

# Comparing Understanding and Memorization in Physicalization and VR Visualization

He Ren

milkywayrh@gmail.com

Bauhaus-Universität Weimar

D-99423 Weimar, Germany

Eva Hornecker

eva@ehornecker.de

Bauhaus-Universität Weimar

D-99423 Weimar, Germany

## ABSTRACT

We investigate whether presenting data in a VR visualization or as a physicalization impacts understanding and recollection. Two equivalent representations of the same data set, one in physical form and one in VR, were created. Participants answered understanding questions while they had access to the model, and were subsequently asked about the data after the model was removed. We recorded time needed to answer understanding questions and correctness rates for recollection questions. The results favour the conclusion that the virtual representation and the technical VR setup significantly inhibit participants' ability to work with the data set. Reflecting on our study setup and participants' comments, we discuss recommendations for future studies aiming at a systematic and comprehensive comparison of the differences in interacting with purely virtual and with physical data representations.

## CCS CONCEPTS

• **Human-centered computing** → **User studies; Empirical studies in HCI; Empirical studies in visualization.**

## KEYWORDS

InfoViz; VR; Immersive Visualization; user study; Information Visualization; data understanding

### ACM Reference Format:

He Ren and Eva Hornecker. 2021. Comparing Understanding and Memorization in Physicalization and VR Visualization. In *Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '21)*, February 14–17, 2021, Salzburg, Austria. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3430524.3442446>

## 1 INTRODUCTION

The young research area of Data Physicalization investigates the presentation of data in physical form, where traditional Information Visualisation focuses on purely virtual and screen-based representations. Jansen et. al. defined a Physicalization as "a physical artifact whose geometry or material properties encode data" [5]. How physicalization supports cognition, communication, learning, and problem solving is an active area of research [2, 5, 6]. Prior research has revealed that there are advantages of using physicalizations.

Jansen et al. [4] compared the differences between 3D physical and 3D on-screen bar chart visualizations, measuring time on task and error rates. The ability to touch the physical model supported understanding. Stusak et al. [11] compared the memorability of physical visualization and on-screen 2D visualizations and found that the physical bar chart model led to better retention after 2 weeks, possibly because it supported spatial memory.

While a number of studies have investigated on-screen representations, some in 3D, none have compared whether immersive 3D representations (i.e. VR) performs similarly to physicalizations. Clearly, a fundamental difference between the two types of representations is that physicalizations allow touch and manual exploration. This is still difficult to replicate in VR, but may heavily bias performance, given it allows for different interaction and exploration strategies. Nevertheless, Immersive Visualisation [1] has the potential to support spatial understanding and memory in similar ways as physical representations, allowing for vivid and increasingly realistic imagery and a feeling of presence. 'Immersive Analytics' is the "use of engaging embodied analysis tools to support data understanding and decision making" [3], and is often tied to immersive human-computer interfaces [10]. As a prior step to investigating the role of physically touch for understanding data representations, we decided to study whether there is a difference to experiencing the representation in physical 3D or as a 3D visualisation, i.e. in VR.

## 2 DATA SET, DESIGN CONSIDERATIONS AND MODEL CREATION

The criteria and requirements for the choice of a data set were that the data set should 1) have numerical attributes that can be converted to spatial representations, 2) should be non-trivial, allowing complex questions, and 3) concern a topic that participants would find interesting, but not know much about. We selected a data set on meteorite landings, where available data from NASA Open Data Portal [8, 9] covers the year 860 to year 2013.

