Review

Encoding Variables, Evaluation Criteria, and Evaluation Methods for Data Physicalisations: A Review

Champika Ranasinghe [1,*] and Auriol Degbelo [2]

1 Data Management & Biometrics, Faculty of Electrical Engineering, Mathematics & Computer Science, University of Twente, 7500 AE Enschede, The Netherlands
2 Chair of Geoinformatics, Faculty of Environmental Sciences, TU Dresden, 10162 Dresden, Germany; auriol.degbelo@tu-dresden.de
* Correspondence: c.m.eparanasinghe@utwente.nl

Abstract: Data physicalisations, or physical visualisations, represent data physically, using variable properties of physical media. As an emerging area, Data physicalisation research needs conceptual foundations to support thinking about and designing new physical representations of data and evaluating them. Yet, it remains unclear at the moment (i) what encoding variables are at the designer's disposal during the creation of physicalisations, (ii) what evaluation criteria could be useful, and (iii) what methods can be used to evaluate physicalisations. This article addresses these three questions through a narrative review and a systematic review. The narrative review draws on the literature from Information Visualisation, HCI and Cartography to provide a holistic view of encoding variables for data. The systematic review looks closely into the evaluation criteria and methods that can be used to evaluate data physicalisations. Both reviews offer a conceptual framework for researchers and designers interested in designing and evaluating data physicalisations. The framework can be used as a common vocabulary to describe physicalisations and to identify design opportunities. We also proposed a seven-stage model for designing and evaluating physical data representations. The model can be used to guide the design of physicalisations and ideate along the stages identified. The evaluation criteria and methods extracted during the work can inform the assessment of existing and future data physicalisation artefacts.

Keywords: data physicalisation; encoding variables; evaluation criteria; evaluation methods; physical variables; human–data interaction; embodied interaction with data; physical interaction; data physicalisation design model; design process

1. Introduction

Data physicalisation or the physical visualisation of data focuses on representing data using geometric or material properties of physical media [1]. While data visualisations primarily focus on the sense of vision and creating data representations that can be "seen", data physicalisations have the potential to create data representations that can not only be seen, but also can be touched, smelled, heard, or tasted. Thus they enable new ways of interacting with data and multisensory data experiences [2]. They are becoming a means for narrowing the gap between people and data and have shown cognitive benefits that come along with their tangible nature (e.g., be effective for self-reflection, attention, and access to data) [3–8]. Furthermore, data physicalisations have the potential to reach audiences (such as for example people with disabilities) that are difficult to reach with traditional data visualisations. Although physical representations of data have existed for many years, data physicalisation is still emerging as a research area [1]. In recent years, there has been a growing interest in establishing theoretical and design foundations for data physicalisation. For example, design principles and guidelines for physicalizing data have started to emerge (e.g., Hogan and Hornecker [9], Sauvé et al. [10], Bae et al. [11], Sosa

et al. [12], Hogan [13], Willett et al. [14]). Yet, the development of design guidelines and theoretical foundations for data physicalisation research still has a long way to go. In this research, we aim to address two important dimensions of data physicalisation research: encoding variables and evaluation.

Why encoding variables and evaluation: Encoding variables (i.e., the properties of the material used to encode data) are a key design dimension of any data communication activity. This is especially important for designing multisensory and immersive data experiences. Design guidelines for data visualisations assume non-disabled populations; thus, the resulting visualisations inhibit people with intellectual and developmental disabilities (IDD) from accessing and effectively engaging with data [15]. Understanding encoding variables that are perceivable through different human sensory channels is important for designing inclusive and accessible data representations for these user groups. Although visual variables have been well explored and are established in visualisation research, a shared understanding and a common vocabulary for variables of other perceptual modalities (beyond the sense of vision) still need to be established, especially to develop guidelines for designing multisensory and immersive data experiences [9,13,16,17]. There have been some efforts in the past to provide a partial inventory of encoding variables and develop a grammar about experiential encoding variables (e.g., [18–22]). However, a compilation of encoding variables for all human sensory modalities and their usage is not available yet. Thus, finding out what multisensory encoding variables are available and which ones are practically used in data physicalisations is useful for filling this research gap.

In addition, evaluation is important for researchers to assess the quality and impact of their data physicalisations (for example, to ensure that the users perceive the data embedded in physical representations and what short-term and long-term impact they could have on people). Methods and criteria for evaluating data visualisations, especially evaluating their ability to effectively and efficiently analyse and discover information, are well established. Physicalisations substantially differ from visualisations, for example, in their ability to engage people, spark interest, trigger interaction, and stimulate emotions. Therefore, new criteria for describing and evaluating the value of physicalisations are emerging. For example, Wang et al. [23] introduced a model for describing the value of a physicalisation based on its creativity and its ability to engage beyond the raw information content (engagement related to affective, physical, intellectual, and social). However, the evaluation methods and criteria that are currently used, as well as the aspects that are evaluated, are not sufficiently known. For example, there is no common knowledge about what evaluation methods are available and used for evaluating aspects related to information discovery/analysis, hedonic aspects, or more open-ended aspects (such as, for example, behavioural stimulation or initiating a social dialogue). This paper tries to fill these two research gaps by seeking answers to the following research questions (RQs):

- RQ1: Which encoding variables can be used to create data physicalisations?
- RQ2: Which evaluation criteria are relevant to the study of data physicalisations?
- RQ3: Which evaluation methods are relevant to the study of data physicalisations?

Method and contributions: Methodically, these questions are examined through two complementary reviews. Encoding variables (also called 'perceptual variables' [16]) are mentioned at different places in the literature, often using different terms to refer to the same notion, and the same term to refer to different notions, due to the interdisciplinary nature of data physicalisation research. For this reason, a systematic review is not appropriate to answer RQ1. Instead, a narrative review holds more potential to cover the breadth of ideas originating from the overlapping fields with data physicalisation research. A narrative review (see e.g., [24,25]) identifies potentially relevant research that has implications for a topic and synthesizes these using meta-narratives. This narrative review uses knowledge from the Visualisation, HCI and Cartography literature and yields a synthesis of the scattered literature on encoding variables into a coherent framework (Contribution 1). The Cartography literature was included in the narrative review because the two fields of Information Visualisation and Cartography (i) share an object of study (i.e., maps),

and (ii) there is evidence of their mutual interplay. Most importantly, the use of non-visual modalities to communicate geographic information has been extensively studied by cartographers (e.g., for the design of tactile [26] and sonic [27] maps), and some of the insights in that context can benefit data physicalisation research. Answering RQ2 and RQ3 is done through a systematic review of papers published between 2009 and 2022. The systematic review helps to learn about evaluation criteria and evaluation methods relevant to data physicalisation research (Contribution 2).

2. Existing Design Spaces for Physicalisations

Previous work has suggested several design spaces/concepts that describe the dimensions that characterize data physicalisations. Thus, these design dimensions can be used to guide the design and evaluation of data physicalisations. We first analysed these design spaces to understand the extent to which they covered the two aspects of our focus: encoding variables and evaluation. Although each individual design framework uses different terms, they cover a total of 13 different design dimensions, as outlined in Table 1. The exact terms used by each design framework are summarised in Appendix A and Table A1. These distinct dimensions include *Data* that describes the nature of data represented by the physicalisation, *Audience*, which refers to the type of the target audience, *Representational Intent* that describes the purpose of the physicalisation (e.g., the effective and efficient discovery of information, evoking specific feelings, and initiating social dialogue), *Representational Material* that refers to the material used for the physicalisation, *Sensory Modality* that refers to the human sensory channel used to perceive data, *Encoding Variables* that describes the physical variables used to encode data, *Representational Fidelity* that describes the metaphorical relationship between data and the materials used to encode data), *Interaction* that describes the type and the nature of interactions that the physicalisation allows, *Proximity to the Data Referent*, which refers to the degree of embodiment (i.e., proximity/situatedness) of the data physicalisation with respect to the data they represent (data referent), *Proximity to the User* that describes the degree of embodiment (proximity/situatedness) of the data physicalisation with respect to the user/user's environment, *Physical Setup* that details the distribution of components of the physicalisation (i.e., the physical setup of components), *Mobility* that indicates whether the physicalisation is bound to a specific location or not, and *Narrative formulation* that describes how data physicalisation facilitates the discovery of information through its external physical form and through any interactive affordances it provides [28].

Table 1. Summary of design dimensions covered by existing design spaces for data physicalisation.

	Design Dimensions													
Design Space/Framework	Data	Audience	Representational Intent	Representational Material	Sensory Modalities	Encoding Variables	Representational Fidelity	Interaction	Proximity—Data Ref.	Proximity—User	Physical Setup	Mobility	Narrative Formulation	Evaluation
Mutisensory Design Space [22]					✓	✓								
Data Sculpture Domain Model [29]			✓											
Embodiment Model [29]							✓							
Data Sculpture Design Taxonomy [28]							✓						✓	
Framework for Situated and Embedded Data Representations [14]									✓					
Framework for Multisensory Data Representation [9]	✓		✓	✓	✓			✓						
Framework for Multisensorial Immersive Analytics [30]	✓		✓		✓	✓					✓			
Physecology [10]	✓					✓		✓		✓	✓	✓		
Cross-Disciplinary Design Space [11]	✓	✓	✓	✓	✓	✓		✓		✓				
Design Elements in Data Physicalisation [31]	✓		✓	✓	✓	✓	✓	✓						
This Paper						✓								✓
N	5	1	5	3	5	6	3	4	1	2	2	1	1	1

As Table 1 shows, existing design frameworks, especially the full-scale design spaces for data physicalisation that were introduced recently (e.g., [10,11]) identify *Encoding Variables* as a key design dimension. Nonetheless, they left it under-specified. In particular, the encoding variables that are available for each perceptual modality and how these variables have been practically used in existing data physicalisations are not fully discussed. Furthermore, none of the frameworks covers evaluation aspects (see Table 1). That is, the criteria to evaluate the merits of a physicalisation or the methods that can be used to evaluate them remain largely unexplored. Previous research on data physicalisation (e.g., [1,13]) also recommends these two aspects as important aspects that need further exploration and detail. The two gaps can now be addressed through a narrative and a systematic review.

3. Narrative Review: Encoding Variables for Physicalisations

Which encoding variables can be used to create physicalisations (RQ1)? The starting point for the review conducted to answer this question is Jansen et al. [1]'s definition: "A data physicalisation (or simply physicalisation) is a physical artefact whose geometry or material properties encode data". This definition suggests one important axis for data physicalisation research, namely, that of data encoding. The data encoding axis is referred to in the literature as the *representation* dimension (see [32,33]). Representation happens through a *representational medium*, i.e., an artefact that is used to encode and store information. Representational media make use of one or more *representational material* and have different *information channels* (i.e., "perceptual aspect of some medium which can be used to carry information" [34]). Colour, shape, and orientation are examples of information channels. Information channels are manipulated through one or more *variables*: visual variables (properties of visual information channels), haptic variables (properties of haptic information channels), olfactory variables (properties of olfactory information channels), and so on.

The choice of a material inevitably restricts the space of possibilities regarding the encoding variables. For instance, the choice of sound as the material for a scenario precludes the use of visual variables to encode information for that scenario. The number of materials that can be used to encode data is potentially infinite. Hogan and Hornecker [9] have provided 37 examples (e.g., glass, water, bread, electronic motors, infrared light, and many more) based on a review of 154 physicalisations. Once a material is chosen, several variables (physically, these can include five variable types related to sensory channels, and one variable type related to change) are at the designer's disposal. These are briefly reviewed below and summarised in Table 2. Most definitions for the variables start intentionally with 'variations/changes of...' to stress the fact that a property by itself is not a variable, it is used as a variable, when changes in this property communicate information. The number of potential variables per sensory channel is put in brackets next to each encoding variable.

Physical variables (∞): Physical variables are variations in material properties that are used to encode information. An exhaustive listing of these variables is still an area of ongoing research, but a few candidates were brought forth in previous work. Hence, the number of physical variables is initialized to infinity for now. Jansen et al. [1] mentioned smoothness, hardness (called compliance in [19]) and sponginess as examples of physical variables. Additional examples include viscosity [35], permeability [35], slipperiness [19], weight [19,20], reflectance [20], density [20], thermal diffusivity [20], stiffness [20], pyrotechnic color [20], tensile strength [20], electrical resistance [20], and thermal expansion [20]. An important remark about physical variables is that, while the encoding activity (i.e., what [29] calls data mapping) is done using material properties, the decoding can only be done using sensory information channels. Consider, for instance, viscosity. Though it is a property of the material, information encoded using it can be perceived through the haptic and the visual information channels.

Table 2. A synthesis of encoding variables for data physicalisations. Variables not highlighted were extracted from textbooks as well as the scientific and grey literature during the narrative review. Variables highlighted in **bold** are additional variables that were identified while annotating papers during the systematic review.

Variable Type	Options
Physical variables	density, electrical resistance, hardness/compliance, permeability, pyrotechnic colour, reflectance, slipperiness, smoothness, sponginess, stiffness, tensile strength, thermal diffusivity, thermal expansion, viscosity, weight, **material**
Visual variables	visual location, colour hue, colour value, colour saturation, visual size, visual shape, visual orientation, visual arrangement, visual texture, crispness, resolution, visual numerousness
Haptic variables	vibration amplitude, vibration frequency, pressure/force–strength, temperature, resistance, friction, kinesthetic location, tangible size, tangible elevation, tangible shape, tangible texture, tangible orientation, tangible location, **tangible arrangement**, **tangible numerousness**
Sonic variables	sound source location, loudness, pitch, register, timbre, attack/decay, rhythmic patterns
Olfactory variables	scent type, scent direction, scent saturation, airflow rate, air quality
Gustatory variables	taste type, temperature of the taste carrier
Dynamic variables	perception time, temporal order, duration, temporal frequency, rate of change, synchronization, **change pattern**

Visual variables (13): Seven visual variables were originally proposed by [36]. These were extended to a list of 12 by [37], and recently synthesized in [38,39]. The following definitions are largely taken from [39]: *visual size* (variations in the length, area, volume or repetitions of a symbol); *visual shape* (variations in the appearance or form of a symbol); *color hue* (variations in the dominant wavelength of visible light, e.g., red, blue, and green); *color value* (light or dark variations of a single hue); *color saturation* (the intensity of a single hue); *visual orientation* (variations in the direction or angle of rotation of a symbol); *(visual) pattern arrangement* (variations in the distribution of individual marks that make up a symbol); *(visual) pattern texture* (variations in coarseness of the pattern within a symbol); *transparency* (variations in the blend level of a symbol and a background layer); *crispness* (variations in the sharpness of boundaries); *resolution* (variations in the level of detail at which the map symbol is displayed); and *visual location* (variations in the x, y position of a symbol relative to a frame of reference). Next to these 'atomic' visual variables, previous work has also pondered the question of 'composite' visual variables. MacEachren [37] proposed to consider 'pattern' as a higher-level visual variable consisting of units that have shape, size, orientation, texture, and arrangement. This is already reflected in the naming of the variables above (e.g., pattern arrangement, pattern texture). Caivano [40] proposed three dimensions of texture, namely, directionality (i.e., dimension that depends on the proportionality of units of texture), size (i.e., surface of the texturing element) and density (i.e., relation of the texturing elements to the background). This suggests that texture itself is a composite variable. Finally, Kraak et al. [41] mentioned *'numerousness'* (arrangement combined with size) as a composite variable used in dot density maps. Hence, numerousness was included as the 13th visual variable in Table 2.

