with the data and the tasks they are trying to accomplish. Exploring alternative abstractions and their usefulness is much simpler at the beginning of the design process before significant time and resources have been invested. Often there is not a single correct abstraction; instead, abstractions must be designed [26, 29] to best suit the user's needs.

Many authors have identified that difficulties exist in communicating effectively about data abstractions [33, 37, 42]. Trees and graphs can be especially hazardous abstractions to work with, in terms of their potential for miscommunication [31], especially when dealing with edge cases or when people use mathematically imprecise language to discuss graphs [13]. Bigelow et al. found that introducing a *data abstraction typology*, a model that describes the space of possible data abstractions and/or data wrangling operations, can spark discussion and elicit more specific communication about the dataset and abstraction, even when the typology is imperfect [2]. Similar to Bigelow et al. [2], we conduct a study of data abstractions; however, we seek to answer a different set of questions. Bigelow et al. focused on the utility of considering a change in the dataset type of an existing data abstraction. We seek to understand how multi-abstraction datasets can be interpreted and represented.

Tension naturally arises when trying to work with data: tension between the internal data abstraction and the external data abstraction, tension between the imagined visualization and the constraints of the system [2], and tension between the provided data and the desired data. Tension between users and visualization designers may also arise. Even visualization designers and developers may have difficulty communicating about data mappings, anticipating changes to the data, and elucidating technical challenges [42].

2.4 Sketching

Sketching is used in different ways in visualization, often for prototype demonstrations by designers, but also in understanding how people create visualizations for their own personal use. Data sketching is a simple way to show personal mental models, such as students' concepts of time [12] or homeowners' concepts of their home wireless network [32]. Understanding the language of diagrams and how we visualize our thoughts [40] enables us to successfully collaborate and share visualizations. Communication and gestures help augment what is on the page [5].

Walny et al. used data sketching to examine *external representations* people created from a novel dataset [43]. They examined the diversity of data representations and the relationship between sketches and people's understanding of that data. Participants were given a table of ratings of human behaviors in social settings as a dataset.

While our study shares similarities with Walny et al. [43], there are significant distinctions between the two. For methodology, we presented our dataset in paragraph form, rather than in a table, to minimize influencing the data abstraction with a prior data abstraction. Text is not a typical format of the data, but the paragraphs were rather list-like (see subsection 5.2). Our study had a pre- and post-sketching discussion, rather than writing a free-response answer to a question. For research questions, Walny et al. examined the range of visualizations that were created, placing the participants' sketches on a numeracy to abstractness continuum. We use this numeracy to abstractness continuum to code our results, see subsection 4.6. However, rather than the encodings and contents of the sketches themselves, we are more interested in what sketched representations can reveal about the data abstractions that participants assume or construct in their minds. As we were interested in mental models and views about data, our semi-structured interviews allowed us to delve into these discussions.

3 STUDY METHODOLOGY

To elicit data related to data abstractions, we conducted interview sessions where participants were asked to sketch a small dataset and then discuss their sketch and mental model through a semi-structured interview. We used three datasets designed for the potential to elicit different data abstractions, with each participant being shown one. We piloted the study with five participants, after which we iterated on the designed datasets and the interview questions.

Three authors participated in coding interview transcripts and sketch photographs and met regularly to develop codes further. We continued collecting data until we reached saturation regarding our research questions. We describe the details of this study below. An overview of our procedure is shown in Figure 2.

3.1 Participants

We recruited 28 participants, listed in Table 1 by occupation and dataset prompt. We sent out recruitment requests to five organizations, of which we recruited participants from a university's computer science (CS) Discord server and undergrad CS mailing list as well as posting fliers and word-of-mouth in the local YMCA community. Of the participants, 20 had computer science-related work (16 were CS or Information Science students, 4 were computing professionals—3 developers, 1 project manager in an IT department) and 8 had other occupations. Participants ranged in age from 18 to 77 years old, with the mean age being 30.7 years old.

We did not conduct a visualization literacy test but asked participants how frequently they visualized data. Most (21) participants reported "sometimes" and seven reported "always." However, in subsequent discussions, we discovered a wide interpretation of "visualizing data" from imagining data to varied frequencies of plotting. We further discuss these results in the analysis.

Table 1: Participants

Participant	Occupation	Dataset
006	Student (CS + Math minor)	File System
007	Student (CS + Management and Information Systems minor)	Junk Drawer
008	Student (CS + Math minor)	Power Station
009	Student (CS), Software Developer	File System
010	Student (CS)	Junk Drawer
011	Student (CS)	Power Station
012	Student (CS, Biochemistry), CS Teaching Assistant	File System
013	Student (CS, Information Science)	Junk Drawer
014	Student (CS, Information Science)	Power Station
015	Student (CS)	File System
016	Sales	Junk Drawer
017	Substance Abuse Counselor, Swim Instructor	Power Station
018	Web Developer	File System
019	Editor (Retired)	Junk Drawer
020	Software Engineer	Power Station
021	Project Manager (IT)	File System
022	Student (CS)	Junk Drawer
023	Nurse	Power Station
024	Student (CS), Research Assistant	Power Station
025	Student (CS), Research Assistant	Junk Drawer
026	Research Analyst	Power Station
027	Student (CS), Research Assistant	File System
028	Student (CS)	Junk Drawer
029	Student (Chemical Engineering, Information Science)	Power Station
030	Data Scientist, Programmer	File System
031	Army Wife	Junk Drawer
032	Student (Medicine)	Power Station
033	Financial Consultant	File System

3.2 Setup and Materials

All sessions were conducted through video-conferencing software. Participants were instructed to bring a pen or pencil and a sheet of printer paper to the virtual meeting, although seven participants used lined paper and three participants used some form of electronic drawing software (e.g., tablet). Each participant was asked to angle their camera toward the paper as they sketched. At the end of the session, participants were told to take a digital (phone) photograph of their sketch and submit it. Sessions typically lasted around 20 minutes, lasting no longer than 30 minutes.

3.3 Datasets

We created three (3) datasets which we refer to as File System, Junk Drawer, and Power Station. These names were not shared with the participants. Our goal was to design datasets that afforded multiple data abstractions, based on prior research exploring the facility of changing the data type of an existing data abstraction [2].

We created relatively elementary and sparse datasets with the intention of (1) being accessible to people with a broad range of backgrounds, (2) allowing wide interpretations if they existed, and (3) limiting the need for revising the drawing and thus increasing the likelihood we were observing the initial mental model. We recognize that many datasets are often provided to visualization designers and collaborators "as-is". However, we see the value in *discovering, capturing, curating, designing, and creating* [29] the data and wanted to understand if and how our participants explore data abstractions, in this case, for example, the dataset itself is under construction and thus in flux.

We prioritized keeping the datasets short and understandable, though not necessarily comprehensive. All three datasets were presented in paragraph form rather than as a table so as not to influence the mental models toward tables [1].

We chose not to include tasks with our datasets. In visualization design, tasks are often unclear from the beginning, so in addition to using paragraph form, we provided no additional purpose or tasks to the participants so as not to further influence toward a particular data abstraction.

We discuss the limitations of our choices in the use of paragraphs and the omission of tasks in subsection 5.2.

3.3.1 File system. You have two folders. In the first folder are 2 text files and 3 images. In the second folder are 4 text files, 2 code files, and 1 folder. In this folder are 1 text file and 1 image.

The file system dataset was inspired by discussions of file system formats [13] and research regarding difficulties first-year computer science students have with navigating file systems [6]. We collected the age of the participant to see whether we would also notice this

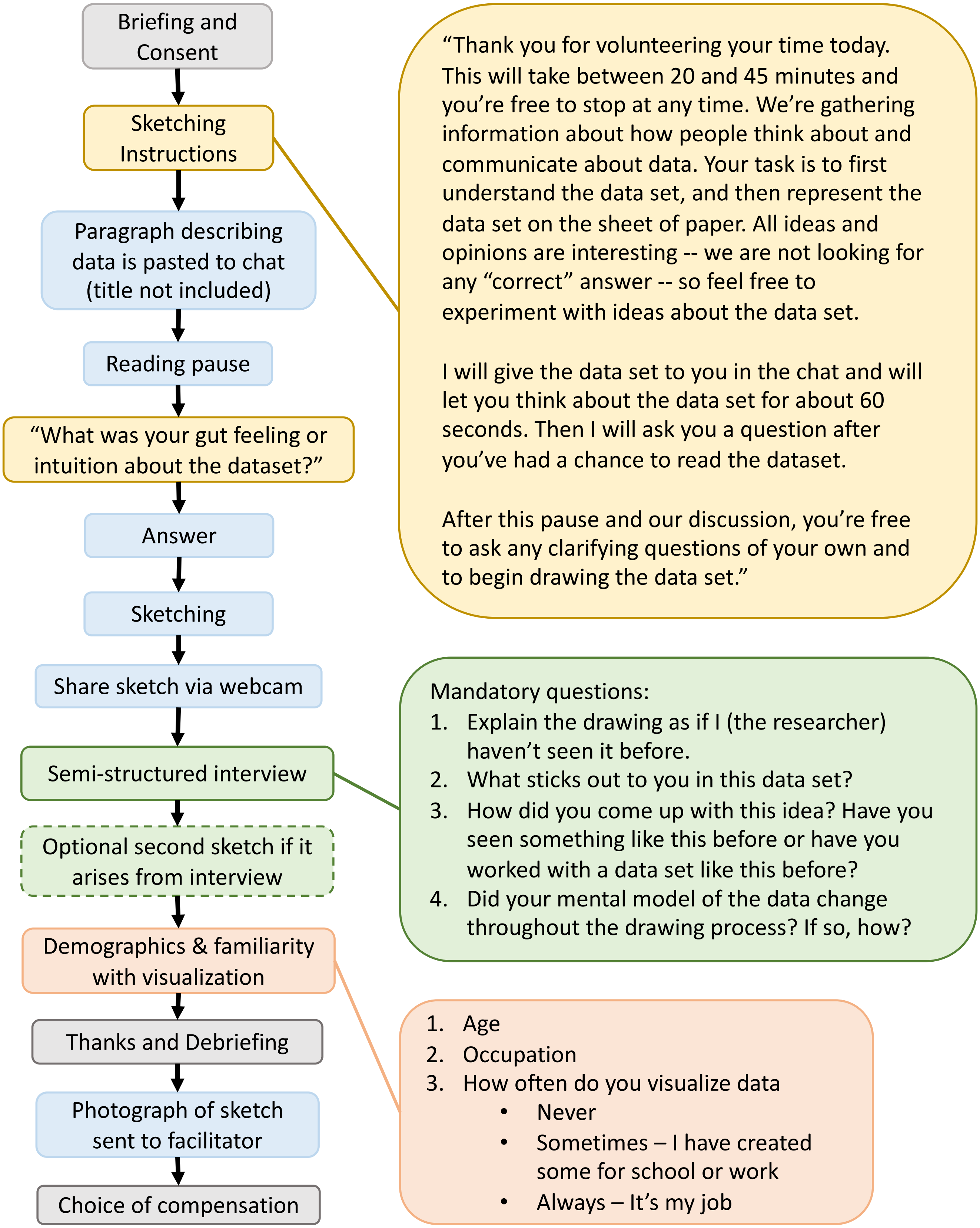


Figure 2: Overview of the interview procedure.

phenomenon. We made this a hierarchical set of files to see how participants handled the nested folder. The resulting sketches are shown in Figure 3.

3.3.2 Junk drawer. *You have 6 rubber bands, 4 tacks, 3 unused envelopes, a roll of stamps, 4 pens, 3 pencils, 2 sharpies, a small basket, a pencil pouch, and a long plastic basket.*

We designed this dataset to be lacking an obvious (non-list) structure, but with several options for imposing one. We included a possible "container" for different groups of items, e.g., the pencils could go in the pencil pouch. We were curious to see if the participants extracted sets or groups from the dataset and how these sets might differ. The resulting sketches are shown in Figure 4.