In order to have a model that could be both produced and handled easily, we restricted it to landing records between the years 2000 and 2009 and only included meteorites that had been observed (not found). The key attributes that appeared interesting for laypeople were the year, the size or mass of the meteorite, and its landing location. Representing this on a map would have made it difficult to also indicate year and mass. Thus, to keep the model reasonably simple, but also to increase relevance for participants, we summarized location in terms of the distance to our own town, Weimar, thereby indicating 'closeness to us'. This leaves three parameters per data point to represent.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).  
TEI '21, February 14–17, 2021, Salzburg, Austria  
© 2021 Copyright held by the owner/author(s).  
ACM ISBN 978-1-4503-8213-7/21/02.  
<https://doi.org/10.1145/3430524.3442446>

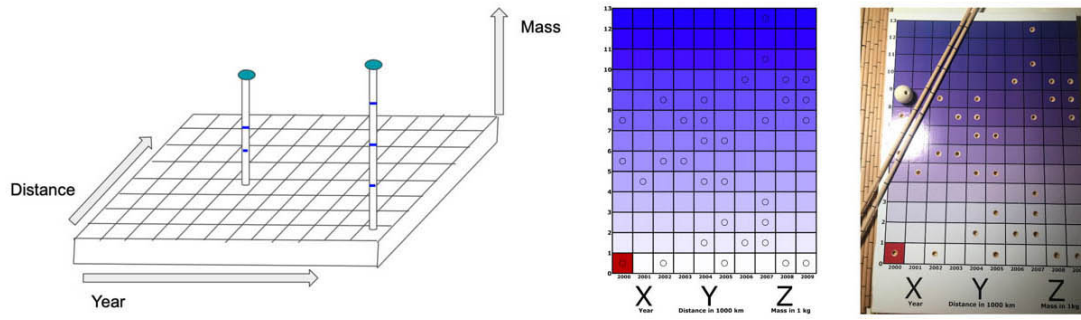


Figure 1: Principle of representation and visual design of grid surface – board after drilling and marking

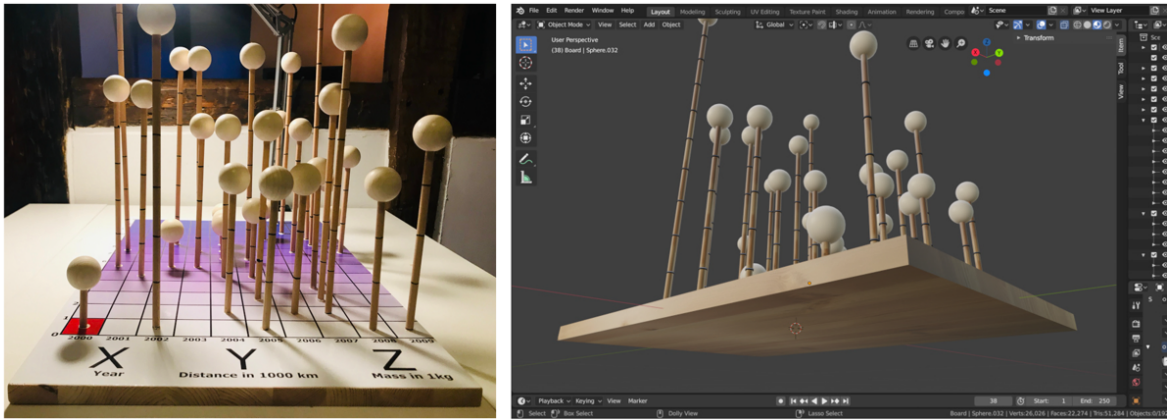


Figure 2: Final Physicalization and Virtual model in Blender

It was important for the visual design to be plausible and functional in the physical world as well as in a virtual environment. A key consideration was to ensure that both models would be as similar as possible. As in VR almost anything is possible, we first created the physical model and then replicated it in VR. The final design uses the x axis of a grid to indicate the year of landing and the y-axes for distance to our city (Figure 1, left). Wooden poles inserted in the grid represent a landing (year/distance). Two options for conveying mass were considered, the height of a pole or the size of an object on top. In some cases, the objects on adjacent poles would have been so big, they would have gotten in each other's way. This practical issue along with Jansen's finding [6] that people find it hard to compare the size of 3D-objects, had us settle for height. Our representational design thus follows Munzner's principles and design recommendations for InfoViz [7]. Munzner lists 'position on a common scale' as the most effective and expressive for conveying data. We thus utilize three spatial axis to encode three attributes.

