Proceedings of the ASME 2022 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2022
August 14–17, 2022, St. Louis, Missouri

DETC2022-88844

CONCREATE: USING DATA PHYSICALIZATION TO INCREASE THE UNDERSTANDING AND INSPIRATIONAL USE OF QUANTITATIVE DATA IN DATA-DRIVEN DESIGN SCENARIOS

Tiara Spalburg[1], **Senthil Chandrasegaran**[1], **Nicole Eikelenberg**[2], **Milene Gonçalves**[1*]
[1]Faculty of Industrial Design Engineering, Delft University of Technology, the Netherlands
[2]Ford Research and Advanced Engineering, Aachen, Germany
email: {t.n.f.p.spalburg, r.s.k.chandrasegaran, m.guerreirogoncalves}@tudelft.nl, neikelen@ford.com

ABSTRACT
In today's world, the advantages of data-enabled design are undeniable, increasing the performance of organisations drastically by informing and inspiring the design process. While organisations seem to be more experienced with quantitative data for evaluative purposes, they do struggle to use data as creative material to inspire the design process. Choosing the right type of data representation is critical for using data for creative purposes. Data visualization has proven to be highly effective in increasing understanding of data, as it is fast, accurate and flexible. Data physicalization, on the other hand, remains unexplored in comparison, especially its effect on creativity. This paper presents the results of two studies (one preliminary and one follow-up study), which explored the use of data physicalization in creative settings. The preliminary study enabled to collect initial requirements for the development of a physicalization toolkit, while the follow-up study investigated its impact on the design process, in comparison to data visualization. From the studies, we developed Concreate, a collaborative data physicalization toolkit designed to lead to creative insights from quantitative data. Our results show that Concreate can potentially stimulate creative thinking, by encouraging intense, tangible interaction with data leading to increased reflection-in-action and a deeper understanding of data. The two studies and toolkit development were carried out at a multinational automotive company, interested in innovating by incorporating data as creative material. Besides the immediate practical implications, we conclude this paper with a discussion on future recommendations for using data physicalization in the design process.

1 INTRODUCTION
With recent advances in data-powered technologies like artificial intelligence (AI), machine learning (ML) and the Internet of Things (IoT), the role of data in the design process is steadily expanding. Organisations engaged in product and service design are experienced in using data to test and evaluate ideas and products, especially using quantitative approaches such as A/B testing. However, 'data-enabled design'—defined as a design methodology that *"embeds data in the design process and capitalizes on the strengths of data, while remaining respectful of designerly ways of knowing"* [1, p. 12]—is yet to reach its full potential in most such organisations. One of the main characteristics of data-enabled design is to use data in a more exploratory way, rather than to solely inform the design process [2].

Organisations tend to struggle with using data to inspire the entire design process for two reasons. Firstly, designers and data scientists have different mindsets, balancing between contextual understanding based on intuition against a strict interpretation of what is and is not supported by the available data. Secondly, there is a lack of clarity on data literacy and skills that may be relevant to designers [3]. The problem is magnified by the overwhelming volume of (big) data, often gathered without a clear vision for its

[*]Address all correspondence to this author.

use [4]. Nevertheless, using data as creative material has shown to lead to more valuable insights from data, eventually leading to better products and services [2, 5].

Data representation plays a critical role in enabling quantitative data to be used as creative material to inspire the design process [6], as it forms the bridge between interpretation performed by the data scientist and inspiration used by the designer. Data visualization and physicalization are two such representation formats. While the former has been widely investigated in the context of design, the influence of data physicalization on creativity remains largely unexplored. Considering that we are slowly returning from a period where we were deprived of physical interaction in professional settings, caused by the restrictions imposed by the Covid-19 pandemic, it is particularly interesting to investigate the potential of data physicalization. Therefore, in this paper, we focus on data physicalization as a data representation format, to explore its ability to stimulate creativity and aid the quest of using data as creative material.

2 BACKGROUND

In data-enabled design, qualitative data should ideally be used to supplement quantitative data and vice versa, through data blending. Data blending is the process of combining data from multiple sources (quantitative and qualitative) into a single data sheet to reveal deeper insights [7]. Quantitative data is also referred to as thin data, whereas qualitative data is classified as thick data. Thin data usually holds little to no contextual information, only giving information about 'what' is happening. Thick or qualitative data adds context and can give information about 'why' something is happening [8]. As portrayed in Figure 1, both thin and thick data can either be big or small. Big data is the term used for complex data sets sourced from large samples that generally requires machine learning to reveal insights, whereas small data is smaller in volume, making it more accessible and workable for human comprehension [9]. To understand how data can be used as creative material in the design process, we need to know more about the design process itself and how data can be represented within this process in a way that stimulates creative thinking.

2.1 Data as creative material in the design process

Designers have always used data in the design process. Specifically, they are familiar with the collection of thick data through various methods like interviews, observation and focus group discussions [10]. Such data is collected with the purpose of understanding the design context, evaluating prototypes and stimulating creativity. The last purpose—stimulating creativity—is of special interest to use in the light of evidence that access to stimuli can enhance creativity [11], as well as reflection [12]. Sarkar and Chakrabarti [13] define stimulus as: *'An agent that activates exploration and search in design'*. Based on this definition, both thick and thin data could be classified as external stimuli and, as such, are able to trigger creative insights [11, 14]. However, far

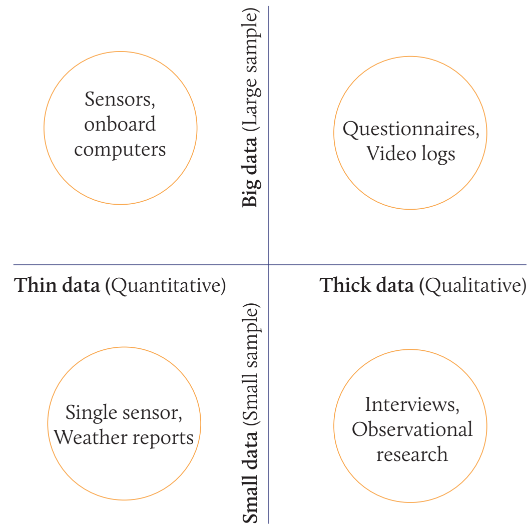

FIGURE 1. Classification of Data from the point of view of its relevance to design. Thin data (e.g. data from sensors) tends to be quantitative with little or no contextual information, while thick data (e.g. from observational studies) describes context and thus tends to be qualitative.

too little attention has been paid to understand the effect of thin data on creativity, defined as the production of novel, original, and valuable ideas in response to an open-ended task [15]. Thin data—such as those obtained by sensors and internet-of-things frameworks—are currently mostly being used to inform and evaluate decisions but implementing thin data as creative material to support idea generation is still relatively new [16].

