Revealing Incomplete Data through Scientific Visualizations in an Immersive Dome Experience

Jakub Stepanovic
jakub.stepanovic@kuleuven.be
Department of Physics and Astronomy
KU Leuven
Leuven, Belgium
Mediated Environments/Intermedia
LUCA School of Arts
Schaarbeek, Belgium

Jan Sermeus
jan.sermeus@kuleuven.be
Department of Physics and Astronomy
Faculty of Psychology and Educational Sciences
KU Leuven
Leuven, Belgium
Royal Observatory of Belgium
Laken, Belgium

Sandy Claes
sandy.claes@kuleuven.be
Institute for Media Studies
KU Leuven
Leuven, Belgium
Mediated Environments/Intermedia
LUCA School of Arts
Schaarbeek, Belgium

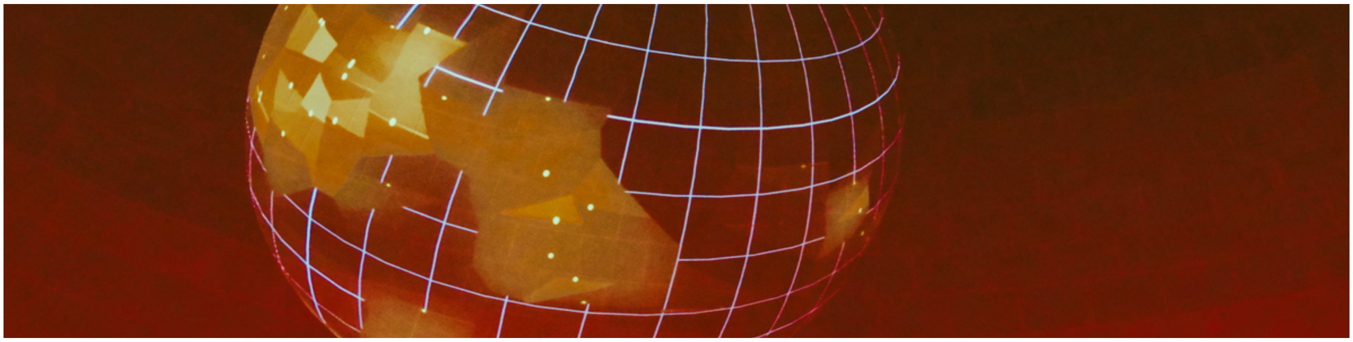

Figure 1: Visualization detail displaying unevenly distributed data in Venus's atmosphere, utilizing Voronoi cells and Value Suppressing Uncertainty Palettes.

ABSTRACT

Missing and sparse data and the associated uncertainty are inevitable in science, and their accurate portrayal in media is essential for upholding scientific transparency and credibility. Yet, in an era of conflicting information and deceptive sources, revealing uncertainty can be seen as unfavorable, in particular for science engagement media. Our study focused on conveying incomplete data on Venus's upper atmosphere to an adolescent audience using a scientific visualization designed for a planetarium's dome. Through a comparative study of visualizations with an unprocessed versus a processed dataset, we found that translating data into a Voronoi diagram can make the concept of sparse data understandable and aesthetically pleasing to a broader audience, yet it can come at costs of lower perceived details and accuracy. Additional results hint that participant's preference for visualization can differ from their perceptions of clarity and that neither the preference nor the clarity appears to be linked to participants' science literacy. Finally, we discuss the potential to design immersive science media experiences in planetariums and dome settings.

IMX '24, June 12–14, 2024, Stockholm, Sweden

ACM ISBN 979-8-4007-0503-8/24/06
https://doi.org/10.1145/3639701.3656305

CCS CONCEPTS

• **Human-centered computing → Visualization design and evaluation methods**; **Scientific visualization**.

KEYWORDS

Planetarium, Science Communication, Scientific Visualization,

ACM Reference Format:
Jakub Stepanovic, Jan Sermeus, and Sandy Claes. 2024. Revealing Incomplete Data through Scientific Visualizations in an Immersive Dome Experience . In *ACM International Conference on Interactive Media Experiences (IMX '24), June 12–14, 2024, Stockholm, Sweden.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3639701.3656305

1 INTRODUCTION

Uncertainty is a core aspect of science. Factors like measurement precision, experimental design limitations, simplification of complex phenomena, and incomplete data are just a few aspects that introduce uncertainty within a single scientific study. Furthermore, because scientific theories are revised or refined, the interpretations of data may change over time. Scientists are aware of this uncertainty; they recognize it and embrace it as a fundamental element of the scientific process. Yet, despite the awareness, there are several reasons why communicating uncertainty to the public remains a challenge [72], of which the lack of expression formats is one.

Poorly communicated uncertainties are often susceptible to misinterpretation by the general public and can lead to overconfidence and distrust in science [45]. Given that the current information landscape is saturated with inconsistent messages (which can magnify yearnings for a sense of definite clarity [60]), it is not surprising that scientists are concerned that the portrayal of uncertainty can lead to inducing negative feelings and reducing in trustworthiness [38]. Yet, effective communication of uncertainty is argued to play a crucial role in maintaining public trust in science [16]. This was highlighted during the COVID-19 pandemic when communicating uncertainties was reported to mitigate the negative impact of exposure to conflicting information [12], whereas not mentioning uncertainty led to decreased trust in science and a decreased willingness towards vaccinations and nonpharmaceutical interventions [4].

While there has been an increasing interest of scholars to reveal missing data and uncertainty through traditional, single screen-based science engagement media, such as science news visualizations and weather forecasts [42, 74, 77], its application in other types of science engagement media, such as digital dome theaters that are often used in planetariums, are under-explored. Such digital domes, however, have seen a growth in applications [71], and with the uptake of portable lower-cost dome projections, it is expected to further expand [34]. The levels of audience engagement in digital domes can be remarkably high due to the immediacy of the mediated, nonlinear exploration of more or less real-time data [84]. Indeed, because the user is surrounded by the screen on most sides, the experience is more immersive than the display of visualization on a traditional computer or projection screen [37].

In this multifaceted study, we bridge design, media experience, science communication, and science education in the context of planetary science, to investigate how to reveal incomplete data through scientific visualizations. Incomplete data, a particularly relevant aspect of scientific uncertainty, comprises both missing data (unexpected gaps in the dataset) and sparse data (a low density of data points). To the best of our knowledge, no studies exist that ask similar questions due to the interdisciplinary nature of our work. We explore the visualizations within a planetary science project that researches the atmosphere of Venus and aims to communicate its findings to an adolescent audience in the dome of a public planetarium (Figure 1). Following calls from scholars emphasizing the need for evidence-based, audience-driven insights in science engagement media [44], we adopted a Research through Design approach [86], and developed two dynamic scientific visualizations that are part of the narrative of the planetarium's lecture. After several iterative design cycles informed by feedback from the audience and scientists, our contributions include: 1) implications for the design of visualizations that reveal uncertainty in science engagement media settings, 2) empirical evidence of how two different types of uncertainty representations influence perceptions of clarity, accuracy, and overall attractiveness, and 3) insights and reflections on our design process to create science media content for immersive dome experiences.

2 RELATED WORK

In the following, we discuss the challenges towards i) the design of content (i.e. revealing missing data in scientific visualization and the specific requirements of astronomical visualizations) science engagement media experiences, ii) the role of the setting, and iii) the influence of context.

2.1 Visualizing Missing Data and Uncertainty

Visualizing missing data and uncertainty is a complex challenge in data representation, requiring a delicate balance between conveying the dataset's limitations and maintaining clarity in the presentation [77]. While the inclusion of missing data and uncertain information in scientific visualizations is on the rise across domains [42], there is no universal approach. Various techniques such as emptiness, fuzziness, and explanation are employed to signify missing values [5], while indicators like error bars, confidence intervals, or color gradients communicate the level of uncertainty [61]. Still, designing the visualizations so that they are understandable to the lay audience remains a challenge [61, 74], and authors might omit to communicate uncertainty for various reasons: 1) scientists might assume that uncertainty is already widely recognized as common knowledge, 2) uncertainty may be seen as unnecessary detail or as too difficult for a lay audience, 3) scientists might be worried that discussing uncertainty opens the door to criticism or 4) they might not know how to express uncertainty [30].

2.2 Presenting Astronomical Visualizations

Astronomical data are captured by instruments operating across the entire electromagnetic spectrum, from x-rays over the visual spectrum to radio waves. Data from all of these wavelengths can be used to create scientific visualizations [68], often requiring a transformation of information that is beyond the visible spectrum to a visually attractive image [69]. This process can involve artistic decisions to appeal to a non-expert audience [53]. Space agencies supply their images with captions to detail the visualization content and how was it created [81]; however, this additional information can get lost over time. For example, it has recently come to the public's attention [33] that the depiction of Neptune, deep blue in contrast to the pale blue color of Uranus, is not accurate [41], something that was well-known within the planetary science community. Lay audiences are not often exposed to the interpretations that lie behind the construction of astronomical images [75]. Consequently, these constructed depictions may be perceived by viewers as true to nature [27]. This perception is partly rooted in a photographic heritage often linked with unaltered documentation [6], it may also be because the individual creators of these constructed depictions are often not credited [27].

In planetariums, visuals depict celestial objects based on limited available data. For example, the map of Venus's visible atmosphere released by the National Aeronautics and Space Administration (NASA) and used by planetariums, including European Southern Observatory's Supernova planetarium [57] is based on a single visible-frequency image, captured in 1974 by Mariner 10 mission [79], which was cloned to cover the entire sphere [73]. This illustrates that, while these renderings are done based on the best available scientific data at the date of making, they often involve artistic decisions to fill up incomplete data.