Haptic variables (13): A few haptic variables were mentioned in [42]. These were the following: actuator position, force–strength, vibration frequency, and surface texture. A similar list is found in [30]: force, position, vibration, texture, and temperature. An earlier, much more comprehensive suggestion of haptic variables was proposed by [43]. She proposed that haptic sensations can be decomposed into three categories of variables: those derived from touch (tactile), those derived from kinesthesia (kinesthetic), and those derived from visual analogues (i.e., variables that can be perceived by both vision and touch). Tactile sensations are perceived when the skin comes into contact with an object; kinesthetic sensations are stimulated by bodily movements and tensions. Based on these

lists and the summary of [39], the following haptic variables can be mentioned. There are four *tactile* variables: *vibration amplitude* (also called force, see [44]), *vibration frequency* (also called flutter, see [39,43], or speed), *pressure* (changes in the perceived physical force exerted upon a surface or body), and *temperature* (changes in the perceived temperature of a surface). The perceived intensity of vibration patterns is a function of both their amplitude and frequency (see [45]). There are three *kinesthetic* variables: *resistance* (felt when attempting to deform a surface, e.g., push a button), *friction* (felt when the hand moves across or through a surface), and *kinesthetic location* (changes in the location of the hand in relation to the body). Haptic variables derived from their *visual analogues* include the following: *tangible size* (changes in length, area, or volume), *tangible elevation* (changes in z locations), *tangible shape* (changes in form), *tangible texture/grain* (changes in patterns), *tangible orientation* (changes in alignment), and *tangible location* (changes in x,y locations) (Tangible location generalizes what [42] called the actuator position; tangible texture is synonymous with surface texture from [42]; the terms 'pressure' [43] and 'force-strength' [42] describe the same reality from different perspectives: the encoder uses the strength of the force to communicate data, and the decoder perceives a pressure). The adjective 'tangible' is added to make clear that the information can be perceived by the haptic senses. For instance, a bar chart printed on a T-shirt [46] can be perceived by the eyes (visual size), but not perceived through the hands or kinesthetic. Thus, in that example, visual size is used to communicate information while tangible size is not.

Olfactory variables (5): Patnaik et al. [47] discussed how scent can be used to convey data using introduced olfactory marks and a few olfactory variables. Olfactory marks (i.e., glyph, bouquet, or burst) are analogous to visual marks (i.e., points, lines, or polygons) and refer to the most primitive blocks that can be used to encode scent. Attributes of these marks form the olfactory variables and include the following: the *scent type* (i.e., the signature of the mark), the *direction of the mark* (e.g., changes in the position in space where the scent originates), the *saturation*, a.k.a. chemo-intensity (changes in the concentration of odour molecules in the air), the *airflow rate*, a.k.a. kinetic intensity, the *air quality* (e.g., humidity, temperature, and other non-olfactory properties of the air that can be used to encode information), and the temporal pattern, a.k.a. scent animation. Since dynamic variables are discussed separately as an orthogonal dimension to all other variables, the temporal pattern is not listed here as an olfactory variable. For a related discussion on the olfactory design space, see [48]. The four key dimensions identified in [48]—namely, chemical, emotional, spatial, and temporal—overlap to a great extent with the variables from [47].

Gustatory variables (2): How can the gustatory channel be used to encode information? This is a slightly different question from "which properties of food can be used to encode information?" (For example, one may use the food's shape and colour to communicate information as discussed in [49]. Shape and colour are *visual* variables, not gustatory variables) This has been discussed, for example, in [49,50]. This is also different from the question "how do people describe taste sensations?", which was discussed in previous work (e.g., [51]). In that respect, two gustatory variables can be mentioned: the *signature of the taste carrier* (i.e., changes in the taste type) and the *temperature of the taste carrier* (e.g., hot or cold). Note that 'taste type' here does not only refer to the basic taste types mentioned in the literature (sweet, salty, sour, bitter, and umami, see e.g., [52]), but more broadly to any taste that can be uniquely distinguished from another one. For instance, nominal data values can be mapped to different types of unique tastes, while ordinal values can be mapped to different types of unique temperatures (e.g., the hotter, the higher). In practice, taste variables are used in conjunction with other modalities (e.g., smell and sight) during food consumption, and there is documented evidence in the literature that inputs from other modalities (e.g., visual) to affect gustatory perception [53–55].

Sonic variables (6): Sonification is the transformation of data relations into perceived relations in an acoustic signal for communication or interpretation purposes (see [56]). The following properties of sound mentioned in [39,57] can be used to this end: *sound*

source location (variations in the perception of the placement of the sound's source in a two/three-dimensional space (that perception depends on the physical location of the sound source, the environmental acoustics, and the shape of the ear, see [58])), *loudness* (variations in the magnitude of the sound), *pitch* (variations in the frequency of the sound, i.e., highness or lowness), *register* (variations in the location of a pitch within a range of pitches), *timbre* (variations in the general prevailing characteristic or quality of the sound), and *attack/decay* (variations in the time needed by the sound to reach its maximum or minimum). Duration (variations in the length of time during which a sound or silence is heard), rate of change (variations in relation between the duration of sound and silence over time), and order (variations in the sequence of sounds over time) were also mentioned in [39,57] as sonic variables, but are not included here, because these are dynamic variables discussed below. Finally, rhythmic patterns, mentioned for example in [30], can be used to encode information. Rhythms result from grouping separated sounds into periodic patterns, see e.g., [59,60]. In principle, they can be generated through a combination of other variables (e.g., pitch, timbre, duration, and order). Nonetheless, since they can be used on their own to communicate information (e.g., at least theoretically, nothing prevents the use of different patterns to communicate variations in the data), they are mentioned in the table. Rhythmic patterns are of a composite sonic variable. Melodic patterns, which are the basic units for musification (see [61]), are an example of rhythmic patterns.

Dynamic variables (6): Representing change over time is a recurrent need during the creation of artefacts encoding information, and dynamic variables are useful to this end. Dynamic variables are helpful when designing animations and self-reconfigurable physicalisations. Discussions on dynamic variables in the literature have focused separately on visualisations [37,62–65], sound [39,57], and scent [47]. Nonetheless, given that dynamic variables are orthogonal to all other variables, an account that abstracts from the specifics of sensory modalities is needed. The review of previous descriptions led to the observation that a unifying notion across all modalities is currently missing. We propose the concept of the *representational state* (or state for short) to fill this gap. A representational state refers to the particular condition of a representation (i.e., visualisation, tactile/kinesthetic sensation, scent, taste sensation, and sound) at a given point in time. Then, similarly to other variables (visual variables are properties of visual marks, sonic variables are properties of sound, and so on), dynamic variables could be conceptualised as properties of representational states. Six variables inspired from the works mentioned above and illustrated in Figure 1 are relevant: *perception time* ((variations in the moments in time (a.k.a. temporal locations) the user perceives representational states); *temporal order* (variations in the sequences in which the representational states are perceived); *duration* (variations in the temporal life of the representational states, i.e., how long a representational state is perceived); and *temporal frequency* (variations in the temporal distances between representational states, i.e., how fast/slow new representational states are communicated to users) (This variable can also be called 'rate of occurrences of representational states' or 'number of identifiable representational states per unit time'); *rate of change* (variations in the difference in magnitude of change per unit time for a sequence of representational states); and *synchronization* or phase correspondence (variations in the temporal correspondences of two or more time series). Synchronization is useful to highlight the potential relationship between two phenomena (see e.g., [37]).
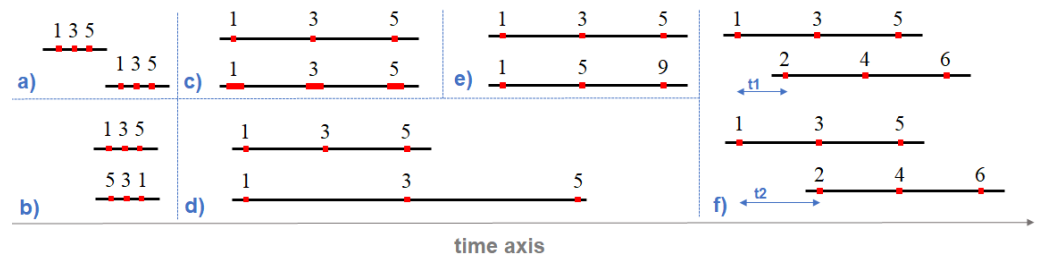
Figure 1. Dynamic variables illustrated: numbers (e.g., 1, 3, and 5) stand for examples of representational states, and the space between them stands for a time interval. (**a**) Two examples of perception times; (**b**) two examples of temporal orders (chronological, reverse chronological); (**c**) two examples of duration; (**d**) two examples of temporal frequency; (**e**) two examples of rates of changes; (**f**) two examples of synchronizations (lags t1 and t2) between two time series.

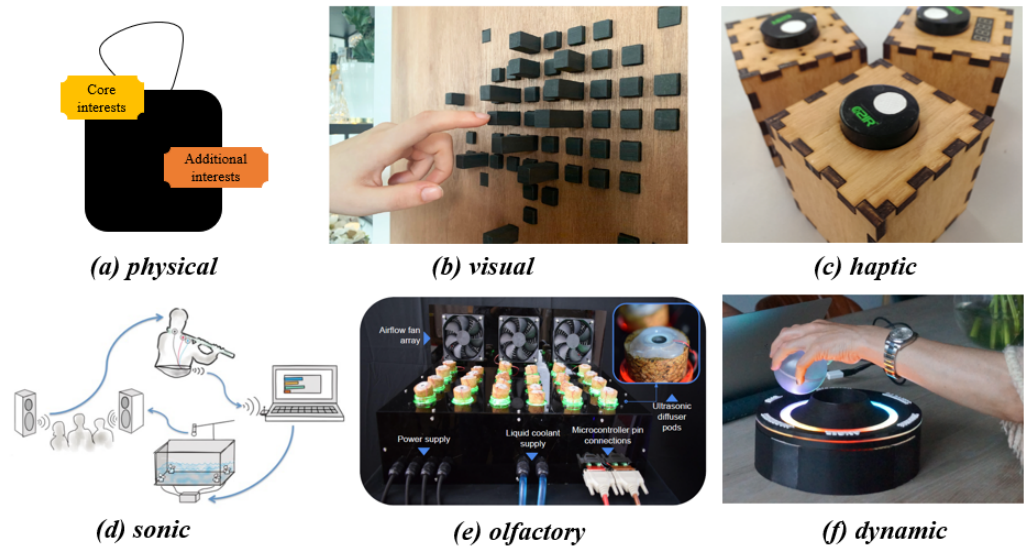Figure 2 shows examples for different variable types (examples selected from or corpus).

(a) physical   *(b) visual*   *(c) haptic*

(d) sonic   *(e) olfactory*   *(f) dynamic*

Figure 2. Examples of encoding variables from papers of the systematic review. *Physical*: different types of material are used to represent the users' core academic interests (Yellow stands here for 'folding paper') and their additional research interests (Orange stands here for 'acrylic'). For the original figure, see [66]. *Visual*: the average effort of users during a running segment is encoded as the length of a pin on the board [67]. *Haptic*: indoor air quality data is encoded as vibration in the haptic probe from [68]. *Sonic*: the muscle tension of flutists is used to create live water sounds as they play their flutes [69]. *Olfactory*: the fan's speed is used to control the airflow rate [70]. *Dynamic*: the LED ring encircling the device fades in/out slowly or quickly to convey if the overall emotional experience of a participant is positive or negative [71].

4. Systematic Review: Encoding Variables, Evaluation Criteria, and Methods

The systematic review presented in this section is an attempt to understand the evaluation criteria and methods used to assess data physicalisations (RQ2 and RQ3). In addition to the annotation of articles with evaluation criteria/methods, and since the encoding variables are a dimension related to representation (see Section 3), we have included all dimensions from Table 1 that touch on an aspect of data representation (as opposed to interaction) in the annotation process. This is useful to assess how representational dimensions relate to each other on the one hand and to the evaluation dimension on the other hand. Hence, the annotation focused on the following dimensions: data type represented, representational material, representational intent, representational fidelity, and the encoding variables. A by-product of the review is to learn about the completeness of the encoding variables derived from the current textbooks (mostly originating from work on

visualisation). The remainder of this section describes the procedure used for the literature search, the screening criteria, the annotation of the papers, and the coding schemes.

4.1. Searching and Retrieving Publications

We employed an analytical approach using a representative sample of publications on empirical work on data physicalisations and followed a systematic procedure similar to previous CHI reviews [72–76]. We used the ACM Digital Library and Scopus as the scientific repositories for our search, as many of the publications related to data physicalisation are included in these outlets. We limited our search to articles written in English and published from 2009 to 2022. The search was carried out in February 2022. We used *data physicalization* and *physical visualization* as keywords for the search and we searched within article title, abstract, and author keywords. The following search queries were used:

Search query for the ACM full-text collection:

"query": Title:("data physicalization"; "physical visualization") OR Abstract:("data physicalization"; "physical visualization") OR Keyword:("data physicalization"; "physical visualization") "filter": Publication Date: (1 January 2009 TO 31 March 2022), ACM Content: DL

Search query for Scopus:

(TITLE-ABS-KEY ("Data Physicalization") OR TITLE-ABS-KEY ("Physical Visualization")) AND PUBYEAR > 2008 AND PUBYEAR < 2023

This initial search yielded 228 articles (*n* = 228): 77 from ACM and 151 from Scopus.

4.2. Screening and Paper Selection

All articles retrieved went through a screening using inclusion/exclusion criteria:

- *Criteria 1*: Articles that were not original peer-reviewed articles or that were not full papers (to ensure that the papers had a complete full scale evaluation of a data physicalisation) (e.g., late breaking works, workshops, pictorials, posters, speeches, doctoral consortium papers, etc.) were excluded.
- *Criteria 2*: Only the articles that discussed an artefact of data physicalisation and empirically evaluated that physicalisation were selected. Therefore, publications that introduced frameworks, theories, processes, opinions, methodologies, concepts, and reviews, as well as publications that did not empirically evaluate a physicalisation, were excluded.
- *Criteria 3*: Articles that discussed augmented physicalisations (for example, [77]) were excluded from the analysis.
- *Criteria 4*: Articles that discussed the same data physicalisation discussed in another article were removed, as our objective was to review different data physicalisation artefacts.