3.3.3 Power station. *There are 6 power stations, labeled A through F. Power station A powers 100 homes. Power station B powers 150 homes. Power station C powers 1 warehouse and 100 homes. Power station D powers 4 apartments, each housing 100 residents. Power station E powers 50 homes and 2 apartments, each housing 100 residents. Power station F powers 50 homes.*

For the power station dataset, we wanted a variety of classes of data items to allow for different mark types or icons. We also were curious if participants would tie in geographic attributes to the dataset or if we would see any networks, allowing for different visualizations from the previous two datasets. The resulting sketches are shown in Figure 5.

3.4 Procedure

We first briefed participants and obtained the study and recording consent. Each participant was then given an overview of the sketching activity verbally and the text of one dataset through the videoconferencing application's chat feature. Our name for the dataset was not included. See Figure 2 for the overview script. Participants independently read and considered the dataset, then informed the facilitator once they were through. The approximate time most participants took to read and consider the dataset was under 30 seconds. After this, the facilitator asked, *"What was your gut reaction or intuition about the dataset?"*

After the ensuing discussion, the participant was asked to angle their camera and sketch the dataset. Participants were allowed to draw until they felt satisfied with their drawing, with most participants completing their sketches in under 4 minutes. [2] We then conducted a semi-structured interview with the following pre-set questions:

(1) Explain the drawing as if I (the researcher) haven't seen it before.
(2) What sticks out to you in this dataset?
(3) How did you come up with this idea? Have you seen something like this before or have you worked with a dataset like this before?
(4) Did your mental model of the data change throughout the drawing process? If so, how?

The first question (explanation of the drawing) is designed to help disambiguate the sketched representation as the authors might interpret it from the participant's view of the sketch. The intent is to separate the sketched visual form from the data abstraction that matches the participant's mental model, providing a way for participants to clarify their representation of their mental model when inhibited by their sketching capabilities. The combination of drawing and interviews is a technique used in mental model research [21].

The second and third questions probe possible influences. The fourth question is designed to provide insight into the possible evolution of mental models, both during initial formation and possibly due to the study design.

In these discussions, some participants augmented their responses by making a second sketch, sometimes prompted by the interviewer to better understand their words. These bonus sketches occurred in eight of the 28 sessions, bringing the total sketch count to 36 sketches.

We concluded with demographic questions and a short debriefing. The procedure was designed to take no longer than 30 minutes. Due to a logging error, exact times are missing for five participants, though all finished within their 30-minute slot. The remaining participants finished within 15-25 minutes with a median finish time of 20 minutes. Participants were compensated with their choice of plush toys, a $10 gift card, or a $10 donation.

3.5 Thematic Analysis

We took an inductive thematic analysis approach. During data collection, three authors individually noted *codes* and thoughts regarding the transcripts and sketches, initially following an unconstrained *open coding* [28] practice. These codes were recorded as *memos* on a shared GitHub repository[3] to facilitate remote collaboration and to track the provenance of codes.

Though we could have chosen a deductive coding approach for the data abstraction using an existing typology, we deliberately chose to exclusively use inductive coding to not limit, bias, or constrain the data abstractions discovered or our interpretation of the ways the participants spoke about their mental models.

The authors met regularly to discuss the codes, limiting the discussions to the sessions where all authors had had a chance to code. Typically 3-5 sessions occurred between each meeting to discuss initial codes. In total, we coded 28 transcripts and 36 sketches.

As these discussions took place, we moved to axial coding to develop hierarchical concepts. The authors used Google Jamboard to cluster, merge, and split their initial codes and to identify concepts arising from multiple codes. We arrived at 24 consensus codes. The identified concept groupings were then discussed, distilled, and refined into the shared document in the GitHub repository. This permitted asynchronous discussions regarding concepts as they progressed.

The discussions and refinement of concepts led to the discovery of common observations that reinforced codes and also helped us refine our data collection. For example, after observing that participants tended to draw items in the order of reading, we wondered if the alphabetical order of the power station dataset might be influencing this phenomenon. Thus, in sessions 026 and 029, we

[2] Participant 013 continued to draw and add detail to their sketch for 11 minutes, at which point the facilitator asked them to stop so that they had time for the discussion questions.

[3] https://github.com/kawilliams/mental-models-codes/blob/main/codes.md

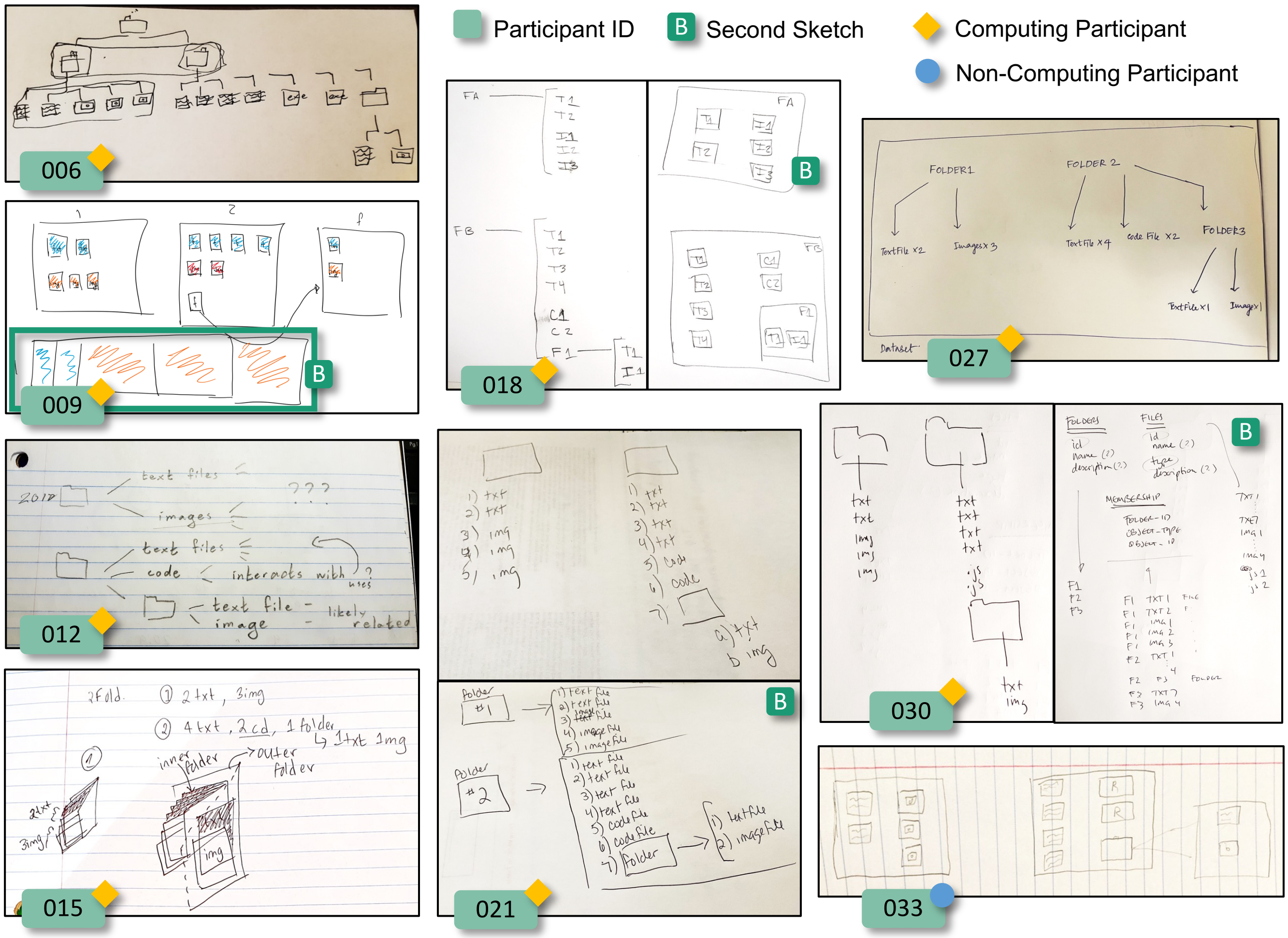


Figure 3: Sketches that participants made of the File System dataset. Large versions are in the supplemental archive.

presented the power stations in non-alphabetical order; however, the participants continued to fill in the data in the order of reading.

The authors initially developed a set of themes from research questions and hierarchical groupings of codes. After external feedback, two authors reconsidered the codes and initial themes, determining the themes had become too broad. Codes and concepts were then reorganized into six themes (described below in section 4) elaborating on the research questions and one secondary theme regarding perceptions of data.

4 THEMES AND CODES

We arrive at three clusters of themes relating to mental models of data: mental model content, mental model elicitation, and mental model formation as well as a secondary theme regarding beliefs about data. Below, we explore the themes in each cluster in the context of our study and explain select codes that made up these themes. For detailed supporting evidence for each theme and code, see the supplemental archive.

After presenting our main and secondary themes, we follow up with a discussion of themes regarding our computing and non-computing populations (subsection 4.5) and our mental model characterization in discussion with the model of Walny et al. [43] (subsection 4.6).

4.1 Themes about Mental Model Content

During our thematic analysis, we developed two themes regarding mental model content. The codes comprising these themes have to do with the breadth and composition of mental models. While this cluster contains our best effort in understanding the form of the participants' mental models, it does not contain codes relating to how participants depicted or otherwise communicated that mental model. We discuss those latter codes in the cluster Themes about Mental Model Elicitation (see subsection 4.2). The list of codes relating to mental model content can be found in Table 2 with the code label listed as "(C#)"; the complete list of corresponding codes and their definitions and supporting data can be found at https://osf.io/kvnb9/ and in the supplemental material.

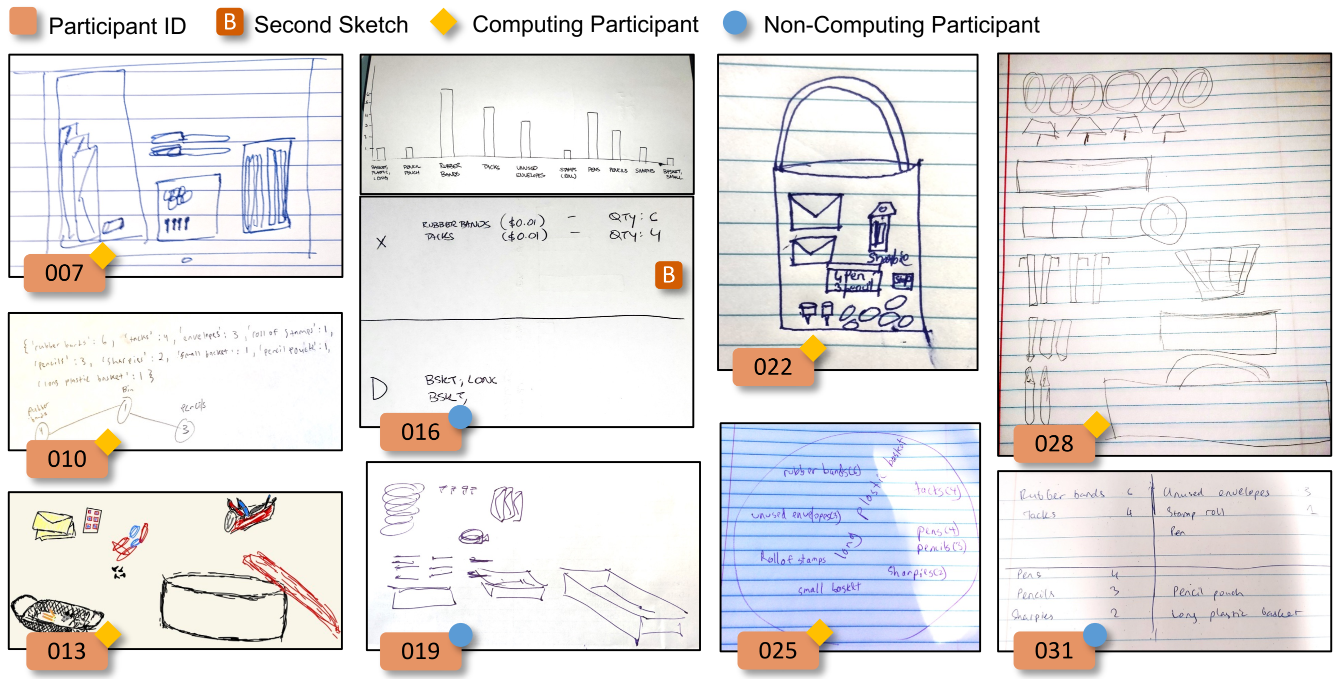


Figure 4: Sketches that participants made of the Junk Drawer dataset. Larger versions are in the supplemental archive.

