Bringing Data to Real Life

Context and Practices of Data Physicalization

SHING-YUN CHIANG
Bringing Data to Real Life
Context and Practices of Data Physicalization

Thesis presented by Shing-Yun Chiang

To The Department of Art + Design

In Partial Fulfillment of The Requirements for The Degree of Master of Fine Arts in Information Design and Visualization

Thesis committee
Pedro Cruz (advisor)
Miso Kim (chair)
Dietmar Offenhuber
Kate Terrado

Northeastern University
Boston, Massachusetts

December, 2018
Acknowledgements

To my advisor Pedro Cruz, for the guidances and helps along this journey.

To my thesis reader Dietmar Offenhuber, for the suggestions and directions.

To the IDV cohort, for the company, joy, and inspirations.

To my family, for the support from all aspects.

To Li-Xuan, for always being there.

To all my friends, for the adventures and food in the States.
Abstract

A physical object takes on new meanings when it connects to the experience or knowledge from individuals. Similarly, data physicalization refers to the process of giving data tangible shapes by building data-based physical artifacts. Through the transformation of data into a physical representation, data physicalization conveys perceivable information to the audience. The active perception, depth perception and non-visual perception provided by a data physicalization are the keys for leveraging humans’ perceptual exploratory skills. With the capability of engaging people through a multisensory experience, data physicalization is potentially an impactful approach to tell stories about social issues. Poverty and flood is an existing issue that is the topic of the physical project in this thesis. A 3D physical map is designed and produced to display the poverty levels and flood damages from Tropical Storm Allison, a representative flood event occurred in Houston, Texas in 2001. The map portraits the terrain of poverty levels with flood damages painted on top. “Depth” is adopted as the metaphor for social status. Higher poverty rates result in lower social status that in turn are represented as deeper areas in the map. An embodiment model is adopted to analyze the performance of the metaphor applied in the physical representation in this thesis. This thesis follows a physical visualization pipeline to explain the design and production process. A new idea of physicalization and memorability experiment are suggested in this thesis.
Content

11 Introduction

17 State of the art
   On physicalization
   Related work

33 Data physicalization of poverty levels and flood: a case study
   Social content and data
   Reasons for physicalization
   Embodiment and metaphor
   From data to physical presentation

49 Discussion
   Metaphorization and design process
   Audiences and potential impacts
   Limitations
   Future work

56 References
Introduction

In humans’ experience, a physical object takes on new meaning when it connects to the experience or knowledge from individuals. A souvenir from a trip; a ring from fiancé; or a toy that your cat loved to play with before passing away. The same concept exists in the field of information visualization: “data physicalization” is the domain that deals with the process that designers use to embed data values into physical objects, as well as with the contextualization of the embedded objects themselves (Jansen et al. 2015). Through the transformation of raw data into physical represen (Jansen and Dragicevic 2013), artifacts possess meaningful information connected to the audience. Data physicalization gives data tangible shapes, therefore enables data to “live” in the physical environment with the audience.
Overview of data physicalization

Data physicalization is an emerging research field that uses physical forms to represent data (Jansen et al. 2015). The concept of data physicalization has been referred to “physical visualization” or “data sculpture”. The term of “data physicalization” was proposed by Jansen et al. (2015). Data physicalization as an artifact, refers to “a physical artifact whose geometry or material properties encode data”. Data physicalization as a research area, focuses on the physical visual form, and the process of giving data physical appearance. This thesis presents an exploratory experiment on how data can be physicalized and how data connects with the audience through physical forms.

Physical data-based artifacts

Data physicalization has been developed as long as the human’s civilization. Sumerians used clay token to express economic data (Schmandt-Besserat 2014). At the early stage of technology development, people tried to build models to visualize and understand complex systems with data. For example: the MONIAC and The Mississippi River Basin Model (Atlas Obscura). Along with the development of technology and statistics, data visualization was applied to scientific research and then extended its boundary to the art, design and journalism fields. Visualizations nowadays mostly live on paper or screen.

Yet there is a new practice of design that encodes data into physical forms. Benefiting from the development and accessibility of certain technologies, designers are able to create physical visualizations with vast datasets at a large scale. Mount Fear East London that sculpted violent crimes by Abigail Reynolds (2003) is one of the well-known projects. Some physical artifacts combine electronic devices and tangible interfaces. The Emoto Project is an interactive installation that physicalized the volume and emotions of tweets during Olympic Game days in London 2012 (Stefaner, Hemment, and Studio NAND 2012).
Research on data physicalization

Data physicalization is beneficial in many ways. Due to the physical form, it can be observed by the viewer from multiple angles; it provides a multisensory experience other than visual perception; and it embeds data in tangible materials from the real world (Jansen et al. 2015). Researchers also have developed studies on evaluating the physicality of visualization using 3D physical bar charts. Work has shown that the physical form of a visualization is able to increase the memorability when the data has: 1) dramatical changes in values and 2) connection with the audience (Stusak, Schwarz, and Butz 2015). In terms of efficiency, 3D visualizations perform better in physical modality than on screen (Jansen, Dragicevic, and Fekete 2013).

Data physicalization possesses physicality and affordance that humans know how to perceive and interact with (Zhao and Moere 2008). They have both functional and artistic qualities, resulting in the capability of conveying data-driven insights. People would be able to reflect on social and cultural impacts through experiencing a physical representation of data. Therefore, aside from digital data-driven storytelling, data physicalization is proposed to be an alternative approach to tell informational stories to a mass audience.