We decided to show the grid to help orientation and legibility (see Figure 1, middle). The axes X, Y and Z and their corresponding units are labelled on the board surface. The Y axis is marked with 0 - 13, referring to distance in units of 1000 km. Because people would navigate around the model (unlike on a 2D screen), we furthermore marked the origin point in red. A colour gradient supports the perception of distance: the further away the landing site is, the

darker blue. The mass of meteorites is represented in units of 1 kg. To support reading mass value, poles are marked with a black line along units of 1 kg. In addition, an indexing pole was added at the origin point of the board (marked in red), with height 1 (1 kg mass), to provide a comparison point.

The physical model was built first (Figure 2, left). We settled for wood given its sturdiness and ease of working. After deciding on the board's size, the grid design was printed at the exact size of the board onto sticker paper, which was then glued to the board. Imprinted circles on the sticker marked the position of the poles; we could then drill holes at the exact location. The board is 60x40cm, poles are 0,8cm wide and a maximum of 50 cm high. On top of each pole, representing the meteor, is a ball of 3.5cm diameter. Deciding on the sizes took some iteration, so the board would not be too densely populated and not get too large, and the diameter of poles and balls fitting the proportions.

The virtual representation uses a 3D model and a HMD setup to present the data in VR. The free software Blender was used to generate the 3D model; Steam VR and Unity were employed to power the VR lab. The texture put on the model was identical with the printed sticker for the physicalization. To match the quality of the material exactly, photographed sections of the board were used as texture that was mapped onto these parts in the 3D model. This ensures the virtual model looks and feels identical to the physical

model. The poles also received black markings at regular intervals. There are 192 objects in the scene in total. However, to represent a round shape, such as the balls, many polygons are needed in Blender. This ultimately leads to a potential risk of heavy geometric payload in the scene, which can decrease rendering performance. Figure 2 shows both models.

To present the VR representation, an HTC Vive Pro head-mounted display was used. The Vive supports tracking spaces up to 3.5m x 3.5m, with a resolution of 1440 x 1600 pixels per eye (2880 x 1600 pixels combined), and the field of view is 110 degrees. The HTC Vive Pro was powered with Unity, which has high compatibility with the Vive and connects with SteamVR. Because size and weight of the Physicalization meant it was impossible for participants to move it, the Vive controllers were not utilized in the study, meaning the model would remain immobile in place. Teleportation was disabled to ensure participants had identical means of navigation. Since the 3D model should not fly in the air, we moreover imported a table which was adjusted to the same height as how the model was positioned for the physicalization condition of the study. Figure 3 shows how participants experienced the 3D model.

### 3 USER STUDY

As there was only one data set in two variations, a between-subjects study design was used, where two groups of participants experience either the physical or the virtual representation. This is clearly a limitation of our study. The key questions were: 1. Do humans answer questions about unknown data faster when presented with a physicalized or virtual data set? 2. Do they memorize data better when presented in a physicalized or virtual data set?

First, all participants were informed of the procedure and overall purpose of the study, and signed a consent form. However, the precise nature of research questions and recorded data was not disclosed. The exact same wording was used to describe the data set in both conditions and English was used to communicate with participants. Participants were told that they were neither allowed to touch or pick up the Physicalization nor the Virtualization.

For each participant, the study consists of four phases. In Phase 1 (familiarization), participants were exposed to the model, the coordinate system was explained, they could ask questions, and had about 5 minutes time to investigate the representation. In Phase 2, participants were asked increasingly complex questions about the data set, and the time taken until a correct answer was given was recorded. The participant was informed if their answer was incorrect and asked to check again. At the beginning of Phase 3, the model was removed, and participants were distracted with questions about an unrelated topic for several minutes (how they coped with the pandemic situation). Then, memorization questions of increasing complexity were asked, and correct answers noted. This time, the participant was not informed about correctness of answers. Lastly, Phase 4 consists of a short semi-structured interview, during which the participant was asked to reflect upon the difficulty of the experience.