4.1.1 Theme: Diversity of mental models. Across each of the three datasets, the participants chose different abstractions and representations. We classified both the data abstraction (e.g. hierarchy) and the representation used (e.g. node-link) for all sketches. While it is hard to disassociate the typology from the representation in some cases, we took a best-effort approach based on both the sketch and the way the participant spoke about the sketch and their mental model. This led us to classify some mental models as multiple concepts. Figure 6 shows our mental model classifications for all three datasets.

Within the same dataset, we further observed diverse groupings and orderings of the data that had personal meaning to the participant. Participants grouped the data by type, logical association, size, function, and even by the attribute "price" that the participant added based on personal experience. For the file system dataset, participants expressed a desire to reorganize the folders to homogenize file types. The junk drawer dataset was often organized by functionality, by logical associations (e.g. writing implements in the pencil pouch), or by a participant-selected category (e.g. Participant 016 organized by the "durability" of the items, recognizing disposable items might be less valuable). Participants frequently explained their reasoning for grouping the data, with less explanation for the logical and functional groupings in the junk drawer dataset and more explanation for the desire to modify the file system dataset structure, often hypothesizing about reasons for the existing file structure. No two participants grouped their junk drawer items in the same way, except for the no-grouping list order.

One caveat to the ordering: despite the different orderings we observed, most participants still drew the data in the order of reading. We observed 22 participants draw their dataset in the order in which they read the dataset, and 4 participants draw the dataset in a way that did not reflect the order presented in the dataset (2 participants were not able to easily display their sketch to the camera while drawing, so we did not consider their sessions for this code). All participants who had the file system dataset drew it in read-order. Most participants who had the junk drawer drew in read-order (7/9 participants), and most participants with the power station dataset drew in read-order (9/11 participants). Those 4 drawings that were not in read-order were drawn in order of some internal mental grouping or categorization: the 2 participants who had the junk drawer dataset discussed logically grouping the items, the 2 participants with the power station dataset drew representations for the categories of power (home, apartments, warehouse) rather than sketching a representation of the first power station and its recipients.

4.1.2 Theme: Components of mental models. We developed three codes regarding the components of mental models, in particular, regarding the presence of physical objects, ambiguity in mental models involving relations such as trees and sets, and the presence of affordances.

Physical objects were prevalent in the mental models we observed. The drawings of the objects mirrored their appearance, affordances, and orientations in the real world. This theme cross-references codes under Mental Model Content and Mental Model Elicitation, but as these physical objects were what the participant thought of as their mental model, we placed this code under Mental Model Content. The appearances of the objects were tied to memories: the cooling towers drawn by Participant 008 were based on power stations the participant had seen in their hometown and

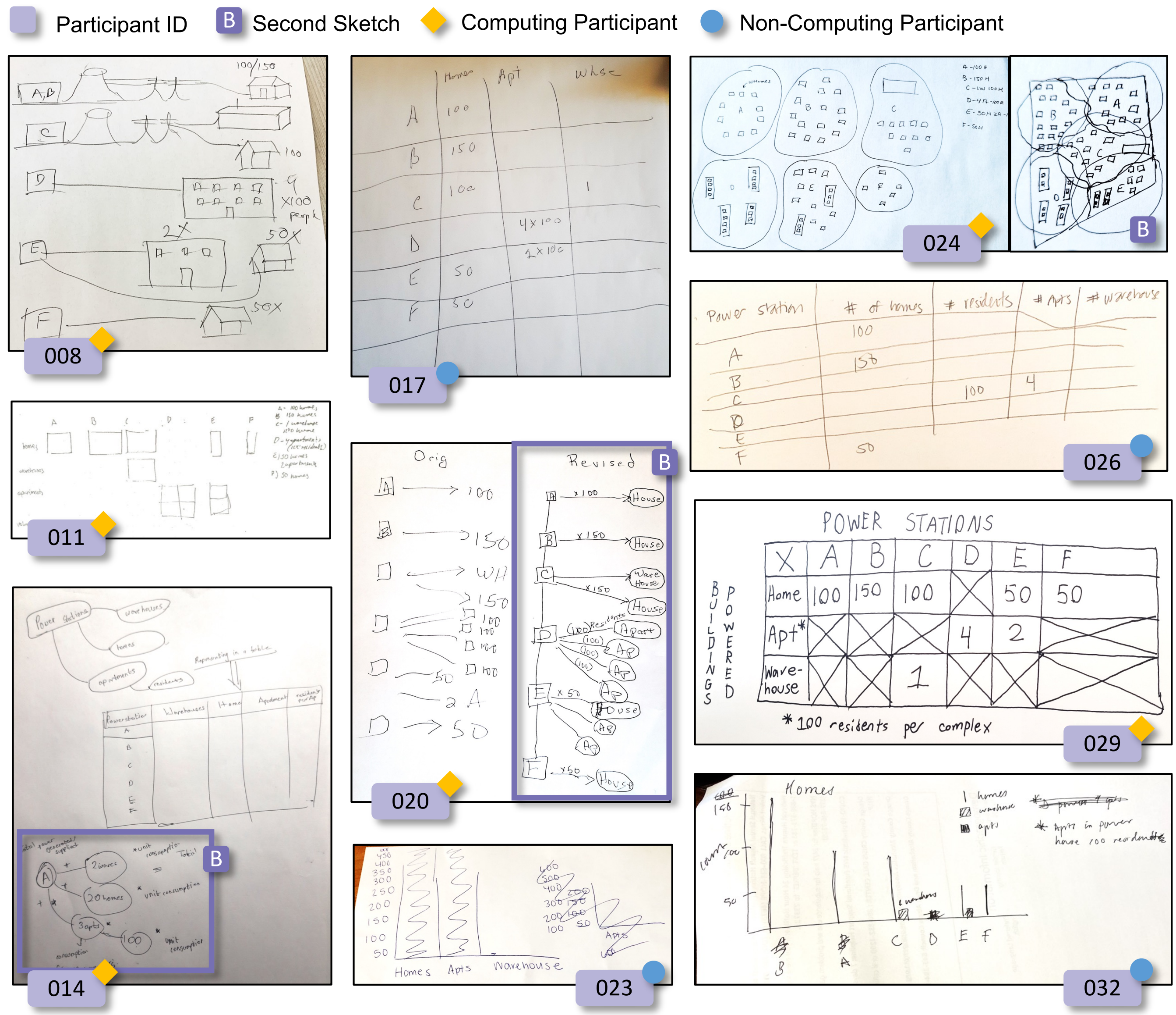


Figure 5: Sketches that participants made of the Power Station dataset. Larger versions are in the supplemental archive.

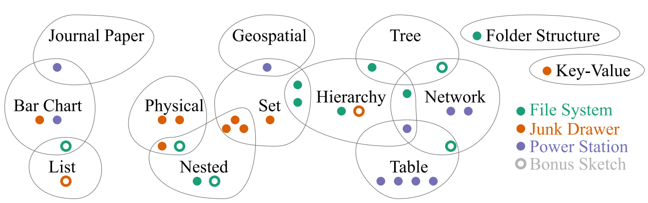


Figure 6: Our best-effort classification of mental models expressed by participants. Open circles indicate the second sketch made. Participants expressed a variety of mental models, many of which were ambiguous between multiple categories. Some mental models aligned well with data typologies, while others, like "Journal Paper", did not.

engineering textbooks, while nearly all of the junk drawer items drawn by Participant 013 had a story or memory tied to them. We note the strong semantic connection between our datasets and concrete objects may have influenced these observations and discuss this further in our Limitations section (subsection 5.2).

We found difficulty disambiguating and naming mental models that involved relationships between data items. For example, mental models similar to data abstractions typically classified as trees, hierarchies, and sets. We carefully considered language cues—terms such as "levels," "branches," "associations," "nesting," "underneath," "inside," "hierarchy," and "graph." There were no clear boundaries in how representations were used, and sometimes the terms vocalized were associated with multiple different data abstractions. For

Theme	Code	Representative example/Evidence
Diversity of mental models	Diversity of abstractions and representations (C1)	For the power station data, we saw tables (4), set/geospatial (1), bar charts (2), node-link networks (2), set (1), table and node-link (1), and a multi-figure "journal paper"-like representation (1) that included captions and text.
	Ordering diverse, personal (C2) * Caveat: participants drew in order of reading (C3)	Participant 016 (JD) organized by desired category of "durability" based on personal experience.
	Diversity of groupings (C4)	Grouped by functionality (2), participant-selected category (2), grouped only the writing implements (2), list order (i.e. no grouping) (4).
Components of mental models	Physical objects represent data (C5)	Participant 008 (PS): "I've seen a lot of power plants back home... that's why I drew the cooling towers." Participant 013 (JD): "There's a red pencil case that I had during my last year of high school and these are the pens that I have right now in college."
	Tree/Network/Set ambiguity (C6)	Participant 018 (FS) used terms "set" but also "level" and "nesting." Drew a node-link initially but said they considered a nested drawing (shown in bonus for 018).
	Mental models include affordances (C7)	Participant 022 (JD) drew a basket with a handle "so it's organized in a way and you can carry it around." 5/9 participants who had the FS dataset spoke about interactions.

Table 2: *Themes about Mental Model Content.* This table contains our themes and codes about mental model content and some representative examples for each code. The codes are labeled as "C#". The complete list of codes and all supporting evidence can be found at https://osf.io/kvnb9/ and in the supplemental material.

example, Participant 018 used the term "set" and the term "level" in describing their node-link sketch.

Some participants described interactions, or *mental affordances* [29], within their mental model. Six participants who had the file system dataset explained how they would interact with it, specifically how they would navigate it (Participants 006 and 030), or even actually drawing an inset to show this interaction (Participant 033). Even though we were discussing an abstract concept (i.e., their mental model) in a static medium (i.e., paper), participants referred to interactions with their visualization and mental affordances, or internal interactions, that they used with their mental model.

4.2 Themes about Mental Model Elicitation

We developed two themes regarding mental model elicitation, encompassing how the mental models were drawn on paper and how they were verbally described by participants. These codes solely relate to the choice of representation and encodings on the paper and what verbiage the participant used to describe their drawing. The list of codes relating to mental model elicitation can be found in Table 3 with the code label listed as "(E#)"; the complete list of corresponding codes, their definitions, and backing data can be found at https://osf.io/kvnb9/ and in the supplemental material.

4.2.1 Theme: Depictions of mental models. We formed several codes regarding how participants depicted their mental models, such as their use of text, legends, details, and abstractions, as well as where they were constrained by the sketch format.

We only noted the use of text when the participant commented on their use of text. Participants specified that they would use words to communicate with another person (Participant 019, 027), with Participant 027 noting, "When I use icons, unless it's mutually understood by both people, it might confuse; or even I might forget what the notation actually stood for... You can't go wrong with text, and it's [the file extension] not long either." To help their understanding of the file system dataset, Participant 021 decided to put "2 code [files]" since they did not know what code meant.