The hand-on practices

Considering the advantages of data physicalizations for storytelling, this thesis describes hands-on practices of creating a physical artifact based on data related to social issues. Poverty and flood is chosen as the theme of the physicalization. The datasets in use are the poverty data in 2000 in combination with flood damages from Tropical Storm Allison in 2001. A 3D physical terrain map is produced as the outcome. The landscape of the map portrays poverty levels, and blue dots on top represent flood damages. Poverty levels indicate social status, and are expressed by “depth”. The deeper the areas, the lower the social status. “Depth” acts as the metaphor of social status.
Introduction

The design process is described through the physical visualization pipeline from Jansen and Dragicevic (Jansen and Dragicevic 2013) in this thesis. On the purpose of provoking awareness towards a certain social issue, it is critical to analyze how well the data is embodied as well as how the metaphor associates data with the audience. The artifact produced for this thesis is evaluated using the embodiment model from Zhao and Moere (2008).

The physical map made out of this thesis combines the practical work and theoretical analysis. In order to convey stronger message to the audience, a new idea of the physicalization about poverty and flood is proposed. This thesis also suggests a memorability experiment as a future direction.
This chapter will discuss the theories and studies of data physicalization as well as the related work. The theoretical discussion starts from the concept of data physicalization and data sculpture, the most common type of data physicalization. The embodiment model that examines the metaphorization, the physical visualization pipeline, and the evaluations on memorability, efficiency and interaction are also introduced in this chapter. The exploration of the academic research and related artistic work contribute to the practice of data physicalization on a public issue in this thesis.
ON PHYSICALIZATION

Data physicalization refers to physical representations of data and a research field focusing on the process of transform data into physical forms (Jansen et al. 2015). As an artifact, a data physicalization results in a set of artifacts that encodes data with their geometry and/or material properties. As a research field, data physicalization “examines how computer-supported, physical representations of data can support cognition, communication, learning, problem solving, and decision.” The data sculpture is the most popular and common type of physical visualization (Jansen, Dragicevic, and Fekete 2013). As the direct externalization of the data, the data sculpture focuses on the physical embodiment of data (Moere 2008).

Zhao and Moere (2008) proposed a domain model (Figure 2.1) to demonstrate the role of data sculpture. Data sculpture is related closely to four research areas including information visualization, tangible user interfaces, visualization art, and interactive art. Data sculpture is located in between artistic focus and functional focus, and leans to the physical manifestation side (Moere 2008). In other words, data sculpture possesses both artistic and functional features of data visualization, emphasizing on physical representation.
State of the art

**Physical visualization Pipeline**

Jansen and Dragicevic (2013) proposed a physical visualization pipeline (Figure 2.2) to describe the sequential steps from raw data to physical presentation. The pipeline includes the visualization system on the design end, and perception process on the user end. This thesis focuses on the visualization process that includes 4 transformations:

1) Data transformation: it is the transformation from raw data into a suitable format for visual experiments. This process contains data cleaning, filtering or aggregation that make data compatible with visualization techniques.

2) Visual mapping: this transformation gives data graphical properties using visualization techniques (e.g. bar chart, scatter plot) and produces the initial visual form for data (Card, Mackinlay, and Shneiderman 1999). The visual mapping transformation produces the abstract visual form that the visualization properties (e.g. color, scale) have not been defined yet.

3) Presentation mapping: this transformation turns the abstract visual form to visual presentation by setting up visualization properties. The visual presentation is ready for production either for print, digital display or physical artifacts.

4) Rendering: This transformation produces the perceivable physical presentation from visual representation. The physical presentation is the physical object that can be observed in the physical world.

*Figure 2.2*

Physical visualization pipeline (Jansen and Dragicevic 2013).
State of the art

Embodiment Model

Embodiment in data sculptures refers to the “expression of abstract data in physical representation through the process of data mapping” (Zhao and Moere 2008). Metaphor is used in the process of mapping that translates data into physical representations. The embodiment model (Figure 2.3) from Zhao and Moere (2008) is the technique to evaluate the performance of embodiment in physical representation of data. The model is mapped in two dimensions: 1) the metaphorical distance from data and 2) the metaphorical distance from reality.

The metaphorical distance from data

The concept of signification in semiotic theory (Saussure 1983; Chandler 1999) is adopted to describe the metaphorical distance from data (signified) to physical representation (signifier). The signifier is the material or physical form of signified. The signification indicates the relationship between signifier and signified. Chandler (1999) classified the signification into three modes with examples:

1) Symbolic: the signifier does not resemble the signified. Human have to learn the relationship from conventions. For example: languages, national flags. The symbolic relationship indicates the furthest distance from data.

2) Indexical: the signifier is associated with signified directly in a physical or casual way. For example: a phone ring is the signifier of the situation (signified) that someone is calling. The indexical relationship has the shortest metaphorical distance from data.

3) Iconic: the signifier is the resemblance or imitation of signified by possessing some qualities from signified. For example: a portrait. The iconic relationship has the distance longer than indexical relationship but shorter than symbolic relationship.
State of the art

The metaphorical distance from reality

A tangible interaction system (Wozny 1989) is used for measuring the metaphorical distance from physical representations to reality (Zhao and Moere 2008). A metaphor should be recognizable in order to be interpreted correctly. A successful metaphor should possess all the three basic-level concepts from tangible computing field, the more criteria it qualifies, the shorter the metaphorical distance from reality.