2.3 Designing Science Engagement Media

Social media took over the traditional media channels as the public's primary reference concerning science and technology. While this trend enhances information accessibility, it also poses challenges for non-experts in making informed science-related decisions [10]. As such, public spaces such as planetariums prove to be vital venues where the lay audience gets to interact and be immersed in curated learning resources for astronomy and space exploration education [47, 59]. They employ the 'exploranation' concept, a combination of exploration and explanation [84]. Notable examples of exploranatory designs include interactive visualization tables in museums, e.g. to explore the inside of an ancient mummy [84], interactive exhibits to visualize data on climate change [15] and interactive public virtual reality experiences to convey the concept of nanotechnology in relation to data [40]. Furthermore, the projection dome (as the central tool planetariums use to deliver content as immersive visuals [85]) is deployed to illustrate complex astronomical concepts in an attractive [43], and relatable [64] manner. When applying the closed, sensory experiences domes provide, this environment enhances the educational content [78]. Despite the growth in dome use [71], few studies exist that focus on the unique design challenges in creating content that effectively utilizes the dome's experiential properties [70].

2.4 Revealing uncertainty in Science Communication and Education

Scientific visualizations serve as tools in both science communication and science education; therefore, it is important to contextualize both settings. Science Communication is, as an area of research, younger and less established than science education from which it could learn [9]. Unfortunately, uncertainty tends to be neglected in both fields.

In Science Education, uncertainty is often presented in a vague form or even deflated [1], making it difficult for students to grasp how 'real' science operates [35, 45, 55]. One noteworthy exception is the acknowledgment of the tentativeness of science (i.e., scientific knowledge remains subject to change and is never absolute or certain). Tentativeness is a part of the Science Education concept 'Nature-of-Science' (NoS) [51], encompassing what constitutes science, how it operates, and its limitations.

From science education literature [20, 50] it is clear that both students and teachers [2] mostly have a naive understanding of NoS. Yet, getting students beyond a naive understanding of NoS is essential for countering scientism and skepticism [29]. However, while tentativeness represents a part of uncertainty in science, there are many more. Therefore, further clarification and contextualization of NoS are needed [32], incorporating philosophical and historical perspectives [56].

3 METHODOLOGY

We designed two distinct segments to visualize the identical dataset of Venus's atmosphere and contrasted them in a comparative study. One showed the data in a scientific visualization that is relatively unprocessed, referred to as the Unprocessed Segment - USeg, whereas the other contains processing, referred to as the Processed Segment - PSeg. Regarding the designation, it is essential to underscore that raw data, of course, necessitates processing for effective visualization; in our approach, the USeg retains a comparatively lower degree of processing and interpretation compared to the PSeg.

To develop the visualization segments, the study adopted a Research through Design approach [86], i.e., the design was presented through iterative design cycles (DC) to relevant stakeholders, including adolescents and teachers (n = 139; DC1-2, DC4), and scientists (n = 8; DC3), and their input was evaluated and adapted to address shortcomings and build upon previous successes to improve the design. We selected adolescents as the primary demographic target due to curriculum-related factors (scientific disciplines are typically introduced during the later stages of secondary education, thereby NoS introduction becomes relevant). Moreover, adolescents who participate as part of school visits include individuals who may not have otherwise attended independently, which broadens the visualizations' reach. The DC1-2 and DC4 were shown in the planetarium to high school classes, and we switched the order of presenting USeg and PSeg between schools to mitigate possible anchoring biases. DC3 was screened remotely to a group of scientists.

The study is descriptive with a large qualitative component. Data in DC1-2 and DC4 were collected through a survey that evaluated 1) the participants' preference, 2) their perception of clarity in conveying the message "There is missing data in the data collected on Venus," and 3) their understanding of missing data in science. Because the latter might be confounded by a prior interest and familiarity with science, we also measured 4) the participants' Science Capital [7]. Following calls to evaluate whether to gather additional information such as gender or ethnicity [18], we omitted to collect the data because we focused on entire school classes without excluding anyone based on demographic factors. We also aimed to create an inclusive environment where all participants could engage without concerns about being labeled or stereotyped. The survey questions and the Science Capital analysis logic and workflow can be found in Appendix A.1 and A.2. DC3 data were collected through interviews, asking for perceptions of the visualizations' clarity and overall attractiveness. The process resulted in an optimized design and insights contributing to theory. An overview of the cycles, including the location, method of gathering the feedback, and the composition of participants, is depicted in Table 1.

The survey (DC1-2 and DC4) was administered online using Microsoft Forms and made accessible by a QR code, which allowed participants to use their mobile devices to complete the survey, bypassing the low light conditions of the planetarium. We thematically analyzed the responses through inductive open coding and removed entries that were either completed before the screening concluded, or did not provide qualitative reasoning behind their choices. To adhere to the European data protection regulations, we restricted the data collection only to those aged 16 and over and had participants give their informed consent at the start of the survey. The study was examined and approved by the ethics and privacy committees of our institution before the start of the data collection.

Table 1: Overview of the design cycles process.

Design Cycle	Location	Audience	Method
DC1	Planet. lobby	Students (n = 5), Teachers (n = 1)	Survey
DC2	Planet. dome	Students (n = 10), Teachers (n = 1)	Survey
DC3	Remote	Scientists (n = 8)	Interview
DC4	Planet. dome	Students (n = 111), Teachers (n = 11)	Survey

4 PLANETARIUM CONTEXT

4.1 Environment

The planetarium for which we designed the science visualizations has a level dome with a 23-meter diameter, a concentric seating arrangement for up to 300 people, and a unidirectional digital projection system. Fifty percent of its visitors are students attending as part of school groups to experience an educational activity [58].

4.2 Content

The current content offered in the Planetarium's dome consists either of film projections or live lectures supported by software simulation, both featuring smoothly rendered realistic visuals. The astronomical software used for the lectures is `SkyExplorer 2021` by the RSA Cosmos. The developer markets SkyExplorer as having "the most realistic rendering in the industry" [22]. The software allows the presenters to showcase the night sky and photorealistically illustrate celestial objects like stars, planets, and satellites; however, it does not explicitly show missing data (Figure 2). The designs that are used in our study depart from the existing realistic imagery of the planetarium.

4.3 Interactivity

Currently, the narrators at the Brussels Planetarium interact with the audience through questions and answers, where the questions can be initiated both by the presenter or, to a lesser degree, by a member of the audience. The Q&As are typically limited to directive questions to advance the lecture or transition to the next topic. Our study was built to expand upon the existing operations, so it includes narration but does not allow for advanced forms of interactivity that stimulate discussions and directly influence the content that is presented.

4.4 Structural Limitations

The visualization design for planetariums faces an additional challenge due to the architectural properties of their dome shape: projecting rectilinear images onto the domed ceiling introduces hemispherical distortion. The Planetarium employs directional projection to mitigate this effect; nevertheless, the distortion is particularly apparent with images including straight lines, which is more pronounced for viewers seated towards the periphery of the theater, as illustrated in Figure 3. We applied subtle pincushion distortion to compensate for the dome curvature.

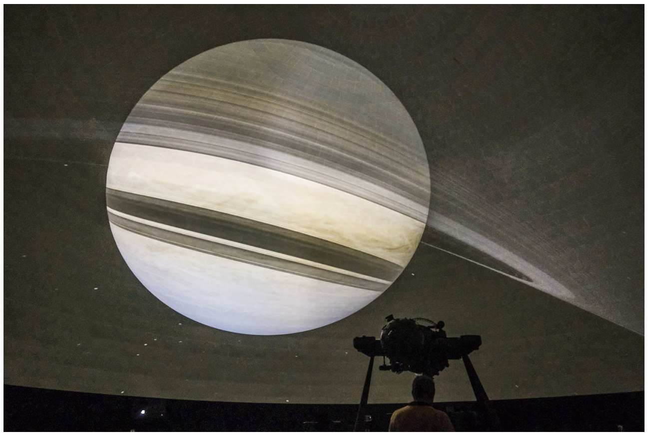

Figure 2: Planetariums tend to use photorealistic imagery that does not explicitly show missing data. Pictured is a depiction of Saturn during a lecture of the Brussels Planetarium.

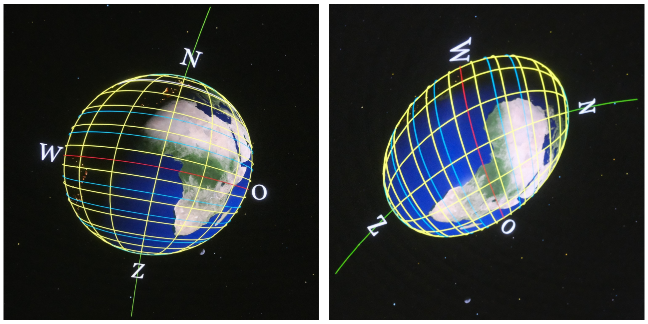

Figure 3: An example of the impact of the viewer's position on the distortion of a rectilinear image in the dome: The left image is captured from the center of the seating, while the right image is taken from the left edge of the auditorium.