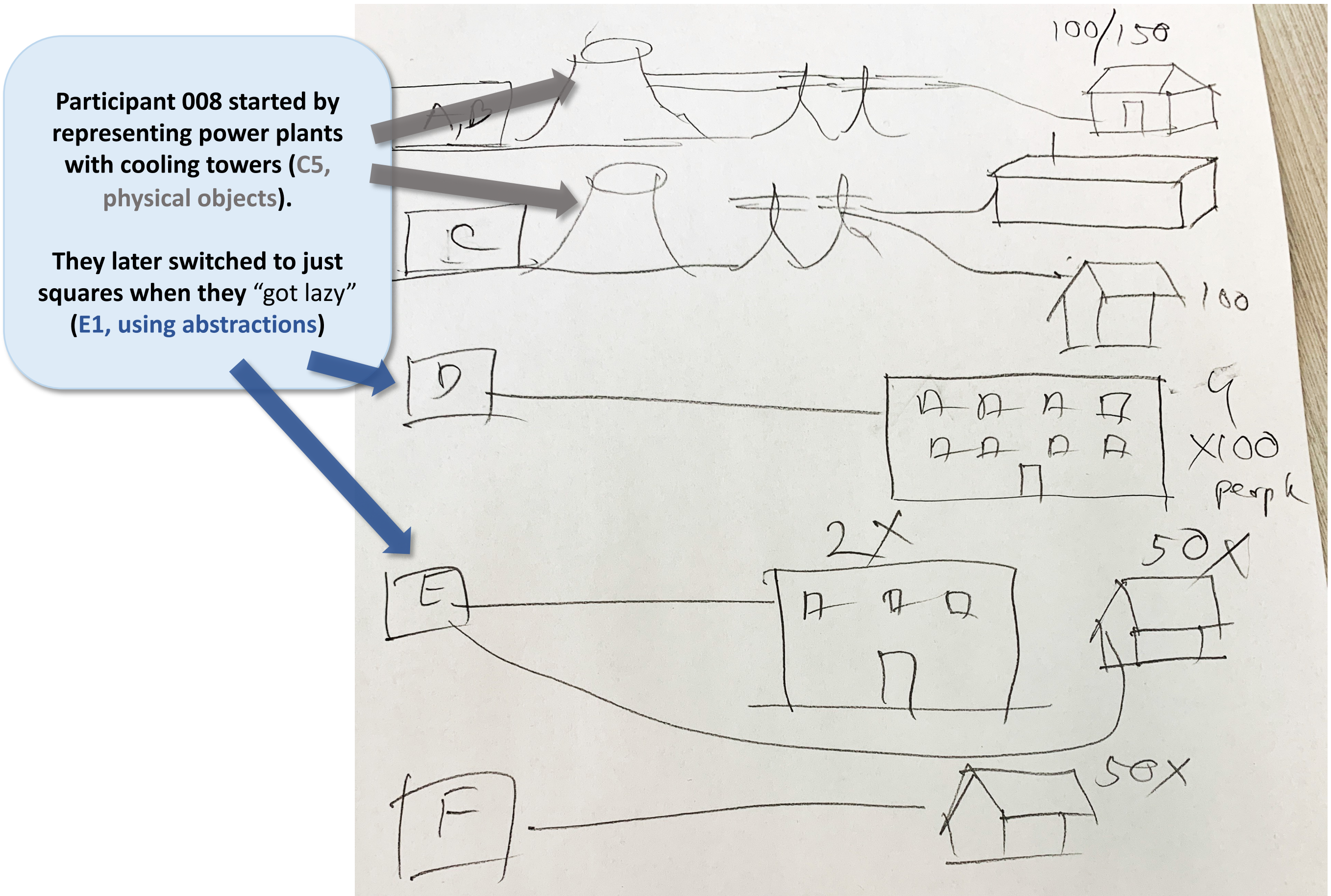


Figure 7: The sketch by Participant 008, with annotations explaining how this sketch exemplifies our codes about physical objects and about using abstractions (C5 and E1, respectively).

Other participants chose a code type for clarity—Participant 009 used ".java" so that "we can be more explicit" and Participant 030 used "common file extensions" but recognized they used a mix of "tokens" for the file types.

Similarly, the use of detail was only noted if the participant commented on adding detail. Some participants wished to add detail to distinguish between the junk drawer items (Participants 019, 031). Other participants wanted to add arrows and labels: Participant 012 added labels and arrows to suggest hypotheses about the relationships between the files, Participant 020 wanted to add weights to directed arrows for the power stations and said that the addition "would be an improvement."

Only one participant drew a legend on their sketch (Participant 032). Other participants verbally described what the icons meant, like explaining the icons for the types of files (Participant 006) or explaining that the small squares represent homes (Participant 024).

Some participants either vocalized their use of an abstract mark or switched to a more simple mark during sketching. Participant 011 drew boxes because "drawing houses would be too difficult;" Participant 014 sketched the idea of a table rather than the full one; Participant 024 said their mental model was geographical but chose to draw without geographical marks. Participant 008 started drawing buildings in 3D, but then switched to 2D icons. To show their reaction to the file system data, Participant 012 said "I marked it with a bunch of question marks to the right because I don't have any idea what [this folder] was for; it's just there."

Other participants left out details altogether. Participant 009 originally labeled the text files with "txt" but stopped labeling them because "I'm gonna be lazy." In their bonus drawing, Participant 030 added an ellipsis for the "OBJECT_TYPE" attribute after writing one complete data table entry since the rows beneath were all the same type. Participant 023 used a squiggly line instead of a rectangle for the bar graph; the lack of detail is possibly related to their level of math literacy (code F7).

Some participants ran out of space while drawing and verbally noted it. To adapt, some added their marks to a different location (Participant 007 drew the sharpies outside of the pencil pouch, and Participant 016 skipped back to the left side of their x-axis since they ran out of space going left to right). Others continued with the existing drawing and expressed regret (Participant 015 said "it's hard to draw this. I should've brought a pencil." Participant 022 wished they "made the basket a little bigger.").

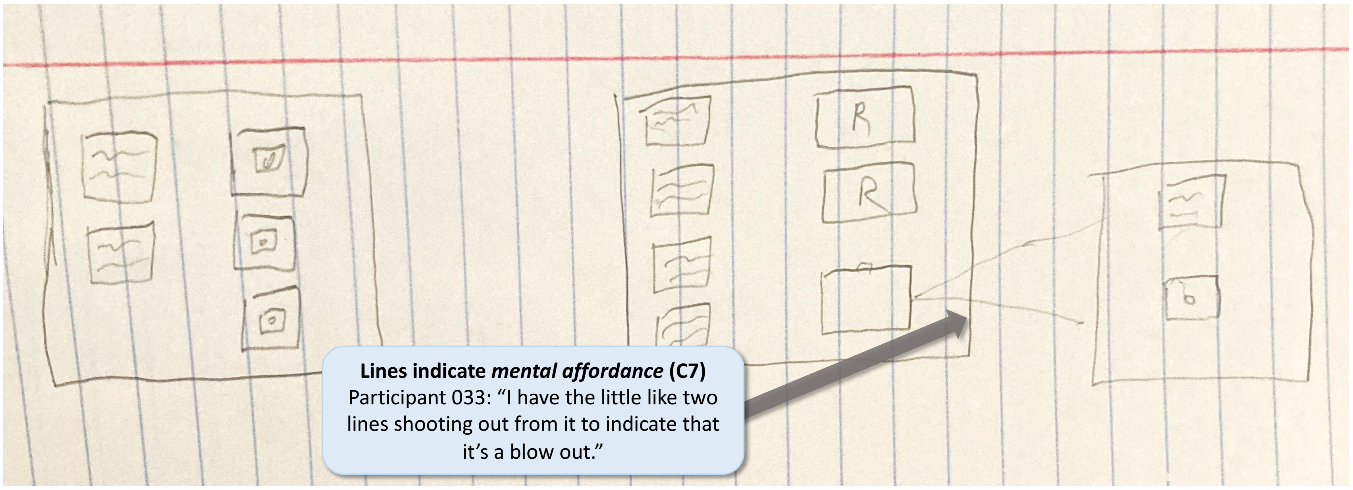


Figure 8: The sketch by Participant 033 of the file system dataset. The participant drew lines to indicate interaction in their mental model (code C7).

Some participants were constrained by the encoding schema they chose, rather than space. This happened to Participants 030 and 033 when they encountered the code files in the file system dataset since they did not anticipate the file type and had not prepared a way to encode it.

4.2.2 Theme: Communication with others. We observed participants making conscious choices about how they represented the data when communicating with others (code E6[4]), but varied in their level of detail and abstraction, and use of terms.

Some participants sketched at a high-level abstraction but added detail, like colors (Participant 009) or item detail (Participant 031), to clarify for others. One participant re-oriented their tree from top-down to left-right and attempted to add interaction indications (Participant 021). Participant 031 recognized which aspects of their drawings could be confusing and said, "Between the pens and the pencils and the Sharpies, you can't really tell what they are. If I were to give it to somebody, they probably wouldn't be able to tell—to differentiate between those groups...I probably should have written 'envelopes' on them or some type of—you know, if somebody were to look at this, I don't think they would know what I drew."

Two participants said they would choose a different data abstraction, depending on the audience's "quantitative literacy" (Participant 026), or they would find a "better way" to represent the data, possibly by adding a table or other figures and captions (Participant 032).

When communicating with the facilitator, participants added annotations when discussing their sketches. Participant 006 added encompassing circles around the top-level folder of their file tree and the children under folder 1. To explain how they would solve for the total power generated, Participant 014 added a graph with root node 'A' at the bottom of the page.

[4]The complete list of codes can be found at https://osf.io/kvnb9/ and in the supplementary material.

Sometimes participants used terminology in conflict with visualization community concepts for dataset abstractions. One participant drew a table, even though their description and interaction with the dataset focused more on data item relations (Participant 014), reinforcing the code about ambiguity when using trees/networks/sets from Table 2. In particular, they used the terms "endpoint", "layers", and "map" and relied on their other drawing of a node-link graph to augment their description of how they would solve for the amount of power produced. When referring only to the data (no longer problem-solving), they said what stuck out to them was the "layers" and "sublayers" in the dataset.

One participant used set-like terminology to describe their node-link diagrams. Participant 006 described, "In my head, I'm oddly enough in the folder that those two folders are within," and often used "within" and "in" to describe the location.

In response to the interview prompt "Describe your sketch", we observed a range in the level of descriptive detail. We categorized the levels of detail in the verbal descriptions : (1) individual data points, (2) individual icons, (3) relations of icons or positions of icons, and (4) data abstraction. By "individual data points," we mean the participant nearly restated the dataset and did not describe the drawing. Five participants stuck to this individual data point level of detail (Participants 008, 014, 018, 027, 032).

The next level of detail, "individual icons," means the participant gave visual descriptions of the icons or marks used in the dataset. These verbal descriptions ranged in detail, with six participants matching this level (Participants 010, 011, 013, 019, 030, 031). Some participants named every type of mark, while others got distracted midway.

The third level of detail, "relations of icons or positions of icons," means the participant stated where the icon was on the page or in relation to other icons (e.g. "in a folder," "next to the files," "roll of stamps down there and tacks to the right"). Nine participants referred to relation/positioning when describing their sketch (Participants 006, 007, 012, 015, 021, 022, 024, 025, 033).

Theme	Code	Representative example/Evidence
Depictions of mental models	Used text to clarify (E3)	Participant 009 (FS) calls the files ".java" so that "we can be more explicit."
	Legends/verbal legends (E5)	Participant 006 (FS) verbally explained the icons used for file types. Participant 032 (PS) made a legend.
	Used abstractions in depiction (sometimes laziness, sometimes deliberate) (E1)	Participant 008 (PS): "...just repeating the picture of the power plant but then I got lazy and then just drew a square for the power plants."
	Constrained by sketches (E2)	Participant 007 (JD): added Sharpies outside of pencil pouch because "I didn't make the pouch big enough." Participant 016 (JD): "I had to skip back here [to the left] to fill out the rest of the space."
	Added details to clarify (E4)	Participant 020 (PS): suggests adding weights to directed arrows, "it would be an improvement."
Communication with others	Conflict of the terminology used (E7)	Participant 006 (FS) drew something closer to a node-link diagram but their description was more set-like ("levels", "within", "in the folder").
	Range of description detail from literal to abstract/overview (E8)	5 participants did not describe the drawing and instead restated the dataset (4/5 were CS-related participants). 15 participants gave a visual description of the icons and marks (13/15 were CS-related participants). 8 participants named a data abstraction or data chart type (4/8 were CS-related participants).
	Changes to depiction for communication (E6)	Participant 021 (FS) said they would draw the tree left-to-right to better communicate with others and attempted to add interaction indications.