1) Single mental image: the metaphor should be identified easily and associated with a single mental image. For example, the term “chair” generates a single mental image of a four-leg sitting device.

2) Affordance: the metaphor should possess the affordance that people know how to interact with it. For example, the handle of the cup possesses the affordance that invites people to hold the cup through the handle.

3) Intuitiveness: the metaphor satisfies the concept of intuitiveness when people are familiar with it without prior training or education needed.

---

**Figure 2.3**
Embodiment model. Redraw from Zhao and Moere's analysis model (2008).
Evaluation on memorability, efficiency and interaction

With an emphasis on functionality, work was done by measuring the memorability (Stusak, Schwarz, and Butz 2015; Stusak, Hobe, and Butz 2016), efficiency (Jansen, Dragicevic, and Fekete 2013) and interaction (Taher et al. 2015; Taher et al. 2017), as well as how well physical objects embody virtual data (Zhao and Moere 2008). 3D visualizations are known to be problematic in digital format, but are common among physical visualization (Jansen, Dragicevic, and Fekete 2013). Most of the evaluation research was conducted by comparing screen-based visualizations with 3D physical data-based artifacts, as well as 2D form with 3D form of physical artifacts. In terms of visual mapping, bar chart is the most common used among memorability, efficiency and interaction evaluation.

The memorability of bar charts in 2D digital and 3D physical format

One of the benefits of data physicalization is haptic perception that is a lack in screen-based visualization (Jansen et al. 2015). Previous studies indicated that haptic experience would affect visually-acquired memory and physical objects would be recalled more frequently than pictures or words (Kelly, Avraamides, and Giudice 2011; Bevan and Steger 1971). Haptic perception and physical modality generally help memorability. Based on this hypothesis, Stusak et al. (2015) conducted an experiment about memorability on a static vertical bar chart in 2D digital and 3D physical representations (Figure 2.4).
State of the art

Forty participants were divided into two groups; one group read digital bar chart on tablets, the other read physical bar chart (Stusak, Schwarz, and Butz 2015). Participants filled out a questionnaire immediately after the exploration and another questionnaire after two weeks. Two questionnaires had the same questions about extreme values, specific numeric values and facts of the content. The results suggested that physical 3D bar chart performed a better memorability of extreme values (e.g. minimum and maximum). It could be concluded that physicality is helpful for memorability for a certain kind of information.

The memorability of 2D and 3D physical bar chart

Another evaluation on memorability was also done by Stusak et al., and compared a 2D and a 3D version of token-based physical visualization with an emphasis on datasets (Stusak, Hobe, and Butz 2016). A token-based approach (Huron, Jansen, and Carpendale 2014) was adopted in the experiment, so that each data point was represented by an independent physical token. The 2D version was bar charts made of paper stripes, and the 3D one was built from wooden blocks (Figure 2.5). The data used was population and economic datasets at country level, both datasets had a 2D and a 3D version.

Sixteen participants worked with 2D and 3D version of two datasets, and assembled the bar charts manually. Subjects had to take the first online quiz immediately after exploring the visualization haptically, and the second one after one week. Both quizzes contained the questions about the underlying data, the latter one had additional questions about data that were not asked in the
first quiz. The second quiz also included image recognition that participants had to recognize the picture of a certain group of assembled physical bars (Stusak, Hobe, and Butz 2016).

Stusak et al. (2016) observed that 3D modality outperformed 2D version with population data, especially for extreme values. However, there was no obvious difference between 2D and 3D bar charts with economic data. The results suggested that data played a crucial role in this experiment in two ways: 1) population data had a wider range of values, which led to a distinct visual difference and could be memorized better potentially; 2) population data could be will related to participants and was more interesting than economic, according to the feedback from participants.

The two evaluations of memorability of physical bar charts had conclusions aligned with each other: with extreme values in data, 3D physical bar charts were more memorable than 2D bar charts in digital and in physical modalities (Stusak, Schwarz, and Butz 2015; Stusak, Hobe, and Butz 2016). It suggested that 3D physical visualization has the potential to tell a story better when the data has dramatic rise and fall in values. The results of the latter experiment suggested that the context of data is critical as well. An abstract and un-interesting dataset is not likely to impress audience in either 2D or 3D visualization (Stusak, Hobe, and Butz 2016).

Efficiency of bar charts across modalities

Jansen, Dragicevic and Fekete (2013) designed experiments to evaluate the efficiency of physical visualization. There were four techniques representing datasets via bar chart: physical 3D bar chart, on-screen stereoscopic 3D bar chart, on-screen monoscopic 3D bar chart and on-screen 2D matrix with 2D bar chart (Figure 2.6). 16 participants were presented the four visualizations one after the other, then saw 3 datasets in order and had to answer 3 questions for each dataset. Efficiency is about time and accuracy, and was therefore measured by 1) time that participants spent on completing tasks and 2) error rate related to datasets.
The error rates of the four techniques were low in general, the focus of analysis turned to time on task. Subjects spend the least time on 2D matrix with 2D bar chart on-screen, and then physical 3D bar chart. Stereoscopic and monoscopic 3D bar chart in digital took the most time. The results suggested that 3D bar chart in physical form is more efficient than 3D bar chart on screen (Jansen, Dragicevic, and Fekete 2013). In other words, physical modality would be a better solution under the condition that 3D visual technique is necessary.