The first three questions in phase 2 aim to help familiarize with the data model: Q1. What is the first and last year we are considering in the model? Q2. What is the minimum and maximum distance in the model? Q3. What is the minimum and maximum mass in

the model? The next set asks more complex questions: Q4. Which years have 4kg meteorite landings? Q5. Which distance has the most landings? Q6. How many landings are presented in total? Q7. Which year has the most landings? Q8. Which meteorite size is the most common? Q9. Which year has the 5kg meteorite landing which is closest to Weimar?

Phase 3 also starts with three simpler recall questions: Q10. What is the first and last year we are considering in the model? Q11. What is the minimum and maximum distance in the model? Q12. What is the minimum and maximum mass in the model? This is followed by questions requiring more detailed memory: Q13. Which year has the biggest meteorite landing? Q14. Which year has three landings of the same mass? Q15. Which year has the landing which is furthest away from Weimar? Q16. Is there a distance that has no landings? Q17. Is there a year which has no landings? Q18. What size is the smallest landing?

Time taken for the understanding questions in phase 2 and number of correct responses in phase 3 were recorded. Moreover, approximate position movements of all participants was sketched while they were interacting with the model during Phase 1 and 2, as indicators from which angles and how closely they observed the model.

As the models and structure of the study was similar, and the VR setup appeared to run the highest risk of issues, the VR setup had been pilot tested with one participant. Based on this, both sets of questions (understanding and memory) were expanded, in particular with three familiarization questions at the start of phase 2, designed to guide the participants attention to the dimensions and their ranges encoded in the model.

The user study was conducted during a time of social distancing rules due to the Covid-19 pandemic, and had to be postponed initially. Once official regulations allowed to meet with small groups from a maximum of two households outside, the physicalization part of the study was conducted outside, in a public park, with one participant at a time (the pandemic also prohibited touching the model, another reason not to include manual interaction). A quiet location was found, and 10 participants were booked for different time slots on 3 afternoons, so they would not meet each other. All were required to wear a mask and no companion was allowed.

Once the situation improved, and social distancing rules were relaxed, visitors at one's home were allowed again. Participants were invited to come to a make-shift VR lab in the researcher's home. All were booked for different time slots on different days, spaced apart, and windows were kept open for air circulation. Furthermore, multiple replacements of the removable parts of the HTC Vive Pro HMD that touch participants' face (foam that touches the skin around the eyes) were purchased, so that this could be replaced for each new participant to ensure hygiene. The HTC Vive Pro controllers were not used during the study. Figure 4 shows both study environments.

A convenience sample from the researchers' social circle was recruited. Each representation was experienced by 10 people (in total 20), in each case 6 male and 4 female. All participants were between 24-35 years old. None had photographic memory or were familiar with meteorite landings. Prior experience with VR was relatively low, as only one of the participants in the Virtualization group used Virtual Reality more than twice per year in gaming.

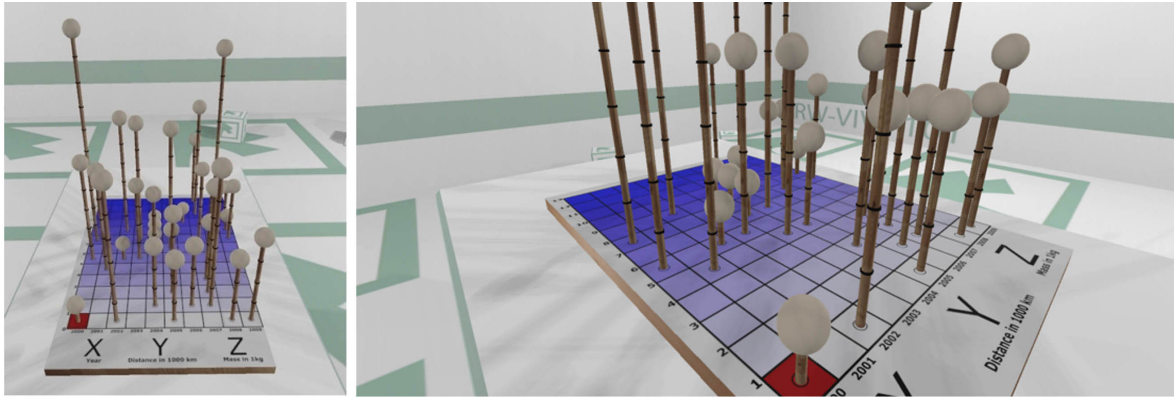


Figure 3: Virtualization from front view and side as experienced by participants in the study



