Delivering data differently[*]

gwynn sturdevant

Laboratory of Innovation Science, Harvard University

A. Jonathan R. Godfrey

School of Mathematical and Computational Sciences, Massey University

Andrew Gelman

Departments of Statistics and Political Science, Columbia University

14 Apr 2022

Abstract

Human-computer interaction relies on mouse/touchpad, keyboard, and screen, but tools have recently been developed that engage sound, smell, touch, muscular resistance, voice dialogue, balance, and multiple senses at once. How might these improvements impact upon the practice of statistics and data science? People with low vision may be better able to grasp and explore data. More generally, methods developed to enable this have the potential to allow sighted people to use more senses and become better analysts. We would like to adapt some of the wide range of available computer and sensory input/output technologies to transform data science workflows. Here is a vision of what this synthesis might accomplish.

Keywords: Accessibility, Sensification, Statistical graphics, Visualization, Vivification

"Part of the aesthetic orientation is a perceptual openness, a willingness to notice and believe in connections and meanings that may not be instantly apparent." — Elizabeth Hellmuth Margulis, *On Repeat.*

1 Introduction

In traditional Andean South America, people recorded data using knotted strings called khipu. After their arrival, the Spanish incorporated khipu to both ecclesiastical and administrative purposes [Brokaw, 2010]. Like many tools that were once useful, knots are no longer used to store or communicate data in South America.

Another data visualization tool was invented fifty years ago to display multivariate data using a coding that produced cartoon faces [Chernoff, 1973]. Because a large portion of the human brain is reserved for processing facial features, scholars believed that this presentation made it easier for humans to understand data. The practical value of this idea did not match its theoretical appeal.

These ways of delivering data are of interest because they are not widely used. What are the similarities between these data delivery methods and other methods that are not used? What data delivery methods have recently been advanced that will not be used in twenty years? And most importantly, what methods will be used?

From the other direction, some familiar visualization tools are surprisingly recent. For example, the scatterplot appears to be less than two hundred years old, appearing centuries after the development of Cartesian coordinates; see Friendly and Denis [2005]. We can consider this relatively late appearance as frustrating or as a source of hope that perhaps other valuable tools remain waiting to be developed.

1.1   From visualization to vivification

In this article, the authors consider a number of ways that data can be delivered differently. These methods have the potential to enhance data delivery for all people, and could be monumental for people with low vision for whom literal visualization of data is difficult or impossible.

Terminology is important, and the authors have spent time searching for words that encapsulate this work. Unfortunately, we have not found any single ideal term. Candidates have been visualize, visceralize, sensify, Umweltbeziehung, Umwelterfahrung, and vivify. Visualize does not include the multiple senses that we hope to engage and may not be relatable to people with low vision. Visceralize has strong resonances of "viscera," the internal organs of the body, not necessarily just the senses. The root word of sensify is sense and we will not be limited to the five senses that are traditionally considered. German has words for everything. Exploring resulted in Umweltbeziehung or Umwelterfahrung which are challenging for English speakers. Although nothing actually comes to life, for this article the authors have settled on vivification.

Related is the biosemiotic term Umwelt, first used by Jakob von Uexküll [Sebeok, 1976, Tønnessen et al., 2016]. The Encyclopaedia Britannica defines Umwelt as "perceptual environment" and the Oxford English Dictionary defines it as "The outer world, or reality, as it affects the organisms inhabiting it." Each organism experiences the world through a perceptual bubble. Dogs experience much of the world through scents. Humans generally use sight as our primary cue. When you go for a walk with your dog, your Umwelts are completely different due to your differing senses. The movie Superpower Dogs uses visual cues to present the dog Umwelt to humans.

Umwelts differ within as well as between species. When a biologist and a geologist go for a walk, they can have completely different Umwelts due to different education and interests. One impact

of different Umwelts is differing internal experiences; two people may have different imaginations of a box. To accommodate and supersede these differences scientists could consider delivering data differently.

The authors can formally define data vivification to mean a mapping of a dataset to the human Umwelt. Historically, statisticians have focused largely on linking visual representations of data to the brain. We propose that a multitude of other representations can assist the data pipeline. These include auditory, tactile, and muscular resistance senses. As Wilkinson [2010] put it, "The graphs we have been talking about cannot be seen, heard, tasted, or otherwise perceived. It is the job of the Aesthetic object to act as a function on the elements of a CGraph and translate them into real numbers or characters that control some display device, such as a CRT, plotter, sound generator, or even odor generator. Once we have done this, we have what I call a graphic. A graphic is a perceivable graph."

The options presented in this article are restricted to today's thinking and must surely be a subset of what a wider collective might dream up. To that end, the authors think of this work as a conversation starter, and it is a conversation worth having, and worth involving as many people as possible, among engineers, developers of statistical methods, and users of statistics. At times, the only difference between an awesome novelty and a foolish notion is the uptake by users; we need inventiveness to create innovation. Until we know how well received an innovation ends up being, all ideas are only potential solutions and opportunities.