The aforementioned literature has mainly addressed the impact of contextual stimuli on creativity and, as such, could be considered thick data. Likewise, we posit that presenting thin data as an external stimulus could potentially encourage reflection and increase creative thinking. Schön [17] outlines two types of reflection: reflection-on-action and reflection-in-action. Reflection-on-action happens after a design activity is finished, fitting well within the evaluation phase of the design process [18]. Reflection-in-action requires designers to consciously reflect while designing, resulting in ideas higher in originality [19] as new perspectives are gained, problems are reframed and new actions added to the design process [12]. Therefore, to get an idea of how to implement thin data as creative material in the design process, we need to investigate how data can be represented in a way that stimulates both reflection-on- and -in-action.

2.2 Data visualization vs. Data physicalization

The role of design on data visualization—not only in terms of utility and soundness of representation but also in terms of social engagement and impact—has long been acknowledged [20]. On the utilitarian side, data visualization is seen as means of extending the human cognitive ability to make sense of, find patterns and anomalies, and characterise data [21]. On the side of human engagement, it has the power to transform data without context into interactive visual representations that can stimulate thought, inspire ideas, and call for action [22, 23]. Data visualization has in recent years gained immense popularity in its application to diverse domains, with tools to author visualizations ranging from interactive, easy-to-use systems to high-performing libraries requiring substantial software engineering knowledge [24]. These are especially helpful for communication; for instance, when data scientists and data journalists are interested in conveying information to the increasingly data-literate public in an effective and digestible manner.

In large organisations requiring interdisciplinary work, data scientists need to collaborate closely with other domain practitioners such as designers, leading to certain tensions about the interpretation one can make from a given data set. Lu et al. [3] report that while designers are expected to make data-driven decisions, their expertise and interpretation of the context may clash with data scientists' stricter definitions of what can be interpreted, which is often seen by the designers as reductive. Designers are also not trained in data analysis and visualization, reducing their independence toward making data-driven decisions.

A possible mitigation of this issue could stem from data physicalization, defined as *"a research area that examines how computer-supported, physical representations of data (i.e., physicalizations), can support cognition, communication, learning, problem solving, and decision making"* [25, p. 3230]. Although the principle of data physicalization is not new, with presumed examples dating back to 5500BC, research in data physicalization is a fairly recent phenomenon [26]. Data physicalization has been gaining popularity over the past years, which might be explained by the growing emergence of fast prototyping techniques, but also the increase of virtual fatigue, leading to a rediscovery of the physical world [27].

Although slower and less suitable for big data and quick adjustments of datasets than data visualization [25], the required effort and awareness of the data physicalization process has the potential to encourage more reflection-in-action moments that stimulate creativity [28], next to reflection-on-action occurring after finishing the physicalization. Research has shown that abstract physicalizations of data that are both playful and enjoyable can also benefit task performance while engaging users and stimulating active perception by providing a multi-sensory experience [29]. Duarte [30] stresses that the dynamic relationship that develops between the user and data increases critical reflection and understanding of data. Thus, data physicalizations can allow users to explore data using their natural abilities to understand the world around them by perceiving and manipulating physical objects and materials [26]. However, the question remains: how can organisations implement a more data-driven approach to trigger innovative insights and solutions? And considering the potential benefits of data physicalization for creativity, how could such approaches be developed and implemented in data-driven organisations? Combining these questions leads to the following research question: *How can data physicalization approaches be developed and implemented into data-enabled design processes to inspire creative insights?*

3 CONTEXT

This research took place in the context of Ford Motor Company, specifically, within Ford Research and Innovation Center, in Aachen, Germany. In the last eight years—having recognised the benefits of data-enabled approaches to design—Ford has taken steps towards integrating data into their design practices. Ford has been incorporating *big data*—collected from sensors and other measuring devices—and *thick data*—observational data collected through ethnographic and other qualitative studies—in their processes. Ford's Global Data, Insights and Analytics department in particular—henceforth referred to as the '*data team*'—is responsible for the use of data within the company, allowing different teams to share knowledge and optimise the use of available data science expertise throughout all departments. Besides this centralised department, another unit is relevant for the context of this research: the multi-disciplinary design team called 'Innovation Management for smart vehicle concepts' focused on finding opportunities for innovation (from now on referred to as the 'design team'). The design team consists of experts with backgrounds ranging from industrial design and mechanical engineering to electrotechnical, business, marketing and more. They use a design thinking approach [31], to which creative sessions are included in each stage of the design process, allowing for the opportunity to introduce data as creative material to stimulate the generation of valuable insights and solutions. However, the design team consists of non-data experts with limited data analysis experience. Thin data is currently only used to inform or test decisions, as opposed to implementing it to inspire the entire design process. The design team observed the following challenges to using quantitative data as creative material:

C1 Having an overload of data gathered without a clear goal in mind and no plan on how to implement it.
C2 Not having a clear communication structure with the data team.
C3 Not having a set data analysis or data implementation process or stage in place within the design team.

Tackling these challenges can help the design introduce thin data as creative material in their design process, and combine thin data with the more familiar thick data resulting in richer insights. In Sec. 2.2 we discussed prior work showing how the multi-sensory environment provided by data physicalization can improve task performance and stimulate perception [29], and how physicalization can promote reflection [28]. Based on these findings, we propose that data physicalization, in the form of a toolkit, can help design and data teams address the challenges listed above.