The removal of the duplicates and the application of inclusion/exclusion criteria 1 resulted in 64 articles. The subsequent screening using criteria 2, 3, and 4, which resulted in 36, 34, and 31 articles, respectively (Figure 3). Therefore, our final set of papers selected for the analysis consisted of 31 articles (n = 31). The majority of the selected publications were from CHI (n = 9), followed by TEI (n = 6), IEEE TVCG (Transactions on Visualisation and Computer Graphics) (n = 4), IEEE CG&A (Computer Graphics and Applications) (n = 3), DIS (n = 3), NordiCHI (n = 2), and AI & Society, Elsevier C&G (Computers and Graphics), ASSETS and SVR (n = 1 each). These 31 articles included 50 different data physicalisation artefacts (some papers contributed to more than one data physicalisation, see Table 6). For example, [78] contains four physicalisations that used four different physical modalities (light, vibration, movement, and air) to encode data.
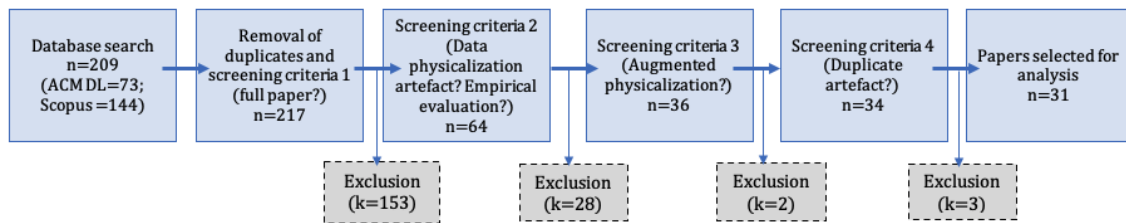
Figure 3. Paper screening procedure.

4.3. Paper Annotation

Every paper was coded independently by two coders, and the coding results of all papers were discussed afterwards among the two coders. When there were conflicting codes, the reasons for the individual decisions were discussed before resolving the inconsistencies. We looked not at papers, but at artefacts mentioned in these papers. One paper could, therefore, present more than one physicalisation. That is, even if an evaluation criterion was used multiple times in one paper to assess several physicalisations, we still counted that the criterion had been used multiple times. The rationale was simple: given the exploratory nature of the work, the number of artefacts over which a criterion has been used matters more at this stage than the number of different authors/research groups who used that criterion/method.

4.4. Coding Schemes

We used the following coding schemes to annotate the artefacts with respect to their evaluation criteria/methods and all the dimensions of Table 1 that touch upon an aspect of data representation.

Data scale: This refers to the type of data that is encoded. Drawing on Stevens [79]'s seminal taxonomy, a three-fold classification for data has become widely used in Information Visualisation and HCI research: nominal (categorical data without a natural order or rank); ordinal (ranked categories); and numerical (quantitative data).

Type of representational material: Bae et al. [11] draw a useful distinction between electronic and non-electronic material: an electronic material has a least one electronic component, while a non-electronic material has none. An electronic component is an entity that has the ability to control electric current (e.g., microcontrollers, sensors, computers).

Representational fidelity: This refers to the metaphorical distance between the physicalisation and the data it represents. Vande Moere and Patel [28] proposed three types of representational fidelity: iconic (the physicalisation bears some relationship to the data being represented through a defined metaphorical relationship); indexical (the physicalisation bears a direct relationship [either physical or causal] to the data being represented); and symbolic (the physicalisation bears no resemblance to the data being represented, and the relationship between the two must be learned using a defined convention). A detailed discussion of the concept of metaphorical distance and examples for each type of representation fidelity are available in [29]. Though only three types of representational fidelity were proposed in [28], a fourth type was identified during the annotation of the articles. The representational fidelity is dynamic if it can vary between iconic, indexical, and symbolic within the physicalisation (not necessarily automatically). For example, in PhysiAir [78], air (coming from a fan) was used to represent ambient air (e.g., CO_2 level or NO_2 level) quality (thus, it can be considered "iconic"). However, PhysiAir can also be configured so that air can also represent other ambient parameters such as humidity and temperature (thus, in that case the fidelity becomes "symbolic"). Therefore, the representation fidelity in PhysiAir [78] can vary between iconic and symbolic depending on the configuration; thus, the fidelity is "dynamic" (in this case via manual reconfiguration).

Representational intent: This refers to the system designer's intention for encoding the data (see [9]). Utilitarian representations were defined as those that 'target a specific audience to reveal data insight related to an explicit task' [9]; casual representations instead

are 'intended for a much broader audience and the exploration of data may be more open-ended and not related to a work task' [9]. The dimension of intent was also mentioned in [8,80]. Dragicevic et al. [8] distinguished between the motivation to discover/present and the motivation to enjoy. Nonetheless, the utilitarian/casual distinction was preferred in this work because, as Dragicevic et al. [8] noted, it is challenging to determine in hindsight whether a physicalisation was created for the purpose of analysing data (discovery) or for the sole purpose of communicating/teaching data insights (presentation). A similar argument applies to Djavaherpour et al. [80]'s distinction between physicalisations with a pragmatic goal (present information in a way that allows the user to thoroughly understand the data) or artistic intent (communicate a concern, rather than show data). The two goals are not mutually exclusive, which makes it challenging to know in hindsight if the motivation of the designer was one solely or the other. In the work, a simple rule was used to classify a physicalisation as utilitarian or casual a posteriori. The intent is classified as utilitarian if the physicalisation is designed to support a specific task and the evaluation (e.g., efficiency, effectiveness/understanding) concerning that task has been done. Otherwise, it is classified as casual.

Evaluation criteria and methods: For the evaluation criteria, we used the performance-related criteria mentioned in [81] and UX-related criteria mentioned in [82] as a starting point for the annotation. As for evaluation techniques, the methods to evaluate the UX identified in [74] were used as a starting point.

Encoding variables: We used the dimensions identified and described in Section 3 to annotate the articles with the encoding variables. The guiding questions used to identify the presence/absence of a variable type in a physicalisation a posteriori are presented in Appendix B.

5. Systematic Review: Results

The main objective of the systematic review was to explore how the data physicalisations were evaluated (i.e., what evaluation criteria are relevant (RQ2) and what evaluation methods can be used (RQ3)). In addition, we also wanted to explore the representation dimension (i.e., what encoding variables have been used in practice, how the dimensions related to representation relate to each other and to the evaluation dimension). The following sections thus present the results of the systematic review along these lines: evaluation criteria/methods (Section 5.1), connection between evaluation criteria and the intention of the data physicalisations from Section 5.2, and the lessons learned about the encoding variables (Section 5.3.1) and the interrelationships observed between the dimensions related to representation and evaluation (Section 5.3.2).

5.1. Evaluation Criteria and Methods

The systematic analysis of the physicalisation artefacts revealed several evaluation criteria that can be used to assess the impact of data physicalisations (Table 3) and a wide variety of methods to collect data about these criteria (Table 4). This subsection summarises these evaluation criteria and the methods used to implement them.

Table 3 summarises the evaluation criteria used to assess the data physicalisations in our sample, ordered based on the frequency of use (most to less frequent). There is the intuition that thecriteria used in HCI/Information Visualisation can be used to evaluate data physicalisations when appropriate (see e.g., [9]), but an open question is whether there are some criteria that could be distinctive to data physicalisation research. It can be seen that UX and performance-related criteria that are widely used in HCI/Information Visualisation were also used to evaluate data physicalisations. We also discovered several evaluation criteria that seemed particular to data physicalisations. Thus, we grouped the evaluation criteria into those that wer enot particular to, and those that seemed particular to data physicalisation research (the dashed line in Table 3 serves that purpose).

Table 3. Summary of evaluation criteria (N = number of physicalisations that were evaluated using the criterion). Criteria above the dashed line (except physical engagement) can be used to evaluate both visualisations and physicalisations; criteria below the dashed line seem unique to data physicalisation research.

Evaluation Criteria	Example Papers	N	%
engagement		16	32
physical engagement	[5,83–86]	5	10
intellectual engagement	[3,4,78,83,84,87]	9	18
social engagement	[83,84,87]	3	6
affective engagement	[3,4,46,83,84,88]	8	16
engagement over time	[3–5,78,86,87]	9	18
user experience	[5,67–69,85,89–94]	15	30
utility	[4,71,78,89,91,95–97]	15	30
effectiveness (question answering)	[7,17,67,70,85,89,91,94,98]	13	26
efficiency (question answering)	[17,67,70,85,91,94,98]	9	18
potential for self-reflection	[3–5,87,95]	8	16
understanding (qualitative)		7	14
personal understanding	[83,87,91,92]	6	12
collaborative understanding	[90]	1	2
attitude change/behavioral stimulation	[4,5,83,86,90,92]	7	14
memorability	[6,7,83,85,89]	6	12
enjoyment/satisfaction	[6,70,71]	4	8
motivational potential	[95]	4	8
ease of use	[6,70,71]	4	8
design parameters	[89,92]	3	6
learning curve/ease of learning	[70,71]	2	4
social acceptance/ease of adoption	[70,88]	2	4
size judgment	[17]	2	4
confidence	[70]	1	2
creativity	[83]	1	2
users' reactions	[46,68,88,96]	8	16
orientation consistency	[99]	6	12
quality of the design	[88,95]	5	10
potential for self-expression	[66,95]	5	10
representational possibilities	[66]	1	2
representational precision	[66]	1	2
quality of the information content	[95]	4	8
aesthetics of the physicalisation	[95]	4	8
remote awareness of physiological states	[71,88]	2	4

Table 4 summarises the evaluation methods used to evaluate data physicalisations, which are presented in order of frequency of use (most to less frequently used). It can be seen that the methods that were widely used in HCI/Information Visualisation were frequently used in data physicalisation research. While both lab-based and field-based experiments were used in equal frequencies in our sample, we discovered that the percentage of longitudinal/repeated studies was significantly low compared to one-time studies.

It is useful for a data physicalisation researcher to understand the methods that have been employed to evaluate each criterion. Table 5 presents the evaluation methods used to assess a criterion (the evaluation criteria are presented in their order of appearance in Table 3 (i.e., most to less frequent in the sample)). A reference next to an evaluation method stands for an example of an article implementing it. A detailed description of the meaning of the evaluation criteria is provided in Appendix C. A broad discussion on the takeaways and implications of the results of our review of evaluation criteria and methods is available in Section 6.

Table 4. Summary of evaluation methods (N = number of physicalisations that were evaluated using a method).

Evaluation Methods	Example Papers	N	%
field-based	[3–5,46,66–69,71,78,84,86–89,92,95]	28	56
lab-based	[6,7,17,67,70,83,85,89,91,92,94,96,97,99]	27	54
semi-structured interviews	[3–6,46,69,71,78,83–87,89–91,95–98]	32	64
self-developed questionnaires	[6,7,17,66,67,69,70,83,90–92,94–96]	21	42
video recording	[6,67–69,83,89,90,96–99]	19	38
information retrieval tasks	[6,17,70,85,89,91,94,96,98,99]	18	36
audio recording	[5,68,83,87,90,92,96,99]	16	32
live user observation	[6,46,85,88,93,99]	14	28
interaction logging	[3,4,78,86]	7	14
experience sampling/diary study	[5,78,87]	6	12
interaction tasks	[6,67,90,94,96]	6	12
standardized questionnaires	[71,83,89,92]	5	10
unstructured interviews	[17,70,88]	4	8
contextual inquiry	[78]	4	8
focus group	[68,90]	4	8
micro-phenomenological interview	[68]	3	6
repGrid technique	[68]	3	6
ratio estimation	[17]	2	4
constant sum	[17]	2	4
sketch of participant's movements	[85]	1	2
social interaction with a confederate of the researcher	[83]	1	2
think aloud	[90]	1	2
post-it note feedback	[84]	1	2
one-time	[17,67,69,83–85,90,93–95,98,99]	35	70
longitudinal/repeated	[3–7,68,78,86]	15	30

5.2. Connecting Evaluation Criteria and Utilitarian/Casual Intents

Previous work has identified the evaluation of physicalisations with open-ended intents as an open research challenge. For instance, Hogan and Hornecker [9] commented that "more attention is needed to evaluate representations whose purpose is more open-ended" and Jansen et al. [1] pointed out that finding appropriate ways of studying how people engage in data exploration when no clear task is defined is a pending evaluation-specific challenge for data physicalisation research. Thus, we also analysed how the criteria from our sample related to the utilitarian and casual representational intents (Figure 4). The findings outlined in Figure 4 can inform designers about how they can evaluate their physicalisations, should these have a similar type of intent. It shows how the criteria were used so far to evaluate one type of intent (either utilitarian or casual) or both:

- Criteria used for physicalisations with a casual intent: intellectual engagement, social engagement, affective engagement, the potential for self-reflection, motivational potential, creativity, user's reactions, quality of the design, potential for self-expression, quality of the information content, aesthetics, and remote awareness of physiological states.
- Criteria used for physicalisations with a utilitarian intent: effectiveness, efficiency, size judgement, confidence, and orientation consistency.
- Criteria used for both types of physicalisations: user experience, utility, understanding (qualitative), attitude change/behavioural stimulation, memorability, enjoyment/satisfaction, ease of use, design parameters, learning curve/ease of learning, social acceptance/ease of adoption, and physical engagement.