Interaction

The progress of technology development allows designers to make interactive data physicalizations. Taher et al. (2015) built the dynamic physical bar chart called “EMERGE” to explore the interaction with physical representation of data. EMERGE (Figure 2.7) was a 10 X 10 grids of bar chart and illuminated by RGB LED. 17 participants conducted the interaction tasks under 4 data exploration scenarios: annotation, filtering, organization and navigation. In each scenario, participants were asked to use different ways to interact with 14 interaction techniques (e.g. press, swipe away, point, drag and drop). The study found that whether an interaction technique is helpful or not, depends on the context of use. For example, a direct press is highly effective for the participants to annotate a data point than other techniques. Another study investigated the behaviour when participants interacted with EMERGE (F. Taher et al. 2017). The study found that participant were likely to explore the artifact with body movement. It suggested that the physicalization should be placed in a location that allows the audience view the artifact from multiple angles.
Applications of data physicalization

Data physicalization is adopted to many fields. Moere and Patel (2009) implemented the creation of data physicalization in a design class in an university. A team from Microsoft designed a set of interactive physical bar charts for community engagement (Regan et al. 2015). Some domains use data physicalization as an approach to help people understand professional knowledge such as statistics (Gwilt, Yoxall, and Sano 2012), science and mathematics (Bourke 2015), and geology (Djavaherpour, Mahdavi-Amiri, and Samavati 2017).

RELATED WORK

This section will be introducing work about data physicalization throughout the history. The data physicalization can be a token that is assigned a data value, it can be a large data-driven model embedded in the physical environment, it can also be a data sculpture with the help from technology.

The beginning of physicalization

The history of data physicalization can be dated back to Sumer, the earliest known civilization. Before paper and writing were invented, Sumerians used clay tokens (Figure 2.8) to count, store and communicate economic data (Schmandt-Besserat 2014). Due to the need of people in that era, the number are about economics and trading. At that time, technology has not been developed, they embedded data into physical objects with their original shape.

Figure 2.8
Tokens from an ancient Sumerian city (Schmandt-Besserat 2014)
Physical Models

Before the technology has been well-developed, people built models in order to understand a complex system. The Monetary National Income Analogue Computer (MONIAC) (Figure 2.9), built by Bill Phillips in 1949, modeled the economic flow of the United Kingdom (Atlas Obscura). Several tanks in the model represented different roles in a nation, such as households, business, government. The water indicating money ran through the system and can be explained as income, spending or GDP depends on where to water passed by. Another model embedded into the physical environment was the The Mississippi River Basin Model (Figure 2.10), built by the Army Corps of Engineers in from 1948-1966 (Atlas Obscura). The model represented 40% of the US and was used to simulate the impact of flood events.

Figure 2.9
MONIAC model (Phillips 1949; Atlas Obscura).

Figure 2.10
The Mississippi River Basin Model (The Army Corps of Engineers 1948; Atlas Obscura).
The combination of 2D and 3D physical visualization

There were several physical visualizations using 3D physical objects in a combination with 2D physical. A 2D map with beads on pins and wires visualized the locations of a class of male students in Harvard University six years after graduation (Figure 2.11). Another physical map showed the passengers of Frankfurt Streetcar in Germany using wood stripes glued on each others (Figure 2.12). The more passengers, the more wood strips glued together.

Figure 2.11
A physical visualization of the locations of a class of male students in Harvard University six years after they graduated (Brinton 1914).

Figure 2.12
A physical map visualized passengers of Frankfurt Streetcar in Germany (Brinton 1914).
Data sculptures

Benefited from the development of technology, designers are able to create large-scale data physicalizations. Mount Fear East London (Figure 2.13) by Abigail Reynolds (2003) is a well-known one, it about 60.6 inches high and 212.6 inches wide. The work is an elevation map of violent crimes in East London. It represents the rise and fall of recorded violent crimes from 2002 to 2003 geographically. The higher the peaks, the more crimes occurred. The sharp peaks express the fear of crimes.

Figure 2.13

Interactive installation

Moreover, some physical artifacts combine electronic devices. The Emoto Project physicalized the volume of tweets and emotions over the Olympic Games London 2012 (Stefaner, Hemment, and Studio NAND 2012). One column of the physical object was one day, one row was an hour, the volume of tweets was represented by the height of the objects (Figure 2.14). Different emotions were projected with different gradations of colors. The artifact was installed with a controller, so that users were able to interact with the visualization and highlight a certain tweet.
In short, data physicalization is a storytelling approach that brings abstract data to real life. With such an approach, the visualization is able to leverage human's perceptions therefore engages people into the context provided by the metaphor. Due to the ability of engagement, data physicalization has the potential to catch people's attention towards public issues. This thesis is an exploratory application of adopting physicalization in the expression of public data.
Data physicalization of poverty levels and flood: a case study

This thesis reports a series of hands-on processes to build a physical visualization, from data transformation to implementation. Data physicalization is a storytelling approach capable of creating artifacts that are beneficial to multisensory experience and engaging the audience. It would be appropriate to apply this approach to social issues. The practice on physicalization in this thesis focuses on poverty levels and flood, visualizing percentage of people in poverty and flood damages from Tropical Storm Allison in Houston, Texas in 2001. There is an underlying relationship between poverty and flood that general public should be aware of. The theme of poverty and flood is the concept, and represented by the data from one of the most devastating flood event caused by Tropical Storm Allison in Houston.