Table 3: *Themes about Mental Model Elicitation.* This table contains our themes and codes about mental model elicitation, as well as representative examples for each code. The codes are labeled as "E#". The complete list of codes and all supporting evidence can be found at https://osf.io/kvnb9/ and in the supplemental material.

Eight participants named a data abstraction (Participants 009, 016, 017, 020, 023, 026, 028, 029). However, their name did not always match the visualization community's name for data abstraction or data representation that they used. After naming the data abstraction or representation, they went on to describe the icons or markings they used, the second level of detail.

4.3 Themes about Mental Model Formation

We describe two themes describing our observations regarding how participants came to their mental model, based on their descriptions. These codes do not attempt to explain how the mental models are formed; instead, they are observations of how a mental model develops in a data- and visualization-related setting. The list of codes relating to mental model formation can be found in Table 4 with the code label listed as "(F#)"; the complete list of corresponding codes, definitions, and supporting data can be found at https://osf.io/kvnb9/ and in the supplemental material.

4.3.1 Mental model formation process. Participants suggested their mental models form quickly, with little change, though we observed they became more detailed during the session. Our first interview question asked about initial impressions and reactions. Participants' responses already expressed ideas for data abstractions and what they would draw, including thoughts about ordering and purpose-seeking. Participant 024 said, "My gut reaction was like an image of—I dunno if you know cell-free MIMO graphs..." and described how they would use the idea to draw the power stations. Other participants immediately tried to find the purpose or context of

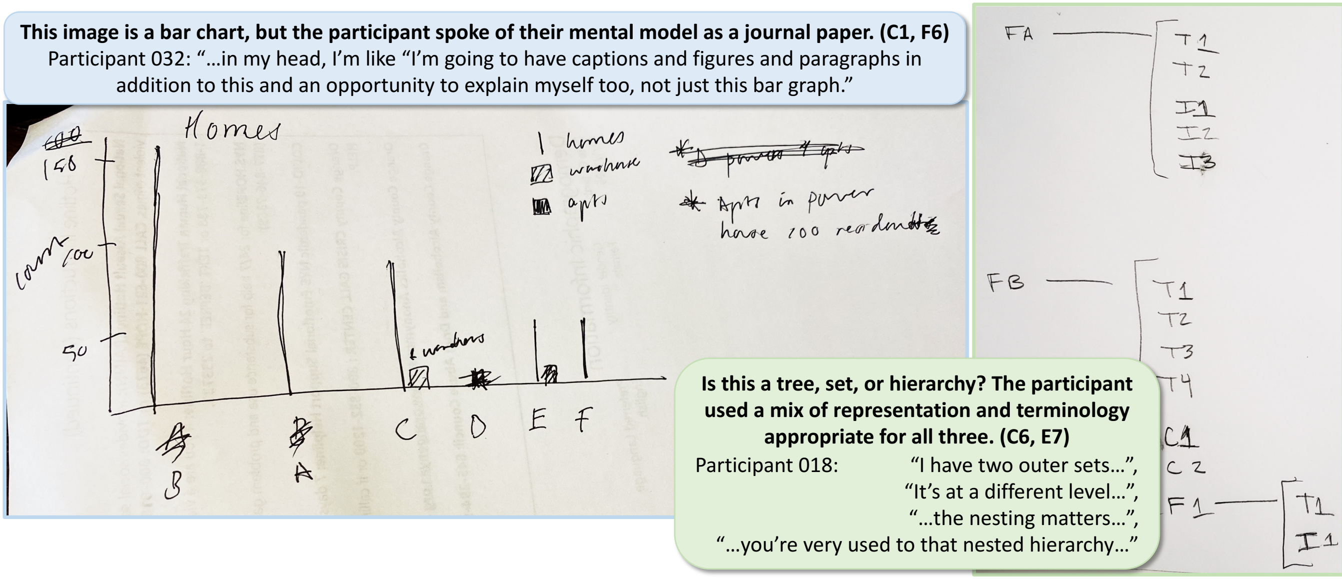


Figure 9: The sketches by Participants 018 and 032. Annotations explain how the verbal description from Participant 032 expanded their representation from solely the bar chart that they sketched to an elaborate multi-figure "journal paper" (C1, F6). The annotation to Participant 018's sketch includes quotes that highlight the mix of terminology that the participant used (codes C6, E7).

the dataset, such as supposing that the junk drawer dataset "it's like a handy toolbox for a home" (Participant 013). Participants also expressed a desire to organize the data by categorizing or by finding a more "efficient" way.

We later asked participants if their mental models had changed. About 60% (17/28) of participants said it did not. Several participants mentioned aspects of their mental model that were obvious, such as "it's obviously a folder structure" (Participant 009) or "I feel like in my head it's the simplest conclusion" (Participant 032). For the participants whose mental models had significantly changed, they often cited trying to find a "better" or "best" way to display the data (Participants 010, 016, 023, and 032).

We observed some participants vocalizing their revisions, adding details to their sketches and mental models as they drew, or adding clarifying information afterward when describing how their mental model changed throughout the interview. Participants 030 and 033 chose icons to represent different file types during the drawing. Participant 030 explained when asked about mental model changes that this was a "minor hiccup... trying to choose tokens to represent the category of files: text, image, code." While explaining their sketch, Participant 006 added a root to their file system.

4.3.2 *Influences on the mental model formation.*

When asked where their idea for their mental model originated, many participants explained where they had seen something similar or hypothesized the source of their inspiration. All but one participant who drew the file system dataset cited an operating system or software for either the structure or icons, including Participant 015 who drew the nested manila folders similar to a Windows icon. Despite the common inspiration, there were several different types of depictions. Across all datasets, participants cited a real-life example (e.g. a drawer in their home, power plants, cell-free MIMO) or their work as the reason for the choice of data abstraction.

Participants with less math literacy had a limited representation and mental model. Two participants expressed math hesitancy, 023, "Yeah, that's a lot of math. I'm not - I've only taken high school math", and 031 "I am not very good in math." Both were part of our non-computing population of participants. Participant 023 had difficulty finding a concise way to express the power station dataset. Their sketch was missing a data dimension (the power stations A-F). Participant 031 had the junk drawer data and drew the physical objects as given. While the evidence for this code is lighter as we suspect our IRB-approved advertisements may have dissuaded people not comfortable with data (see Limitations, subsection 5.2), we noted this code for future investigation.

Since we only provided a dataset to the participants, participants often wanted to add additional data or information to the dataset or they added additional context or a hypothetical source for the dataset. We classify such behavior as *purpose-seeking* behavior, or an attempt to connect with the dataset in an imagined real-life setting. Six participants suggested other data item attributes they would want, and four participants added relationships between items. This was most prevalent in the power stations dataset with requests for the number of people per house, power requirements per building, and one instance of geospatial coordinates. In the file system, additional data requests manifested as the wish for file sizes, code file types, and suggested relationships beyond the folder structure. In the junk drawer, price, durability, and nesting structures were suggested or imposed.

Participants also suggested a task associated with the data. With the power stations, three participants wanted to solve which station

Theme	Code	Representative example/Evidence
Mental model formation process	Immediate mental model formation (F1)	Participant 024 (PS): in response to the first question, "My gut reaction was like an image of—I dunno if you know cell-free MIMO graphs..."
	Mental model did not vary much or at all for 17/28 participants (F1) * Caveat: changed significantly for 4 participants due to changes in data abstraction representation (F1)	Participant 018 (FS): "No, I mean I kind of saw it for how I was gonna do it right away and stuck with that." Participant 016 (JD) considered "maybe there's a better way to accurately display—because it sounds like this is someone inventorying the items... I wish there was a way that I could highlight that, or draw attention to that this [the stamp roll] is probably more important than rubber bands and tacks."
	Mental models became more detailed (F2)	Participant 013 (JD): "In the beginning, I was just thinking about the basket and then I started to remember how things were more clearly, so I started drawing slightly more elaborately and really thinking about what I wanted to draw."
Influences on the mental model formation	Explicit mental model origins (F3)	Participant 006 (FS): "The [Windows] file system, the file structure, has definitely left a mark on me."
	Purpose-seeking: Participants added or assumed tasks (F5)	Participant 019 (JD) associated the dataset with cleaning or organizing their desk. Participants 011, 014, 032 (PS) all sought to discover ultima (e.g. the maximum power produced).
	Purpose-seeking: Participants suggested data sources (F6)	Participant 017: "I dunno, it's a power station, it's probably a municipal guide or a power company's guide to how to distribute power."
	Purpose-seeking: Participants desired to add data/information (F4)	Add data attribute: Participant 029 (PS) wanted to add people per house.
		Add relationships: Participant 013 (JD) wanted to add relationships between items.
		Add naming schema: Participant 018 (FS) wanted to add folder names.
	Lower math literacy works against the mental model (F7)	Participant 023 had difficulty with multidimensional aspect of power station data, dropped the power stations' label dimension (i.e. A–F labels).

Table 4: *Themes about Mental Model Formation.* This table contains our themes and codes about mental model formation, as well as some representative examples for each code. The codes are labeled as "F#". The complete list of codes and all supporting evidence can be found at https://osf.io/kvnb9/ and in the supplemental material.

produced the most or least energy. Other assumed tasks included taking inventory (Participant 016), cleaning (Participant 019), carrying (Participant 022), determining affected people or buildings (Participant 029), solving a math problem (Participants 014, 031) or presenting in a scientific journal (Participant 032).

Half of the participants (14/28) contextualized their dataset with a suggested source or generator of the data. The junk drawer dataset was suggested to be a "stationary drawer" (1 participant), "office supplies" (4 participants), and "an electrician's toolbox" (1 participant). Participants who received the file system dataset imagined a new program or provided reasons why a program would be organized in the given manner. Participants with the power station dataset supposed that the data was for a "residential part of the city" (2 participants) or a "municipal or power company guide" (1 participant).

4.4 Beliefs about Data

In addition to our main themes regarding the content, elicitation, and formation of mental models, we developed a secondary theme describing our observations about participants' beliefs about data and data analysis.

The idea that data relates to tables is prevalent (code D1). Four participants who did not use tables mentioned expectations involving tables. The facilitator asked Participant 007 if they were surprised that the junk drawer dataset was given as a dataset and they replied, "I was expecting something more structured, maybe like a table or something. I guess I was expecting a table. Because that's the most common form of storing data, like a spreadsheet or table, something like that." Participant 019 also received the junk drawer dataset and said, "I didn't work with Excel very much, so I don't think of datasets, but when I heard the term 'datasets', I really thought about the analysts I worked with and Excel data, and I thought of big datasets and grouping people by demographics, that kind of thing. I refer to datasets and I was familiar with them but I never thought of stuff like this as a dataset."

Two other participants mentioned using tables to organize the data via relational tables (Participant 030, bonus sketch) or to answer questions about the data (Participant 014).

There was hesitation regarding whether the given dataset was truly data (code D2). Initial impressions of the dataset included impressions on the term "data" itself. Participant 007 concluded, "I guess this [the junk drawer dataset] is a valid dataset, it's got objects and quantities for those objects." Participant 010 associated the word "data" with relating to computers, and drew a Python-like dictionary of the junk drawer dataset.

When asked about how often they visualize data, Participant 012 considered that it "depends on what you consider data." Participant 019 responded to this question by relating data analysis to grouping people by demographics: "I thought of big datasets and grouping people by demographics, that kind of thing."

Two participants made us question the distinction between the dataset and the data. Participant 014 distinguished the dataset and the data within it as different ideas: "like the emphasis should be on the dataset, not the data containing [sic] in it, right?" Participant 027 didn't think of the items in the file system dataset as the data; instead, they assumed the data was inside those items (files) and not explicitly given.

4.5 Differences between computing and non-computing participants

We recognize that the majority of our participants had a computing background. While not part of our original experiment design, we revisited our codes to check if any were heavily computing-biased in their evidence. Of our codes, the following codes had solely computing-based evidence:

- Node-link sketched representations (all 9 node-link sketches were done by computing participants– discussed in subsection 4.6),
- E1 (from 5 computing-related participants and 0 non-computing-related participants): using abstractions in the depiction,
- F5 (4 computing, 0 non-computing): purpose-seeking by adding relationships between items, and
- C5 (3 computing, 0 non-computing): using physical objects that were cited from prior experience to represent data.