Table 5. Evaluation criteria and methods used to evaluate them. Criteria above the dashed line (except physical engagement) can be used to evaluate both visualisations and physicalisations; criteria below the dashed line seem unique to data physicalisation research.

| Criteria	Evaluated Through
Intellectual engagement [23]	semi-structured interviews [4,84], diary studies [78], contextual inquiries [78], and/or self-developed questionnaires [83]
Social engagement [23]	semi-structured interviews [87], self-developed questionnaires [83] and/or the use of a confederate [83]
Affective engagement [23]	semi-structured interviews (e.g., [4,66,84]), user observations [66], self-developed questionnaires [83], and standardized questionnaires (AttrakDiff [83], PANAS-X [83])
Engagement over time	interaction logging [3,4,86], repeated interviews [3,86], diary studies [78] and/or contextual inquiries [78]
User experience [100,101]	standardized questionnaires (UEQ-S [89], AttrakDiff [92]), self-developed questionnaires [69], semi-structured interviews [85,90], a RepGrid study [68], and user observations [93]
Utility [81,102]	semi-interviews [4,78,95,97], self-developed questionnaires [91,95] and a standardized questionnaire (the USE questionnaire [71])
Effectiveness (question answering) [103]	the accuracy with which participants completed information retrieval tasks [6,89] and interaction tasks [96]
Efficiency (question answering) [81,103]	the time taken by participants to complete information retrieval tasks [85] and/or interaction tasks [96]
Potential for self-reflection	self-developed questionnaires [95] and/or semi-structured interviews [3,4]
Understanding (qualitative)	qualitative feedback during an interview [85] or as a rating on a self-developed questionnaire [83,91]
Attitude change/behavioural stimulation	semi-structured interviews [4,95], think-aloud feedback [90], self-developed questionnaires [92], interaction logging [4,86], user observations [83], video recording of the interaction with the physicalisation [90]
Memorability [7,82]	recall tasks [6,89], recall questions [85] and/or self-developed questionnaires [6,7]
Enjoyment/satisfaction [82,103]	self-developed questionnaires (e.g., Likert scales [6,70]) and a standardized questionnaire (the USE questionnaire [71])
Motivational potential	self-developed questionnaires [95]
Ease of use	self-developed questionnaires (e.g., Likert scales [6,70]) and a standardized questionnaire (the USE questionnaire [71])
Design parameters	systematic variation of design parameters (e.g., motion speeds [89] or size [92])
Learning curve/ease of learning	self-developed questionnaires (e.g., Likert scales [70,71])
Social acceptance/ease of adoption [70]	self-developed questionnaires [70] and unstructured interviews [88]
Size judgment	the accuracy of participants on information retrieval tasks [17]
Confidence [70]	self-developed questionnaires [70]
Creativity [23]	the use of a standardized questionnaire (AttrakDiff) [83]
Physical engagement [23]	the sketching of the participants' movement patterns in [85], semi-structured interviews [84,86], self-developed questionnaires [83] and a standardized questionnaire (NASA TLX [83])
Users' reactions	semi-structured interviews [46], unstructured interviews [88], a micro-phenomenological interview [68] and user observations [46,96]
Orientation consistency [99]	information retrieval tasks [99]
Quality of the design	self-developed questionnaires [95], post-it notes feedback [84] and unstructured interviews [88]
Potential for self-expression	self-developed questionnaires [66,95]
Quality of the information content	self-developed questionnaires [95]
Aesthetics of the physicalisation	self-developed questionnaires [95]
Remote awareness of physiological states	user observations and an unstructured interview in [88], and a standardized questionnaire (the emotional awareness survey) and a semi-structured interview in [71]
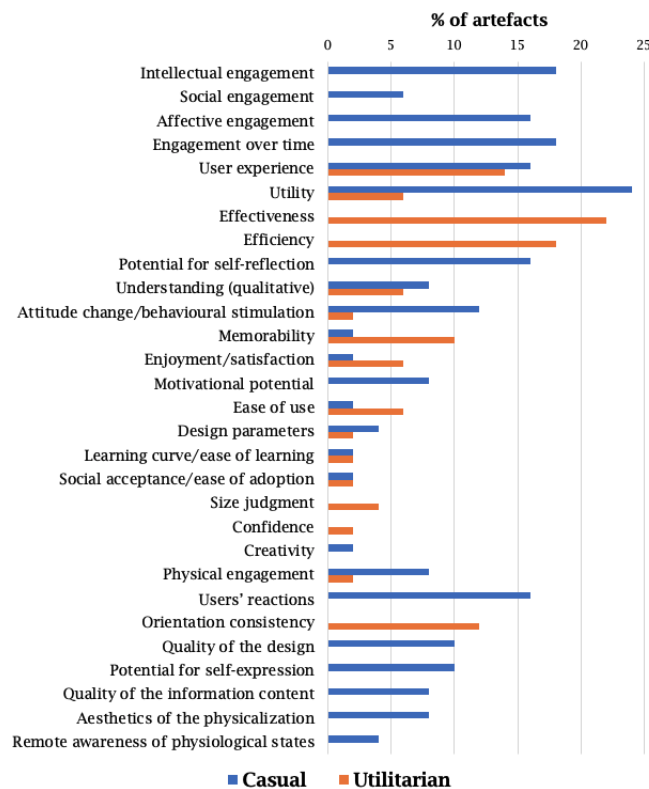
Figure 4. Evaluation criteria used to evaluate physicalisations with casual and utilitarian intents.

5.3. Representation Dimensions

A second objective of this review is to explore how encoding variables have been used in practice and how dimensions related to representation relate to each other. This section presents the results: lessons learned about the encoding variables (Section 5.3.1) and interrelationships between the dimensions (Section 5.3.2). Table 6 summarises the coding of the dimensions related to representation and Table 7 shows the frequencies of the encoding variables found in the sample.

5.3.1. Lessons Learned about the Encoding Variables

During the coding process, we discovered the following extensions to the initial list of encoding variables:

- Physical variables: *Material* should be added to the list in addition to the properties of the material. A nice example can be found in [66], which used the tokens' material (folding paper vs acrylic) to differently encode information related to the core academic background and the additional academic interests of the users.
- Haptic variables: The list of haptic variables that are derived from visual analogues can be extended with at least two variables: *Tangible arrangement* (variations of the distribution of individual marks that make up a symbol) and *tangible numerousness* (arrangement combined with size), as both can be perceived through touch. For instance, the number of squares and their size were used in the 'Dressed in Data' clothes to communicate data about indoor air chemicals, and this resulted in a lace pattern [46]. That lace pattern (both the arrangement of the squares and their numbers) can be perceived by touch. This is an example of both tangible arrangement and tangible numerousness.
- Dynamic variables: The list of dynamic variables should be extended with *change pattern* (variations in animation/movement patterns used to communicate change) as a new variable. For instance, the PhysiMove physicalisation [78] used counterclockwise movements to indicate decreases in value, clockwise movements for increases in value,

and no movement for the lack of change; Keefe et al. [93] used different animation effects to communicate the occurrence of different weather events (i.e., rain, snow, and cloud cover) and Pepping et al. [71] used the slow/fast fading of LED lights to communicate whether or not an emotional experience was positive/negative.

Table 6. An overview of the 50 data physicalisations included in the sample list of 31 academic publications (dimensions related to representation).

Physicalisation	Venue	Reference	Data Type			Material		Intent		Fidelity				Variables						
			Categorical	Ordinal	Numerical	Electronic	Non-Electronic	Casual	Utilitarian	Iconic	Indexical	Symbolic	Dynamic	Visual	Haptic	Olfactory	Gustatory	Sonic	Dynamic	Physical
PhysiLight	CHI	[78]	1	1	1	1		1					1	1					1	
PhysiBuzz	CHI	[78]	1	1	1	1		1				1			1					
PhysiMove	CHI	[78]	1	1	1	1		1				1		1	1				1	
PhysiAir	CHI	[78]	1	1	1	1		1					1		1				1	
Spheres	TVCG	[17]					1		1			1		1	1					
Bars	TVCG	[17]					1		1			1		1	1					
Figure	TVCG	[95]			1		1	1				1		1	1					
Necklace	TVCG	[95]			1		1	1				1		1	1					
Lamp	TVCG	[95]			1		1	1				1		1	1					
Jar	TVCG	[95]			1		1	1				1		1	1					
Bookly	CHI	[4]			1	1		1				1		1	1					
CairnFORM	TEI	[89]		1			1		1			1		1	1				1	
Chemicals in the Creek	TVCG	[84]	1	1	1	1		1				1		1					1	
CoDa	TEI	[90]	1		1	1			1			1		1	1					
Meteorite landings physicalisation	TEI	[85]			1		1		1	1				1	1					
Data Badges	TVCG	[66]	1	1	1		1	1				1		1	1					1
BigBarChart	CG&A	[46]	1	1	1	1		1				1		1	1					
DressedInData	CG&A	[46]	1	1	1		1	1				1		1	1					
DataShirts	CG&A	[46]	1	1	1		1	1				1		1	1					
DayClo	DIS	[3]	1	1	1	1		1			1			1						
Glyph Model	CG	[91]	1		1		1		1	1				1	1					
Streamline Model	CG	[91]	1		1		1		1	1				1	1					
Phys1	CHI	[99]					1		1			1		1	1					
Phys2	CHI	[99]					1		1			1		1	1					
Phys3	CHI	[99]					1		1			1		1	1					
Phys4	CHI	[99]					1		1			1		1	1					
Phys5	CHI	[99]					1		1			1		1	1					
Phys6	CHI	[99]					1		1			1		1	1					
White threads	TEI	[92]		1	1		1	1				1		1	1					
Hoop	TEI	[92]		1	1		1	1				1		1	1					
ViScent 2.0	CHI	[70]	1	1	1	1			1				1		1	1			1	
Auditory Probe	DIS	[68]			1	1		1				1						1		
Haptic Probe	DIS	[68]			1	1		1				1			1					
Visual Probe	DIS	[68]			1	1		1				1		1						
Physical 3D Bar chart	CHI	[98]	1		1		1		1			1		1	1					
EMERGE	CHI	[96]	1		1	1		1				1		1	1					
FluxMarker	ASSETS	[97]			1	1			1			1		1	1					
2D Bar Chart	TEI	[6]	1		1		1		1			1		1	1					
3D Bar Chart	TEI	[6]	1		1		1		1			1		1	1					1
Vital + Morph	AI & Soc	[88]			1	1		1				1		1	1					
Loop	NordiCHI	[87]	1		1	1		1				1		1	1					
Motiis	NordiCHI	[71]	1		1	1		1				1		1	1				1	
Move&Find	CG&A	[83]			1	1		1			1				1					
Torrent	TEI	[69]			1	1		1				1		1				1	1	
Weather Report	CG&A	[93]	1	1	1	1		1		1				1	1				1	
Physical bar chart	CHI	[7]	1		1		1		1			1		1	1					
ADIO	CHI	[86]	1		1	1		1				1		1	1					
Laina	DIS	[5]			1	1		1		1				1	1					
Physical graph	CHI	[67]					1		1		1			1	1					
Visuo-haptic interface	SVR	[94]			1	1			1			1			1					

Table 7. Number of times (N) the variables were found in the sample used.

visual	N	haptic	N	sonic	N
visual location	29	tangible location	29	pitch	1
visual size	28	tangible size	26	timbre	1
colour hue	17	tangible arrangement	7	rythmic patterns	1
visual arrangement	9	tangible numerousness	7	olfactory	N
visual numerousness	8	tangible shape	4	air quality	1
visual shape	7	tangible orientation	4	scent saturation	1
colour value	4	vibration amplitude	3	airflow rate	1
visual orientation	4	vibration frequency	2	scent type	1
visual texture	1	force-strength	3	dynamic	N
physical	N	tangible texture	2	perception time	9
material type	1	temperature	1	change pattern	6
weight	1	resistance	1	temporal frequency	2

In addition, though previous work (e.g., [17]) mentioned the ambiguity surrounding the use of 'size' as an encoding variable (as size can have different aspects), a systematic account of possible usage is still lacking. Our annotations led to the following dimensions of "size" for data physicalisation research: length [66,99], height [7,17,46,85,95,96,98,99], diameter [17,85,87,89,95], area [46], surface area [66], and volume [4,6,91]. Size was also used, not as an encoding variable, but to denote the overall size of the physicalisation. This use of size in the sense of a design parameter that influences the user experience was investigated in [92] (the authors used 'scale' as a synonym for size in their work). The multiplicity of interpretations for 'size' suggests a necessary precisification by the authors investigating it: either as a variable or as a design parameter.

5.3.2. Interrelationships between Dimensions

Several pipelines describing the process of creating physicalisations were proposed in previous work, which include, for example, the extended version of the infovis pipeline to accommodate the physical rendering of data [104], the data sensification workflow that focuses on encoding data in the experience people have with representations [13], and the pipeline for the digital fabrication of physicalisations from [80]. While these pipelines are valuable, none has explicity linked the dimensions we have examined in our systematic review. Hence, we looked into possible connections between the dimensions examined as a first step towards a theory of *representation* in data physicalisation research. Such a theory would inform researchers and designers about the consequences of their choices during the process of building and evaluating physicalisations (e.g., how the choices made at early stages impact the options available at later stages). There are four important elements of theory development according to [105]: (1) extract key concepts, (2) identify the (causal) relationships between these concepts, (3) elaborate on the rationales for these relationships, and (4) clarify the range of application of the theory. We address the four elements in turn.

Key concepts: These are the dimensions considered during the annotation of the articles: all dimensions from existing design spaces touching on data representation, plus the evaluation dimension (see Section 4). The interaction concept is key to the design of data physicalisations (see Table 1) and is hence included in the model, even if it was not explicitly examined during the work.

Relationships: We proposed to link the dimensions considered sequentially into a seven-stage model for designing and evaluating data physicalisations, as shown in Figure 5.
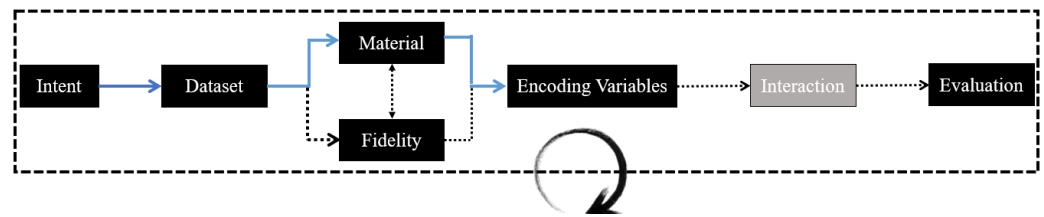
Figure 5. A model connecting the dimensions investigated during the systematic review. Blue arrows indicate a statistically significant association between two dimensions. The interaction dimension is coloured grey because we did not study this dimension in our systematic review. The process is iterative, but arrows describing iterations are omitted in the figure to ease readability.

The steps in this seven-stage model are the following:

1. Form the intention (goal/purpose) of the representation (casual, utilitarian, see Section 4.4);
2. Select a dataset (categorical, ordinal, numerical, or a mix of these);
3. Choose the representational fidelity (iconic, indexical, symbolic, or dynamic, see Section 4.4);
4. Choose a material (examples in Section 3);
5. Pick the encoding variables (examples in Section 3). The choice of the encoding variables entails the choice of sensory modalities.
6. Design the interaction (not discussed in this article, but useful references can be found in [10,35]);
7. Evaluate the artefact (examples in Section 5.1).