The output as a physical artifact is a 3D terrain map of poverty (Figure 3.1); the deeper the areas, the higher percentage of people in poverty. The depth of the map acts as the metaphor for social status where low-income people have lower social status and the wealth have higher. This chapter discusses: 1) social context and data; 2) reasons for physicalization; 3) the process of building the artifact; 4) discussion on embodiment and 5) limitations.
Figure 3.1
The 3D physical map visualizes the poverty levels and flood from Tropical Storm Allison in 2001.
Low-income people suffer from flood damages more. It is an existing issue that scholars in related fields such as environmental justice and flood management have been investigating for years. Sarmiento and Miller (2006) explored the relationship between poverty and flood risk in the U.S., and found both households in poverty and higher income families face greater flood risk. They explained that higher income households are facing higher flood risk because they often live by water for beautiful views, and the wealth are able to afford flood insurance. On the other hand, people in poverty live in risky areas due to lower rent.

Houston in Texas is a flood-prone city. Seventeen years ago, Tropical Storm Allison blew through Houston and caused fatal flood (Figure 3.2). The flood took 23 people’s lives (Hlavaty 2018) and damaged more than 70,000 households (Harris County Flood Control District). The flooded households were located more in the areas where poverty rates were higher. After almost two decades, the issue of poverty and flood is brought into discussion again due to the flood from Hurricane Harvey in 2017 (Reeves 2017). The distress that low-income people have been facing never changed.
Data physicalization of poverty levels and flood

Poverty levels and flood damages from Tropical Storm Allison as an example in this thesis, telling a story about the vulnerability that people in poverty are suffering. The flooded households dataset is obtained from Harris County Flood Control District, containing the coordinates of 73,255 households that were affected by flood in Harris County. Since the flood happened in 2001, the poverty data in use is the dataset released in 2000 from U.S. Census Bureau.

The Decennial Census provides domestic survey data at five geographical levels from the biggest to the smallest: State, County, Census Tract, Block Group, Census Block. In order to represent the details and preventing the difficulty from production in parallel, Census Tract is selected to be the geographical level in poverty data. The poverty data is shown as the ratio of income to poverty threshold. Families with an income below poverty threshold have a ratio under 1.00. The percentage of population with ratio under 1.00 per Census Tract is represented in this thesis.

Poverty threshold varies with family size and number of children under 18 years old (US Census Bureau). For example, the poverty threshold for a family of four with no child younger than 18 is $17,761. If a family meets the criteria and has an income less than $17,761, the four people in this family live under the poverty line.

REASONS FOR PHYSICALIZATION

Poverty and flood is the theme of the data sculpture, the outcome is a 3D physical terrain map made of wood. Poverty levels are represented by a red color scale and different depths at the same time. Flooded households are painted in blue using thin wood stripes. The 3D physicalization approach is adopted for the following reasons: 1) multi-layer visualization, 2) the metaphor and depth perception, and 3) the theme is in a social context.
**Multi-layer visualization**

Geography is the first layer of this visualization. The data on poverty distribution and flood damages contains geographical features and is strongly related to geography. For example, people live along the coast would face a higher risk of coastal flood. Therefore, a map is an appropriate visual mapping for this topic, in order to demonstrate the spatial pattern of poverty and flood from a flood event.

The second layer of this visualization is poverty levels while the third layer displaying flooded households. Some digital visualizations of poverty used 2D choropleth maps as the visual approach. A 2D choropleth map is good for displaying one variable at a time through a graduated color scale. When coloring two variables in one map, it could cause confusions due to the overlapping. Considering the flood damages as the third layer will be added and colored blue, the depth dimension is adopted to prevent from the confusion of overlapping colors from the two data layers.

**The metaphor and depth perception**

The depths indicate the percentage of people in poverty by Census Tracts. There are 5 brackets in the scale: 0-10%, 11-20%, 21-30%, 31-40%, above 40%, meaning that there are 5 levels of depth in the map (Figure 3.3). The lower the Census Tracts are, the higher the poverty rates. Depth acts as the metaphor for social status. Low-income people have lower social status while the wealth has higher. Therefore, the areas with a higher proportion of residents in poverty have lower terrain in the map.

Audience is expected to walk around and view the artifact from multiple angles, perceiving the changing depth. Overlooking the physical map, the lower and deeper places look like the caves and holes down into the ground and provide an image of lives in poverty. Depth perception is one of the benefits from physicalizations. Physical objects provide “rich cues of shape and volume” that are able to better display data in 3D (Jansen et al. 2015). Audience perceives a 3D object in physical modality with less effort and higher accuracy than 3D on screen (Jansen, Dragicevic, and Fekete 2013).
The theme is in a social context

Stepping out from scientific research areas, information visualization nowadays is a common storytelling approach used to communicate socially-relevant messages (Moere and Purchase 2011; Zhao and Moere 2008). As the sub-domain of information visualization, the data sculpture possesses physicality and affordances that help convey information to a mass, lay audience (Zhao and Moere 2008).

Humans have the inherent proficiency of interacting with physical environment using visual as well as non-visual senses. Zhao and Moere (2008) stated that affordances are “tangible properties of an object that influence how it can be used”. Affordances provided by physical objects are capable of being interpreted functionally, therefore conveying “informational meaning”.

The wooden crosses at Field of Remembrance (The Royal British Legion) in UK is a great example of telling stories to mass audience (Figure 3.4). A wooden cross is an externalization of a soldier who sacrificed himself in the Service of UK. More than 120,000 wooden crosses are planted in the fields each year, representing the loss of soldiers and demonstrate the scale of deaths due to wars. This example provides audience the tangible connection between the wooden crosses and their knowledge and experience in the real world (Zhao and Moere 2008).