4.5 Data

We worked with data collected by the SOIR (Solar Occultation Infra Red) instrument [54] aboard the ESA Venus Express spacecraft between 2005-2016. The data are vertical profiles of the upper atmosphere (altitude 60-220 kilometers) consisting of temperature, pressure, and the density of the following chemical species: CO_2, H_2O, HCl, and HF. The data is sparse and it has, due to the elliptical orbit of Venus Express, a prominent gap in the coverage between the +40 and +60 degrees latitude, which is noticeable on a rectangular 2D scatterplot of the data (Figure 4). This form; however, is not suitable for showing in the dome, due to the dome's rounded format. Besides the distortion inherent in the structural limitation described above, presenting a bordered scatterplot on a 360-degree screen

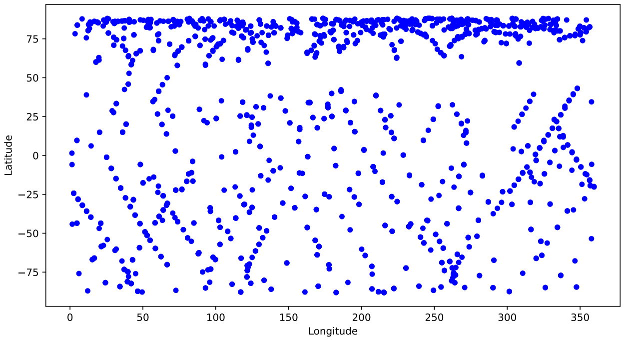

Figure 4: A plate carrée projected scatterplot of the locations where the SOIR instrument gathered data on Venus's Atmosphere.

can reduce the audience immersion, and it further compromises the plot's accuracy and interpretability.

5 DESIGN CHOICES

Two distinct visualizations were designed: the USeg approach provided an unaltered depiction of the data points in their original spatial arrangement, the PSeg enhanced interpretability by revealing the dataset's underlying patterns. We focused on revealing the data on a rotating photorealistic representation of Venus, a decision directed by the existing visuals and conditions of the planetarium, as well as insights from our preparatory interviews and observations that informed the initial design. As the starting point, we utilized NASA's cylindrical map of Venus's visible atmosphere derived from the Mariner 10 mission [73], warped it on a sphere programmed to spin around its Y-axis, and positioned the camera perpendicular to the equator.

USeg. Our first version of the scientific visualization, running on the `Processing` software [67], removed the visible cloud layer to show Venus's surface based on NASA's Magellan scans [49]. Above the surface, we rendered columns in the data's locations to show the vertical profiles (Figure 5a). However, through discussions between the authors emerged that this prototype could be susceptible to misinterpretation: When the columns had sufficient scale to hold any details to distinguish the species, their proportion was too large relative to the planet; moreover, the gap between the columns and the ground was small enough to imply that the data came from Venus's surface. If the columns covered the appropriate area where the data were collected, they became invisible. These findings led to abandoning the representation through the columns and the surface imagery.

The following version used a Python code to scatterplot locations on a transparent 3D sphere by using Matplotlib [39], NumPy [36], and Pandas [62] libraries. Pandas read spherical coordinates from CSV files, NumPy converted them to Cartesian coordinates, and Matplotlib plotted them on the sphere. Initially, we plotted each element as a point with corresponding colors; however, similarly to the first version, this suffered from a spatial proximity saturation: the data collection locations were so close to each other that it rendered the points unclear. We then simplified it by plotting the location data for all species in one color and introducing an outline for locations that overlap (Figure 5b).

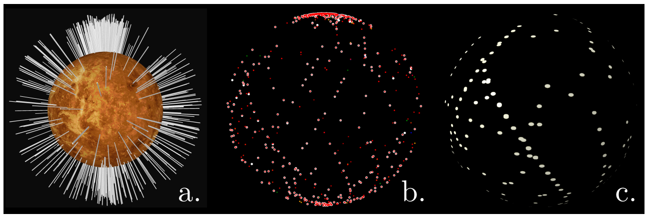

Figure 5: Initial iterations of the unprocessed segment show simplifications from vertical profiles with surface texture (a) through scatterplot distinguishing species (b) to a one-color version that reduces possible misinterpretations (c).

The transparent sphere showed points simultaneously from both the front and back sides, reducing clarity. Hence, we opted to make the sphere opaque with a black fill color to maintain a neutral appearance for the sphere and plotted the locations without additional information (Figure 5c). This ensured the visualization remains comprehensible without the need for a legend, as introducing one could disrupt the immersive experience by being susceptible to the dome distortion, and text would introduce further constraint by the limited number of characters viewers can read per second (whereas graphic visualization enables the simultaneous processing of the entire two-dimensional information space) [31].

PSeg. A common cartographic practice to display spatial data on Earth is through Choropleth maps, which use pseudocolor corresponding with an aggregate summary of geographic data. Choropleth maps are also frequently featured in school materials [82]. However, they need the plane to be divided into distinct districts, usually defined by geographical or political boundaries, so they cannot visualize data without divisions, as is required here. Heat maps, on the other hand, might be promising, as they present values and ranges as a continuous spectrum that does not rely on discrete cells [83]. We deployed heat maps in spatial and grid layouts, rendered through geometric and mapping Python libraries Cartopy [17], Matplotlib [39], NumPy [36], Pandas [62], and SciPy [80]. However, with the sparse dataset we worked with, we found that both spatial and grid heat maps did not offer the resolution to have a narrative quality (Figure 6a and 6b).

To address the challenges associated with sparse spatial data on the undivided plane, we adapted a Voronoi diagram [8], which provided cells that correspond to the area closest to a particular coordinate, indicating the areas where the data are distributed. Regions with higher population density result in smaller Voronoi cells, while areas with lower population density manifest as larger cells. We utilized the same Python libraries as for the heat maps to visualize the data using Voronoi cells: Matplotlib customized the plot layout, Pandas loaded latitude and longitude data, NumPy processed them, Cartopy formed a PlateCarree Projection, and SciPy calculated the Voronoi diagram. We assigned a color to each cell where the saturation is inversely proportional to its size and thus proportional to the data density. This allows viewers to interpret and comprehend the spatial relationships and divisions without

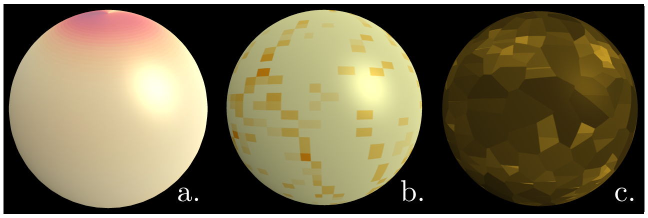

Figure 6: Initial iterations of the unprocessed segment: The spatial (a) and grid (b) heat maps did not offer the resolution to show the data distribution on a sphere, so we adopted Voronoi cells (c).

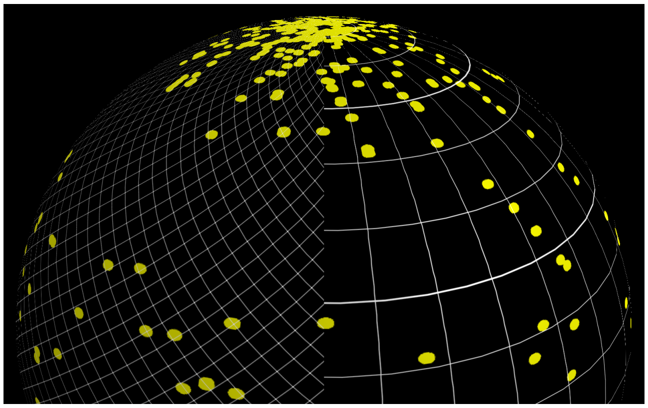

Figure 7: The effect of different grids on the missing space: Angled lines (left) clarify the volume of the sphere, and using meridians and parallels (right) provide additional information on the locations of the poles.

understanding the mathematical principles underlying Voronoi diagrams (Figure 6c). Then, we adapted Value Suppressing Uncertainty Palettes, known for enhanced interpretability compared to bivariate maps [21], to emphasize the data distribution further. To do so, we suppressed the diagram's cells with low-density values with black color, which also created a neutral base comparable to USeg.

Refining the Segments. Plotting the data on a sphere made the coverage gap between +40 and +60 degrees less noticeable. To keep the focus on this gap, we adjusted the camera position to be perpendicular to the 35th parallel north latitude. We also incorporated a white grid on the neutral sphere to emphasize the space without data. This design choice aligns with the principles of wabi-sabi aesthetics in the digital realm, embracing the beauty of imperfections and the transience of data representation [48]. The grid serves as a subtle yet intentional element, accentuating the spaces where the data are absent, aiding in clarity for the viewer. The first draft composed lines in 45 degrees; however, driven by the aesthetic decisions of authors and a recommendation of a planetary researcher, we ended up using a grid of traditional parallels and meridians. The application of longitude and latitude lines also clarified the location of the poles without the need for introducing an axis (Figure 7).

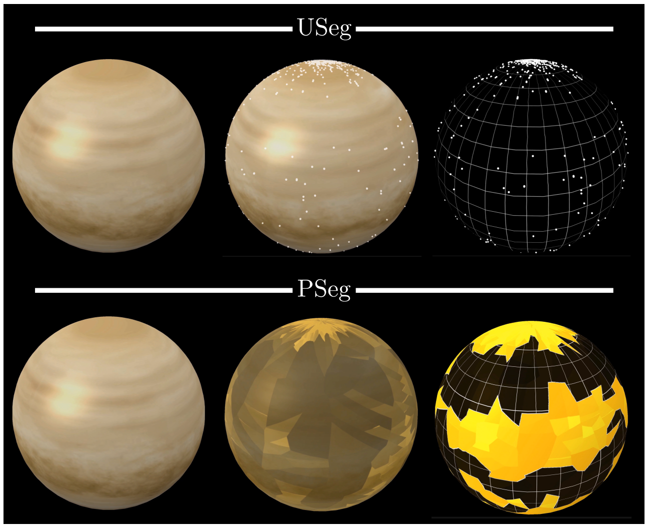

Figure 8: Animation phases of the segments during the DC2.