In a nutshell, the researcher interested in studying data physicalisations starts with a purpose and selects a dataset in line with that purpose. Afterwards, they choose a representational material, which is a choice that is strongly tied to the choice of the representational metaphor. Since both representational material and fidelity strongly determine each other, they are given an equal footing on the diagram. The choice of the encoding variables follows that of the material. In practice, the design of the interaction happens concurrently with the design of other aspects of the representation, but since the choice of the encoding variables (e.g., visual vs sonic) constrains the interaction possibilities, they are shown sequentially. The systematic evaluation of the artefact happens last (and is the step that distinguishes the researcher from the designer in this model). The meaning of the arrow $\longrightarrow$ is 'precedes and constrains'. The whole process is **iterative**, which means that designers can come back to any stage from any stage, but, for simplicity, arrows representing iterations are not shown in Figure 5.

Range of application: The relationships proposed above and tested below are based on the operationalizations of the concepts related to data physicalisation described in Section 4.4. They may not be valid for other operationalisations (e.g., other taxonomies for data type [80] or representational intent [8,80]).

Quantitative analysis: We computed Fisher's exact test [106] (and, when appropriate, used Pearson's chi-squared test instead) and the Cramér's V for all pairs of dimensions (Table 8). A significant value for the Fisher's exact test or chi-squared test indicates a nonrandom association between two categorical variables, while Cramér's V indicates the strength of the association (0 = no association; 1 = complete association). We now report on the findings for all pairs of consecutive dimensions of the model:

- Intent—dataset (*p*-value < 0.001; Cramér's V = 0.67): There were differences in proportions for nearly all types of datasets. Most notably, physicalisations with a casual intent used the combination of categorical and ordinal and numerical datasets more often than those with a utilitarian intent; they also used numerical data much more often than those with a utilitarian intent. The nonrandom association observed here could be due to some bias in the sample: all physicalisations where the type of dataset was 'not documented' were those having a utilitarian intent (these physicalisations

were used to investigate the theoretical properties of physicalisation in [17,67,99]: orientation consistency, size judgment, and graph physicalisation).

- Dataset—material type (*p*-value < 0.001; Cramér's V = 0.65): Physicalisations encoding three types of datasets (categorical and ordinal and numerical) all used electronic material. The nonrandom association observed here could also be due to some bias in the sample: all physicalisations where the type of dataset was 'not documented' were those using non-electronic material (investigation of theoretical properties).
- Dataset—representational fidelity: The Fisher's exact test between the data type and the representational fidelity was not significant. Nonetheless, the association between the number of datasets and the fidelity was significant (*p*-value = 0.01; Cramér's V = 0.46). In particular, there was no physicalisation with two/three datasets that had an indexical fidelity (i.e., the physicalisation bore a direct relationship [physical or causal] to the data being represented) in our sample.
- Material type—encoding variables (*p*-value < 0.001; Cramér's V = 0.69): Physicalisations combining variables beyond the visual and haptic dimensions (e.g., visual and sonic and haptic and olfactory) all used electronic material.
- Representational fidelity—encoding variables: The Fisher's exact test was not significant.
- Encoding variables—evaluation criteria: We grouped the evaluation criteria into three categories: traditional, novel and traditional and novel. 'Traditional' refers to the criteria above the dashed line (except physical engagement), whereas 'novel' refers to physical engagement and criteria below the dashed line. The Fisher's exact test was not significant.

Table 8. Relationships between the different dimensions. A number within a cell is the Cramér's V (strength of the association) for a statistically significant association (i.e., a possible non-random association between two dimensions). A '-' indicates statistically non-significant associations. To improve readability, some names were abbreviated in the table: evaluation (evaluation criteria), n_modalities (number of modalities), n_datasets (number of datasets), material (material type).

	Intent	Fidelity	Evaluation	Variables	n_Modalities	Dynamicity	Data_Type	n_Datasets	Material
intent		-	0.57	-	0.36	-	0.67	-	0.49
fidelity	-		-	-	-	-	-	0.46	-
evaluation	0.57	-		-	-	-	0.62	-	0.40
variables	-	-	-		1	1	0.55	0.72	0.69
n_modalities	0.36	-	-	1		-	-	-	0.49
dynamicity	-	-	-	1	-		0.79	0.71	0.47
data_type	0.67	-	0.62	0.55	-	0.79		1	0.65
n_datasets	-	0.46	-	0.72	-	0.71	1		0.41
material	0.49	-	0.40	0.69	0.49	0.47	0.65	0.41	

Table 8 summarises the results from the analysis. The key observations from Table 8 are the following:

- The Cramer's V between n_modalities/variables, n_datasets/data_type, and dynamicity/variables was 1 because the dimensions were derived from one another. In particular, n_modalities counted the number of encoding variables used, n_datasets counted the number of data types used, and 'dynamicity' documented whether (or not) dynamic variables were part of the encoding variables.
- Only non-random associations between two consecutive dimensions are highlighted in Figure 5. Nonetheless, the data suggests that there were more non-random associations (e.g., intent/evaluation, intent/data, and data/material). Overall, the material dimension exhibited significant correlations with other non-derived dimensions most often (4/5: intent, evaluation, variables, and data type), followed by the data type dimension (4/5: intent, evaluation, variables, and material) and the intent dimension (3/5: evaluation, data type, and material). The fidelity dimension correlated with other non-derived dimensions the least often.

Overall, the fact that the material/data type/intent dimensions exhibited nonrandom associations with other dimensions is in line with intuition. The number of datasets to encode and the combination of encoding variables emerged as determinants to watch, but since this study is the first to assess the interrelationships between these different representational dimensions and given the size of the sample, more work is needed to unveil the exact nature of the influences between dimensions. Thus, the observations above should be taken as working hypotheses [107] about the relationships between the representation dimensions in physicalisation research.

6. Discussion

So far, this work has provided a synthesis of encoding variables for physicalisations (Section 3 and Table 2), a snapshot of evaluation criteria, as well as examples of methods to apply these criteria (Section 5.1) and working hypotheses about the relationships between different representational aspects of data physicalisation (representational intent, material type, representational fidelity, data type, and encoding variables, see Section 5.3.2). We now discuss general observations made about encoding variables and the evaluation criteria/methods, as well as their implications.

6.1. Encoding Variables

Takeaways: One takeaway from the narrative review is that data encoding as an object of study is a fertile ground for interdisciplinary research. Indeed, several variables synthesized in Table 2 were mentioned separately (and sometimes under slightly different names) in the literature on Information Visualisation, Cartography, Human–Computer Interaction, Sonification, Immersive Analytics, and Neuroscience. A case in point is dynamicity (the representation of change), for which the variables were 'rediscovered' separately for the visual, sonic, and olfactory modalities. As for the systematic review, one takeaway is that inclusiveness is always realized to a certain extent, namely to the extent to which a given sensory modality is supported. In that sense, none of the physicalisations in the sample was fully inclusive (Table 6). Finally, we have observed that only a few physicalisations actually used physical variables (i.e., changes in material properties) to convey messages about phenomena. A similar observation was made in [10], who reported that information communicated through (a change in) physical or material form has been so far rare in practice.

Implications: Looking forward, researchers can use the framework from Table 2 as a vocabulary to describe experiments assessing the effectiveness of variables (across disciplines). That is, the variables can serve as 'boundary objects' [108] between data physicalisation researchers and researchers from the fields mentioned just above (Information Visualisation, Cartography, Sonification, and so on). Boundary objects are concepts *shared* by different communities, which can be viewed or used differently by each. For instance, a subject that can benefit from a plurality of perspectives is the study of users' perception of time-varying representations across different modalities (visual, haptic, aural, etc.). Here, the dynamic variables from Table 2 can serve as boundary object between the different communities investigating the user experience of time-varying representations. Another subject that can benefit from a plurality of perspectives is the notion of variable syntactics (Interpretive flexibility is only one distinguishing characteristic of boundary objects. Another important characteristic is the arrangement of how to operate and collaborate [109]. As Vuillemot et al. [110] put it, "Groups can work on common objects locally, making them more tailored to their local use and needs, i.e., something that is not interdisciplinary, and then share it back in a way that works across the various groups". It is challenging to exactly predict how arrangement will look like, as several disciplines use the variables. Hence, arrangement is not further specified in this article). As indicated in [38], variable syntactics prescribe the use of a variable given a type of dataset (e.g., nominal, ordinal, or numerical). That is, variable syntactics tell how effective/ineffective a given variable is with respect to encoding a given data type. While variable syntactics have been sug-

gested (mostly for the visual [37], aural [57], and dynamic variables [37]), researchers have so far been using different schemes while relating variables to data types. For example, 'unacceptable/acceptable' was used for Bertin's original visual variable syntactics [37], 'not effective/effective' was used for auditory variables in [57], and 'poor/marginally effective/good' was used for visual and dynamic variables in [37]. Data physicalisation research will benefit from the harmonization of these schemes so that they abstract from the specifics of sensory modalities in a similar way to that done for dynamic variables in Section 3. Next to researchers, designers of physicalisations can use the framework to identify design opportunities (e.g., through unexplored variables or an untried mix of variables). In that sense, the framework can be useful to support their goal, discussed in [111], of creating the not-yet-existing.

6.2. Evaluation Criteria and Methods

Takeaways: Jansen et al. [1] identified several key challenges for evaluating data physicalisations: (i) finding appropriate ways of studying how people engage in data exploration when no clear task is defined, (ii) assessing the merits of data representations that go beyond pure time and error metrics, (iii) exploring methodologies to understand how people reason, collaborate, and communicate with physicalisations, and (iv) finding fair alternative representations to use as a baseline for comparison. Our sample and analysis suggest that research that provides answers to challenges (i) and (ii) is ongoing. Notably, several criteria for evaluating aspects that go beyond the traditional time and error metrics have emerged (see Section 5.1). However, we found fewer answers relevant to challenges (iii) and (iv) in our sample. For example, many of the papers in our corpus used on-screen visualisations (e.g., [98]), paper representations (e.g., [6]) and VR representations (e.g., [85]) as a baseline for comparing the effects of data physicalisations. The extent to which these baselines provide a fair ground for comparison still needs to be systematically assessed and discussed.

Furthermore, the majority of the studies in our corpus (70%, Table 4) were one-time studies. The evaluation of variables such as behavioural change and the impact on learning/skills development requires long-term studies to understand the long-term effects. This long-term assessment would be relevant, for instance, to work using personal data physicalisations for teaching (e.g., [112]), or personal data physicalisations in real-world contexts (e.g., [113]). Besides, the evaluation of some aspects of physicalisations has not appeared in our sample, and this suggests that they could be under-explored or not explored at all. Since the systematic review has focused on representational aspects, we mention here a few, related to representation primarily. These include the following: strategies to communicate uncertainties in the underlying dataset; the impact of the material and representation fidelity on meaning-making and memorability; the connection between material properties and data types (e.g., would viscosity be a good material to represent numerical/ordinal/categorical data?); evaluating aspects related to affordances (cognitive affordances, physical affordances, sensory affordances, and functional affordances) in relation to data physicalisation; evaluating the multisensory perception of data; evaluating the interplay between representation and situatedness (e.g., physicalisations that are situated in close spatial proximity to their data referents [14] compared to non-situated data physicalisations); and evaluating representation strategies and their adequacy for diverse user groups (e.g., children or elderly).

Implications: We have already mentioned above that, despite the progress, there are still many unanswered questions. In particular, challenges (iii) and (iv) mentioned above deserve more attention. To these, the gap related to the long-term assessment of physicalisations' impacts and the need for more systematic accounts of the impact of representational features of physicalisations on users can be added.

6.3. Relationships between Representational Dimensions

Takeaways: Though the precedence links connecting the dimensions remain conjectural at this point, our analysis has highlighted that there is a non-random association between several dimensions touching on representational aspects for physicalisations. The results of the quantitative analysis suggests plausible relationships between the most important dimensions related to representation (i.e., material, data type, and intent). We were, however, surprised to see that the fidelity dimension did not seem to strongly connect with other dimensions related to representation. This may be due to the fact that the majority of the physicalisations in our sample (78%) had a symbolic intent, and, hence, bore no resemblance to the data represented.

Implications: Looking forward, the observations of the non-random associations encourage further research towards structural equation models for data physicalisation research. The exact nature of these non-random associations will be uncovered with more examples (and the working hypotheses mentioned in Section 5.3.2 about the associations that can be used as a starting point). Models that describe the expected consequences of design choices during the process of building and evaluating physicalisations will benefit researchers and designers alike. These models will need to provide an account of the indirect relationships between dimensions. For instance, there was no significant association between encoding variables and evaluation, but there still was a non-random association between intent and evaluation (Table 8). The documentation of researchers' work, using a consistent vocabulary to facilitate cross-comparison (e.g., the coding scheme from Section 4.4), will be needed to catalyze progress along these lines.

6.4. Reflections on the Methodical Approach

As Roberts and Walker [16] pointed out, we need a body of research that helps researchers tackle questions such as 'what are the perceptual variables that are available?', 'what are their limitations?', and 'what guidelines are there for each variable?'. While it is clear that these questions still deserve attention, what is less clear are the methodical steps to arrive at general answers. This article has addressed the first question through a combination of a narrative and a systematic review. The narrative review has given the flexibility to draw ideas from different disciplines and reconcile differences in terminologies where appropriate. The systematic review has highlighted how work done in Information Visualisation and Human–Computer Interaction has been implementing the framework from the narrative review (Table 2). We anticipate that the framework from the narrative review can serve as a starting point for answering the other questions above. We also anticipate that replicating the study using articles from other communities will help to progressively extend that framework. For example, given the current corpus with predominant papers from the ACM Digital Library as input for the analysis, physicalisations from the Geography and Cartography communities (e.g., [114,115]), the variables that they use, and the lessons learned about them were not taken into account. Furthermore, since we wanted to learn about evaluation criteria and methods, we restricted ourselves to articles that evaluated their physicalisations in some way. Hence, some articles that could have possibly been useful to extend the list of variables (e.g., [116]) were excluded. Replicating the study by removing this constraint would also be useful as we expand our understanding of perceptual variables for physicalisation research. In summary, though not without flaws, the combination of narrative and systematic review seems promising as we seek answers to the questions mentioned above.

6.5. Limitations

Our search criteria might have excluded some relevant papers, such as those that did not contain the search keywords that we used (e.g., data sculpture, composite physicalisation, constructive visualisation) or if they did not contain an empirical evaluation of a physicalisation. Also, we did not search for particular forms of physicalisations such as, for example, 'sonification', 'haptification' or 'olfaction'. Consequently, our findings are depen-

dent on the sample of papers we selected. Hence, the work does not claim to be exhaustive with respect to the evaluation criteria/methods collected. For instance, one may draw a distinction between physicalisations as designed artefacts, and physicalisations as printed artefacts. The former are physicalisations that were outcomes of a design process (some components of these physicalisations may be 3D printed and some not), while the latter denote physicalisations that were created entirely through printing (i.e., data is rendered as a physically fabricated object, see [80]). Our sample is biased towards the former. Thus, we likely missed criteria relevant to the evaluation of physicalisations as printed artefacts (e.g., the accuracy of the printed artefact was proposed in [117] to document the errors introduced by the printing process and does not appear in Table 3).