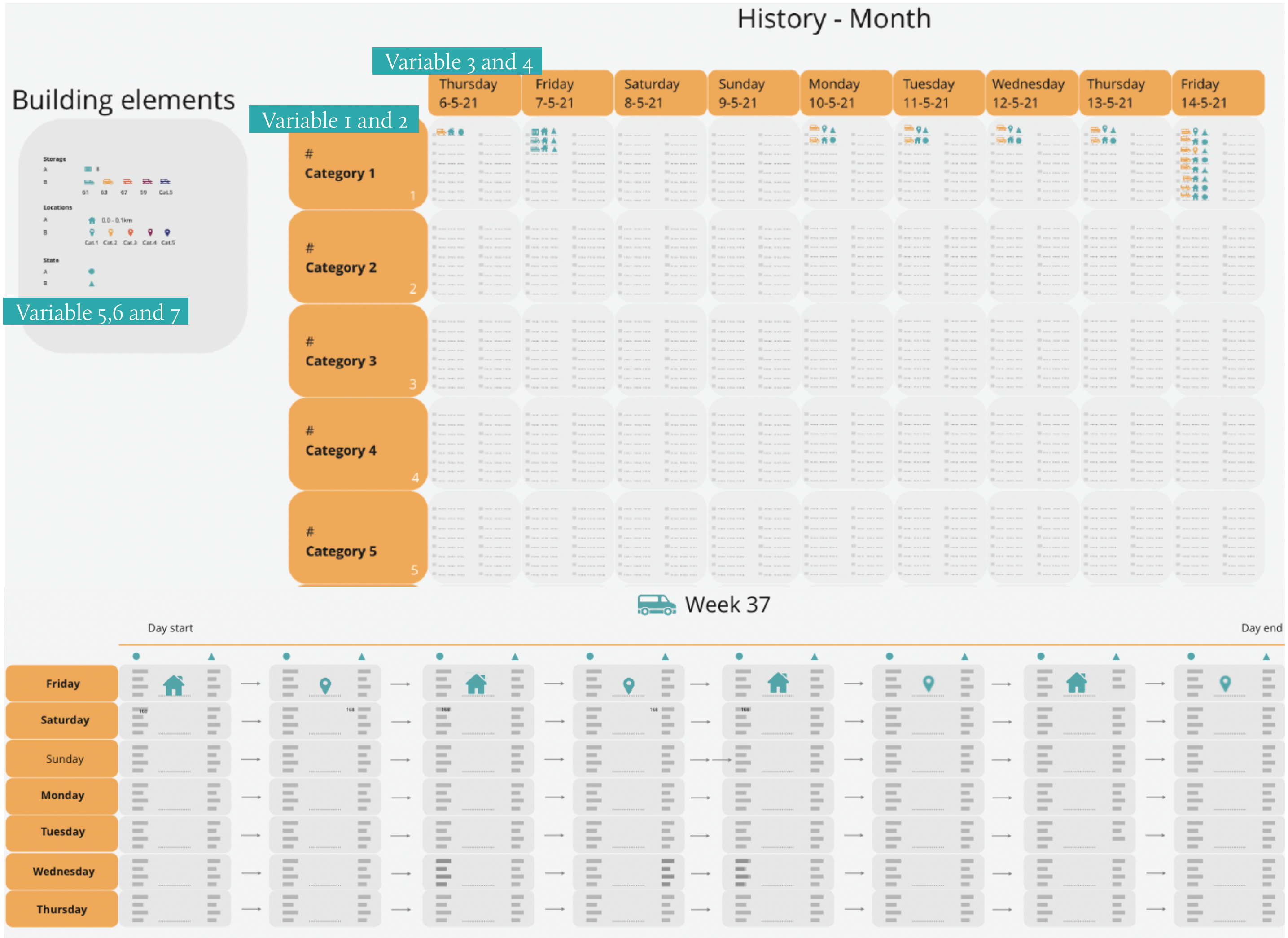

FIGURE 2. Overview of the online data physicalization prototype created in Miro, with highlighted variables from the data set (1-name category, 2-number category, 3-date, 4-day, 5-storage, 6-location, 7-state).

With the goal of gaining a better understanding of the process of physicalizing data, we first conduct a preliminary, formative study (Sec. 4) with a simplified online (digital) version of a "data physicalization" prototype. We then develop and evaluate a data physicalization toolkit, Concreate (Sec. 6) based on the insights of the preliminary study.

4 PRELIMINARY STUDY

The goal of this study was to get a better understanding of the data physicalization process and find opportunities for optimisation of the aforementioned prototype to implement in the final data physicalization toolkit. Three assumptions, supported by literature in data physicalization design [25] and operationalisation [32] guided this preliminary study:

A1 Building out a data set (online) helps encourage moments for reflection.

A2 Although the task of building out the data set could be time-consuming, the experience is seen as pleasant or satisfying.

A3 The created physicalizations are perceived by the target group as valuable.

A creative session of ninety minutes was chosen as a suitable format for this study as it closely resembles the actual design process of the design team. The creative session was split into three sections: an introduction session, a "data building" exercise where participants used digital representations of physical "tokens" to build representations of given data, followed by a collective evaluation interview. The first author functioned as the session facilitator and did not take an active role in the data building phase other than assisting. Participants were asked to discuss out loud to identify problems in the process and opportunities for optimisation. The full session was recorded and evaluated.

4.1 Participants

The study was conducted with four participants, among which three were members of Ford's design team. Participant 1 was a design and engineering researcher within the team, with an industrial design education background. Participant 2 was a research engineer with an engineering background, and Participant 3 had a marketing background and was the user experience expert within the design team. These three participants had varying, limited experience with data analysis. The last participant was a Master's student in data visualization, graduating at Ford.

4.2 Prototype

The prototype consisted of a set of materials designed to represent a data set through physicalization, as pictured in Figure 2. As this study was conducted during the Covid-19 pandemic and participants were not able to come together physically, the tool was designed to be used and investigated online. The prototype was created and shared across the group of participants using Miro[1], an online collaboration tool. A set of building elements—digital representations of physical shapes or tokens—was designed to be placed on the mapping boards by dragging or copy-pasting them on. These building elements would represent aspects, dimensions, and/or units of the given data in any way that the participants deem appropriate. Although this preliminary study was conducted virtually, the exercise can still be considered a valid simulation of data physicalisation, based on the direct actions (pick up, move, place, rearrange etc.) and interaction with the data representations.

4.3 Procedure

The data that was physicalized in this creative session was part of an ongoing project and can therefore not be disclosed. A selection of seven variables from this larger data set was made, presented in Figure 2. The session facilitator—an author of this work—cleaned and modified the data based on relevance and the available time of the participants. All participants were familiar with the data to a certain degree, using this study to build on previously gathered insights. The different phases of the study and their respective duration are portrayed in Figure 3.

During the introduction, the session facilitator explained the goal and setup of the session, introduced the data set and prototype, and explained the data building phase. This was followed by two building exercises in which parts of the data set were built using the prototype created in Miro. They had to complete this assignment as a group within a given timeframe but could decide how to divide roles or tasks among themselves. While the participants were performing the exercises, the session facilitator noted down observations within the same collaborative online environment (Miro). Finally, during the evaluation, the participants were collectively asked about their experience and assumptions were reflected upon. An interview outline was used to guide the evaluation and included such questions as 'How did you experience the session overall?', 'How would you describe the interaction with the prototype?' and 'How do you think the resulting physicalization could be optimised?'