The code about math literacy (F7) was from 2 non-computing participants. All other codes contained supporting evidence from both non-computing and computing participants. See subsection 5.2 for more discussion on our participants' relationships to data.

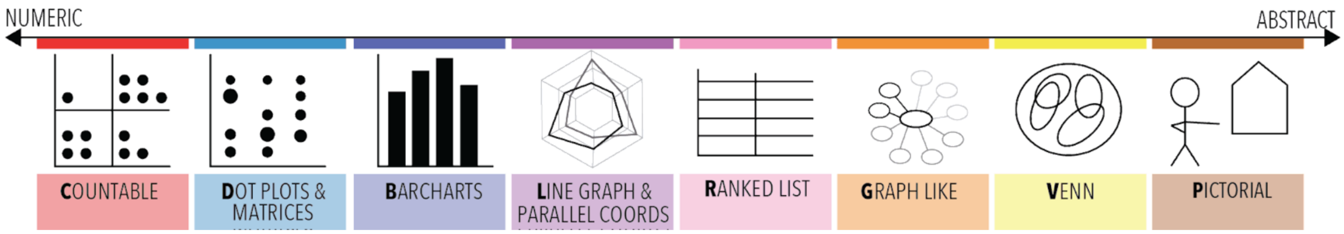
4.6 Comparison to prior work on sketching and data reports

Given the similarities between our study and Walny et al.'s sketching study [43], we examined our sketches on their numeracy-to-abstractness continuum. Figure 10 shows our best-effort categorization, given that we used a different dataset and therefore saw different data representations than Walny et al. (their dataset was a table of behaviors in social scenarios). The session number for each sketch is placed within the data representation category along the continuum. Possibly due to the nature of our experimental datasets and population, we did not observe any sketches in the "line graphs and parallel coordinates" category, nor any in the "ranked lists" category. Of the 28 sketches (not including bonus sketches), 8 leaned toward the numeric side of the continuum and 20 leaned toward the abstract side. The category with the largest number of sketches for our participants is node-link representations, whereas the most common representation for Walny et al. was bar charts. This effect is likely due to the qualities of the datasets given to the participants.

Within each category of abstraction, there was a mix of computing and non-computing participants represented, except for the "node-link/node-link hierarchy" category (9 computing participants, 0 non-computing), "bar charts" (0 computing participants, 3 non-computing participants), and "table with symbols" (1 computing participant, 0 non-computing participants). The prevalence of computing participants in the node-link category can partly be explained by the dataset they received: 6 of the 9 participants coincidentally received the file system dataset, and the other 3 received the power station dataset. These participants may be more familiar with node-link diagrams, especially related representations commonly used in file systems.

Walny et al. also examined the participants' written reflections about what they had discovered in their datasets, which the authors call *data reports*, and the authors developed a *data reports spectrum*, which placed responses that contained direct readings of individual data values at one end and higher-level conjectures and hypotheses at the opposite end. A major finding from intersecting their participants' sketches with the data reports was how "the participants who submitted the most abstract sketches were among the participants whose data reports tended to be in categories E3 (including extrinsic information) and F (statements with analytic potential)."

To test this finding in our work, one author reviewed the interview transcripts for such statements. The author chose to exclude statements in category E3 because the interview question, "How did you come up with this idea? Have you seen something like this before or have you worked with a dataset like this before?" prompts the participant to relate the data to external information, which would not be an organic source for such statements. Therefore, only F statements, statements that offer fledgling hypotheses or



Numerica/Countable/Table	Table with symbols	Bar charts	Line Graph & Parallel Coords	Ranked List	Node link / NL hierarchy	Containment / Sets/ Venn/ hierarchy via enclosure	Pictorial / Physical (with and without nesting)
10	11	16			6	9	7
17		23			8	24	13
26		32			12	25	15
29					14	28	19
					18	33	22
					20		31
					21		
					27		
					30		

Figure 10: Our categorization of data representations from the participants' sketches, placed along the numeracy-to-abstractness continuum of Walny et al. [43]. Yellow shading indicates the participant has a computing background, blue shading indicates a non-computing background.

conjectures about reasons for values, were included. For example, an F statement from Participant 008 is, "I mean, I don't know what their fuels are, I'm assuming they're maybe coal-powered power plants." Another example is from Participant 009: "I don't know how big it [the file] is, at the end of the day, right, so the text file could be bigger... could be super big, could be smaller."

We found such F statements (statements with analytic potential) from 14 participants. Of those 14 participants, 11 of the participants' sketches fell on the more abstract, pictorial side of the continuum, supporting Walny et al.'s finding that participants with more abstract sketches included statements that were more analytic on the spectrum of data report statements. Our results thus show reasonable agreement with those of Walny et al., given the differences in study design.

4.7 Comparison to prior work on mental models

Our findings bear similarities to the Paris map study of Milgram and Jodelet [27]. In that study, an individual's mental model of their hometown included important locations and roads connecting them, and their personal background influenced the order and size they drew these locations. In our study with datasets, we observed specific data points had personal connections, like the locations in Milgram and Jodelet, but also outliers were of interest. We further observed that people drew from their personal knowledge to communicate "rules" of the dataset, which they utilized to determine outliers or suggest new data. For example, some of the File System participants suggested alternate organizations based on file types.

Other prior work evaluated the accuracy of people's mental models when learning new phenomena via text and visual elements, only text, or only visual elements [4, 17, 35]. However, most datasets start as text-only collections and are frequently not human-readable, so we did not present visuals and provided only text. Datasets may not necessarily have relationships explicit, but rather are for the individual to determine, leading to more ambiguity as we observed.

Across studies, participants choose to emphasize elements and connections in the dataset that they have a special connection to, that reflect aspects of their demographics and background, and draw inspiration for their drawings from other maps and data visualizations that they had seen. Like eliciting the Parisian maps and our dataset sketches, people have a wealth of knowledge and expertise that they may not realize. They have an intuition about the dataset and have an idea of what "makes sense" in the data, even if some of these ideas may be inaccurate. This knowledge may be subconscious yet useful for visualization designers, especially the unwritten "rules" of a dataset. Further work in this area can help visualization designers work more efficiently and produce more useful visualizations for their users.

5 DISCUSSION

We discuss our findings regarding the research questions we initially posed. We then discuss the limitations of the study. Finally, we discuss the implications of our findings for visualization design.

5.1 Revisiting our Research Questions

Our study was inspired by the research questions listed in section 1. We revisit those questions and discuss our discoveries towards them.

What factors influence people's initial mental models of data? We observed that participants quickly came to their mental models, with several (11/28) expressing how they would represent their mental models right after reading the dataset. Most (17/28) remained consistent in the high-level data types they sketched and discussed, though details regarding the particular encodings required further consideration. This consistency leads us to believe that the sketches and descriptions were close representations of the participant's mental model in many cases.

In a few cases, however, a participant realized their mental model required revision. While drawing they realized their approach did not permit them to add all of the data from the paragraph (code E2). Some also expressed the desire to approach it differently after they had completed their first sketch.

When discussing how they arrived at their mental model, many participants (16/28) related their sketch to something they had seen before, some directly applicable to the data, such as operating system file browsers for the file system (7 computing participants, 1 non-computing participant), and some less direct. Two participants discussed recent sources of inspiration such as coursework. We hypothesize that our participants were able to connect their mental models to existing visualizations and data representations because of the accessibility of the datasets and the data-literate background of our participants. Further study is needed on less-accessible datasets and people with low data literacy.

Ideas about data organization and structure also influenced how our participants sketched the datasets. Three participants suggested expectations regarding how "data" should be organized. Four participants suggested their model was obvious.

Our participants' mental models were also influenced by inferred purposes, which had implications for the corresponding data abstractions. Participants added purpose and context to the dataset by suggesting a generator for the dataset, problems to solve with the data, or insights they wanted to glean from the data (see codes F4, F5, F6). The inferred purposes led them to further suggest more data or attributes that could help achieve these imagined purposes and create more elaborate mental models than ones based strictly on the data alone.

What encodings and visualizations do people commonly use to communicate their mental model? We observed a variety of encodings and visual representations. Tables, node-link diagrams, containment/enclosure, indented nesting, icons/physical depictions, proximity for grouping, and bar charts were each seen multiple times (see Figure 10). Beyond bar charts, there were no common statistical charts. This may be an artifact of our dataset design, which was designed to enable the use of one of several data abstractions from data typologies.

The least diverse set came from the file system prompt, where node-link diagrams and enclosure were common approaches, though one participant drew a physical depiction inspired by Windows icons (Participant 015).

Participants used a range of specificity and generality in their marks, for different reasons. Many participants used abstract marks or elided details, some from the beginning of their sketch and some changing to more abstract marks along the way for efficiency during drawing (5 participants remarked when they deliberately made this choice). We saw text used in tables, bar charts, file system icons, and labels, and exclusively text in three of the junk drawer depictions rather than physical icons. Participants explained that their use of text was to clarify (when representations were unclear or they wanted to specify) or to simplify their representations (rather than drawing items). These explanations may indicate that the participants were unconsciously aware of their communication efforts: they used a shorthand version of the representation and skipped drawing details when they felt a viewer could understand what they meant, yet they added text to clarify when they felt a viewer may misunderstand a sketched representation.

How do they describe how they think about the data? How do people describe their sketches? We observed a spectrum in the level of detail participants expressed when describing their sketches. The level of detail ranged from essentially repeating the dataset, to describing a verbal legend of the marks, to naming the data abstraction or representation. This differing specificity has implications for how people communicate about datasets and how they emphasize important aspects of datasets or visualizations. Two people with the same dataset may value different levels of granularity in the data (e.g. one may care about individual data point values while another may only want to see regression lines). A visualization designer must be sensitive to both perspectives and weigh how or if they want to encode the data to support these views.

Participants made deliberate presentation choices with their sketches while presenting their sketches to the facilitator. Some participants added detail either to clarify the depiction for the facilitator or to emphasize parts of the sketch to the facilitator, e.g., circling the part they were explaining. Some participants also suggested changes they would make for another audience, such as re-orienting the sketch to make it easier to read, changing the data abstraction entirely, or adding more explanation or a legend. These changes in presentation pose interesting questions on the transferability of mental models of datasets and how well a person can communicate their mental model to one another to create a shared understanding of the data.

In the post-study demographic questions, participants described how often they visualize data by citing software (e.g. Excel, Tableau, D3), types of charts (e.g. box plots, radar graph, bar graph), but also describing mental imagery (internal visualization) and plain visual representations (e.g. a list, note-taking). The participants who mentioned software described interacting and analyzing data through the use of spreadsheets and statistical software. The participants who interpreted "visualize" to mean "imagine internally" gave more descriptive answers about how they mentally interact with and think about data. These participants described estimations, relationships, and trends in data, and used more inward vocabulary ("organize in my brain", "imagine in my mind").

How difficult is it for people to sketch and/or describe their mental model? How difficult is it for us to understand? Nearly all participants began sketching right after our discussion, with some attempting to draw before the initial discussion question. Participants paused during their sketching when they ran into a space constraint, schema constraint, or an outlier (e.g., the warehouse in the power station dataset), but otherwise, drawing was uninhibited. Occasionally they paused to evaluate their current sketch and then modified it. Often participants paused when asked what sticks out in the dataset, suggesting they were thinking and evaluating the dataset against previous experience or searching for outliers. The participants who quickly expressed an answer to what sticks out had typically mentioned the phenomenon earlier (e.g., how to draw the code files, the structure of the file system, the warehouse or apartments in the power station dataset).

When asked to describe their sketch, depending on the level of detail given, the facilitator asked questions to try to better understand their mental model and to get verbal descriptions of the visual phenomena. Some participants willingly launched into more detail about their sketches with minimal prompting. The participants that were less willing to discuss details of their sketch may have felt their sketch was self-explanatory—a trend with some of the sketches of tables and pictorial sketches of the junk drawer. However, those participants did have ideas about the data itself, hypothesizing about the source or asking for different attributes or more data.