Transitions. To transition from the photorealistic representations to our uncertainty visualization, we initially created an animated transition through a scanning stripe that revealed the data. However, we dropped it in favor of a crossfade transition, which improved simplicity. For USeg, the plotted points initially appeared atop the cloud layer; subsequently, the cloud layer faded out to reveal the neutral sphere with the white grid lines of parallels and meridians. In PSeg, the cloud layer crossfaded into the original Voronoi diagram and followed with increasing contrast and suppressing the low-density polygons, which revealed an identical grid to the USeg. We also accentuated the border between highlighted and suppressed areas with a white outline. The animation steps were programmed in `Arena` [25] and are shown in Figure 8.

Narration. Our narration complemented the existing narrative of the Planetarium's lecture. The limitations of visual observation when studying Venus were emphasized, as well as the importance of satellite missions to collect comprehensive data. It introduced the locations where the Venus Express gathered atmospheric data, their uneven distribution, and the challenge scientists face to construct a cohesive model of Venus' atmosphere even if there are gaps in available data. The full script is available in Appendix A.3.

6 RESULTS

We will first discuss the quantitative data from DC1, DC2, and DC4, and then present the qualitative data on segments' preference and clarity for each DC, together with DC-specific discussions and implications for the design. Detailed distribution of quantitative data on visualization preference and perceived clarity across each cycle can be found in Appendix A.4.

In DC1, DC2, and DC4, we collected valid responses from 139 participants; 126 were 16-24 years old, and 13 were 25 and over. The majority (n = 74) saw the message, that there is missing data in the data collected on Venus, as somewhat or very clear, compared to 29 somewhat or very unclear. Most participants indicated PSeg as both the segment that they prefer (n = 87) and perceived as clearest

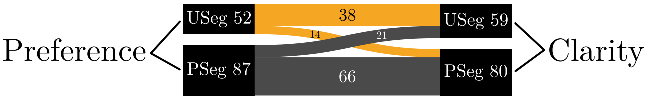

Figure 9: Alluvial diagram of the participants (n=139, DC1, DC2, DC4) preference and clarity. While most are congruent, 35 participants prefer a different segment to the one they think is clearer.

(n = 80); 35 people preferred a different segment than the one they saw as clearest (Figure 9).

Most participants (n = 88) reported that missing data in science, in general, is 'very frequent' or 'frequent,' 33 answered 'common,' and 7 answered 'rare' or 'very rare.' 89 participants saw the gaps in the collected data about the atmosphere on Venus as 'enormous' or 'large,' and 21 as 'small' or 'minimal.' 97 participants said that the available data on the atmosphere of Venus, even with the gaps, is 'very useful' or 'useful;' 11 participants said that it is somewhat or very useless. Participants' Science Capital levels were as follows 'high' = 4, 'medium' = 98, and 'low' = 37. There was no observed correlation between participants' Science Capital levels and their preferred segment, or the segment they perceived as clearer. The Science Capital levels also did not correspond with participants' understanding of the message.

6.1 DC1:pilot study

Four students and the teacher indicated their preference for PSeg, citing clarity as the motive, while one student preferred USeg. Three students reported that they preferred a different segment than the one they perceived as clearer. Student #8 preferred USeg, because "*[it] is more detailed. It gives off a professional, scientific view on the topic...;*" however, Student #8 also reported that "*[PSeg] is more obvious with the coloured areas, so you need to focus less on the dots.*" Student #12 preferred PSeg, because "*[it] made it more clear... where the points made it harder to imagine; [however, USeg] only showed the parts where data has really been collected and didn't average it out, so the surface area looks much more empty... [PSeg] looks much more filled.*" In total, three students stated that PSeg is clearer in showing missing data, while two students and the teacher stated it about USeg.

Three students pointed out that PSeg appears to have more data than USeg: "*The [PSeg] gives the impression of an higher average of exploration because it covers whole areas,*" - Student #9; "*You see... that we have more data [in PSeg] than in [USeg]... I disliked that in [PSeg] it looked like we have almost more than a half investigated, but if you see [USeg] you see it's not the case,*" - Student #11. One student (#10) referred to the visualization as "ground data."

Discussion and Implications for the Next Design Cycle. We noted the remarks about the difference in the data amount between USeg and PSeg and the association with the ground as worrying; however, we felt like the group was too small to draw conclusions from. As this iteration exhibited no discernible failures in the survey functionality, we proceeded to transition the study into the dome.

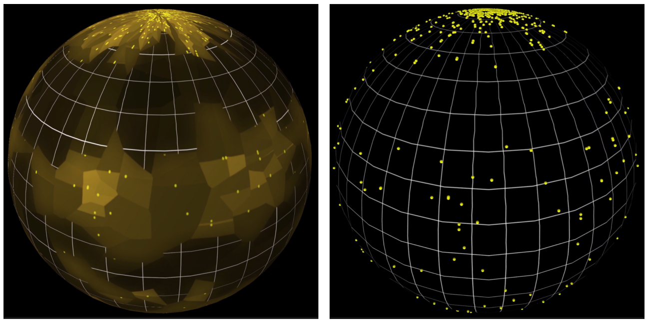

Figure 10: Scientific visualization in DC3-DC4. Left: Pseg; Right: USeg.

6.2 DC2: Possible Misconception

All (n = 11) participants cited clarity and visual appeal as their reason for preference; all but one indicated a preference for PSeg. Every participant stated that PSeg was clearer: "*[PSeg] gives a clearer representation because... Fields are easier to comprehend than a lot of single dots,*" - Student #18; "*...It is a lot easier to convey how much is missing because of the highlighted area compared to just the location points,*" - Student #23. Similar to the statements in DC1, two students in DC2 perceived that PSeg shows more data than USeg.

Discussion and Implications for the Design. As we recognized that students continued to overestimate the amount of data in PSeg, we discontinued the research data collection in this DC to refine the design of the segments. We superimposed the unprocessed location points on the Voronoi diagram, underscoring the congruence of the dataset with the USeg dataset. We also simplified the transition from the cloud layer to the data layer of both USeg and Pseg by skipping the middle step and going directly to the "highlighted stage," and optimized the colors to have the segments as comparable as possible. Furthermore, while making adjustments, we also addressed the student's comment from DC1, who linked the data to the ground. We removed PSeg's lines accentuating the border between highlighted and darkened areas, decreased the contrast, and applied light Gaussian blur to smoothen the edges instead. These changes made the visualization gradual to maintain the non-rigid nature of the atmosphere. The resulting scientific visualizations are depicted in Figure 10.

6.3 DC3: scientists' input

Scientists (n = 8) reported that the clarity of USeg surpasses that of PSeg; therefore, they would not use PSeg to communicate their work to other scientists: "*Scientifically I don't know how we could use [PSeg] ... I have a feeling that, for other scientists, we'd not use this,*" - Scientist #2. This was due to the complexity of Voronoi's calculation; three scientists said that it was new to them. However, they also found that PSeg as "*... a good visualization to make [the message] accessible,*" and that the version with the points superimposed on Voronoi "*carries the message pretty well,*" - Scientist #1.

Discussion and Implications for the Next Design Cycle. Since the scientists did not express worries about accuracy, we proceeded to include the visualizations in DC4.

6.4 DC4: emerging preferences

When reasoning for their preferred segment and perceived clarity, both USeg and PSeg decisions were driven by visual appearance and clarity. The aesthetics of the USeg was described as "*beautiful to look at*," - Student #175, and that "*[USeg] appealed to me more and it's more fun because in the atlas or other things we only look at images [like PSeg]*," - Student #142. PSeg was described as "*[the message is] shown in a bigger and more obvious way,*" - Student #173, and that "*The color gives more clarification, visual aid*," - Student #192. "*...it is more visual and is better to understand,*" - Student #184. Teacher #172 called PSeg as "*a bit old-fashioned*."

The clarity of USeg was described as that it "*allows you to view and analyze the data more easily*," - Student #159, and that it "*provides a better overview of the specific points examined. The planes are too general*," - Student #163. "*You don't have any other things [in USeg] to confuse with the dots*," - Student #154 added. On the other hand, PSeg was described as that "*It feels clearer and more organized. I understand the concept better*," - Student #197. "*I think [PSeg] is better to understand and see [the] pat[t]ern where they still have to do research,*" - Student #155. "*You have more correct insight of what [PSeg] says,*" - Student #108. Comparing the segments, Student #154 said, "*I find the [PSeg] a bit chaotic, the points are more accurate.*." "*It is a lot more detailed*," Student #94 added. At the same time, Students said that "*[PSeg] was more accurate than [USeg] and I understood the overall theme better,*" - Student #201, and "*With polygons, [PSeg] is clearer*," - Student #168.

The qualitative data also hinted at other patterns without any specific inquiry prompting students' remarks. Students were surprised at the small amount of data: "*I realized how little data we have about the atmosphere of Venus*," - Student #68. "*...I thought there was more [data]*," Student #194. For Student #173, the surprise was emotional: "*I kinda learned that [Venus] is very much not known and that's sad.*" One student (#160); however, reported that the amount of data was more than expected: "*We already know a whole lot about Venus, mor[e] than expected*". Students also recognized the temporality and that there is still progress to be made: "*...we haven't learnt all there is to learn*," - Student #90, and that "*...Science ensures that we can achieve more in the future*," - Student #92, and that "*It is important that we know that we do not know everything...*," - Student #103. Student #109 highlighted that one could see the distribution more clearly because there was a grid. Only one student (#46) said that PSeg shows more data than USeg.