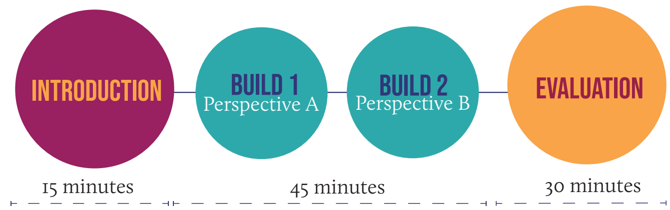

FIGURE 3. Overview of the preliminary study with its phases.

4.4 Results

The observations from the data physicalization activity performed by the participants—completed through observation during the study and analysis of the created data physicalization—and the responses from the evaluation interview were both analysed. The results of this analysis are organised below under *data & insights*, *process*, *communication*, *understanding of data*, and *reflection*.

4.4.1 Data & insights Right at the start of the building process, it became clear that the data set included errors caused by wrongly installed and inaccurate sensors that were used to collect the data. This in itself was significant: though participants were already familiar with the data and were aware of some of the errors, others were discovered only when they started building their representations of the data. This prompted questioning of the data itself:

> *"Well, I think for me there are two levels. It makes it very explicit that there is some data that needs checking and cleaning. So that's one thing. The other thing is that it makes me think about what other kinds of parameters are relevant."*
>
> – Participant 1

Although the participants were not specifically instructed to formulate insights, multiple topics and data that could be investigated further were discussed among the team. Insights were found by discussing the created physicalizations. Again, although the team was familiar with the data, these insights were not found before this physicalization exercise, suggesting that the prototype could be valuable to arrive at new insights.

4.4.2 Process & Communication Participants started by quickly dividing roles, generally looking for the most efficient way to perform the tasks. They did not initially decide what token was going to be used to represent what dimension or unit of the

[1] Miro: https://miro.com

data. Part of the reason was that their perceived familiarity with the data, and partly because they felt they needed to concentrate on counting, aligning, and building their representations.

> *"You need your concentration to build it, if you're discussing it, you'll make mistakes."*
>
> – Participant 2

This quote was one of a few comments about the process being prone to mistakes, with one participant working in the wrong field and others needing to shift back and forth between the data set and the tool. During the data building exercises very little communication between participants was observed. The high level of concentration required to perform the building task correctly obstructed participants from simultaneous, in-action-reflection and free discussion while engaging in the physicalization process.

4.4.3 Tool The created physicalization was well-liked by the participants although they identified several elements that could be improved to make the result more explicit and better readable. For example, participants suggested the inclusion of ways to highlight important areas and the implementation of more visual cues like icons and colours. As the tool is highly focused on data exploration by triggering questions and less on offering direct insights through data representation, one participant suggested that the tool would be more successful as a preparation tool to be used next to data visualization software rather than a tool used during creative sessions. However, they also mentioned that the tool does stimulate discussion after the building phase, which can lead to collective insights.

4.4.4 Evaluation interview The post-session interview was used to discuss the creative session and assumptions, which are presented here. The findings from this interview are presented in the context of the assumptions made at the start of Sec. 4.

Building out a data set (online) helps encourage moments for reflection (A1): The high level of concentration required to correctly perform the building task obstructed participants from having shared moments of reflection with their team. However, it is expected that this was caused by the online session format and could be reduced in an offline environment. Reflection-on-action did occur after the physicalization was finished.

Although the task of building out the data set could be time-consuming, the experience is seen as pleasant or satisfying (A2): Participants questioned if the physicalization process was worth the time investment or if the same result could also be achieved by an automated visualization process. Although gaining new insights was perceived as satisfying, working with an online tool, especially after a certain amount of time, was considered tedious. As one participant said:

> *"I think that the way you ask us to represent it has value because it enables us to identify patterns in an easy way. But I would actually doubt if it's maybe more efficient to already automatically generate such a graph, versus manually working with it."*
>
> – Participant 2

The created physicalizations are perceived by the target group as valuable (A3): Participants agreed that working consciously and intensively with the data revealed interesting patterns and led to valuable insights. They achieved a deeper understanding of the data and saw connections they did not see before, which was considered valuable.

4.5 Discussion

The prototype helped determine elements within the data set that would be interesting for the design team to investigate further, filter out errors in the data set and see new connections between variables. These findings suggest that physicalizing data can be valuable. Although the exact effects of the used data set on the generated insights were unclear, it can be assumed that a more accurate data set would have led to more valuable insights. The results also showed that the conscious, intense interaction with data kept them from reflecting-in-action but the finalised physicalization did allow for relevant reflection-on-action that led to new insights. However, it could not yet be determined if these insights could have also been found after automated data visualizations were presented and therefore if data physicalization is worth the time investment. As such, we defined two requirements to supplement the initial assumptions A1–A3 for further development of the data physicalization toolkit:

R1 The data physicalization toolkit should make the building process less prone to mistakes.
R2 The toolkit should encourage reflection.

5 THE CONCREATE TOOLKIT

Based on the findings from the preliminary study, we developed Concreate, a collaborative data physicalization toolkit designed to stimulate reflection and creative thinking. Through physically engaging with data using Concreate, we anticipate that users would gain a better understanding of the data. This in turn would prompt reflection and lead to new insights. The complete toolkit (Figure 2) consists of physical tokens to represent units and dimensions of data, along with a mapping board to ease the process of aligning and arranging the tokens, and a set of "reflection cards" to prompt reflection during the process. The toolkit also includes a detailed set of instructions outlining steps to ensure optimal use of the toolkit. However, as this instruction process was not part of the follow-up study, it will not be further elaborated upon.