7. Conclusions and Future Work

This research provides two contributions to data physicalisation research: (i) a synthesis of the scattered literature on perceptual variables into a coherent framework, and (ii) a snapshot of evaluation criteria and methods relevant to the study of physicalisations. These two contributions can serve as a starting point for further work on the theories and guidelines for data physicalisations such as, notably, the empirical effectiveness of encoding variables for physicalisations and the applicability of perceptual variables to data communication/analysis scenarios.

A question that could guide follow-up reviews to this article is the following: 'what do we know to be true of all perceptual variables, empirically?' In addition to including more examples of physicalisations, follow-up reviews could also cover more dimensions (e.g., reconfigurability discussed in [1], the interaction discussed in [1,9], and the audience mentioned in [10]) and the relationships, if any, with encoding variables, as well as evaluation criteria/methods.

Regarding encoding variables, there is a need for a more systematic investigation of which of these are atomic and which are composite. For instance, air quality is currently listed as an olfactory variable, but is in fact an umbrella term for many variables (e.g., air temperature and air humidity). Another direction for future research is the investigation of the effect of redundant sensorization (i.e., the combined use of several modalities) on user experience. There are works in the literature documenting the positive effects of redundant symbolization for the visual channel (e.g., [118,119]), as well as the visual and haptic channels used in combination (e.g., [67]), and more work along these lines is needed to increase our understanding of the use of redundancy during data encoding more broadly.

Finally, some criteria will benefit from a breaking down of factors that constitute them. This is the case, for example, for "users' reactions", "design quality", and the "aesthetics of the physicalisation". Developing standardized questionnaires that support the evaluation of these criteria, and more generally of criteria unique to data physicalisation research, is also an interesting direction for future work.

Data Availability Statement: The data presented in this study are available in the article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Dimensions of Existing Design Spaces

The dimensions of existing design spaces along with their original names are summarised in Table A1.

Table A1. Dimensions of existing design spaces

Design Space/Framework	Dimensions
Multi-Sensory design space [22]	Sensory modalities, Encoding Variables
Data sculpture domain model [29]	Focus, Manifestation
Embodiment model [29]	Metaphorical distance from data, Metaphorical data from reality
Data Sculpture Design Taxonomy [28]	Representational fidelity, Narrative formulation fidelity
Framework for multi-sensory data representation [9]	Use of modalities (material, sensory modality), Representational intent (utilitarian, casual), Human data relations (interaction mode, type of data)
Framework for multi-sensorial Immersive Analytics [30]	Data (type of data, analytics possible), Sensory Mapping (encoding variables), Devices, Human (human sensory channel)
Physecology [10]	Data type, Information communication, Interaction mechanisms, Spatial coupling, Physical setup, Audience.
Cross-disciplinary Design Space [11]	Context (task, audience, location, data source), Structure (embodiment, material, encoding channel, mobility, data scalability, data duration), Interactions (interaction mediator, sense modality, data interactions)
Design Elements in Data Physicalisation [31]	Design objective (Data form and property, Data theme and topic, design purpose, researched impact of physicalisations), Aesthetics (Design metaphor), Appearance (geometry, material), User experience (interaction, use of technology)

Appendix B. Guideline: Identifying When a Variable Type Has Been Used

How to recognize the presence of a variable a posteriori: a variable is used if the sensory modality can be used *independently* to perceive differences in data.
Guiding questions:

- Imagine I were blind; would I still perceive differences in the data?
- Imagine I could not touch; would I still perceive differences in the data?
- Imagine I could not smell; would I still perceive differences in the data?
- Imagine I could not hear; would I still perceive differences in the data?
- Imagine I could not taste; would I still perceive differences in the data?

If any of these questions is answered by "No", then it is evidence that *only* the variable type corresponding to the sensory encoding channel mentioned in the question (visual, haptic, olfactory, sonic, and gustatory) has been used. If the question is answered by "Yes", it is evidence that a different sensory encoding channel from the one mentioned in the question has been used to encode data. Finding out whether or not dynamic variables are used can be done by asking the question: is animation or self-reconfiguration implemented?

Appendix C. Definitions of Evaluation Criteria

This supplementary material provides additional details (e.g., definitions) about the evaluation criteria mentioned in the article. Some evaluation criteria were defined explicitly in the literature, and, for these, we add the references of the original articles next to their names. The remaining criteria were either (i) mentioned without explicit definition in the articles annotated, or (ii) needed relabelling to reflect the deeper notion they point at. For these criteria, we provide a tentative definition congruent with the article(s) annotated. The criteria are mentioned in their order of appearance in the article (most frequent to least frequent).

Criteria not particular to data physicalisation research.
- Intellectual engagement [23]: Refers to the ability to engage the user in intellectual activities such as recognition, analysis, and contemplation.

- Social engagement [23]: This is present when observers talk with companions, but also when laughing, gesturing, and mimicking the body postures of others. It was assessed, for instance, in [83] through the use of a confederate.

  – *Confederates* are individuals recruited by lead experimenters to play the role of a bystander, participant, or teammate (see e.g., [120]).

- Affective engagement [23]: Refers to the emotional experience of users. The arousing of feelings such as awe, respect, wonder, concern, fear, disgust, anger, or intimidation are indicators of an affective engagement.
- Engagement over time: The evolution of engagement over a given time period.
- User experience [100,101]: The review of definitions by Law et al. [101] pointed out that the ISO definition of UX, "A person's perceptions and responses that result from the use or anticipated use of a product, system or service", is in line with what most UX researchers associate to the concept. In essence, UX refers to all aspects of the users' interaction with a product. It has pragmatic attributes and hedonic attributes [100].
- Utility [81,102]: It is the usefulness of an interface for completing the user's desired set of objectives [102].
- Effectiveness (question answering) [103]: In the sample analyzed, effectiveness was measured through the accuracy with which participants completed information retrieval tasks [6,89] and interaction tasks [96].

  – Information retrieval tasks are specifically directed at retrieving information (e.g., cluster, maxima, or minima of a dataset), whereas interaction tasks are more open-ended (e.g., data analysis tasks such as annotation, filtering or navigation). Hence, not every interaction task is an information retrieval task.

- Efficiency (question answering) [81,103]: This is the time taken by participants to complete information retrieval tasks or interaction tasks.
- Potential for self-reflection: This is the ability of the physicalisations to prompt users to think about themselves. Thudt et al. [113] identified four types of personal reflection in the context of data physicalisation: reflection on (their) data, reflection on (their) context, reflection on (their) action, and reflection on (their) values.
- Understanding (qualitative): This refers to the assessment of the understanding of datasets through qualitative feedback during an interview [85] or as a rating on a self-developed questionnaire [83,91].

  – This assessment may touch upon the understanding by an individual (in that case we talk about personal understanding, see [87,91]), or a group of people (in that case, we talk about collaborative understanding, see [90]).

- Attitude change/behavioural stimulation: This refers to the extent to which a physicalisation can change the attitudes of users (e.g., do they care more about a given subject?) or inspire them to take some action [83].
- Memorability [7,82]: Memorability has different facets, for instance, recognition or recall (see [121]), explicit or implicit memorability (see [7]), and the storage of information in short-term memory or long-term memory (see [122]). It is the capability of maintaining and retrieving information [82].
- Enjoyment/satisfaction [82,103]: Enjoyment is a feeling that causes a person to experience pleasure [82]. Satisfaction denotes the freedom from discomfort and positive attitudes towards the use of the product [103].
- Motivational potential: The ability of the physicalisation to promote gradual changes in individuals' behaviour or sustain the changes over time. It was evaluated through self-developed questionnaires [95].
- Ease of use: The perceived ease of use.
- Design parameters: sSme studies intended to find optimal design parameters and conducted a systematic evaluation of these parameters to that end. For instance, Daniel et al. [89] systematically varied motion speeds to find out the best speed to animate the CairnFORM physicalisation. López García and Hornecker [92] systematically varied

the size of two physicalisations and assessed the impact of these changes on ease of viewing and understanding.

- Learning curve/ease of learning: This refers to the perceived learning curve.
- Social acceptance/ease of adoption [70]: this refers to participants' opinions about the possible introduction of the physicalisation in their lives or sentiments regarding the ease of adoption of the physicalisation.
- Size judgment: Although this was assessed primarily through the accuracy of participants on information retrieval tasks in [17] (and, hence, could have been said to belong to the assessment of the effectiveness of the physicalisation), we still kept this criterion as separate, because it is important for the development of theories of perceptual effectiveness of variables. Ratio estimation [123] and constant sum [124,125] are two methods to collect data about participants' judgments.
- Confidence [70]: This refers to the self-reported confidence levels of users.
- Creativity [23]: The ability of the physicalisation to support the introduction of new and original ideas.

Criteria that seem particular to data physicalisation research.
- Physical engagement [23]: It invites people to spend time touching and interacting with the data (even if just in imagination), moving around it to take different perspectives, bending down to read a label, and employing senses including smell and hearing.
- Users' reactions: Some articles used the term 'user reaction' [88,96] or 'ad-hoc impression' [68] to refer to how the users react to a physicalisation. While there are overlaps with engagement (e.g., the user reactions mentioned in [96] could be classified as an assessment of physical engagement, and part of the reactions documented in [88] could be classified as an assessment of affective engagement), we still keep this evaluation criterion as distinct, because it could be useful for exploratory studies.
- Orientation consistency [99]: The consistency of user responses to information retrieval tasks across different orientations.
- Quality of the design: This touches upon participants' general feedback about design decisions and material choices. It was evaluated, for instance, through self-developed questionnaires [95], post-it note feedback, [84] and unstructured interviews [88].
- Potential for self-expression: The extent to which the physicalisation can help users express some personal characteristics (e.g., academic profile or running performance). It has at least two components mentioned in [66]: representational possibilities (what the user can say through the physicalisation) and representational precision (how accurately they can say what they intend to say).
- Quality of the information content: Evaluated in [95] through self-developed questionnaires.
- Aesthetics of the physicalisation: This touches upon the appearance of the physicalisation. It was evaluated using self-developed questionnaires in [95].
- Remote awareness of physiological states: Some studies [71,88] explored the use of physicalisations as a means for remote monitoring. That is, a user uses a physicalisation to infer the physiological state (e.g., emotional state [71] or arterial blood pressure [88]) of another distant user.

References

1. Jansen, Y.; Dragicevic, P.; Isenberg, P.; Alexander, J.; Karnik, A.; Kildal, J.; Subramanian, S.; Hornbæk, K. Opportunities and challenges for data physicalization. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems—CHI '15, Seoul, Republic of Korea, 23–28 April 2015; Begole, B., Kim, J., Inkpen, K., Woo, W., Eds.; pp. 3227–3236. [CrossRef]
2. van Loenhout, R.; Ranasinghe, C.; Degbelo, A.; Bouali, N. Physicalizing Sustainable Development Goals Data: An Example with SDG 7 (Affordable and Clean Energy). In Proceedings of the CHI Conference on Human Factors in Computing Systems Extended Abstracts, New Orleans, LA, USA, 30 April–6 May 2022; pp. 1–7.

3.   Lee, K.R.; Ju, S.; Dzhoroev, T.; Goh, G.i.; Lee, M.H.; Park, Y.W. DayClo: An everyday table clock providing interaction with personal schedule data for self-reflection. In Proceedings of the DIS'20: Designing Interactive Systems Conference 2020, Eindhoven, The Netherlands, 6–20 July 2020; Wakkary, R., Andersen, K., Odom, W., Desjardins, A., Petersen, M.G., Eds.; 2020; pp. 1793–1806. [CrossRef]

4.   Ju, S.; Lee, K.R.; Kim, S.; Park, Y.W. Bookly: An interactive everyday artifact showing the time of physically accumulated reading activity. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; Brewster, S.A., Fitzpatrick, G., Cox, A.L., Kostakos, V., Eds.; pp. 1–8. [CrossRef]

5.   Menheere, D.; Van Hartingsveldt, E.; Birkebæk, M.; Vos, S.; Lallemand, C. Laina: Dynamic data physicalization for slow exercising feedback. In Proceedings of the Designing Interactive Systems Conference 2021, virtual event, 28 June–2 July 2021; pp. 1015–1030.

6.   Stusak, S.; Hobe, M.; Butz, A. If your mind can grasp it, your hands will help. In Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction, Eindhoven, The Netherlands, 14–17 February 2016; Bakker, S., Hummels, C., Ullmer, B., Geurts, L., Hengeveld, B., Saakes, D., Broekhuijsen, M., Eds.; pp. 92–99.

7.   Stusak, S.; Schwarz, J.; Butz, A. Evaluating the memorability of physical Visualizations. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015), Seoul, Republic of Korea, 23–28 April 2015; Begole, B., Kim, J., Inkpen, K., Woo, W., Eds.; pp. 3247–3250. [CrossRef]

8.   Dragicevic, P.; Jansen, Y.; Moere, A.V. Data physicalization. In *Springer Handbook of Human Computer Interaction*; Springer: Berlin/Heidelberg, Germany, 2021.

9.   Hogan, T.; Hornecker, E. Towards a design space for multisensory data representation. *Interact. Comput.* **2017**, *29*, 147–167. [CrossRef]

10.   Sauvé, K.; Sturdee, M.; Houben, S. Physecology: A Conceptual Framework to Describe Data Physicalizations in their Real-World Context. *ACM Trans. -Comput.-Hum. Interact.* **2022**, *29*, 1–33. [CrossRef]

11.   Bae, S.S.; Zheng, C.; West, M.E.; Do, E.Y.L.; Huron, S.; Szafir, D.A. Making Data Tangible: A Cross-disciplinary Design Space for Data Physicalization. In Proceedings of the CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 30 April–6 May 2022; pp. 1–18.

12.   Sosa, R.; Gerrard, V.; Esparza, A.; Torres, R.; Napper, R. Data objects: Design principles for data physicalisation. In Proceedings of the DS 92: Proceedings of the DESIGN 2018 15th International Design Conference, Dubrovnik, Croatia, 24 May 2018; pp. 1685–1696.