7 DISCUSSION

7.1 Visualization Challenges

7.1.1 Perceiving Usefulness of Sparse Data. Previous research shows that authors might fear that showing missing data can undermine the audience's trust in the presented data [38]. However, our results showed that the majority of students in our study saw even small amounts of data as useful for contributing to scientific understanding. This viewpoint aligns with studies indicating that emphasizing missing values in graphs is generally preferable to minimizing or omitting information [76], and that transparency about uncertainty is desirable [16]. Nevertheless, some students expressed surprise at the extent of missing data in our scientific visualization, highlighting a potential lack of awareness among the general public regarding the prevalence of missing data. This underscores the need for heightened efforts in communicating incomplete and uncertain data to the public.

7.1.2 Deploying Voronoi Diagrams for Sparse Datasets. We learned that common cartographic methods, such as heat maps or clusters, may not effectively convey the underlying patterns when data points are sparse. We worked with a Voronoi diagram as a possible method to provide spatial partitioning based on the given data points, maintaining physical meaning. According to our results, it also introduced an attractive visual form, which concurs with prior research that has demonstrated the Voronoi diagrams as useful artistic tools [46]. However, the complexity of Voronoi diagrams can lead to misinterpretation and misconceptions. The audience might expect that bigger equals more (such as the size-weight illusion [63]); yet, the Voronoi diagram partitions the most data-dense areas into the smallest cells, whilst the largest cells represent the least populated areas. We applied choropleth color coding [82] to the Voronoi cells to ensure that cell size is not the sole indicator of data density, and increased the scientific visualization's relatability to a lay audience. The overall preference for the Voronoi diagram (PSeg) in our study underscores that it was successful, and that overlapping the Voronoi diagram with a scatterplot reduces the misinterpretation of the data quantity (from five out of fifteen students in DC1 and DC2 to one out of 111 students in DC4). However, while Voronoi diagrams retain the precision and accuracy of the data, our study also showed that converting them into a geometric representation can mask the clarity offered by the scatterplots: certain individuals from both the public and scientific communities expressed a belief that Voronoi diagrams are a less clear and less detailed representation than scatterplots.

Still, the Voronoi diagram appears to be an advantageous option to make the concept more understandable and aesthetically pleasing to the broader general audience than the scatterplot, which can also overwhelm or confuse the adolescent audience. The overall higher perceived clarity for the Voronoi diagram (PSeg) in our study shows its potential for conveying complex data and invites further research to help mitigate the challenges of perceived clarity and details.

7.1.3 Visualizing the Element of Absence. Our findings indicate that the utilization of a grid effectively communicated the concept of incompleteness, concurrently amplifying the visual representation of missing data and providing supplementary context, including the positioning of the poles. This contrasts previous research findings that grid lines can clutter graphs [3]. It highlights the usefulness of further research into visualization techniques tailored for the unique setting of the planetarium which differs from the traditional 2D visualizations.

7.1.4 Link Between Clarity and Preference. Previous research on information visualization design suggests clarity and preference are linked [66], and laypeople attach more value to the clarity of designs than to their attractiveness [65]. While the majority of participants in our study confirmed this by selecting the visualization with higher perceived clarity as their preference, the results also show that almost a fourth does not equate these two (Figure 9). It invites more research on the link between aesthetics and clarity and how each influences lay audience's engagement and learning outcomes.

7.2 Scalability and Application

Discussing missing and sparse data is relevant in all science communication settings, and our findings primarily address science museums, science centers, and other institutions dedicated to science education and engagement. While the Voronoi application could be utilized for conventional 2D-based displays, we focused on delivering visualization where common 2D plots fail – a dome in our case. The visualization characteristic that allows use on curved surfaces expands the application to cultural venues and productions using prominence-gaining immersive environments, including 3D omniglobes, AR/VR/XR devices, and video mapping, further broadening the reach and impact [28].

7.3 Designing for Planetariums

Addressing the needs of the audience, institution, and the technical constraints of the dome structure requires a specific design approach to ensure a meaningful and informative experience. We outline challenges below as those that should be further developed.

7.3.1 Catering to Expectations. Designing scientific visualizations for science engagement media presents unique challenges due to the interplay of visitor expectations of education and entertainment [23]. As such, in our design process, we perceived little opportunity to create an experience that diverges substantially from science communication conventions, as we might consider for the setting of an art museum, for instance. However, we recognize this study to be one of the first to push the boundaries of these conventions and see opportunities for future research to explore artistic approaches. In fact, humans seem to intentionally seek novel and stimulating sensory experiences [19], which can make for more memorable experiences [52]. While we did not measure novelty in our study, our results hinted that the ways we presented incomplete data were not expected.

Visitors anticipate seeing truths in museums [26], which, in planetarium settings, can translate to expecting the depictions of the planets to be true – even more so if they are visualized as photorealistic renderings [27]. Revealing the gaps in data that those depictions are based on can create a novel angle on the truth. Efforts to embrace missing data in planetariums are sparse and only recently explored [24]. We argue that unveiling the incomplete datasets and the scientific journey behind the depictions can provide insight into the scientific process, which can form a more memorable experience than seeing only the photorealistic imagery and help progress the visitors' understanding of NoS.

7.3.2 Deploying a Dome to Enhance Immersivity. Our study departed from the existing realistic imagery of the Brussels Planetarium that shows planets "as seen from space," which limits the immersion to a directional view of the screen. Yet, when the planetarium shows the night sky, it adopts the perspective "as seen from Earth," which fully utilizes the dome and creates 360-degree immersion. Perhaps this could also be applied to displaying planetary atmospheric data. After all, for spherical data, like data on the atmosphere of planets, the curved dome allows for a direct mapping of the data on the dome screen. It thus eradicates the distortion associated with common mapping projections of curved data on a flat plane [11]. Moreover, it also allows visitors to be immersed in scientific visualizations.

7.3.3 Supporting Interactivity. Our observations of the Brussels Planetarium lectures showed a gap between the interactive opportunities and the interactions provided. The current engagement with the audience through directive questions is a low level of interactivity. Preparatory notes from this study pointed to a demand for interactive experiences, which corresponds with other researchers advocating for the audience to be able to steer the content [14], and identifying domes as a unique immersive environment supporting drop-in play interactions and blended physical-virtual play [34].

Incorporating open-ended questions, such as some that were part of our survey, could be an attainable start to adapting the existing lectures into a more interactive form. Still, we suggest future research may look into novel and advanced interactivity methods that retain planetariums' collective and immersive experience while supporting critical reflection on the information shown.

7.3.4 Overcoming technical and operational challenges when creating content. In the Brussels Planetarium, the presenters need to write a script in the built-in software to prepare the visualizations' trajectories and order to accordingly illustrate their lectures. They use a tablet to trigger this script. We observed the interface's complexity makes the software a challenging tool for the presenters to master. Furthermore, the utilization of tablets relies on delicate input and uninterrupted connectivity with the master computer; during our observations, experienced presenters had to troubleshoot its functionality. Such an environment means that implementing new visualization ideas into everyday use is challenging. As digital domes' applications are increasingly deployed [71], as well as portable domes [13], future research should explore how prototyping can be easily facilitated within such settings, and how software can allow for flexibility.

7.4 Portraying Nature of Science

We identified that recommendations for NoS instruction do not explicitly deal with uncertainties and incomplete data, even though the research on NoS is in line with the emphasis on transparency in scientific methods and results. After all, science is a reliable source of knowledge that, with its self-correcting nature, offers trustworthy tools to navigate uncertainties [45]; we suggest more refinements of NoS categories to include these elements in NoS' recommendations for instruction.

Planetariums could serve as a prime venue for introducing the public to NoS, as they have a large educational function and are regularly attended by schools, bringing in adolescent audiences who might be not intrinsically interested in science. The omnipresent incomplete data in astronomy could facilitate discussions on missing and sparse data, uncertainty, and their integral part in astronomy and science in general. However, introducing elements of NoS into science engagement media is not well documented in literature. Our work offers first steps and insights in this regard; however, portraying NoS in venues such as planetariums requires further research and clarification.

7.5 Capturing Science Capital

The mean Science Capital score of students from our survey was 43.04, which is in line with the mean of 43.65 from the study by Archer et al. [7]. Contrary to our hypothesis, the levels aligned neither with perceived clarity nor with students' overall preference of the visualization in our study. We also noted that the qualitative data indicated differences in scientific literacy across the participants who had similar Science Capital scores. As such, we found the score division into three levels (high, medium, low) insufficient, and we encourage further research on the conceptualization of Science Capital. For now, we think Science Capital is not the best framework to use for studying to what extent scientific background, interest, or literacy impacts the relations with scientific visualizations.

7.6 Limitations

The study was done with high school groups at the end of a planetarium program with lengths ranging from 1.5 to two hours, which could have impaired the responses by the declining attention span of students. Furthermore, while the online survey was offered in three languages (Dutch, French, and English), the narration was done in English. We made efforts to make the narration easy to understand; yet, we recognize that language could have introduced a barrier for some students to participate. We checked whether this was the case by going through the qualitative data to identify responses that did not recognize the main message of the animation to see if they correlated with the language used for writing responses. The understanding levels, however, were equal across the languages.