We encountered difficulties in classifying the mental models. There were ambiguities in the terms used. Understanding the relations between data items and the structures that the participants imagined was especially difficult due to the ambiguity of terms. Another ambiguity was that participants suggested their model was obvious without a line of reasoning. Some used language that read the data back rather than describing the data in a new way, whether it was the individual items or more holistic descriptions like "a folder structure."

While we are not confident we fully understood any participant's mental model, or we ever could, the combination of sketching and semi-structured interview did help us gain a significant understanding in a relatively small amount of time. Perhaps a second step where the facilitator and participant draw a second dataset independently, with the facilitator trying to mimic their understanding of the participant's mental model, may serve as validation of our understanding.

5.2 Limitations

We discuss the limitations of this study concerning the participant population, the dataset prompts, and the study procedure.

5.2.1 Limitations in participant population. Our participant population was likely biased toward more data-literate people. This is probable for three reasons: (1) all sessions were organized and conducted online over Zoom; (2) the header of the advertisement for the study was "How do you imagine a dataset?", which may have attracted people interested in data; (3) due to the ongoing pandemic, we did not recruit participants in person in common high-traffic areas, aside from a trio of flyers posted at a local YMCA.

In particular, we had a high proportion of participants with computing-related occupations (16 students, 3 professional developers, and 1 manager), which may have influenced the breadth of data abstractions we identified. Computing participants were the sole users of node-link diagrams, though this is in part due to their high concentration among File System participants. However, even with this almost homogeneous set of participants, we observed diversity in their representations, data abstractions, and ways of speaking about the File System dataset. This diversity could be magnified with a more diverse group. Our analysis of data abstractions used across all participants (code C1, figures 3, 4, 5) showed that other than the node-link diagrams, visual forms used by several participants were mixed between computing and not-computing participants.

We also collected the age of our participants to see if there was a relationship between the participant's mental model of the file system dataset and their age. The younger participants tended to be computer science students or similar so they were familiar with file systems. Thus, we saw no relationship between age and understanding of file systems.

5.2.2 Limitations in the study dataset prompts. The datasets we designed do not represent the full spectrum of data we see across the field of visualization. For example, continuous values are not directly represented in these datasets, which focus more on counts and relations. Still, one participant considered geographical location as a missing column in the Power Station dataset and showed how they would include it in their thinking.

All three datasets had data items that have relations to concrete objects (e.g., files and folders, office items, buildings). This was done to make the data more accessible to a wide audience. More abstract items are not handled in this study. We did not include the names of these datasets when sharing them with participants, to avoid further bias. However, the strong semantic meanings of the data items may have influenced our findings in ways that we would not observe with more abstract data.

Regarding the semantic forms associated with the data, some of the authors initially had strong semantic notions, like "a file system is obviously a tree," but informal discussions revealed that these notions are far from universal, which was an impetus for the research questions and later the dataset prompts.

We decided against providing purpose, a context of use, or tasks with the datasets. Our rationale was that early in the design process, these tasks and data are often in flux. However, there is typically *some* notion of purpose for a visualization (e.g., to analyze, compare, or predict) that will influence tasks. Data abstractions are often task-sensitive, so by omitting purpose or a set of tasks, we may have observed a more diverse set of data abstractions than in a purpose-influenced study. We hope purpose and task influence can be further explored from this study's baseline.

Our dataset prompts are small in the number of data items. We chose the small size so participants could consider and sketch the full dataset. When communicating with visualization designers, collaborators often start by offering a toy dataset to aid their explanation, which is closer to the dataset size we use. Further research is needed to understand the strategies people have in forming a mental model of bigger datasets. Thus, our study does not answer questions as to how participants' mental models might change between a toy and a full-size dataset, how they communicate datasets too large to draw, and how they might choose to describe and represent the data, for example in terms of overviews, details, or aggregations.

In addition, the datasets and design of this study do not cover uncertainty in data, dynamic data, or data that necessarily forces multiple abstractions. Therefore, we do not report on these cases but note that even in our simple case, we observed diversity in data abstractions and difficulty in describing relations, thus we would expect more pronounced indications of these phenomena in more complicated data.

We presented the datasets as a text paragraph to avoid adherence to a given data abstraction or representation, such as a table, observed in prior studies [1, 42]. However, we did observe a tendency to observe the listed order in the paragraph rather than re-order the data among some participants, especially in the Junk Drawer dataset. This could indicate the participants were basing their model on the implicit list. This was not universal though, as other participants re-organized the same dataset.

On a more specific note, there was some confusion with the "long plastic basket" in the junk drawer dataset. We envisioned this to be a holder for envelopes or writing implements, but we received many different interpretations that revealed the ambiguity of the word "long" (Participants 013, 019, 022).

5.2.3 Limitations in the study procedure. Our aim with this study was to capture the initial mental models of people regarding data form and abstractions, before being presented with one. We acknowledge that mental models are ever-evolving and that mental model elicitation is difficult [21]. We chose a direct elicitation technique via interviews and drawing as it is as effective as others. Like other mental model elicitation methods, it can only provide a representation of the mental model and is dependent on the participant's drawing and verbal descriptive abilities, as well as the skill of the interviewer and their ability to build trust for productive communication. Our choice of paragraph representation, the initial gut reaction question, and the interview question regarding whether participants felt their mental model had changed were all designed to help assess whether we were achieving our target and to provide more data about early changes in the data mental model.

Through our semi-structured interview discussions, we found a mix of participants, some who claimed their mental model was unchanged and others where the changes were apparent from their words.

Some participants expressed confusion in response to the initial question after they read the dataset: "What was your gut reaction or intuition about the dataset?" While we intended to capture open-ended responses, participants sometimes asked us to clarify what we meant by "gut reaction." This confusion may have influenced their responses or the responses of others who were confused but did not vocalize their confusion.

5.3 Implications for the Data Visualization Design Process

We share implications about mental models and participants' thoughts on data for the visualization community.

5.3.1 Personal Mental Models.

Beware of data abstraction elephants. We observed diverse data types and visual representations arising from our study datasets (code C1). Even in the file system case, where most participants cited a similar source of inspiration, participants sketched a variety of concepts and imparted differing grouping biases. Across all datasets, some participants were influenced by recent events in their life (e.g., labmate's talk) or by expectations of what 'data' in general should be, rather than the dataset at hand (code F3, code D2). These realities present hazards in choosing effective data abstractions. We suggest that designers sample multiple target users, potentially multiple times, so that our interpretation of our users' mental models can solidify. Once solid, we can better identify what data abstractions best align with these mental models.

Visualizing and discussing help elicit a person's mental model comprehensively. The varying levels of verbal descriptions of their sketches (code E8), the assortment of terminology used about their data abstractions (code E7), and the range of data abstractions (code C1) suggest that people can generally explain their mental model well but need multiple avenues to externalize it. The discussion with Participant 014 (see code E7) and prior work with visualizing genome sequences [31] show that observing problem-solving can expose underlying aspects of the mental model.

We observed that participants tended to overestimate what is "obvious" in their mental model, a psychological phenomenon shown in a visualization setting by Xiong et al. via a controlled study [44]. The lack of legends (code E5) suggests that the sketch is truly the user's mental model of the dataset, but makes understanding the sketches difficult for anyone but the sketch's author. We suggest visualization designers solicit conversations and sketches about the dataset—*not* of chart or representation types—from their users. Centering the conversation on the dataset, rather than a representation, will focus the conversation.

Different abstractions support different mental affordances, indicating tasks. Although visualization researchers tend to consider interactivity in the context of a visual design, people readily described their mental models in interactive terms, often with only loose—if any—association to specific visual encodings (code C7).

We suggest designers consider the interplay between the tasks that different data types readily afford and note the interactions that the users describe. For example, a table lends itself naturally to sorting, whereas a graph lends itself more naturally to navigation. The affordances of the specific data abstraction that a person latches onto may betray the tasks they most need to perform. Conversely, a given data abstraction may inspire specific, predictable forms of purpose-seeking (code F5). Ensuring task and data abstractions are aligned may translate to more intuitive interactions and more effective visualization designs.

People's views about what data is *and what it* isn't *may limit ideas during data creation.* Several participants related data to tables, computers, Excel, or database tables (code D1). They had definitions for what "data" is and what it isn't (code D2). This may limit or expand the data abstraction. These ideas are important to discuss during data reconnaissance [9] and throughout the creation of the data abstraction. We suggest visualization designers provoke users by proposing alternative data abstractions, sources, and formats that may help expand the definition of data and uncover latent data abstractions [2].

5.3.2 *Before Designing Visualizations.*

Visualization design starts with data design. Many of our participants imagined beyond the dataset. They suggested possible sources of the data (code F6), invented tasks to be done with the data (code F5), and wished for additional data or information (code F4). This curiosity may be due to the participants having no relationship to the dataset, but could also be due to inherent curiosity.

This creativity can be useful to visualization designers, as it highlights what aspects of the data the user finds relevant and the tasks and operations the user wants to execute. If possible, when encountering a situation with "no real data available (yet)" [37], visualization designers have a chance to be part of the data design phase. We define the *data design phase* as an unconventional part of the Design phase in the design study methodology where the contents, attributes, and format of the data is still under discussion even though the design study process has begun. In such a scenario, visualization designers should engage in these conversations with their collaborators. By being present in these discussions, designers can better understand the priorities and motivations of their users. An alternative would be to have the users recreate the dataset from memory; the features and entries that the user finds most important will likely be remembered.

Extra care must be taken when eliciting data abstractions with relations between data items. We observed participants used a variety of terms and visual representations when there were relations between data items (code C6). Some used terms that were inconsistent with the visualization community's definitions (code E7). It was difficult to confidently determine the participant's mental model despite their language and sketch. Even when there were similar visual representations among sketches, such as in the file system dataset, the ways they were spoken about were different.

These observations suggest that visualization designers should practice extra care when eliciting data abstraction when relations are present. One example may not be enough to determine the nature of the data described. If we were to probe further, a set of relation assertions (e.g., "connections like this may never occur") may elicit more detail. However, there are some abstractions where the structures may be the same but the meaning and conventions are different, such as trees and hierarchies, where this approach alone would not be enough.

The way people express relations between data items suggests a continuum of data abstractions. Visualization designers may be able to leverage this fluidity. Another way to interpret our observations regarding how people organize relations between data items (code C6) is that the data abstraction classifications represent examples in a continuum of data abstractions. This continuum does not seem to have clearly defined axes but instead seems to be a continuous space of how they organize data without the strong boundaries associated with data type classifications and taxonomies. Datasets do not need to fit one named abstraction. We observed Participant 030 refine their data abstraction from a hierarchy to a network when considering how the dataset might be stored in a database. Participant 018 spoke of sets and hierarchies together. This suggests there may be utility in representations that allow people to leverage these multiple ways of abstracting the relationship structures in the data.

Suggesting multiple audiences can elicit multiple data abstractions. When the topic of communicating with other people was discussed we observed participants changing their sketches, or claiming they would, sometimes to the point of selecting a different data abstraction. This behavior could be employed to explore several useful abstractions of the same data. There may be differences among what they would sketch for themselves, how they would communicate *their mental model* to someone else, and how they would communicate the data to someone else—or potentially multiple such "someone elses."

Multiple audiences may also help identify cases where people bow to expectations about how they are "supposed to" visualize data. We observed expectations regarding what is considered "data," including tables, demographics, and counts. Such conventions may lead to less useful abstractions. Prior work in creative visualization workshops and collaborative prototyping provides a framework for facilitating exploring alternative, useful design ideas [10, 11, 22].

6 CONCLUSION

We presented a sketch-based study to probe people's mental models and their corresponding data abstractions before they are given visual or other structural cues. We observed diversity and fluidity in the mental models that participants described and the data abstractions and visual representations that the participants used. This diversity can be influenced by several factors including examples from their lives, common approaches to the context, things they had seen recently, imagined tasks, and their conceptions of what "data" is and the conventions that come with it. We also observed that participants used a variety of terms and relations to describe the data and their sketches and would reconfigure their model when considering different audiences. These observations suggest that care must be taken when eliciting descriptions of data for the process of data abstraction and visualization design, but also offer options for leveraging the data design process to further probe user needs and possible abstractions, as well as opportunities to use the framework of communication to explore the data exploration space.