13.   Hogan, T. Data sensification: Beyond representation modality, toward encoding data in experience. In Proceedings of the Design as a Catalyst for Change—DRS International Conference 2018, Limerick, Ireland, 25–28 June 2018; Storni, C., Leahy, K., McMahon, M., Lloyd, P., Bohemia, E., Eds. [CrossRef]

14.   Willett, W.; Jansen, Y.; Dragicevic, P. Embedded data representations. *IEEE Trans. Vis. Comput. Graph.* **2016**, *23*, 461–470. [CrossRef]

15.   Wu, K.; Petersen, E.; Ahmad, T.; Burlinson, D.; Tanis, S.; Szafir, D.A. Understanding data accessibility for people with intellectual and developmental disabilities. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; pp. 1–16.

16.   Roberts, J.C.; Walker, R. Using all our senses: the need for a unified theoretical approach to multi-sensory information visualization. In Proceedings of the IEEE VisWeek 2010 Workshop on The Role of Theory in Information Visualization, Salt Lake City, UT, USA, 24–29 October 2010.

17.   Jansen, Y.; Hornbaek, K. A psychophysical investigation of size as a physical variable. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 479–488. [CrossRef] [PubMed]

18.   Hogan, T.; Hinrichs, U.; Alexander, J.; Huron, S.; Carpendale, S.; Hornecker, E. Toward a design language for data physicalization. In Proceedings of the IEEE VIS Workshop, Estrel Hotel Berlin, Germany, 21–26 October 2018.

19.   Stusak, S.; Butz, A.; Tabard, A. Variables for Data Physicalization Units. In Proceedings of the Toward a Design Language for Data Physicalization: Workshop at IEEE VIS, Estrel Hotel Berlin, Germany, 21–26 October 2018.

20.   Oehlberg, L.; Willett, W. Encoding data through experiential material properties. In Proceedings of the Toward a Design Language for Data Physicalization: Workshop at IEEE VIS, Estrel Hotel Berlin, Germany, 21–26 October 2018.

21.   Moere, A.V. Beyond the tyranny of the pixel: Exploring the physicality of information visualization. In Proceedings of the 2008 12th International Conference Information Visualisation, London, UK, 9–11 July 2008; pp. 469–474.

22.   Nesbitt, K.V. Modeling the multi-sensory design space. In Proceedings of the APVis '01: Proceedings of the 2001 Asia-Pacific symposium on Information Visualisation, Sydney, Australia, 1 December 2001; Volume 9, pp. 27–36.

23.   Wang, Y.; Segal, A.; Klatzky, R.; Keefe, D.F.; Isenberg, P.; Hurtienne, J.; Hornecker, E.; Dwyer, T.; Barrass, S.; Rhyne, T.M. An emotional response to the value of visualization. *IEEE Comput. Graph. Appl.* **2019**, *39*, 8–17. [CrossRef]

24.   Snyder, H. Literature review as a research methodology: An overview and guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [CrossRef]

25.   Grant, M.J.; Booth, A. A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Inf. Libr. J.* **2009**, *26*, 91–108. [CrossRef]

26.   Wabiński, J.; Mościcka, A.; Touya, G. Guidelines for standardizing the design of tactile maps: A review of research and best practice. *Cartogr. J.* **2022**, *59*, 239–258. [CrossRef]

27.   Laakso, M.; Tiina Sarjakoski, L. Sonic maps for hiking - use of sound in enhancing the map use experience. *Cartogr. J.* **2010**, *47*, 300–307. [CrossRef]

28. Vande Moere, A.; Patel, S. The physical visualization of information: Designing data sculptures in an educational context. In *Visual Information Communication*; Huang, M.L., Nguyen, Q.V., Zhang, K., Eds.; Springer: Boston, MA, USA, 2009; pp. 1–23. [CrossRef]

29. Zhao, J.; Moere, A.V. Embodiment in data sculpture: A model of the physical visualization of information. In Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts, Athens, Greece, 10–12 September 2008; pp. 343–350.

30. McCormack, J.; Roberts, J.C.; Bach, B.; Freitas, C.D.S.; Itoh, T.; Hurter, C.; Marriott, K. Multisensory immersive analytics. In *Immersive Analytics*; Springer Nature Switzerland AG 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 57–94.

31. Dumičić, Z.; Thoring, K.; Klöckner, H.W.; Joost, G. *Design Elements in Data Physicalization: A Systematic Literature Review*; Lockton, D., Lenzi, S., Hekkert, P., Oak, A., Sadaba, J., Lloyd, P., Eds.; Design Research Society: Bilbao, Spain, 2022. [CrossRef]

32. Yi, J.S.; ah Kang, Y.; Stasko, J.; Jacko, J. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Trans. Vis. Comput. Graph.* **2007**, *13*, 1224–1231. [CrossRef]

33. Roth, R.E. Interactive maps: What we know and what we need to know. *J. Spat. Inf. Sci.* **2013**, *6*, 59–115. [CrossRef]

34. Bernsen, N.O. Foundations of multimodal representations: A taxonomy of representational modalities. *Interact. Comput.* **1994**, *6*, 347–371. [CrossRef]

35. Rasmussen, M.K.; Pedersen, E.W.; Petersen, M.G.; Hornbæk, K. Shape-changing interfaces: A review of the design space and open research questions. In Proceedings of the CHI'12—Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; Konstan, J.A., Chi, E.H., Höök, K., Eds.; pp. 735–744. [CrossRef]

36. Bertin, J. *Semiology of Graphics: Diagrams, Networks, Maps*; Translated by William J. Berg; The University of Wisconsin Press: Madison, WI, USA, 1983.

37. MacEachren, A.M. *How Maps Work: Representation, Visualization, and Design*; Guilford Press: New York, NY, USA, 1995.

38. Roth, R.E. Visual variables. In *International Encyclopedia of Geography: People, the Earth, Environment and Technology*; Richardson, D., Castree, N., Goodchild, M.F., Kobayashi, A., Liu, W., Marston, R.A., Eds.; John Wiley & Sons, Ltd.: Oxford, UK, 2017; pp. 1–11. [CrossRef]

39. White, T. Symbolization and the visual variables. In *Geographic Information Science & Technology Body of Knowledge*; Wilson, J.P., Ed.; Number Q2; University Consortium for Geographic Information Science (UCGIS): Chesapeake, VA, USA, 2017. [CrossRef]

40. Caivano, J.L. Visual texture as a semiotic system. *Semiotica* **1990**, *80*, 239–252. [CrossRef]

41. Kraak, M.J.; Roth, R.E.; Ricker, B.; Kagawa, A.; Sourd, G.L. *Mapping for a Sustainable World*; The United Nations: New York, NY, USA, 2020; p. 79.

42. Paneels, S.; Roberts, J.C. Review of designs for haptic data visualization. *IEEE Trans. Haptics* **2010**, *3*, 119–137. [CrossRef]

43. Griffin, A.L. Feeling it out: The use of haptic visualization for exploratory geographic analysis. *Cartogr. Perspect.* **2001**, 12–29. [CrossRef]

44. Novich, S.D.; Eagleman, D.M. Using space and time to encode vibrotactile information: Toward an estimate of the skin's achievable throughput. *Exp. Brain Res.* **2015**, *233*, 2777–2788. [CrossRef]

45. Chouvardas, V.; Miliou, A.; Hatalis, M. Tactile displays: Overview and recent advances. *Displays* **2008**, *29*, 185–194. [CrossRef]

46. Perovich, L.J.; Cai, P.; Guo, A.; Zimmerman, K.; Paseman, K.; Espinoza Silva, D.; Brody, J.G. Data clothing and BigBarChart: Designing physical data reports on indoor pollutants for individuals and communities. *IEEE Comput. Graph. Appl.* **2021**, *41*, 87–98. [CrossRef] [PubMed]

47. Patnaik, B.; Batch, A.; Elmqvist, N. Information olfactation: Harnessing scent to convey data. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 726–736. [CrossRef]

48. Maggioni, E.; Cobden, R.; Dmitrenko, D.; Hornbæk, K.; Obrist, M. SMELL SPACE: Mapping out the olfactory design space for novel interactions. *ACM Trans. -Comput.-Hum. Interact. (TOCHI)* **2020**, *27*, 1–26. [CrossRef]

49. Mueller, F.F.; Dwyer, T.; Goodwin, S.; Marriott, K.; Deng, J.; Phan, H.D.; Lin, J.; Chen, K.T.; Wang, Y.; Khot, R.A. Data as delight: eating data. In Proceedings of the CHI'21: CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; Kitamura, Y., Quigley, A., Isbister, K., Igarashi, T., Bjørn, P., Drucker, S.M., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 621:1–621:14. [CrossRef]

50. Wang, Y.; Ma, X.; Luo, Q.; Qu, H. Data edibilization: Representing data with food. In Proceedings of the Extended Abstracts of the 2016 CHI Conference on Human Factors in Computing Systems (CHIEA'16), San Jose, CA, USA, 7–12 May 2016; Kaye, J., Druin, A., Lampe, C., Morris, D., Hourcade, J.P., Eds.; Association for Computing Machinery: New York, NY, USA, 2016; pp. 409–422. [CrossRef]

51. Obrist, M.; Comber, R.; Subramanian, S.; Piqueras-Fiszman, B.; Velasco, C.; Spence, C. Temporal, affective, and embodied characteristics of taste experiences: A framework for design. In Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'14), Toronto, ON, Canada, 26 April–1 May 2014; Jones, M., Palanque, P.A., Schmidt, A., Grossman, T., Eds.; Association for Computing Machinery: New York, NY, USA, 2014; pp. 2853–2862. . [CrossRef]

52. Kikut-Ligaj, D.; Trzcielińska-Lorych, J. How taste works: Cells, receptors and gustatory perception. *Cell. Mol. Biol. Lett.* **2015**, *20*, 699–716. [CrossRef] [PubMed]

53. Gal, D.; Wheeler, S.C.; Shiv, B. Cross-Modal Influences on Gustatory Perception. 2007. Social Science Research Network ("SSRN"). Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1030197, (accessed on 30 May 2022).

54. Harrar, V.; Spence, C. The taste of cutlery: How the taste of food is affected by the weight, size, shape, and colour of the cutlery used to eat it. *Flavour* **2013**, *2*, 21. [CrossRef]

55. Van Doorn, G.H.; Wuillemin, D.; Spence, C. Does the colour of the mug influence the taste of the coffee? *Flavour* **2014**, *3*, 10. [CrossRef]

56. Kramer, G.; Walker, B.; Bonebright, T.; Cook, P.; Flowers, J.H.; Miner, N.; Neuhoff, J. *Sonification Report: Status of the Field and Research Agenda*; Faculty Publications, Department of Psychology, DigitalCommons@University of Nebraska: Lincoln, NE, USA, 2010.

57. Krygier, J.B. Sound and geographic visualization. In *Visualization in Modern Cartography*; MacEachren, A., Taylor, F., Eds.; Academic Press: Cambridge, MA, USA, 1994; pp. 149–166. [CrossRef]

58. Madhyastha, T.; Reed, D. Data sonification: Do you see what I hear? *IEEE Softw.* **1995**, *12*, 45–56. [CrossRef]

59. Palomäki, H. Meanings conveyed by simple auditory rhythms. In Proceedings of the 12th International Conference on Auditory Display (ICAD2006), London, UK, 19–24 June 2006.

60. Bernard, C.; Monnoyer, J.; Wiertlewski, M.; Ystad, S. Rhythm perception is shared between audio and haptics. *Sci. Rep.* **2022**, *12*, 4188. [CrossRef]

61. Visi, F.; Dothel, G.; Williams, D.; Miranda, E. Unfolding | Clusters: A Music and Visual Media Model of ALS Pathophysiology. In Proceedings of the SoniHED Conference: Sonification of Health and Environmental Data, York, UK, 12 September 2014.

62. Carpendale, S. *Considering Visual Variables as a Basis for Information Visualisation*; Technical Report; University of Calgary: Calgary, AB, Canada, 2003. . [CrossRef]

63. Köbben, B.; Yaman, M. Evaluating dynamic visual variables. In Proceedings of the Seminar on Teaching Animated Cartography, Madrid, Spain, 30 August–1 September 1995; Ormeling, F., Köbben, B., Pérez-Gómez, R., Eds.; pp. 45–51.

64. DiBiase, D.; MacEachren, A.M.; Krygier, J.B.; Reeves, C. Animation and the role of map design in scientific visualization. *Cartogr. Geogr. Inf. Syst.* **1992**, *19*, 201–214. [CrossRef]

65. Blok, C. Dynamic visualization in a developing framework for the representation of geographic data. *Bull. Com. Fr. Cartogr.* **1998**, *156*, 89–97. [CrossRef]

66. Panagiotidou, G.; Gorucu, S.; Vande Moere, A. Data badges: Making an academic profile through a DIY wearable physicalization. *IEEE Comput. Graph. Appl.* **2020**, *40*, 51–60. [CrossRef]

67. Drogemuller, A.; Cunningham, A.; Walsh, J.A.; Baumeister, J.; Smith, R.T.; Thomas, B.H. Haptic and visual comprehension of a 2D graph layout through physicalisation. In Proceedings of the CHI'21: CHI Conference on Human Factors in Computing Systems, Virtual Event, Yokohama, Japan, 8–13 May 2021; Kitamura, Y., Quigley, A., Isbister, K., Igarashi, T., Bjørn, P., Drucker, S.M., Eds.; pp. 463:1–463:16. [CrossRef]

68. Hogan, T.; Hinrichs, U.; Hornecker, E. The Visual and beyond: Characterizing experiences with auditory, haptic and visual data representations. In Proceedings of the 2017 Conference on Designing Interactive Systems, (DIS'17), Edinburgh, UK, 10–14 June 2017; Mival, O.H., Smyth, M., Dalsgaard, P., Eds.; pp. 797–809. [CrossRef]

69. Pon, A.; Pattison, E.; Fyfe, L.; Radford, L.; Carpendale, S. *Torrent*: Integrating embodiment, physicalization and musification in music-making. In Proceedings of the Tenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI 2017), Yokohama, Japan, 20–23 March 2017; Inakage, M., Ishii, H., Do, E.Y.L., Steimle, J., Shaer, O., Kunze, K., Peiris, R.L., Eds.; Association for Computing Machinery: New York, NY, USA, 2017; pp. 209–216. [CrossRef]

70. Batch, A.; Patnaik, B.; Akazue, M.; Elmqvist, N. Scents and sensibility: Evaluating information olfactation. In Proceedings of the CHI'20: CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Bernhaupt, R., Mueller, F.F., Verweij, D., Andres, J., McGrenere, J., Cockburn, A., Avellino, I., Goguey, A., Bjøn, P., Zhao, S., et al., Eds.; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–14. [CrossRef]

71. Pepping, J.; Scholte, S.; van Wijland, M.; de Meij, M.; Wallner, G.; Bernhaupt, R. Motiis: Fostering parents' awareness of their adolescents emotional experiences during gaming. In Proceedings of the NordiCHI '20: Shaping Experiences, Shaping Society, Proceedings of the 11th Nordic Conference on Human-Computer Interaction, Tallinn, Estonia, 25–29 October 2020; Lamas, D., Sarapuu, H., Lárusdóttir, M., Stage, J., Ardito, C., Eds.; Association for Computing Machinery: New York, NY, USA, 2020; pp. 58:1–58:11. [CrossRef]

72. Baykal, G.E.; Van Mechelen, M.; Eriksson, E. Collaborative technologies for children with special needs: A systematic literature review. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–13.