8 CONCLUSION

In this study, we focused on a scientific visualization of incomplete data in the dome environment of the planetarium; a science engagement media venue combining an immersive experience, an out-of-classroom learning with a lecture format, and the authority of an established science education institution. We examined how to visualize and highlight incomplete data to adolescents. Using a Research through Design approach, we designed and evaluated two scientific visualizations presenting identical datasets on a sphere in raw and processed forms. The former employed a scatterplot, while the latter utilized a Voronoi diagram with choropleth color coding using uncertainty palettes to suppress the lowest-density values.

In a comparison study involving 126 high school students and 13 teachers, we measured understanding, clarity, and preference for the visualizations. We found that the Voronoi diagram can lead to lower perceived accuracy and details, but appears to be advantageous in making the concept more understandable and aesthetically pleasing to the broader general audience. The implications show the Voronoi diagrams as a suitable method for visualizing sparse data across a plane with no divisions, and grids to highlight empty space.

We advocated for further research to refine explicit ways of communicating incomplete data and scientific uncertainty, by portraying it not merely as a negative aspect but as an integral and constructive element of science. Lastly, we delved into the challenges of designing content for planetarium domes, describing the current gaps in the fields of interactivity and science education, with a particular emphasis on addressing uncertainty.

ACKNOWLEDGMENTS

The authors would like to thank the Belgian Science Policy Office for providing funding (Grant number B2/233/P2/VAMOS) that made this research possible.

REFERENCES

[1] Fouad Abd-El-Khalick, Jeremy J Belarmino, Jeanne L Brunner, An-Phong Le, John Y Myers, Ryan G Summers, Nader Wahbeh, Noemi Waight, Mindy Waters, and Ava A Zeineddin. 2017. A longitudinal analysis of the extent and manner of representations of nature of science in US high school chemistry, biology, and physics textbooks. In *Representations of nature of science in school science textbooks*. Routledge, New York, NY, USA, 20–60. https://doi.org/10.4324/9781315650524

[2] Fouad Abd-El-Khalick and Norman G Lederman. 2000. Improving science teachers' conceptions of nature of science: A critical review of the literature. *International journal of science education* 22, 7 (2000), 665–701.

[3] Kiran Ajani, Elsie Lee, Cindy Xiong, Cole Nussbaumer Knaflic, William Kemper, and Steven Franconeri. 2022. Declutter and Focus: Empirically Evaluating Design Guidelines for Effective Data Communication. *IEEE Transactions on Visualization and Computer Graphics* 28, 10 (2022), 3351–3364. https://doi.org/10.1109/TVCG.2021.3068337

[4] Yann Algan, Daniel Cohen, Eva Davoine, Martial Foucault, and Stefanie Stantcheva. 2021. Trust in scientists in times of pandemic: Panel evidence from 12 countries. *Proceedings of the National Academy of Sciences* 118, 40 (2021), 1–8. https://doi.org/10.1073/pnas.2108576118

[5] Rebecca Andreasson and Maria Riveiro. 2014. Effects of Visualizing Missing Data: An Empirical Evaluation. In *2014 18th International Conference on Information Visualisation*. IEEE, Paris, France, 132–138. https://doi.org/10.1109/IV.2014.77

[6] Kimberly Arcand, M. Watzke, Travis Rector, Zoltan Levay, Joseph DePasquale, and Olivia Smarr. 2013. Processing Color in Astronomical Imagery. *Studies in Media and Communication* 1 (08 2013). https://doi.org/10.11114/smc.v1i2.198

[7] Louise Archer, Emily Dawson, Jennifer DeWitt, Amy Seakins, and Billy Wong. 2015. "Science capital": A conceptual, methodological, and empirical argument for extending bourdieusian notions of capital beyond the arts. *Journal of research in science teaching* 52, 7 (2015), 922–948. https://doi.org/10.1002/tea.21227

[8] Franz Aurenhammer and Rolf Klein. 2000. Voronoi Diagrams. *Handbook of computational geometry* 5, 10 (2000), 201–290.

[9] Ayelet Baram-Tsabari and Jonathan Osborne. 2015. Bridging science education and science communication research. *Journal of Research in Science Teaching* 52, 2 (2015), 135–144.

[10] Ayelet Baram-Tsabari and Amit M. Schejter. 2019. *New Media: A Double-Edged Sword in Support of Public Engagement with Science.* Springer International Publishing, Cham, Switzerland, 79–95. https://doi.org/10.1007/978-3-030-14610-8_5

[11] Melih Basaraner and Sinan Cetinkaya. 2019. New measures for analysis and comparison of shape distortion in world map projections. *Cartography and Geographic Information Science* 46, 6 (2019), 518–531.

[12] Eleonore Batteux, Avri Bilovich, Samuel GB Johnson, and David Tuckett. 2022. Negative consequences of failing to communicate uncertainties during a pandemic: an online randomised controlled trial on COVID-19 vaccines. *BMJ open* 12, 9 (2022), 1–8. https://doi.org/10.1136/bmjopen-2021-051352

[13] Meryem Görecek Baybars and Mustafa Çil. 2021. Use of portable planetariums in teaching of astronomy: an activity example. *Science Activities* 58, 2 (2021), 43–51. https://doi.org/10.1080/00368121.2021.1906200

[14] Karrie Berglund. 2019. Live and interactive. https://www.digitaliseducation.com/blog-post-20190530.en.html

[15] Lonni Besançon, Konrad Schönborn, Erik Sundén, He Yin, Samuel Rising, Peter Westerdahl, Patric Ljung, Josef Wideström, Charles Hansen, and Anders Ynnerman. 2022. Exploring and Explaining Climate Change: Exploranation as a Visualization Pedagogy for Societal Action. VIS4GOOD, a workshop on Visualization for Social Good, held as part of IEEE VIS 2022, 16 - 21 Oct (2022), Oklahoma City, USA.

[16] Andrew R Binder, Elliott D Hillback, and Dominique Brossard. 2016. Conflict or caveats? Effects of media portrayals of scientific uncertainty on audience perceptions of new technologies. *Risk analysis* 36, 4 (2016), 831–846.

[17] contributors Cartopy. 2010 - 2015. *Cartopy: a cartographic python library with a Matplotlib interface.* Met Office, Exeter, UK. https://scitools.org.uk/cartopy

[18] Yiqun T. Chen, Angela D. R. Smith, Katharina Reinecke, and Alexandra To. 2023. Why, when, and from whom: considerations for collecting and reporting race and ethnicity data in HCI. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 395, 15 pages. https://doi.org/10.1145/3544548.3581122

[19] Andy Clark. 2018. A nice surprise? Predictive processing and the active pursuit of novelty. *Phenomenology and the Cognitive Sciences* 17, 3 (01 Jul 2018), 521–534. https://doi.org/10.1007/s11097-017-9525-z

[20] Hernán Cofré, Paola Núñez, David Santibáñez, José M Pavez, Martina Valencia, and Claudia Vergara. 2019. A critical review of students' and teachers' understandings of nature of science. *Science & Education* 28 (2019), 205–248.

[21] Michael Correll, Dominik Moritz, and Jeffrey Heer. 2018. Value-Suppressing Uncertainty Palettes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal, QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3173574.3174216

[22] RSA Cosmos. 2023. *SkyExplorer software*. RSA Cosmos S.A.S, Sorbiers, France. https://www.rsacosmos.com/skyexplorer-open-planetarium.html

[23] James Croft. 2008. Planetarium professionals: A balancing act to engage and educate. *Planetarian* 37, 4 (2008), 6–14 pages.

[24] Myrto Dagkouli-Kyriakoglou, Ericka Johnson, and Miriah Meyer. 2023. The Butterfly at the End of the Universe. Presented at the Information+ Conference, 22 – 24 Nov (2023), University of Edinburgh, UK.

[25] Edwin de Koning, Bart van der Ploeg, Andreas Lo a njoe, Christoph Löbl, Danny van Swieten, Jens Schindel, Julian Hofmann, Mark Buitenwerf, Martijn Otto, Menno Vink, Ronan Timinello, Sander Cox, Stijn Frishert, Tim Walther, Youp Mickers, and Zoltán Pálffy. 2023. *Resolume Arena*. Resolume B.V., The Hague, The Netherlands. https://resolume.com/

[26] David Dean and Gary Edson. 2013. *Handbook for museums*. Routledge, New York, NY, USA.

[27] Jayanne English. 2017. Canvas and cosmos: Visual art techniques applied to astronomy data. *International Journal of Modern Physics D* 26, 04 (2017), 1–51. https://doi.org/10.48550/arXiv.1703.04183

[28] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matt Whitlock, and Yalong Yang. 2021. Grand Challenges in Immersive Analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 459, 17 pages. https://doi.org/10.1145/3411764.3446866

[29] Matt Ferkany and Catherine Kendig. 2020. Countering Scientism and Scepticism in Teaching the Nature of Science Through Its Virtues. In *Virtues as Integral to Science Education*. Routledge, New York, NY, USA,, 84–102.

[30] Baruch Fischhoff. 2012. Communicating Uncertainty Fulfilling the Duty to Inform. *Issues in Science and Technology* 28, 4 (2012), 63–70.

[31] Steven L Franconeri, Lace M Padilla, Priti Shah, Jeffrey M Zacks, and Jessica Hullman. 2021. The science of visual data communication: What works. *Psychological Science in the public interest* 22, 3 (2021), 110–161.