Poverty and flood is also an existing social issue that people should pay attention to. The wooden map of poverty levels and flood damages draws the connection between depth and social status that is everywhere in audience’s lives.

Figure 3.4
The wooden crosses represent deaths of soldiers. (The Royal British Legion)
EMBODIMENT AND METAPHOR

Embodiment influences how data sculptures communicate information with the audience, and how the audience interprets the metaphorical presence by perceiving the physical artifacts (Zhao and Moere 2008). Embodiment is measured by: 1) metaphorical distance from data and 2) metaphorical distance from reality. For the physical map of poverty levels and flood, the vertical distance (depth) is the metaphor of social status and the social status is measured by poverty rates.

Metaphorical distance from data

In semiotic theory, metaphorical distance from data is specified by the relationship between physical representation (signifier) and datasets (signified) (Zhao and Moere 2008). In this thesis, there are two steps of transition from poverty data to a physically perceivable vertical distance: 1) transforming poverty data into social status and 2) transforming social status into physical vertical distance. The poverty data indicates social status that is represented as vertical distance.

The relationship between poverty data and vertical distance is better explained as the symbolic relationship. The poverty data in this thesis is used to describe social status. The higher the poverty rate, the lower the social status; therefore those areas with higher poverty rates in the map have lower terrains. This connection is established by the physical legend (Figure 3.4). It is necessary for the audience to read the legend and learn the convention of the physical map, so as to understand the metaphor correctly. The symbolic relationship learned from convention indicates that the physicalization in this thesis has a farther metaphorical distance from data (Zhao and Moere 2008).

Metaphorical distance from reality

As pointed out in chapter 2 of this thesis, there are three criteria to determine the metaphorical distance from reality of a data sculpture: 1) single mental image, 2) affordance and 3) intuitiveness (Zhao and Moere 2008). The more criteria a data sculpture satisfies,
the shorter the metaphorical distance from reality. For this thesis, vertical distance (depth) is the metaphor for social status and the social status, expressed in poverty levels, is the reality. The physical map satisfies two and half of the criteria: single mental image, affordance and part of intuitiveness.

**Single mental image**

Social status is hierarchical in its nature, and thus social statuses are ordered. It is reasonable to describe that people in the upper class of society live on the top layer of the pyramid (Figure 3.6) while the under class is at the bottom (Ozkan 2014). Therefore, the vertical distance is metaphor of social status that possesses familiarity to the audience. The social class pyramid would be the single mental image when people talk about social status.

**Affordance**

Zhao and Moere (2008) stated that a metaphor satisfies this criterion when the audience knows how to interact with it. Since the physical map is a passive physicalization, the interaction here would refer to “interpretation”. Since the vertical distance is an appropriate metaphor for social status, the audience is able to operate on the metaphor and lead to the interpretation from metaphor to information.

**Intuitiveness**

The physical map portraits the landscape of poverty and social status. The terrain-map-like appearance could be interpreted as the rise and fall of land at the first glance. However, the stair-like physical legend helps audience easily associate poverty levels with depth and implies the social status. Once the audience reads the legend, it is not difficult to interpret the depth since the concept of social status exists in everyone’s life experience. Therefore, it is reasonable to say that the physical map meets half of the intuitiveness criterion.
FROM DATA TO PHYSICAL PRESENTATION

This thesis utilizes the physical visualization pipeline from Jansen and Dragicevic (2013) to describe the design and production process of the work. This section will be discussing the process in the following orders: 1) data transformation: the process that transforms data into the workable format for visualization; 2) visual mapping: the process of encoding data into graphical primitives and graphical attributes, generating the abstract visual form as the outcome; 3) presentation mapping: it turns the abstract visual form from visual mapping into the visual presentation that “can be printed, displayed or fabricated”. 4) rendering: this process describes the transformation from visual presentation to physical presentation that lives in the physical environment.

Data transformation

Two datasets are used for the physical artifact in this thesis: 1) poverty status in 2000 and 2) flood damages from Tropical Storm Allison in 2001. The flood damages data come with coordinates in latitude and longitude, therefore most of the data transformation work is on poverty data.

The poverty status is displayed in the ratio of income to poverty. The ratio of income to poverty refers the total family income divided by the poverty threshold (US Census Bureau). The data table shows the population of whose poverty status was determined and the population of nine ratio brackets from 0.0 to 2.0 and over. This thesis focuses on the percentage of people who have the ratio under 1.0, as known as people live under poverty line.

The percentage of people in poverty is called poverty levels in this thesis. With 10 percentage points as an interval, the five poverty levels start from 0-10%, and end at the level of above 40%.
Visual mapping

Since the poverty and flood damages have the feature of spatial distribution, the terrain map visual mapping is adopted for the case study. There are three layers of data: 1) geographical boundary by Census Tracts, 2) depths and graduated colors showing the five poverty levels, 3) blue dots showing the locations of flooded households on top of the map. In short, the abstract visual form generated from this transformation is a 3D terrain map of poverty levels with flooded households painted on top.

Presentation mapping

The presentation mapping process in this work contains: 1) deciding the dimension and range of physical map; 2) defining the color scale for poverty levels; 3) picking a color in blue hue to represent flooded households and the painting tool.

Dimension and the range of the map

Due to the capability of the laser cutter, the size of the physical map is limited to 22 inches in both width and length. With this dimension, the physical map includes most of the flooded households in Harris County from Tropical Storm Allison (Figure 3.7). In order to make sure the dimension work for smaller Census Tracts, a 12-inch prototype was made for experiment. Plywood is used as the material of the physical map.