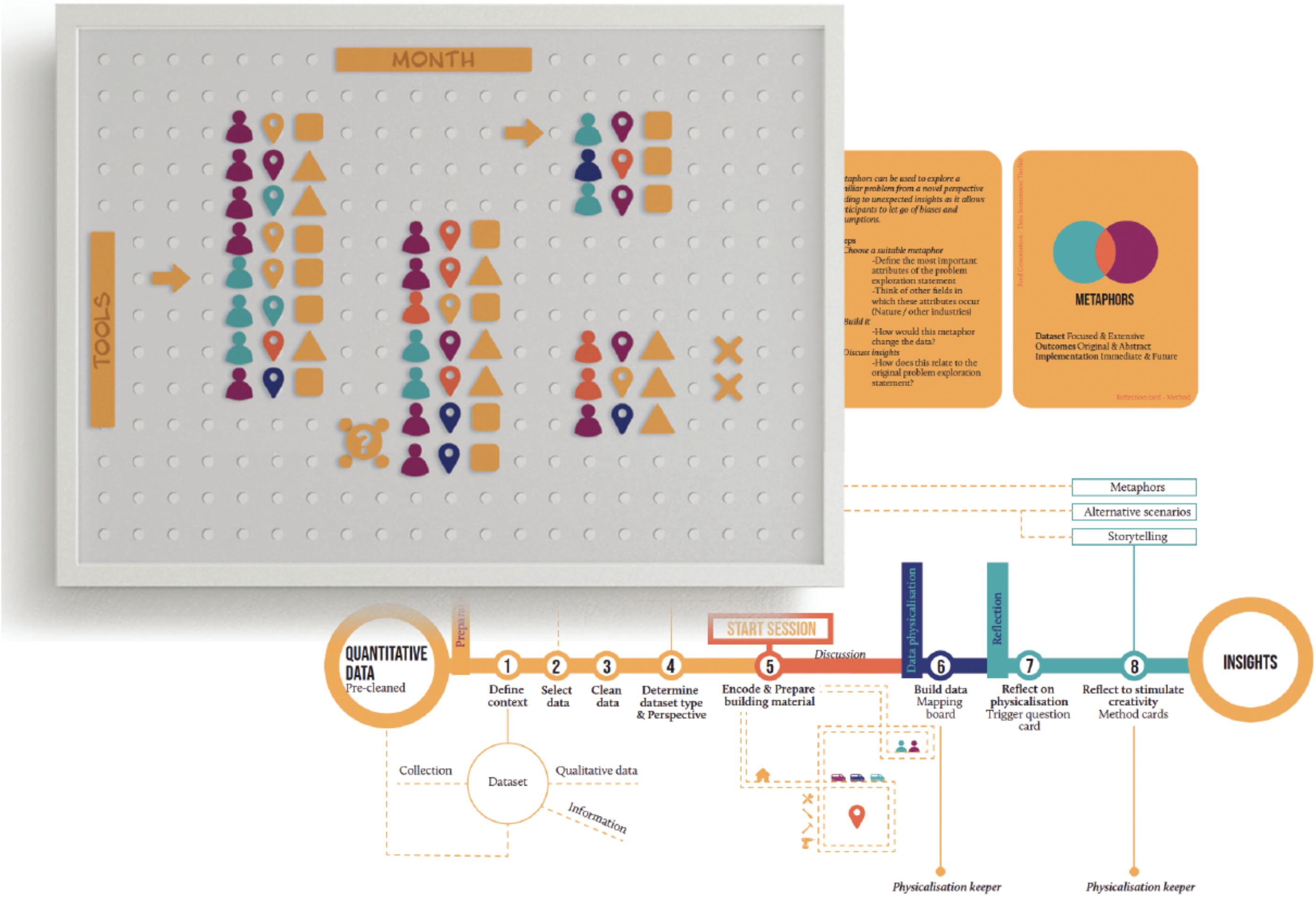

FIGURE 4. Overview of the Concreate toolkit including instruction process, mapping board and reflection cards.

5.1 Mapping board

The mapping board is designed to build a representation of a data set, to facilitate reflection-in-action and physicalizations that reveal insights about the data. The mapping board includes a grid that allows for easy adjustability and building freedom while still providing structure, which is helpful to keep the building elements organised and the physicalization readable *(addressing requirement R1)*. The board also includes a legend that can be filled in after encoding the building elements. The board comes with a set of coloured building elements or tokens that were laser-cut and affixed with magnets. The tokens can be selected and encoded per data set and are used to represent and build out data by placing them on the magnetic mapping board.

5.2 Reflection cards

To explicitly encourage reflection-in and -on-action *(requirement R2)*, two sets of reflection cards were added to the toolkit, to be used after creating the initial physicalization; question cards and method cards.

The question cards include one main question card and multiple sub-question cards designed to formulate and deepen the insights found during and after physicalization. The method cards encourage participants to interact with the finished physicalization by rebuilding new scenarios, stimulating creative thinking. These method cards provide tips toward using such methods as metaphors, alternative scenarios, storytelling etc., with details on their utility in data physicalization along with instructions for effective use. Each method is suitable for a different type of dataset and implementation could lead to different results, which is indicated on the front of the card. A sample of the reflection cards is shown in Fig. 4.

6 FOLLOW-UP STUDY

We decided to evaluate Concreate by comparing it with the use of data visualization as a baseline. We thus compared the insights generated by two teams during a creative session in two

FIGURE 5. The follow-up study setup. Left: participants in condition 1 using the Concreate toolkit, Right: participants in condition 2 using the set of given data visualizations.

conditions: one team used the Concreate toolkit (Condition 1), while participants in the baseline condition had access to printed data visualizations (Condition 2). Again, a creative session was chosen as a suitable format, with a separate session for each condition and no interaction between participants from different conditions. This session took place in an in-person, co-located environment (Figure 5) allowing for tangible interaction with the physicalization toolkit. The two separate sessions each started with an introduction given by the session facilitator, followed by a set of assignments and concluded with an evaluation interview. The facilitator guided the session but did not participate in the assignments. Participants were asked to think aloud as well as discuss among themselves, and the study was video and audio recorded.

6.1 Participants & Conditions

We recruited six participants (three male and three female, between the ages of twenty-one and twenty-eight), all of whom were students.The variety of disciplines was purposeful: each group included either one or two designers and either one or two participants from outside the design field to resemble the design team of Ford. Furthermore, considering the variety of data experience in the design team, non-data experts were selected as participants as well although some experience with data was allowed. Participants were split into two groups of three (see Table 1). One group was assigned the Concreate toolkit (Condition 1) while the other used a set of four data visualizations (representing the same data) printed on sheets of A4 paper (Condition 2). These four visualizations included a histogram, line chart and two scatterplots.

TABLE 1. Participant profiles and conditions in the follow-up study.