[32] Igal Galili. 2019. Towards a refined depiction of nature of science: applications to physics education. *Science & Education* 28, 3-5 (2019), 503–537.

[33] Pallab Ghosh. 2024. Neptune and Uranus seen in true colours for first time. https://www.bbc.com/news/science-environment-67892275

[34] William Goddard, Alexander Muscat, James Manning, and Jussi Holopainen. 2016. Interactive dome experiences: designing astrosurf. In *Proceedings of the 20th International Academic Mindtrek Conference* (Tampere, Finland) *(AcademicMindtrek '16)*. Association for Computing Machinery, New York, NY, USA, 393–402. https://doi.org/10.1145/2994310.2994339

[35] Brendan Hall. 2006. Teaching and learning uncertainty in science: the case of climate change. *Planet* 17, 1 (2006), 48–49.

[36] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. 2020. Array programming with NumPy. *Nature* 585, 7825 (Sep 2020), 357–362. https://doi.org/10.1038/s41586-020-2649-2

[37] Rachel Hendery and Andrew Burrell. 2020. Playful interfaces to the archive and the embodied experience of data. *Journal of Documentation* 76, 2 (01 Jan 2020), 484–501. https://doi.org/10.1108/JD-05-2019-0078

[38] Jessica Hullman. 2019. Why authors don't visualize uncertainty. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 130–139. https://doi.org/10.1109/TVCG.2019.2934287

[39] J. D. Hunter. 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* 9, 3 (2007), 90–95. https://doi.org/10.1109/MCSE.2007.55

[40] Gunnar Höst, Karljohan Palmerius, and Konrad Schönborn. 2020. Nano for the Public: An Exploranation Perspective. *IEEE Computer Graphics and Applications* 40, 2 (2020), 32–42. https://doi.org/10.1109/MCG.2020.2973120

[41] Patrick G J Irwin, Jack Dobinson, Arjuna James, Nicholas A Teanby, Amy A Simon, Leigh N Fletcher, Michael T Roman, Glenn S Orton, Michael H Wong, Daniel Toledo, Santiago Pérez-Hoyos, and Julie Beck. 2023. Modelling the seasonal cycle of Uranus's colour and magnitude, and comparison with Neptune. *Monthly Notices of the Royal Astronomical Society* 527, 4 (09 2023), 11521–11538. https://doi.org/10.1093/mnras/stad3761

[42] Amit Jena, Ulrich Engelke, Tim Dwyer, Venkatesh Raiamanickam, and Cécile Paris. 2020. Uncertainty Visualisation: An Interactive Visual Survey. In *2020 IEEE Pacific Visualization Symposium (PacificVis)*. IEEE, Tianjin, China, 201–205. https://doi.org/10.1109/PacificVis48177.2020.1014

[43] Eric A Jensen, Kalina Maria Borkiewicz, and Jill P Naiman. 2022. A new frontier in science communication? What we know about how public audiences respond to cinematic scientific visualization. *Frontiers in Communication* 7 (2022), 840631. https://doi.org/10.3389/fcomm.2022.840631

[44] Eric A Jensen and Alexander Gerber. 2020. Evidence-based science communication. *Frontiers in Communication* 4 (2020), 78. https://doi.org/10.3389/fcomm.2019.00078

[45] Kostas Kampourakis and Kevin McCain. 2019. *Uncertainty: How it makes science advance*. Oxford University Press, New York, NY, USA.

[46] Craig S Kaplan. 1999. Voronoi diagrams and ornamental design. In *Proceedings of the First Annual Symposium of the International Society for the Arts, Mathematics, and Architecture (ISAMA'99)*. ISAMA, San Sebastian, Spain, 277–283.

[47] Staffan Klashed. 2015. The next 20 years of planetariums: What challenges and opportunities are we facing? *Journal of the International Planetarium Society* 44, 4 (Dec 2015), 8–10.

[48] Aparna Kongot and Alexandra Matz. 2021. Perfectly Imperfect: A Speculation about Wabi-Sabi Inspired User Experience Design: Wabi-Sabi Inspired UX Design. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 18, 6 pages. https://doi.org/10.1145/3411763.3450370

[49] Jet Propulsion Laboratory. 1993. *Revealing the Face of Venus: Magellan*. National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. https://books.google.be/books?id=hdAeAAAAIAAJ

[50] Norman G Lederman. 1992. Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of research in science teaching* 29, 4 (1992), 331–359.

[51] Norman G Lederman, Fouad Abd-El-Khalick, Randy L Bell, and Renee S Schwartz. 2002. Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of research in science teaching* 39, 6 (2002), 497–521.

[52] Noel Scott Liubov Skavronskaya, Brent Moyle and Anna Kralj. 2020. The psychology of novelty in memorable tourism experiences. *Current Issues in Tourism* 23, 21 (2020), 2683–2698. https://doi.org/10.1080/13683500.2019.1664422

[53] Michael Lynch and Samuel Y Edgerton Jr. 1987. Aesthetics and digital image processing: Representational craft in contemporary astronomy. *The Sociological Review* 35, 1_suppl (1987), 184–220. https://doi.org/10.1111/j.1467-954X.1987.tb00087.x

[54] Arnaud Mahieux, Sophie Berkenbosch, Roland Clairquin, Didier Fussen, Nina Mateshvili, Eddy Neefs, Dennis Nevejans, Bojan Ristic, Ann Carine Vandaele, Valérie Wilquet, et al. 2008. In-flight performance and calibration of SPICAV SOIR onboard Venus Express. *Applied optics* 47, 13 (May 2008), 2252–2265. https://doi.org/10.1364/AO.47.002252

[55] Eve Manz. 2015. Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research* 85, 4 (2015), 553–590.

[56] Michael R. Matthews. 2012. *Changing the Focus: From Nature of Science (NOS) to Features of Science (FOS)*. Springer, Dordrecht, the Netherlands, 3–26. https://doi.org/10.1007/978-94-007-2457-0_1

[57] NASA/JPL/Caltech. 2024. Venus map. ESO Supernova Material for Planetariums and Science Centres. https://supernova.eso.org/exhibition/images/Venus-map-Mariner-artwork-10x5k-CC2/

[58] Anke Nooijen, Ivette Oomens, Apolline Terrier, and Geert van der Veen. 2017. Evaluation of the Planetarium Final Report: management summary. Technopolis group.

[59] Mazlan Othman. 1991. Science education in a planetarium. *Publications of the Astronomical Society of Australia* 9, 1 (1991), 69–71.

[60] Jan Overwijk and Daniël de Zeeuw. 2023. The New Clarity: Awakening in the Post-Truth Era. *new formations: a journal of culture/theory/politics* 109, 109 (2023), 129–146.

[61] Lace Padilla, Matthew Kay, and Jessica Hullman. 2022. Uncertainty Visualization. In *Computational Statistics in Data Science*. John Wiley & Sons, Chichester, UK, 405–421. https://doi.org/10.31234/osf.io/ebd6r

[62] The pandas development team. 2023. *pandas-dev/pandas: Pandas*. Zenodo. https://doi.org/10.5281/zenodo.10426137

[63] Myrthe A. Plaisier and Jeroen B.J. Smeets. 2015. Object size can influence perceived weight independent of visual estimates of the volume of material. *Scientific Reports* 5, 1 (02 Dec 2015), 1–9. https://doi.org/10.1038/srep17719

[64] Xiaoli Qiao and Jessica Hullman. 2018. Translating scientific graphics for public audiences. In *Proceedings of the VisGuides: 2nd Workshop on the Creation, Curation, Critique and Conditioning of Principles and Guidelines in Visualization*. IEEE VIS, Berlin, Germany, 1–4 pages.

[65] Annemarie Quispel and Alfons Maes. 2014. Would you prefer pie or cupcakes? Preferences for data visualization designs of professionals and laypeople in graphic design. *Journal of Visual Languages & Computing* 25, 2 (2014), 107–116.

https://doi.org/10.1016/j.jvlc.2013.11.007

[66] Annemarie Quispel, Alfons Maes, and Joost Schilperoord. 2018. Aesthetics and clarity in information visualization: The designer's perspective. *Arts* 7, 4 (2018), 72. https://doi.org/10.3390/arts7040072

[67] Casey Reas and Ben Fry. 2006. Processing: programming for the media arts. *Ai & Society* 20 (2006), 526–538.

[68] Travis A. Rector, Kimberly Arcand, and Megan Watzke. 2015. *Coloring the universe: An insider's look at making spectacular images of space.* University of Alaska Press, Fairbanks, AK, USA.