Figure 4: Participants in the two conditions (in park + in makeshift VR lab)

## 4 FINDINGS

### 4.1 Statistical Analysis

We applied an independent t-test to compare the response times between both groups (Physicalization and Virtual Model). Levene's test shows that the assumption of homogeneity of variance is violated, therefore we report Welch's adjusted t-statistics. Summarizing response times for all questions, the results indicate considerably lower response times in the physicalization (M: 22.8s, SD: 2.0) compared to the VR condition (M: 33.7s, SD: 6.5) ( $t(18) = -5.09$ ,  $p < .001$ ,  $d: -2.3$ ) (see Figure 5 right). Participants looking at the physicalization were almost always faster in answering the questions than the other group, only Q6 is an outlier.

Figure 5 (left) shows response times in detail. For Q1, Q2, Q3, in the Physicalization condition, the average time taken to respond (Q1: M = 6.847, SD = 1.939; Q2: M = 9.308, SD = 3.489; Q3: M = 8.049, SD = 2.521) was lower than for the Virtualization participants (Q1: M = 14.168, SD = 4.965; Q2: M = 15.808, SD = 4.634, Q3: M = 16.348, SD = 6.280). According to Welch's adjusted t-test, the null-hypothesis has to be rejected. Participants investigating the Physicalization thus were significantly faster in responding to questions than those looking at the VR model for Q1, Q2, Q3 ( $p < .05$ ). For Q4, Q8, Q9,

participants in the Physicalization condition needed less time to respond to questions (Q4: M = 42.542, SD = 12.778; Q8: M = 23.581, SD = 7.165; Q9: M = 30.915, SD = 5.196) than the participants in the VR condition (Q4: M = 63.628, SD = 13.837, Q8: M = 51.275, SD = 23.675; Q9: M = 47.497, SD = 14.580). According to Welch's adjusted t-test, the null-hypothesis has to be rejected. Thus, response times in the Physicalization condition were significantly faster than in the VR condition Q4, Q8, Q9 ( $p < .05$ ). Albeit those participants looking at the Physicalization were slightly faster for Q5, Q6, Q7 than the other group, the difference was not significant.

Figure 6 shows the correctness rate of responses for the 8 questions asked in Phase 3. A Fisher exact test showed that there is no significant difference in correctness rates, there is no evidence of differences between Physicalization and Virtualization on short term memorability. As can be seen from the graph, differences are marginal.

### 4.2 Additional Observations

Furthermore, the approximate movement pattern of each participant was sketched during phase 1 (familiarization) and phase 2 (understanding questions). Arrows were drawn to indicate where

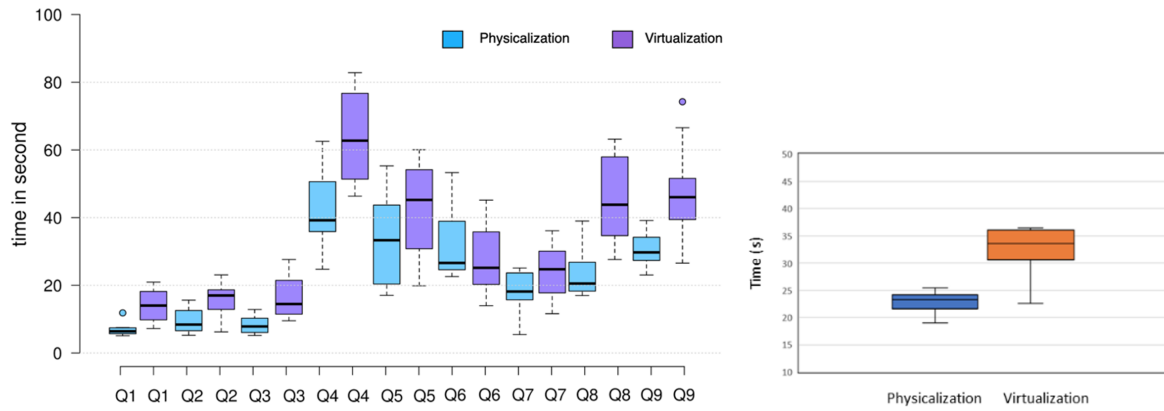


Figure 5: Response times for understanding questions (left: all questions, right: average times)