73. Salminen, J.; Guan, K.; Jung, S.g.; Chowdhury, S.A.; Jansen, B.J. A literature review of quantitative persona creation. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–14.

74. Pettersson, I.; Lachner, F.; Frison, A.K.; Riener, A.; Butz, A. A Bermuda triangle? A Review of method application and triangulation in user experience evaluation. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–16.

75. Bargas-Avila, J.A.; Hornbæk, K. Old wine in new bottles or novel challenges: A critical analysis of empirical studies of user experience. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; pp. 2689–2698.

76. Koelle, M.; Ananthanarayan, S.; Boll, S. Social acceptability in HCI: A survey of methods, measures, and design strategies. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–19.

77. Herman, B.; Omdal, M.; Zeller, S.; Richter, C.A.; Samsel, F.; Abram, G.; Keefe, D.F. Multi-Touch Querying on Data Physicalizations in Immersive AR. *Proc. ACM -Hum.-Comput. Interact.* **2021**, *5*, 1–20. [CrossRef]

78. Houben, S.; Golsteijn, C.; Gallacher, S.; Johnson, R.; Bakker, S.; Marquardt, N.; Capra, L.; Rogers, Y. Physikit: Data engagement through physical ambient visualizations in the home. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems—CHI '16, San Jose, CA, USA, 7–12 May 2016; Kaye, J., Druin, A., Lampe, C., Morris, D., Hourcade, J.P., Eds.; Association for Computing Machinery: New York, NY, USA, 2016; pp. 1608–1619. [CrossRef]

79. Stevens, S.S. On the theory of scales of measurement. *Science* **1946**, *103*, 677–680. [CrossRef]

80. Djavaherpour, H.; Samavati, F.; Mahdavi-Amiri, A.; Yazdanbakhsh, F.; Huron, S.; Levy, R.; Jansen, Y.; Oehlberg, L. Data to physicalization: A survey of the physical rendering process. *Comput. Graph. Forum* **2021**, *40*, 569–598. [CrossRef]

81. Nielsen, J. Usability 101: Introduction to Usability. NN/g Nielsen Norman Group, 3 January 2012. Available online: https://www.nngroup.com/articles/usability-101-introduction-to-usability/ (accessed on 31 August 2022).

82. Saket, B.; Endert, A.; Stasko, J. Beyond usability and performance: A review of user experience-focused evaluations in visualization. In Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization—BELIV '16, Baltimore, MD, USA, 24 October 2016; Sedlmair, M., Isenberg, P., Isenberg, T., Mahyar, N., Lam, H., Eds.; Association for Computing Machinery: New York, NY, USA, 2016; pp. 133–142. [CrossRef]

83. Hurtienne, J.; Maas, F.; Carolus, A.; Reinhardt, D.; Baur, C.; Wienrich, C. Move&Find: The value of kinaesthetic experience in a casual data representation. *IEEE Comput. Graph. Appl.* **2020**, *40*, 61–75. [CrossRef] [PubMed]

84. Perovich, L.J.; Wylie, S.A.; Bongiovanni, R. Chemicals in the creek: Designing a situated data physicalization of open government data with the community. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 913–923. [CrossRef] [PubMed]

85. Ren, H.; Hornecker, E. Comparing understanding and memorization in physicalization and VR visualization. In Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction, Salzburg, Austria, 14–17 February 2021; Wimmer, R., Kaltenbrunner, M., Murer, M., Wolf, K., Oakley, I., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 1–7. [CrossRef]

86. Lee, K.R.; Kim, B.; Kim, J.; Hong, H.; Park, Y.W. ADIO: An interactive artifact physically representing the intangible digital audiobook listening experience in everyday living spaces. In Proceedings of the CHI'21: CHI Conference on Human Factors in Computing Systems, Virtual Event, Yokohama, Japan, 8–13 May 2021; Kitamura, Y., Quigley, A., Isbister, K., Igarashi, T., Bjørn, P., Drucker, S.M., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 164:1–164:12. [CrossRef]

87. Sauvé, K.; Bakker, S.; Marquardt, N.; Houben, S. LOOP: Exploring physicalization of activity tracking data. In Proceedings of the NordiCHI '20: Shaping Experiences, Shaping Society, Proceedings of the 11th Nordic Conference on Human-Computer Interaction, Tallinn, Estonia, 25–29 October 2020; Lamas, D., Sarapuu, H., Lárusdóttir, M., Stage, J., Ardito, C., Eds.; Association for Computing Machinery: New York, NY, USA, 2020; pp. 52:1–52:12. [CrossRef]

88. Boem, A.; Iwata, H. "It's like holding a human heart": The design of Vital + Morph, a shape-changing interface for remote monitoring. *AI Soc.* **2018**, *33*, 599–619. [CrossRef]

89. Daniel, M.; Rivière, G.; Couture, N. CairnFORM: A shape-changing ring chart notifying renewable energy availability in peripheral locations. In Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction, Tempe, AZ, USA, 17–20 March 2019; pp. 275–286. [CrossRef]

90. Veldhuis, A.; Liang, R.H.; Bekker, T. CoDa: Collaborative data interpretation through an interactive tangible scatterplot. In Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction, Sydney, Australia, 9–12 February 2020; van den Hoven, E., Loke, L., Shaer, O., van Dijk, J., Kun, A.L., Eds.; Association for Computing Machinery: New York, NY, USA, 2020; pp. 323–336. [CrossRef]

91. Ang, K.D.; Samavati, F.F.; Sabokrohiyeh, S.; Garcia, J.; Elbaz, M.S. Physicalizing cardiac blood flow data via 3D printing. *Comput. Graph.* **2019**, *85*, 42–54. [CrossRef]

92. López García, I.; Hornecker, E. Scaling data physicalization—How does size influence experience? In Proceedings of the TEI '21: Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction, Online Event, 14–17 February 2021; Wimmer, R., Kaltenbrunner, M., Murer, M., Wolf, K., Oakley, I., Eds.; Association for Computing Machinery: New York, NY, USA, 2021; pp. 8:1–8:14. [CrossRef]

93. Keefe, D.F.; Johnson, S.; Altheimer, R.; Hong, D.G.; Hunter, R.; Johnson, A.J.; Rockcastle, M.; Swackhamer, M.; Wittkamper, A. Weather Report: A site-specific artwork interweaving human experiences and scientific data physicalization. *IEEE Comput. Graph. Appl.* **2018**, *38*, 10–16. [CrossRef]

94. Cuya, F.G.B.; Guarese, R.L.M.; Johansson, C.G.C.; Giambastiani, M.; Iquiapaza, Y.; de Jesus Oliveira, V.A.; Nedel, L.P.; Maciel, A. Vibrotactile data physicalization: Exploratory insights for haptization of low-resolution images. In Proceedings of the SVR'21: 23rd Symposium on Virtual and Augmented Reality, Virtual Event, 18–21 October 2021; pp. 84–91. [CrossRef]

95. Stusak, S.; Tabard, A.; Sauka, F.; Khot, R.A.; Butz, A. Activity sculptures: Exploring the impact of physical visualizations on running activity. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 2201–2210. [CrossRef]

96. Taher, F.; Hardy, J.; Karnik, A.; Weichel, C.; Jansen, Y.; Hornbæk, K.; Alexander, J. Exploring interactions with physically dynamic bar charts. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015), Seoul, Republic of Korea, 18–23 April 2015; Begole, B., Kim, J., Inkpen, K., Woo, W., Eds.; Association for Computing Machinery: New York, NY, USA, 2015; pp. 3237–3246. [CrossRef]

97. Suzuki, R.; Stangl, A.; Gross, M.D.; Yeh, T. FluxMarker: Enhancing tactile graphics with dynamic tactile markers. In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2017), Baltimore, MD, USA, 20 October–1 November 2017 ; Hurst, A., Findlater, L., Morris, M.R., Eds.; Association for Computing Machinery: New York, NY, USA, 2017; pp. 190–199. [CrossRef]

98. Jansen, Y.; Dragicevic, P.; Fekete, J.D. Evaluating the efficiency of physical visualizations. In Proceedings of the 2013 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'13), Paris, France, 27 April–2 May 2013; Mackay, W.E., Brewster, S.A., Bødker, S., Eds.; Association for Computing Machinery: New York, NY, USA, 2013; pp. 2593–2602. [CrossRef]

99. Sauvé, K.; Potts, D.; Alexander, J.; Houben, S. A change of perspective: How user orientation influences the perception of physicalizations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Bernhaupt, R., Mueller, F.F., Verweij, D., Andres, J., McGrenere, J., Cockburn, A., Avellino, I., Goguey, A., Bjøn, P., Zhao, S., et al., Eds.; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–12. [CrossRef]

100. Hassenzahl, M. The interplay of beauty, goodness, and usability in interactive products. *Hum.-Comput. Interact.* **2004**, *19*, 319–349. [CrossRef]

101. Law, E.; Roto, V.; Hassenzahl, M.; Vermeeren, A.; Kort, J. Understanding, Scoping and Defining User eXperience: A Survey Approach. In Proceedings of the CHI'09, ACM SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; pp. 719–728. [CrossRef]

102. Roth, R.E.; Ross, K.; MacEachren, A. User-centered design for interactive maps: A case study in crime analysis. *ISPRS Int. J. -Geo-Inf.* **2015**, *4*, 262–301. [CrossRef]

103. ANSI. *Common Industry Format for Usability Test Reports*; Technical Report; National Institute of Standards and Technology, US Department of Commerce: Gaithersburg, MD, USA, 2001.

104. Jansen, Y.; Dragicevic, P. An interaction model for visualizations beyond the desktop. *IEEE Trans. Vis. Comput. Graph.* **2013**, *19*, 2396–2405. [CrossRef]

105. Whetten, D.A. What constitutes a theoretical contribution? *Acad. Manag. Rev.* **1989**, *14*, 490–495. [CrossRef]

106. Weisstein, E.W. Fisher's exact test (From MathWorld—A Wolfram Web Resource). Wolfram Research, Inc. 2022. Available online: https://mathworld.wolfram.com/FishersExactTest.html (accessed on 10 September 2022).

107. Chamberlin, T.C. The method of multiple working hypotheses. *Science* **1890**, *15*, 92–96. [CrossRef] [PubMed]

108. Star, S.L. The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In *Distributed Artificial Intelligence (Vol. 2)*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1989; pp. 37–54.

109. Star, S.L. This is not a boundary object: Reflections on the origin of a concept. *Sci. Technol. Hum. Values* **2010**, *35*, 601–617. [CrossRef]

110. Vuillemot, R.; Rivière, P.; Beignon, A.; Tabard, A. Boundary objects in design studies: Reflections on the collaborative creation of isochrone maps. *Comput. Graph. Forum* **2021**, *40*, 349–360. [CrossRef]

111. Stolterman, E. The nature of design practice and implications for interaction design research. *Int. J. Des.* **2008**, *2*, 55–65.

112. Perin, C. What students learn with personal data physicalization. *IEEE Comput. Graph. Appl.* **2021**, *41*, 48–58. [CrossRef] [PubMed]

113. Thudt, A.; Hinrichs, U.; Huron, S.; Carpendale, S. Self-reflection and personal physicalization construction. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI 2018), Montreal, QC, Canada, 21–26 April 2018; Mandryk, R.L., Hancock, M., Perry, M., Cox, A.L., Eds.; Association for Computing Machinery: New York, NY, USA, 21 April 2018; p. 154. [CrossRef]

114. Ballatore, A.; Gordon, D.; Boone, A.P. Sonifying data uncertainty with sound dimensions. *Cartogr. Geogr. Inf. Sci.* **2019**, *46*, 385–400. [CrossRef]

115. Moorman, L.; Djavaherpour, H.; Etemad, K.; Samavati, F.F. Geospatial physicalization in geography education. *J. Geogr.* **2020**, *120*, 23–35. [CrossRef]

116. Friske, M.; Wirfs-Brock, J.; Devendorf, L. Entangling the roles of maker and interpreter in interpersonal data narratives: Explorations in yarn and sound. In Proceedings of the DIS'20: Designing Interactive Systems Conference 2020, Eindhoven, The Netherlands, 6–10 July 2020; Wakkary, R., Andersen, K., Odom, W., Desjardins, A., Petersen, M.G., Eds.; Association for Computing Machinery: New York, NY, USA, 2020; pp. 297–310. [CrossRef]

117. Allahverdi, K.; Djavaherpour, H.; Mahdavi-Amiri, A.; Samavati, F.F. Landscaper: A modeling system for 3D printing scale models of landscapes. *Comput. Graph. Forum* **2018**, *37*, 439–451. [CrossRef]

118. Gorte, V.; Degbelo, A. Choriented maps: Visualizing SDG data on mobile devices. *Cartogr. J.* **2022**, *59*, 35–54. [CrossRef]

119. Dobson, M.W. Visual information processing and cartographic communication: The utility of redundant stimulus dimensions. In *Graphic Communication and Design in Contemporary Cartography*; Taylor, D.R.F., Ed.; John Wiley & Sons: Chichester, UK, 1983; pp. 149–175.

120. Leis, R.; Reinerman-Jones, L. Methodological implications of confederate use for experimentation in safety-critical domains. *Procedia Manuf.* **2015**, *3*, 1233–1240. [CrossRef]

121. Brown, J.; Lewis, V.J.; Monk, A.F. Memorability, word frequency and negative recognition. *Q. J. Exp. Psychol.* **1977**, *29*, 461–473. [CrossRef]

122. Camina, E.; Güell, F. The neuroanatomical, neurophysiological and psychological basis of memory: Current models and their origins. *Front. Pharmacol.* **2017**, *8*, 438. [CrossRef] [PubMed]

123. Cleveland, W.S.; McGill, R. An experiment in graphical perception. *Int. J. -Man-Mach. Stud.* **1986**, *25*, 491–500. [CrossRef]

124. Comrey, A.L. A proposed method for absolute ratio scaling. *Psychometrika* **1950**, *15*, 317–325. [CrossRef] [PubMed]

125. Spence, I. Visual psychophysics of simple graphical elements. *J. Exp. Psychol. Hum. Percept. Perform.* **1990**, *16*, 683–692. [CrossRef]