ID	Condition	Gender	Age	Domain	Data Experience (1:novice–5:expert)
1	1 : Concreate	F	25	Design	2
2	1 : Concreate	F	21	Biopharma	2
3	1 : Concreate	M	26	Med. Business	3
4	2 : Baseline	F	26	Design	2
5	2 : Baseline	F	22	Biology	2
6	2 : Baseline	M	27	Design	3

6.2 Procedure

The process that the two teams (one for each condition) went through is illustrated in Fig. 6. Both teams were given a printed spreadsheet of an open data set regarding electric vehicle charging times and duration at a charging location. The data set used was a small selection of five variables from a much larger data set. The chosen attributes are dates and times of charging transactions, how long vehicles stay connected to a charging point, how long they are being charged for, and the total amount of energy transmitted during this time. The session facilitator—an author of this work—chose and cleaned the data based on its suitability to the study. None of the participants was familiar with the data.

Participants in both conditions were instructed to formulate as many insights from the given data set as possible during their creative session. Participants in condition 1 were tasked with using the toolkit—Concreate—to achieve this goal. Their session started with an introduction given by the session facilitator. This introduction included a description of the goal, set-up and dataset, followed by a brief demonstration of Concreate including an explanation of the legend that was created by the facilitator beforehand. During this demonstration the building assignment was described as well. Participants were asked to follow the legend and build out as much of the dataset as possible within the given timeframe, by placing the physical tokens on the mapping board. They could decide how to approach this building exercise among themselves, but were advised to immediately note down insights that were found while physicalizing using given insight forms. The building assignment was followed by two rounds of reflection, using the questions cards in the first round and the method cards in the second. Participants were again asked to formulate new insights after each round, or build on previously generated insights. The creative session was concluded with a

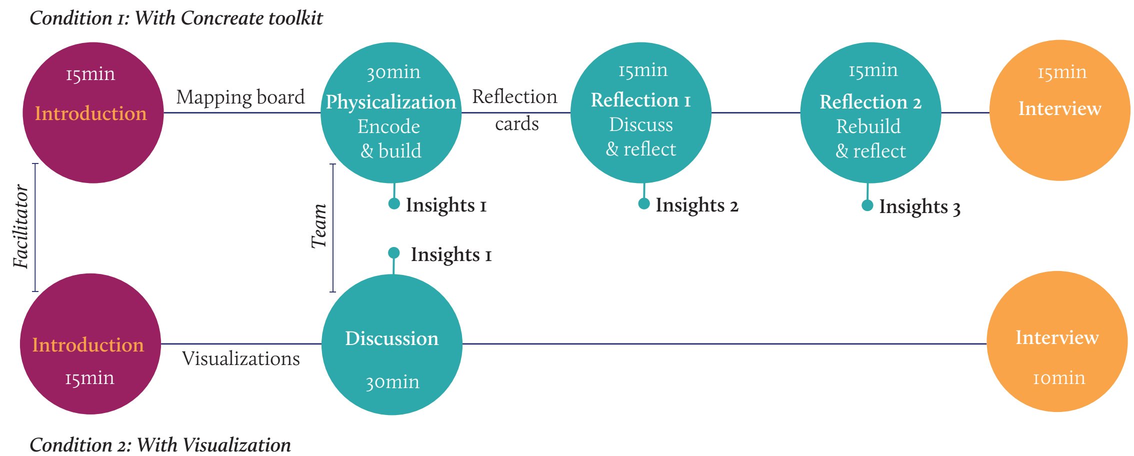

FIGURE 6. Overview of the procedure of the user study outlining the different phases for each condition. Note that the team that followed condition 1 with the Concreate toolkit was different from the team that used the data visualizations (condition 2).

collective evaluation interview led by the session facilitator, immediately after finishing the exercises. The interview outline was used to guide an evaluative discussion and included general questions regarding the process, experience and generated insights, but also specific questions about interaction with the toolkit and communication between the team while physicalizing.

Condition 2 was presented with the same data set and four data visualizations created from this data set and printed on sheets of A4 paper. They were asked to formulate insights by studying the visualizations and discussing them among themselves. Unlike the team using Concreate, the team using data visualizations did not have to do any creation; they were only tasked with discussing and generating insights (see Fig. 6 for a timeline comparison), loosely representing part of the current data analysis process within Ford's design team, which does not at the time include data physicalization. This session was concluded with an evaluation interview as well with similar questions as asked in Condition 1, excluding specific questions about the toolkit.

7 RESULTS

This section presents a comparison of insights generated by the two conditions during the study (Fig. 7), followed by the results of analysed observations and comments made during the evaluation interview. Observations made during the creative sessions were noted down by the session facilitator and later matched to comments made during the evaluation interview. From this analysis four topics emerged: process, communication, understanding of data and reflection.

7.1 Insights Generated by Participants

Figure 7 portrays examples of insights generated by the participants during the main study (this is not the full list of found insights). The insights found by Condition 1 using Concreate were built up over the reflection rounds shown in Fig. 6, starting with statements of observations made after physicalizing. In the first round of reflection, these statements were related to other physicalized variables or previously obtained knowledge and experience. In the last round of reflection, potential improvements or solutions were added to the insights. This resulted in final insights that are based on facts extracted from the data and could immediately be implemented in a design process or further researched. Condition 2 used the printed data visualization to find 27 insights, in contrast to the 17 insights found by Condition 1. The insights found by Condition 2 largely occurred in the form of questions, assumptions and speculation, giving no specific direction for implementation or further research.

7.2 Observations and evaluation

Figure 7 shows an overview of the most important observations made during the study, divided over Condition 1 and 2. Results are categorised and presented below, including comments from participants made during the evaluation interview.

7.2.1 Process As the prototype and assignment were quickly understood by Condition 1, they could start the exercises immediately. They enjoyed the interaction with the building elements and could easily come up with insights from the physicalization, making the process pleasant. In contrast, Condition 2 had to spend more time to understand the visualization, causing a lot of uncertainty that was considered less pleasurable, as a participant noted:

> *"It is quite a lot to understand what's happening and what kind of variables are involved. I felt a bit overwhelmed."*

CONDITION 1 - CONCREATE DATA PHYSICALISATION TOOLKIT	CONDITION 2 - DATA VISUALIZATIONS
Insights round 1 - Physicalization •Little correlation between charging time and total energy. **Insights round 2 - Reflection 1** •There is a lot of overcharging happening (cars being connected longer than the time they are actually charging for), it can be assumed that users charge their car before the battery is actually empty. It would be beneficial to reduce overcharging. **Insights round 3 - Reflection 2** •There is not a lot of charging started during the night (00.00-06.00) but many cars are charged overnight that start charging in the evening. It might be because people do not leave their house during the night, however, people who for example to walk their dog after 00.00, could be incentivised to start charging during the night.	**Insights visual 1:** •To accommodate travellers, more charging stations could be placed along highways. •Quite a lot of charging during the day, maybe charging stations are occupied and people are forced to share stations so charge after another user is done. •Does more driving means more charging? **Insights visual 2:** •Uber drivers with electric cars could charge at night to increase charging events during the night. **Insights visual 3:** •Outlying long charging times might be explained by having something in the car on while charging, that drains the battery. **Insights visual 4:** •The dip in efficiency is maybe because of old batteries?