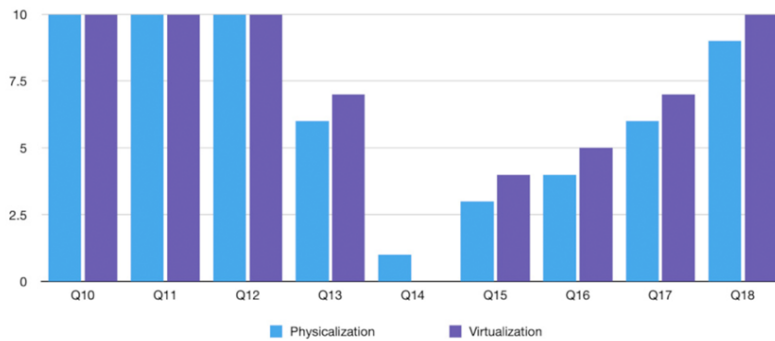


Figure 6: Correctness of answers to memorization questions)

participants moved around the physical or virtual model. For answering questions Q1 to Q9, in phase 2, each arrow was annotated with the question's number. The majority of participants did not exhibit a lot of variety in movement. A considerable number of participants stayed at the original position throughout, without exploiting that they could move around the model freely. It may be the case that they were not aware they were allowed to move (whereas they were told not to touch the model). The following illustrations show some typical movement patterns, beginning with Figure 7 for the Physicalization.

**Movement Pattern 1:** The participant walked to the end of the board around the left during Phase 1 of the study. When answering questions, they mainly stayed put at the starting position. Only for Q5 and Q6, they returned to the left side of the board. Pattern 2 is similar, primarily staying at the starting position and walking around the left side of the model for a few questions and in phase 1. These two patterns were typical for the majority of participants experiencing the Physicalization. Pattern 3 is unique and interesting. The participant barely moved during the user study, despite being told several times that they can walk around.

For the Physicalization, it is noticeable that most participants walked around the model from the left side. This is probably due to the fact that a researcher was standing on the right side of the

model at about 5 meters distance, so they felt more comfortable to go left. This effect might have even been increased in the Corona situation.

With the VR setup, participants walked to both sides of the model (see Figure 8). The participant in VR movement pattern 4 stayed at the starting position for the entire exploration phase of the study. Notably, this participant used both sides of the board to answer questions Q5, Q8 and Q9 as they saw fit. Pattern 5 is very similar; here the ability to move around was utilized for difficult questions, while easy questions were answered from the initial position. The majority of participants seeing the VR representation exhibit movement patterns similar to 4 and 5. The last movement pattern is unique, because this participant moved much more around than all others. It is worth mentioning that this person had played VR games and thus had significantly more VR experience than all other participants.

Some participants in both conditions used the height of their body (kneeling down) to adjust their eye line to judge the configuration of meteorite landings while answering questions, however, this was not recorded in the sketches during the study. Four participants from the physicalization group mentioned in the follow-up interview they found it helpful to squat down to look at the height of the poles. Although participants had been told not to touch the