[69] Travis A. Rector, Zoltan G. Levay, Lisa M. Frattare, Jayanne English, and Kirk Pu'uohau-Pummill. 2007. Image-Processing Techniques for the Creation of Presentation-Quality Astronomical Images. *The Astronomical Journal* 133, 2 (jan 2007), 598. https://doi.org/10.1086/510117

[70] Youness Salame, Pierre-Majorique Léger, Patrick Charland, Maÿlis Merveilleux Du Vignaux, Emmanuel Durand, Nicolas Bouillot, Mylène Pardoen, Marion Deslandes-Martineau, and Sylvain Sénécal. 2022. The Effects of Interactivity on Learners' Experience in a Visually Immersive Display Context. *Computers in the Schools* 39, 1 (2022), 41–60. https://doi.org/10.1080/07380569.2022.2037296

[71] Simone Schnall, Craig Hedge, and Ruth Weaver. 2012. The Immersive Virtual Environment of the digital fulldome: Considerations of relevant psychological processes. *International Journal of Human-Computer Studies* 70, 8 (2012), 561–575. https://doi.org/10.1016/j.ijhcs.2012.04.001

[72] Simon Schneider. 2016. *Communicating Uncertainty: A Challenge for Science Communication.* Springer, Cham, Switzerland, 267–278. https://doi.org/10.1007/978-3-319-20161-0_17

[73] David Seal. 2001. Venus | Solar System Simulator. NASA/JPL/Caltech. https://maps.jpl.nasa.gov/tmaps/venus.html

[74] Jonas Sjöbergh and Yuzuru Tanaka. 2017. Visualizing Missing Values. In *2017 21st International Conference Information Visualisation (IV).* IEEE, London, UK, 242–249. https://doi.org/10.1109/iV.2017.12

[75] Lisa F Smith, Kimberly K Arcand, Jeffrey K Smith, Randall K Smith, and Jay Bookbinder. 2015. Is that real? understanding astronomical images. *Journal of Media and Communication Studies* 7, 5 (2015), 88–100. https://doi.org/10.5897/JMCS2015.0446

[76] Hayeong Song and Danielle Albers Szafir. 2018. Where's my data? evaluating visualizations with missing data. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 914–924.

[77] David Spiegelhalter, Mike Pearson, and Ian Short. 2011. Visualizing Uncertainty About the Future. *Science* 333, 6048 (2011), 1393–1400. https://doi.org/10.1126/science.1191181

[78] Carolyn Sumners, Patricia Reiff, and Wolfgang Weber. 2008. Learning in an immersive digital theater. *Advances in Space Research* 42, 11 (2008), 1848–1854. https://doi.org/10.1016/j.asr.2008.06.018

[79] John Uri. 2019. 45 Years Ago: Mariner 10 flies by Venus. NASA/Johnson Space Center. https://www.nasa.gov/history/45-years-ago-mariner-10-flies-by-venus/

[80] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272. https://doi.org/10.1038/s41592-019-0686-2

[81] Megan Watzke. 2015. Coloring the Universe: An Insider's Look at Making Spectacular Images of Space. https://www.microsoft.com/en-us/research/video/coloring-the-universe-an-insiders-look-at-making-spectacular-images-of-space/.

[82] Patrick Wiegand. 2003. School Students' Understanding of Choropleth Maps: Evidence From Collaborative Mapmaking Using GIS. *Journal of Geography* 102, 6 (2003), 234–242. https://doi.org/10.1080/00221340308978555

[83] Leland Wilkinson and Michael Friendly. 2009. The History of the Cluster Heat Map. *The American Statistician* 63, 2 (2009), 179–184. http://www.jstor.org/stable/25652247

[84] Anders Ynnerman, Jonas Löwgren, and Lena Tibell. 2018. Exploranation: A new science communication paradigm. *IEEE Computer Graphics and Applications* 38, 3 (2018), 13–20. https://doi.org/10.1109/MCG.2018.032421649

[85] Ka Chun Yu. 2023. A guide to communicating astronomy with storytelling in planetariums. *CAPjournal* December, 33 (2023), 13–20.

[86] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through Design as a Method for Interaction Design Research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, CA, USA) *(CHI '07).* Association for Computing Machinery, New York, NY, USA, 493–502. https://doi.org/10.1145/1240624.1240704

A SUPPLEMENTARY MATERIALS

A.1 Survey Questions (DC1, DC2, DC4)

A.1.1 General.

- Indicate Language: [EN, NL, FR]
- If you agree and wish to participate, please check all the boxes below. [Items detailing informed consent.]
- Please state your age:
- Optional: I would like to be informed about the results of this research, the researchers may contact me for this purpose at my e-mail address:

A.1.2 Understanding the visualization.

- You just watched a short segment. What is the main message you take away from it?
- How clear was that message? [5-point Likert scale from 'very unclear' to 'very clear']
- What else is the segment telling?
- The message that we tried to convey is that there is missing data in the data collected on Venus. How clear was this message from the segment? [5-point Likert scale from 'very unclear' to 'very clear']

A.1.3 Perception on missing data.

- The gaps in the collected data about the atmosphere on Venus are: [5-point Likert scale from 'minimal' to 'enormous']
- Missing data in science in general is: [5-point Likert scale from 'very rare' to 'very frequent' or 'I don't know']
- Given that the data has gaps, how useful do you think the data is for developing our scientific understanding of the atmosphere of Venus? [5-point Likert scale from 'very useless' to 'very useful' or 'I don't know']
- Why do you think that?
- What else would you like to add/comment on?

A.1.4 Preferences for the visualization.

- Please mark which version did you prefer: [USeg or PSeg]
- Why do you prefer it?

A.1.5 Clarity of the visualization.

- Which segment was more clear in showing that there is missing data in the data collected on Venus' atmosphere? [USeg or PSeg]
- Why do you think this one is more clear in communicating the message?
- Did you learn anything new after seeing the second segment? If yes, please explain.
- Was there anything you disliked? If so, what and why? (please indicate which segment your comment is about.)

A.1.6 Science Capital.

1. Who do you talk with about science? (Tick as many as appropriate) ['Friends', 'Siblings (brothers or sisters)', 'Parents or guardians', 'Extended family members (grandparents, aunts, uncles, cousins)', 'Scientists', 'Teachers', 'Other (please specify)', or 'No one']
2. Do you know someone who works in a job using science? Who are they? (tick as many as appropriate) ['Parents or

guardians', 'Siblings (brothers or sisters)', 'Extended family members (grandparents, aunts, uncles, cousins)', 'Friends or neighbours', 'Someone I know from my community', 'Other (please specify)', or 'No one']

(3) When you are NOT in school, how often do you talk about science with other people? [5-point Likert scale from 'Never or rarely (once a year)' to 'Almost every day']
(4) When not in school, how often do you read magazines or books about science? [5-point Likert scale from 'Never or rarely (once a year)' to 'Always (every day or every other day)']
(5) When not in school, how often do you go to a science centre, science museum or planetarium? [5-point Likert scale from 'Never' to 'At least once a month']
(6) When not in school, how often do you visit a zoo or aquarium? [5-point Likert scale from 'Never' to 'At least once a month']
(7) How often do you go to an after-school science club? [5-point Likert scale from 'Never' to 'At least once a month']
(8) One or both of my parents/guardians think science is very interesting. [5-point Likert scale from 'Strongly Disagree' to 'Strongly Agree']
(9) One or both of my parents/guardians have explained to me that science is useful for my future. [5-point Likert scale from 'Strongly Disagree' to 'Strongly Agree']
(10) I know how to use scientific evidence to make an argument. [5-point Likert scale from 'Strongly Disagree' to 'Strongly Agree']
(11) It is useful to know about science in my daily life. [5-point Likert scale from 'Strongly Disagree' to 'Strongly Agree']
(12) My teachers have explained to me science is useful for my future. [5-point Likert scale from 'Strongly Disagree' to 'Strongly Agree']
(13) My teachers have specifically encouraged me to continue with science after graduation from high school. [5-point Likert scale from 'Strongly Disagree' to 'Strongly Agree']
(14) A science qualification can help you get many different types of job. [5-point Likert scale from 'Strongly Disagree' to 'Strongly Agree']

A.2 Scoring of Science Capital

In the first question, participants could score +.5 points for each response, unless 'No one' was selected (=0 points). Possible gain: 0-3.5 pts. In the second question, participants could score +2 points for 'Parents or guardians' and +1 point for each other response, unless 'No one' was selected (=0 pts). Possible gain: 0-7. The remaining questions had five possible answers, scored as follows:

- $OptionA = PossiblePts \times -1$
- $OptionB = PossiblePts \times -.5$
- $OptionC = PossiblePts \times 0$
- $OptionD = PossiblePts \times .5$
- $OptionE = PossiblePts \times 1$

PossiblePts for questions 3-5,7,8-10, and 12-14 was 2, and for questions 6,8,11 was 1. Possible gain ±21.

The Science Capital points add up to a scale from -21 to +31.5, which was converted to a scale of 0-105.

$$Converted = \frac{x + 21}{31.5 + 21} \times 105$$

The score is categorized by thirds: range of 0-34.5 = 'low', 35-69.5 = 'medium', and 70-105 = 'high' [7].

A.3 Narration Script

"When we look at the planet Venus from afar, we can tell its color, and we can see some patterns in its atmosphere. However, just by looking, we cannot determine the upper atmosphere's composition and detailed information about the forces there. That is why there have been satellites that have traveled to the planet and collected data for us, to improve our understanding. Here, you can see the locations where a satellite Venus Express collected data about Venus's atmosphere, such as temperature, pressure, and composition of elements. We see that while we have data about many places, we do not have data about many others and they are not equally distributed. And so, scientists face a challenge now to use the data with the gaps, to create a model of Venus' atmosphere. With that model, we would be able to show not just what it looks like but also what the atmosphere is made of."

The narration was identical for both the USeg and the PSeg, with the single difference being that the narration for the PSeg included that: "[Venus Atmosphere data are] calculated into polygons: the brighter the color, the more data is gathered there. The areas with little or no data are darkened, and polygons with higher data density are highlighted."

A.4 Survey Quantitative Results (DC1, DC2, DC4)

See Table 2.

Table 2: Visualization Preference vs. Clarity

# Students + # Teachers	DC1	DC2	DC4
USeg is preferred	1	1	43+7
PSeg is preferred	4+1	9+1	68+4
USeg is clearer	2+1	0	50+6
PSeg is clearer	3	10+1	61+5

Received 9 February 2024; revised 19 March 2024; accepted 27 March 2024