FIGURE 7. Examples of insights generated by participants during the main study

– Participant 4

The participants from Condition 1 agreed that it could be useful to implement a data visualization step after data physicalization, as they were also interested in results on a larger scale. They also thought that incorporating a pre-data visualization step could be valuable to create more optimal categories for encoding the building elements, suggesting a potential iterative loop between data physicalization (for categorization & engagement) and data visualization (for computation and scale).

> *"I think some kind of data visualization pre-process would be helpful to encode the elements, for example, you could create an even distribution graph."*
>
> – Participant 3

7.2.2 Communication Similar to the preliminary study, discussion during building was limited. However, discussion between participants increased overtime which was potentially caused by the learning curve of the building process. In Condition 2 a similar relation was observed, as less communication occurred the more concentration was needed to understand the presented visualization. However, complex visualizations simultaneously triggered more discussion about the right interpretation of the visualization itself, rather than about insights that might be drawn about the data.

7.2.3 Understanding of data Participants in condition 1 reported having a better understanding of the given data set and accompanying context, due to the process of creating the data physicalization themselves. This process enabled them to get well acquainted with that data and created a shared understanding among the team. As participants in condition 2 were presented with a data visualization that was already created for them, similar to what happens in organisations when designers are presented with visualization created by data experts, they had a much harder time to understand the data and context. Participant 6 from Condition 2 commented:

> *"Without being the one to make the graph myself, I cannot fully understand and appreciate the whole picture. I feel like I would need someone to explain every detail of the visualization in order to properly grasp things."*
>
> – Participant 6

7.2.4 Reflection In Condition 1, the combination of the building exercise and reflection cards led to both reflection-in and -on-action. Participants noted what patterns started emerging while building and reflected on outliers that were physicalized. An example of this occurrence is the insight found in round 1 (see Figure 7). Reflection-on-action occurred naturally after the physicalization was finished, but became more explicit and conscious using the reflection cards. As participants in condition 1 did not create the data visualizations themselves, we did not observe reflection-in-action taking place.

8 DISCUSSION

Similar to the preliminary study, participants struggled to communicate during data building as this exercise required much concentration. Nevertheless, participants from Condition 1 perceived it as enjoyable because the building elements were fun to handle and participants liked seeing the physicalization emerge while building. Although communication during the building exercise became easier over time, most insights were found upon explicit reflection after the physicalization was built using the reflection cards (Condition 1). Physicalizing the data set revealed connections between all variables, whereas data visualizations (Condition 2) only showed connections between two variables at the time. This caused Condition 2 to struggle with making sense of the bigger picture and created discussion about how to interpret the data. By being able to reflect on patterns and relations between all variables of the data set, Condition 1 was able to gain a much better understanding of the data itself and the context. This resulted in more robust insights that could be implemented into the design process immediately, whereas insights found by participants in Condition 2 needed verification and specification to be deemed useful. As these insights were found after 30 minutes of physicalization, doubts concerning the time investment being worth the outcome as stated in the preliminary study are assumed to be alleviated.

8.1 Intended Use Case Scenario

Based on the results of the follow-up study an optimal use case scenario was drafted to further clarify the role and suitability of the Concreate data physicalization toolkit in the design process. The toolkit was designed for the implementation of thin data into the design process as this is what the design team and other companies struggle with most. The conducted studies showed that the toolkit is most suitable for data classified as small-thin data, meaning quantitative data collected from sensors or on-board computers but scaled down to structured and smaller data sets. Based on insights found with Concreate, one can then combine them with small-thick data coming from interviews or focus groups. If we consider Figure 1, the use of the Concreate data physicalization toolkit enables the combination of data from multiple quadrants. The follow-up study showed that the toolkit is highly effective in getting users fully immersed in a data set, increasing their understanding of the data and context. In order to allow immersion, smaller data sets seem to be ideal. Furthermore, it is still unclear the effects of Concreate when considering an overall design process (rather than a singular creative session). Therefore it can be concluded that the toolkit is most optimal in the earlier stages of the design process, when getting an understanding of the context and empathising with the targeted user. Concreate could be used prior to data visualization or iteratively in between data visualization steps, giving direction to the team and revealing opportunities to explore on a larger scale.

8.2 Limitations & Future Work

Even though both studies resulted in many valuable insights regarding the concept and validated the value of data physicalization, the studies should be repeated with a larger sample size to determine significant results. The use of static data visualizations on paper were chosen to promote easier engagement with a group that was not skilled in creating data visualizations themselves. However, a future study could compare the use of interactive data visualizations against data physicalizations to get the best of both options. Additionally, it would be valuable to test multiple types of data sets to determine their suitability to data physicalization.

The current prototype of the Concreate toolkit consists of three elements: the instructions, the mapping board and a set of reflection cards. While the focus so far has been on the latter two, the instruction element could be developed into a complete method to ensure usability and implementation in the data analysis process. This could include constructing detailed process steps, a manual for the session facilitator, and accompanying material to expand the toolkit. To determine the most effective place in the data-enabled design process, the relation of data physicalization to other steps in the process should be investigated.

9 CONCLUSION

The study with Concreate shows that data physicalization allows for a greater engagement and understanding of data due to the process of tangibly building data representations. In addition, the reflection cards allow a structured approach to reflecting on what was built, allowing pauses in the physicalization process for reflection. Data physicalization offers a highly engaging way of interacting with data, which seems to allow non-data experts to actively explore and gain a better understanding of thin data and its context. This in turn can lead to valuable insights and creative interpretations of the problem space. Our studies also suggest that an optimal balance between engagement and practicality can be obtained by an iterative switching between data physicalization and data visualization. This work is a stepping stone towards further development of the field of data physicalization to support data-enabled design.