The 12-inch prototype (Figure 3.8) extracted the central areas of Houston where most of the smaller Census Tracts are located. A sheet of plywood has a height of 0.25 inch, the maximum thickness works for the laser cutter. The areas in the lowest bracket of poverty level (0-10%) are made of five sheets of plywood stacked, while the areas in the highest bracket (above 40%) have only one sheet.
Data physicalization of poverty levels and flood

Figure 3.8
The 12-inch prototype.

The prototype displays the dramatic change in the center of Houston. The smallest Census Tract is visible. It is not necessary to adjust the dimension for the final piece. A concern was raised that audience might get lost when there is only depth indicating poverty levels, audience could be confused about which poverty level they are looking at. Therefore, a graduated color scale is adopted and functions as a reference of poverty levels with depth at the same time. In order to make the disparity more visible, the thickness of each poverty level is twice as it was in the prototype.
Data physicalization of poverty levels and flood

*The graduated color scale*

The color-coding experiment is taken with colors in brown and red. Brown color is the natural color of ground and soil; red color could be used to represent risk and danger. Since the color is painted on wood directly, the color scale is generated from mixing a dark color and white proportionally. By comparing two color scales, it is found that the red one (Figure 3.9) works better because it holds a neutral stand without bias while brown color (Figure 3.10) could potentially associate poverty with race.

Figure 3.10
Brown color scale

Figure 3.9
Red color scale
Representation of flood damages

The experiment of color-coding is continued with picking a color in the blue tone that works with the colors for poverty. Water is represented in blue, a color that is a common representation of water. The next step is to figure out what kind of blue works better with a graduated color scale from light to dark. The final decision was to use Cyan. Cyan is a light blue color, which is light enough to contrast with a darker red and also distinguishable from pink (Figure 3.11).

The texture of data sculptures can be perceived by both visual and tactile senses, which is one of the benefits offered by data physicalizations (Jansen et al. 2015). Therefore, how the pigment representing flooded households is painted on the wood is also critical. A dot refers to a flooded household. The shape that indicates flood damages should be rough and messy in order to imitate people’s impression of flood damages. The reason that thin wooden sticks (Figure 3.12) are used for painting is that they create irregular edges and surfaces for dots.

Figure 3.11
Cyan works well with red color scale.

Figure 3.12
Wooden sticks used for painting
Data physicalization of poverty levels and flood

Rendering

The final transformation from raw data to physical presentation is rendering. With the specified colors and painting tools, the rendering process for this thesis includes: 1) laser cutting, 2) assembling and 3) painting.

Plan for laser cutting

There are five levels of poverty; five files are prepared for laser cutting. The rule of layout planning is keeping the Census Tracts that lie in the current poverty level, and cutting off those are at higher level(s) or have no data. Take the lowest level (0-10%) for example. At the layer of 0-10%, Census Tracts with a percentage of people in poverty below 10% are kept at that layer, while the Tracts with a poverty rate more than 10% are left out (Figure 3.13).

![Figure 3.13]
Identify the Tracts below 10%
Merge the boundaries
Generate the format for laser cutting

Prepare files for laser cutting
**Assemble and paint the wood**

The color for each poverty level is painted on the whole plywood board before laser cutting. The flooded households are added layer by layer while assembling the map. The stacked and painted wooden map is put into a case with the physical stair-like legend (Figure 3.14). The case makes the top layer (0-10%) as the ground surface and areas with a poverty rate above 10% are underground. The case helps audience focus on the “how deep the impoverished areas are” instead of “how high the wealthy areas are”.

*Figure 3.14*  
Overview of the map with a case
Discussion

This section will discuss the integrating the metaphorization into the design process, the audience and potential impacts, the limitation of the work and future directions of a new physicalization and memorability testing.
METAPHORIZATION AND DESIGN PROCESS

The physical wooden map used depth as the metaphor to symbolize social status. The step of metaphorization is part of the design process, but it is not included in the physical visualization pipeline. The possible position of the metaphorization could be in the visual mapping process. Visual mapping process gives data an abstract visual form. For example, using bar chart to visualize categorical data, or using a line chart to visualize 10-year change of population in the US.

In the visual mapping process, data is given a graphical shape and dimensions of the data are specified. Here is where the metaphorization comes in. Take the physical wooden map for example, a map is a rough idea of the visualization. After investigating the datasets, it is found that there are three dimensions: geology, poverty level and flood damages. In order to represent three-dimensional data, 3D physical format is adopted. Therefore the height (depth) is used to represent poverty levels that are future explained as social status. The metaphorization of this thesis project appears in visual mapping, the second transformation in the physical visualization pipeline.

AUDIENCES AND POTENTIAL IMPACTS

The physical map is designed for conveying the social issue of poverty and flood, the purpose is to provoke audiences’ awareness. The physicalization approach is adopted since the physical representation is: 1) provide rich perceptual experience than digital format (Jansen et al. 2015); 2) more efficient in visualizing three-dimensional data (Jansen and Dragicevic 2013); 3) more memorable when display data that is related to the audience (Stusak, Schwarz, and Butz 2015; Stusak, Hobe, and Butz 2016).