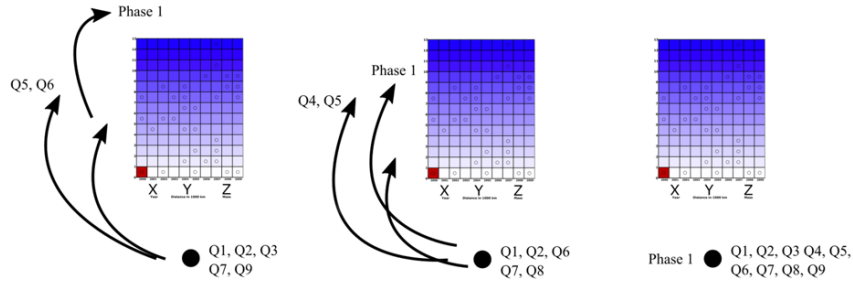


Figure 7: Physicalization movement patterns 1-2-3

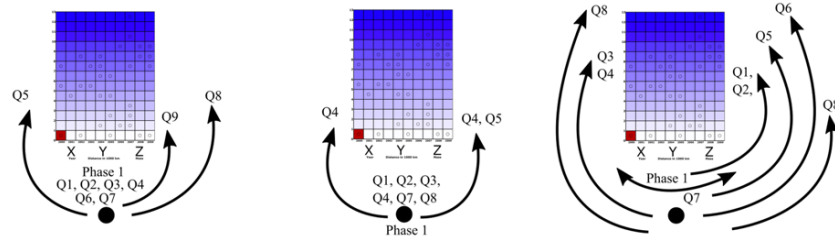


Figure 8: VR movement patterns 4-5-6

models, they tended to point at it (this was not recorded systematically). An interesting observation is that the VR participants also tended to point in mid-air, even though they could not see their hands through the headset.

In the follow-up interview, participants from both groups generally expressed that they found the dataset easy to understand. They also agreed in which questions they found the most difficult (Q14 in phase 2, Q5 in phase 3). All participants found it difficult to answer questions after the representation was removed, with several saying they had to guess answers and could only remember trends. While both groups reported using similar strategies for exploring and understanding the model, the physicalization group reported being more confident. Three participants said that the VR experience was not very smooth for them and they felt it was distracting. One needed to remove his HMD once to re-focus, another said that the VR was too slow.

## 5 DISCUSSION AND RESULTING RECOMMENDATIONS

The findings of our user study indicate a strong tendency for people to answer questions faster when presented with a Physicalization. While participants being slower in VR for the first few questions might be due to them still getting used to VR, the difference persisted for almost all further questions. Two participants from the group Virtualization admitted that they had to guess answers during Phase 3. Our observations can help to explain this. Nine out of ten participants had very little experience with VR. According to the movement patterns observed, participants tended to move with slow head movements in VR, slower than those who explored the physicalization. During interviews, three participants further expressed that they found the VR experience not smooth. One had

to remove his headset to re-focus. This provides reason to assume that participants took longer because the VR experience was not smooth. This was partially because of geometric resolution of the 3D model, where for every meteorite landing, one pole and one sphere are needed, which require a lot of polygons. Moreover, the VR study was powered by a Gigabyte Aero 15 Laptop, which has limited processing power.

An interesting effect found in the study was that participants avoided walking around the side of the model that the researcher was standing on. In VR, other people's location relative to the data representation cannot be perceived while the test subject wears a HMD. This explains why participants in VR moved around both sides. Future studies need to find solutions for this, e.g. using a camera and tracking hardware, and an audio connection, with the researcher not being in the same room to ensure both conditions are equal.

While our study has a number of limitations (in particular we did not allow participants to touch or otherwise interact with the representations) it reveals a number of issues that future studies should attend to that aim to study the potential differences between physical and VR models.

- We recommend to always build the Physicalization first, so it can be constructed in a way that is physically plausible. Nevertheless one should at this stage consider the number of polygons required in VR and to adjust the design so as to minimize rendering effort.)
- Physicalization and Virtualization must be identical in representation, appearance and material properties. From our experience, wood is a good choice, as it is rough and not reflective (mimicking reflective properties in VR takes a lot of extra effort).