ACKNOWLEDGMENTS

Thank you to our reviewers for their thoughtful suggestions and our participants for their creativity and willingness to draw. This work was supported by NSF-1844573.

REFERENCES

[1] Alex Bigelow, Steven Drucker, Danyel Fisher, and Miriah Meyer. 2014. Reflections on How Designers Design with Data. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces (AVI '14)*. ACM, Como, Italy, 17–24. https://doi.org/10.1145/2598153.2598175

[2] Alex Bigelow, Katy Williams, and Katherine E. Isaacs. 2021. Guidelines for Pursuing and Revealing Data Abstractions. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Feb. 2021), 1503–1513. https://doi.org/10.1109/TVCG.2020.3030355

[3] Jeremy E. Block and Eric D. Ragan. 2020. Micro-entries: Encouraging Deeper Evaluation of Mental Models Over Time for Interactive Data Systems. In *2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV)*. IEEE Computer Society, Los Alamitos, CA, USA, 38–47. https://doi.org/10.1109/BELIV51497.2020.00012

[4] Kristen R. Butcher. 2006. Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology* 98, 1 (2006), 182–197. https://doi.org/10.1037/0022-0663.98.1.182

[5] Senthil Chandrasegaran, Devarajan Ramanujan, and Niklas Elmqvist. 2018. How Do Sketching and Non-Sketching Actions Convey Design Intent?. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. ACM, New York, NY, USA, 373–385. https://doi.org/10.1145/3196709.3196723

[6] Monica Chin. 2021. FILE NOT FOUND: A generation that grew up with Google is forcing professors to rethink their lesson plans. The Verge. https://www.theverge.com/22684730/students-file-folder-directory-structure-education-gen-z

[7] Nancy J. Cooke. 1994. Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies* 41, 6 (1994), 801–849. https://doi.org/10.1006/ijhc.1994.1083

[8] Nancy J. Cooke and Anna L. Rowe. 1994. Evaluating Mental Model Elicitation Methods. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 38, 4 (1994), 261–265. https://doi.org/10.1177/154193129403800416

[9] Anamaria Crisan and Tamara Munzner. 2019. Uncovering Data Landscapes through Data Reconnaissance and Task Wrangling. In *Proceedings of the 2019 IEEE Visualization Conference (VIS)* (Vancouver, BC, Canada) *(VIS '19)*. IEEE, New York, NY, 46–50. https://doi.org/10.1109/VISUAL.2019.8933542

[10] Steven Dow, Julie Fortuna, Dan Schwartz, Beth Altringer, Daniel Schwartz, and Scott Klemmer. 2011. Prototyping Dynamics: Sharing Multiple Designs Improves Exploration, Group Rapport, and Results. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 2807–2816. https://doi.org/10.1145/1978942.1979359

[11] Steven P. Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L. Schwartz, and Scott R. Klemmer. 2011. Parallel Prototyping Leads to Better Design Results, More Divergence, and Increased Self-Efficacy. *ACM Trans. Comput.-Hum. Interact.* 17, 4 (dec 2011), 1–24. https://doi.org/10.1145/1879831.1879836

[12] Amy Fox, Erica Vries, Laurent Lima, and Savannah Loker. 2016. Exploring Representations of Student Time-Use. In *Diagrammatic Representation and Interference*, Mateja Jamnik, Yuri Uesaka, and Stephanie Elzer Schwartz (Eds.). Springer International Publishing, Cham, Switzerland, 40–47. https://doi.org/10.1007/978-3-319-42333-3_4

[13] Todd Gamblin. 2021. "Personal OCD: when people use "tree" to refer to a graph that isn't necessarily a tree. Your "dependency tree" is usually a directed acyclic graph (DAG). Do you do this? Why/why not?". Twitter. https://twitter.com/tgamblin/status/1437807851007741952?s=20&t=2JnyjRhAfE7Fx5vI51jPXQ

[14] Katy Ilonka Gero, Zahra Ashktorab, Casey Dugan, Qian Pan, James Johnson, Werner Geyer, Maria Ruiz, Sarah Miller, David R. Millen, Murray Campbell, Sadhana Kumaravel, and Wei Zhang. 2020. Mental Models of AI Agents in a Cooperative Game Setting. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1?12. https://doi.org/10.1145/3313831.3376316

[15] Tina Götschi, Ian Sanders, and Vashti Galpin. 2003. Mental Models of Recursion. *SIGCSE Bull.* 35, 1 (jan 2003), 346 – 350. https://doi.org/10.1145/792548.612004

[16] Samantha Harper and Stephen Dorton. 2020. Towards a Context-Dependent Framework for Visualizing Mental Models. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64, 1 (2020), 308–312. https://doi.org/10.1177/1071181320641071

[17] Mary Hegarty and Marcel A. Just. 1993. Constructing mental models of machines from text and diagrams. *Journal of memory and language* 32, 6 (1993), 717–742.

[18] Edwin L. Hutchins, James D. Hollan, and Donald Arthur Norman. 1986. Direct manipulation interfaces. *Human-Computer Interaction* 1, 4 (1986), 311–338.

[19] Philip N. Johnson-Laird. 1980. Mental models in cognitive science. *Cognitive Science* 4, 1 (1980), 71–115. https://doi.org/10.1016/S0364-0213(81)80005-5

[20] David H. Jonassen and Philip Henning. 1999. Mental Models: Knowledge in the Head and Knowledge in the World. *Educational Technology* 39 (1999), 37–42. Issue 3. https://eric.ed.gov/?id=EJ588188

[21] Natalie A. Jones, Helen Ross, Timothy Lynam, Pascal Perez, and Anne Leitch. 2011. Mental models: an interdisciplinary synthesis of theory and methods. *Ecology and Society* 16 (2011), 46–46. Issue 1. http://www.ecologyandsociety.org/vol16/iss1/art46/

[22] Ethan Kerzner, Sarah Goodwin, Jason Dykes, Sara Jones, and Miriah Meyer. 2019. A Framework for Creative Visualization-Opportunities Workshops. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 748–758. https://doi.org/10.1109/TVCG.2018.2865241

[23] Gary Klein and R.R. Hoffman. 2008. Macrocognition, mental models, and cognitive task analysis methodology. In *Naturalistic Decision Making and Macrocognition*, Jan Maarten Schraagen, Laura G. Militello, Tom Ormerod, and Raanan Lipshitz (Eds.). CRC Press, Boca Raton, Chapter 4, 57–80.

[24] Zhicheng Liu and John T. Stasko. 2010. Mental Models, Visual Reasoning and Interaction in Information Visualization: A Top-down Perspective. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (Nov. 2010), 999–1008. https://doi.org/10.1109/TVCG.2010.177

[25] Eva Mayr, Günther Schreder, Michael Smuc, and Florian Windhager. 2016. Looking at the Representations in Our Mind: Measuring Mental Models of Information Visualizations. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization* (Baltimore, MD, USA) *(BELIV '16)*. ACM, New York, NY, USA, 96–103. https://doi.org/10.1145/2993901.2993914

[26] Miriah Meyer, Michael Sedlmair, P Samuel Quinan, and Tamara Munzner. 2015. The nested blocks and guidelines model. *Information Visualization* 14, 3 (2015), 234–249.

[27] Stanley Milgram and Denise Jodelet. 2014. Psychological Maps of Paris: (1970). In *The People, Place, and Space Reader*. Routledge, Abingdon-on, 45–49.

[28] Michael Muller. 2014. *Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method*. Springer New York, New York, NY. 25–48 pages. https://doi.org/10.1007/978-1-4939-0378-8_2

[29] Michael Muller, Ingrid Lange, Dakuo Wang, David Piorkowski, Jason Tsay, Q Vera Liao, Casey Dugan, and Thomas Erickson. 2019. How data science workers work with data: Discovery, capture, curation, design, creation. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. ACM, New York, NY, USA, 1–15.

[30] Tamara Munzner. 2014. *Visualization Analysis and Design*. CRC Press, Boca Raton, FL. 20–40 pages. https://doi.org/10.1201/b17511

[31] Cydney B. Nielsen, Shaun D. Jackman, Inanç Birol, and Steven J. M. Jones. 2009. ABySS-Explorer: Visualizing genome sequence assemblies. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 881–888. https://doi.org/10.1109/TVCG.2009.116

[32] Erika Shehan Poole, Marshini Chetty, Rebecca E. Grinter, and W. Keith Edwards. 2008. More than Meets the Eye: Transforming the User Experience of Home Network Management. In *Proceedings of the 7th ACM Conference on Designing Interactive Systems* (Cape Town, South Africa) *(DIS '08)*. ACM, New York, NY, USA, 455?464. https://doi.org/10.1145/1394445.1394494

[33] A Johannes Pretorius and Jarke J Van Wijk. 2009. What does the user want to see? What do the data want to be? *Information Visualization* 8, 3 (2009), 153–166.

[34] Ian Sanders, Vashti Galpin, and Tina Götschi. 2006. Mental Models of Recursion Revisited. *SIGCSE Bull.* 38, 3 (jun 2006), 138 –142. https://doi.org/10.1145/1140123.1140162

[35] Wolfgang Schnotz and Maria Bannert. 2003. Construction and interference in learning from multiple representation. *Learning and Instruction* 13, 2 (2003), 141–156. https://doi.org/10.1016/S0959-4752(02)00017-8 External and Internal Representations in Multimedia Learning.

[36] Günther Schreder, Florian Windhager, Michael Smuc, and Eva Mayr. 2016. Supporting Cognition in the Face of Political Data and Discourse: A Mental Models Perspective on Designing Information Visualization Systems. In *2016 Conference for E-Democracy and Open Government (CeDEM)*. IEEE Computer Society, Krems, Austria, 213 – 218. https://doi.org/10.1109/CeDEM.2016.23

[37] Michael Sedlmair, Miriah Meyer, and Tamara Munzner. 2012. Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2431–2440. https://doi.org/10.1109/TVCG.2012.213

[38] Abhay Sukumaran and Clifford Nass. 2010. Socially Cued Mental Models. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI EA '10)*. ACM, New York, NY, USA, 3379 – 3384. https://doi.org/10.1145/1753846.1753988

[39] J. Gregory Trafton, Susan B. Trickett, and Farilee E. Mintz. 2005. Connecting internal and external representations: Spatial transformations of scientific visualizations. *Foundations of Science* 10, 1 (2005), 89–106. https://doi.org/10.1007/s10699-005-3007-4

[40] Barbara Tversky. 2014. Visualizing thought. In *Handbook of human centric visualization*. Springer, Cham, Switzerland, 3–40.

[41] Unknown. 2022. Tittha sutta, Udāna 6.4, Khuddaka Nikaya. E-book. https://suttacentral.net/ud6.4?view=normal

[42] Jagoda Walny, Christian Frisson, Mieka West, Doris Kosminsky, Søren Knudsen, M. Sheelagh T. Carpendale, and Wesley Willett. 2020. Data Changes Everything: Challenges and Opportunities in Data Visualization Design Handoff. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 12–22. https://doi.org/10.1109/TVCG.2019.2934538

[43] Jagoda Walny, Samuel Huron, and Sheelagh Carpendale. 2015. An Exploratory Study of Data Sketching for Visual Representation. *Computer Graphics Forum* 34, 3 (june 2015), 231–240. https://doi.org/10.1111/cgf.12635

[44] Cindy Xiong, Lisanne Van Weelden, and Steven Franconeri. 2020. The Curse of Knowledge in Visual Data Communication. *IEEE Transactions on Visualization and Computer Graphics* 26, 10 (2020), 3051–3062. https://doi.org/10.1109/TVCG.2019.2917689

[45] Cecile Yehezkel, Mordechai Ben-Ari, and Tommy Dreyfus. 2005. Computer Architecture and Mental Models. *SIGCSE Bull.* 37, 1 (feb 2005), 101 – 105. https://doi.org/10.1145/1047124.1047390

[46] Caroline Ziemkiewicz and Robert Kosara. 2008. The Shaping of Information by Visual Metaphors. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1269–1276. https://doi.org/10.1109/TVCG.2008.171