The results from efficiency and memorability evaluations would vary with the design of experiments, the questions asked in the tasks, and numbers of participants. The positive results from previous studies might not be applicable to the physical map created for this thesis. However, the factors of impacts are not limited to efficiency and memorability. Attraction could also be a factor; the
more people interested in and read the physical map, the more impacts the map would make potentially. In Journalism, there are five elements to decide whether a topic is newsworthy to catch attention from the audience. Journalism and Design both need to “communicate” with audiences, therefore it is possible to discuss the attraction of the physical map from journalistic perspectives.

The five elements for newsworthiness are: timing, significance, proximity, prominence and human interest (Media College). Each element has different weights based on topics and audiences. The story delivered from the physical artifact could meet significance, proximity, human interest. The topic of the map is about the poverty levels and flood damages from Tropical Storm Allison in Houston. This flood event is the representative one throughout the history of Houston, clearly it has the significance element. The proximity element is satisfied or not depends on how closely the audience is related to the topic. The human interest element would be satisfied when the audience is interested in the topic.

If the physical map is installed in a library or a museum in Houston, it would have the most attraction due to the proximity. Current residents in Houston are likely to be the people who survived from or witnessed the flooding. The residents are closely related to the story, or it can be said some of them were “in” the story. It has been 17 years from Tropical Storm Allison, people in Houston are still being bothered by flooding. Thus, residents in Houston are the primary audiences of the physical map. With more attraction, more impacts are expected.

Besides the residents in Houston, residents in other flood-prone areas are secondary audiences. For example, residents along Mississippi River, residents in Florida or residents in New Orleans. They would empathize with people in Houston since they share similar experience. Thus people live in flood-prone areas would have higher interest in the physical map than others. A further group of audiences would be those who do not live in flood-prone areas but are interested in the topic. This group of audiences are attracted because their personal interest in poverty and flood. With these two groups, the impacts are expected to be less.
LIMITATIONS

One of the limitations of this physicalization project is the accuracy of location and amount of flooded households. The hand-painting approach is not able to demonstrate the precise coordinate and concentrations of flooded households. However, damages from flood can never be measured precisely and are not limited to households. Moreover, with the aim of provoking awareness from audience, the emphasis of the physicalization is to provide an image of how impoverished people suffer from natural disasters. Thus, the physical presentation with less accuracy on flood damages is acceptable for this thesis.

Another limitation is the lack of quantitative comparison. With the passive physical visualization, it is not easy to aggregate the data values and compare across different poverty levels. However, the main focus of the thesis project is to point out the existing issue and let the audience “feel” the difference between poverty levels. The audience is not necessary to know how many flooded households there are in a certain poverty level. The physical map provides data with context, in order to form the insights that can be remembered by the audience and potentially influence their process of decision-making (Jansen et al. 2015).

FUTURE WORK

The future work proposed in this section contain two aspects: 1) another physicalization on the same theme (poverty and flood), 2) a potential experiment on memorability using the wooden map produced for this thesis.

On practice

Besides the wooden physical map produced out of this thesis, another physicalization idea is also planned. In this plan, the physicalization will be using the same theme — poverty levels and flood — but provides more direct stimulation about flood. In this plan, there will also be a transparent terrain map of poverty levels; water will be poured inside the map to resemble flood. It has static and interactive concepts.
The physical map will be made of 25 transparent acrylic sheets (Figure 4.1). The areas in the lowest bracket of poverty level (0-10%) has a height as twenty five acrylic sheets stacked, while the areas in the highest bracket (above 40%) have five sheets. The acrylic version is five times higher than the wooden one. Since depth is the only reference of poverty levels here, the whole artifact should be tall enough for audience to feel the difference.

A certain amount of water will be poured inside the map. The water could be slightly colored in blue. When the map is filled with a certain amount of water, the areas at a lower poverty level will sink, people in those areas with higher poverty level are above water and able to breath. This version of physical map provide a stronger image of how impoverished people suffer from flood in general. After the water is poured out, some blue stains might be left in the map which could convey a message: damages do not go away with flood. It takes a long time to recover from flood damages, especially for lower-income people.

The interactive version has the same appearance with the static one, but combine with electronic devices. This concept is to
Discussion

On memorability

The purpose of this thesis is not only exploring how data can be physicalized, but also provoking awareness toward a certain social issue. Thus, it is important that how well the issue of poverty and flood can be memorized by communicating through 3D physical display. In terms of memorability, previous studies (Stusak, Schwarz, and Butz 2015; Stusak, Hobe, and Butz 2016) suggest that 3D physical visualization performs better than 2D physical and 2D digital with two features in data values: 1) the values vary dramatically and 2) the data is in the context associated with audience.

The data used in the physical map in this thesis has the two features mentioned above. Firstly, there is a Census Tract in the center of Houston that is much taller than surrounding Tracts (Figure 4.2). It shows a dramatic jump of poverty levels. Second, social status is the concept that general people are familiar with and engaged in, since a person could not prevent himself from being grouped into a social class. Theoretically, the 3D physical map as the results is more memorable than 2D in any modality on the topic of poverty and flood.

Previous studies about memorability were conducted with an emphasis on physical form. Physical bar charts were used in the tasks because they are easy to understand. However, it could have an issue of over-simplification since physical bar chart might not be the most appropriate form to represent the data for testing. When the audience read a visualization, how a physical artifact represents data is possible to affect memorability. Therefore, a direction is
suggested to design an experiment with the physical map produced from this thesis. The physical map is designed for storytelling, combining data, physical form and context. An experiment based on the physical map would potentially provide more comprehensive results of memorability.

Figure 4.2
The center of the map has a dramatical change of values.
References