REFERENCES

[1] Van Kollenburg, J., and Bogers, S. J., 2019. "Data-enabled design: a situated design approach that uses data as creative material when designing for intelligent ecosystems". PhD thesis, Technische Universiteit Eindhoven.

[2] Bogers, S., Frens, J., Van Kollenburg, J., Deckers, E., and Hummels, C., 2016. "Connected baby bottle: A design case study towards a framework for data-enabled design". In Proceedings of the ACM Conference on Designing Interactive Systems, pp. 301–311.

[3] Lu, J., Gomez Ortega, A., Gonçalves, M., and Bourgeois, J., 2021. "The impact of data on the role of designers and their process". *Proceedings of the Design Society,* ***1***, p. 3021–3030.

[4] Sivarajah, U., Kamal, M., Irani, Z., and Weerakkody, V., 2017. "Critical analysis of big data challenges and analytical methods". *Journal of Business Research,* ***70***, pp. 263–286.

[5] Chhabra, A., and Williams, S., 2019. Fusing data and design to supercharge innovation in products and processes. https://www.mckinsey.com/business-functions/

mckinsey-analytics/our-insights/fusing-data-and-design-to-supercharge-innovation-in-products-and-processes. Accessed: 2010-09-30.

[6] Segal, A., 2020. Form follows data: Data physicalization. https://id.iit.edu/id-news/form-followsdata/, April.

[7] Reynolds, J., 2021. Data blending explained: What it is and how to do it. https://blog.panoply.io/data-blending-what-it-is-and-how-to-do-it. Accessed: 2021-06-30.

[8] Bornakke, T., and Due, B. L., 2018. "Big–thick blending: A method for mixing analytical insights from big and thick data sources". *Big Data & Society,* ***5***(1), pp. 1–16.

[9] Dooley, R., 2016. Small data: The next big thing. https://www.forbes.com/sites/rogerdooley/2016/02/16/small-data-lindstrom/?sh=ea7a5d77870b. Accessed: 2022-01-14.

[10] Hennink, M., Hutter, I., and Bailey, A., 2010. *Qualitative research methods*. SAGE publications.

[11] Gonçalves, M., Cardoso, C., and Badke-Schaub, P., 2016. "Inspiration choices that matter: the selection of external stimuli during ideation". *Design Science,* ***2***, p. e10.

[12] Gonçalves, M., 2016. "Decoding designers' inspiration process". PhD thesis, Technische Universiteit Delft, 03.

[13] Sarkar, P. K., and Chakrabarti, A., 2008. "Studying engineering design creativity - developing a common definition and associated measures". In *Studying design creativity*, J. Gero, ed. Springer Verlag.

[14] Vasconcelos, L. A., Cardoso, C., Saaksjarvi, M., Chen, C.-C., and Crilly, N., 2016. "Inspiration and fixation: The influences of example designs and system properties in idea generation". *Journal of Mechanical Design,* ***139***, 12, p. 031101.

[15] Boden, M. A., 2007. "Creativity in a nutshell". *Think,* ***5***(15), pp. 83–96.

[16] Kun, P., Mulder, I., Götzen, A., and Kortuem, G., 2019. "Creative data work in the design process". In Proceedings of the ACM Conference on Creativity and Cognition, pp. 346–358.

[17] Schön, D. A., 1992. "Designing as reflective conversation with the materials of a design situation". *Knowledge-based systems,* ***5***(1), pp. 3–14.

[18] Botella, M., Zenasni, F., and Lubart, T., 2018. "What are the stages of the creative process? what visual art students are saying.". *Frontiers in psychology,* ***9***, p. 2266.

[19] Hao, N., Ku, Y., Liu, M., Hu, Y., Bodner, M., Grabner, R., and Fink, A., 2016. "Reflection enhances creativity: Beneficial effects of idea evaluation on idea generation". *Brain and cognition,* ***103***, 01, pp. 30–37.

[20] Moere, A. V., and Purchase, H., 2011. "On the role of design in information visualization". *Information Visualization,* ***10***(4), pp. 356–371.

[21] Fekete, J.-D., Wijk, J. J. v., Stasko, J. T., and North, C., 2008. "The value of information visualization". In *Information visualization*. Springer, pp. 1–18.

[22] Dove, G., Jones, S., Dykes, J., Brown, A., and Duffy, A., 2013. "Using data visualization in creativity workshops: a new tool in the designer's kit". In Proceedings of the ACM Conference on Creativity & Cognition, pp. 304–307.

[23] Kennedy, H., and Engebretsen, M., 2020. *Introduction: The relationships between graphs, charts, maps and meanings, feelings, engagements*. Amsterdam University Press, pp. 19–32.

[24] Mei, H., Ma, Y., Wei, Y., and Chen, W., 2018. "The design space of construction tools for information visualization: A survey". *Journal of Visual Languages & Computing,* ***44***, pp. 120–132.

[25] Jansen, Y., Dragicevic, P., Isenberg, P., Alexander, J., Karnik, A., Kildal, J., Subramanian, S., and Hornbæk, K., 2015. "Opportunities and challenges for data physicalization". In Proceedings of the ACM Conference on Human Factors in Computing Systems, pp. 3227–3236.

[26] Dragicevic, P., Jansen, Y., and Vande Moere, A., 2020. "Data physicalization". *Handbook of Human Computer Interaction*, pp. 1–51.

[27] Lee, J., 2020. A neuropsychological exploration of zoom fatigue. https://www.psychiatrictimes.com/view/psychological-exploration-zoom-fatigue. Accessed: 2020-09-17.

[28] Thudt, A., Hinrichs, U., Huron, S., and Carpendale, S., 2018. "Self-reflection and personal physicalization construction". In Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, pp. 1–13.

[29] Sosa, R., Gerrard, V., Esparza, A., Torres, R., and Napper, R., 2018. "Data objects design principles for data physicalisation". In DS 92: Proceedings of the DESIGN 2018 15th International Design Conference, pp. 1685–1696.

[30] Duarte, J. All about the #easydataviz project. http://easydataviz.co/about/. Accessed: 2021-04-02.

[31] Dorst, K., 2010. "The nature of design thinking". In The Design Thinking Research Symposium (DTRS8), pp. 131–139.

[32] Huron, S., Gourlet, P., Hinrichs, U., Hogan, T., and Jansen, Y., 2017. "Let's get physical: Promoting data physicalization in workshop formats". In Proceedings of the ACM Conference on Designing Interactive Systems, pp. 1409–1422.