- Only when 90 Hz are achieved continuously, will participants experience a smooth and pleasing VR experience. VR lag heavily impacts performance as well as user experience. Equipment should have suitable processing power for real-time graphics.
- The perceived size of representations should be as equal as possible. The perceived size of models is very important in VR because stereoscopy allows accurate estimation of object sizes. Therefore, a 3D modelling tool should be used that supports real-world measuring units using virtual rulers, e.g. Blender.
- The user study should be conducted in a controlled environment. Ideally, the entire lab (room) for the physical study should be modelled and reproduced in VR. The replication of the study environment includes presence of observers – either a virtual avatar should also be in the room or the participant should be alone with the physicalization and video-recorded.
- Participants' prior experience with Virtual Reality should be taken into consideration. Participants could then be split into several groups that can then be compared. Furthermore, a within-subjects design would result in stronger statistical results and would cancel out the impact of prior experience.

Future work could let people in VR use VR navigation methods, such as teleportation (to see what benefits or drawbacks this might have compared to physical navigation) and also to interact with it. It could be interesting to let people in VR use the Vive controllers to point and to provide them with hand-held poles for pointing when investigating the physical model. Using VR equipment to track pointing actions in both conditions would enable us to more systematically investigate interaction strategies employed in exploring data across both conditions. Furthermore, future work should investigate both short-term memorability and long-term memorability, as well the level of fatigue and cognitive load caused from exploring physicalizations and virtual models.

## ACKNOWLEDGMENTS

Thanks to all participants and to Jan Ehlers for helping with checking the statistics.

## REFERENCES

- [1] Ciro Donalek, S. George Djorgovski, Scott Davidoff, Alex Cioc, Anwell Wang, Giuseppe Longo, Jeffrey S. Norris, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew J. Graham, and Andrew J. Drake. 2014. Immersive and collaborative data visualization using virtual reality platforms. In *2014 IEEE International Conference on Big Data (Big Data)*. 609–614. <https://doi.org/10.1109/BigData.2014.7004282>
- [2] Pierre Dragicevic, Yvonne Jansen, and Andrew Vande Moere. 2019. Data Physicalization. In *Springer Handbook of Human Computer Interaction*, Jean Vanderdonckt (Ed.). Springer.
- [3] Tim Dwyer, Kim Marriott, Tobias Isenberg, Karsten Klein, Nathalie Riche, Falk Schreiber, Wolfgang Stuerzlinger, and Bruce H. Thomas. 2018. *Immersive Analytics: An Introduction*. Springer International Publishing, Cham, 1–23. [https://doi.org/10.1007/978-3-030-01388-2\\_1](https://doi.org/10.1007/978-3-030-01388-2_1)
- [4] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. 2013. Evaluating the Efficiency of Physical Visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Paris, France) (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2593–2602. <https://doi.org/10.1145/2470654.2481359>
- [5] Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. 2015. Opportunities and Challenges for Data Physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (Seoul, Republic of Korea) (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3227–3236. <https://doi.org/10.1145/2702123.2702180>
- [6] Yvonne Jansen and Kasper Hornbæk. 2016. A Psychophysical Investigation of Size as a Physical Variable. *IEEE Transactions on Visualization and Computer Graphics* 22 (2016), 479–488. Issue 1. <https://doi.org/10.1109/TVCG.2015.2467951>
- [7] Tamara Munzner. 2015. *Visualization Analysis and Design*. A K Peters/CRC Press.
- [8] NASA. 2019. Meteorite Landing. NASA Open Data Portal. Retrieved 2020-02-16 from <https://data.nasa.gov/Space-Science/Meteorite-Landings/gh4g-9sfh/data>
- [9] NASA. 2020. About NASA Open Data Portal. Retrieved 2020-02-16 from <https://data.nasa.gov/stories/s/e3dt-gtde>
- [10] Richard Skarbez, Nicholas F. Polys, J. Todd Ogle, Chris North, and Doug A. Bowman. 2019. Immersive Analytics: Theory and Research Agenda. *Frontiers in Robotics and AI* 6 (2019), 82. <https://doi.org/10.3389/frobt.2019.00082>
- [11] Simon Stusak, Jeannette Schwarz, and Andreas Butz. 2015. Evaluating the Memorability of Physical Visualizations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3247–3250. <https://doi.org/10.1145/2702123.2702248>