1.2 Errant backward compatibility

A fable shared on the internet discusses backward compatibility and rocket ships [Precision Content, 2016]. In this story, the first roads were designed for Roman chariots pulled by two war horses. With this standard established, subsequent transportation used the same gauge without reflection. Eventually, this lack of reflection had consequences. Boosters for rockets to be launched in Florida were built in Utah. Transportation from Utah was a limiting factor for the construction of these boosters; the engineers were limited because of the dimensions of the mountain tunnel through which the train carrying the rockets passed. Because of backward compatibility, the width of war horses impacted upon space travel.

This fable illustrates the need for careful reflection on mundane details—or perhaps its popularity demonstrates a desire for a tidy explanation for our organizational oversights.

Human-computer interactions mostly happen via mouse, keyboard, and screen. The mouse mimics infant communication by pointing, while screens and keyboards are largely modeled on paper-based communication. Emails have the same format as letters but travel faster, and spreadsheets have a similar format to Babylonian tablets. Most computer-generated data visualizations are digital updates of paper-based formats. What is a computer's full capacity to support human

interaction with data?

1.3   The overlap between truth and beauty

Statistical vivification falls in the intersection of science, engineering, and art. The authors have practical and statistical concerns about data display and model understanding; how do these connect to questions of graphical beauty? An apparently hard-nosed realistic response would be that, in engineering, there is no beauty beyond utility. But we think that would be too glib, for several reasons. First, aesthetic concerns have in the past often led, rather than followed, practical developments; consider, for example, all the scientists and software designers who have been influenced by the ideas of Tufte [1983] and Tufte [2006]. Second, as noted by Hansen [2013], choices of display are also questions of design, and the same principles that can make data-based art appealing [Hansen and Rubin, 2003] can inform functional vivifications, in the same way that innovative or "cool" engineering can facilitate scientific development.

2   Tools for conveying information through different senses

2.1   Sight

Data visualizations are the most common mapping between data and the human Umwelt. Since the 1970s, scholars have developed both static and dynamic interactions of this type. Trellis plots are still common but rotating 3D plots have fallen out of use. The Grammar of Graphics [Wilkinson, 2005] continues to have a big impact on the current visualization landscape as implemented at Tableau and the R package ggplot2 [Wickham, 2016]. New ideas have been developed to visualize large datasets [Unwin et al., 2006].

Data visualization bridges between statistics, graphic design, and software engineering. Information visualization and statistical graphics have different goals, but both have become an important part of our culture with news outlets displaying them throughout their websites [Gelman and Unwin, 2013].

An extension of data visualization is immersive analytics, using virtual reality headsets and hand sensing devices to explore data. Some of these are FIESTA [Lee et al., 2021], VRIA [Butcher et al., 2021], Uplift [Ens et al., 2020], u2vis [Reipschläger et al., 2020], and Tiltmap [Yang et al., 2021].

The authors see visualization as a starting point when thinking about other vivifications. In particular, the large body of work on research and practice in visual communication should inform progress in the theory and practice of other sorts of vivification.

2.2   Touch

3D printing and tactile graphics are a direct way that people can feel data. Some methods involve Braille printers [Braillo, 2020, American Printing House, 2021], embossing [Viewplus, 2016], and contouring [Li et al., 2011]. Many of the available tools are open source [Morgan-Wall, 2018, Seo, 2020, Murdoch, 2021]. All can be improved. As 3D printings are durable, it makes sense to improve the accessibility of this technology by adding descriptions about axes, tick marks, point characteristics, and other assumed aspects of plots.

Dragicevic and Jansen [2021] maintain an online gallery of physical visualizations, including a 2010 adaption by Hans Rosling who uses Ikea boxes, pebbles, fruit juice, and even toilet paper. The first ever interactive data visualization was physical. Bertin [1969] developed a way to cluster data using physical cubes arranged in a matrix where columns and rows can be interchanged. His method was groundbreaking since computational methods to cluster complex data did not then exist.

Because the production cost of that physical representations of data is decreasing while the fidelity and flexibility continues to increase, some scholars believe physical representations "will eventually support data analysis tasks as complex as performed on today's desktop computers" [Jansen et al., 2015].

2.3   Muscular resistance sense

Researchers are in the early stages of developing haptic technologies and the authors see potential to use this technology to vivify data. Haptic interfaces currently include gloves [Mimu, 2021], hand-tracking devices [Ultraleap, 2021], recognition of mid-air shapes [Hajas et al., 2020], and vests [Eagleman, 2021]. Some scholars have developed tools which allow the public to "exploranate" the nano scale [Höst et al., 2020]. It is even possible to build your own haptic equipment that detects finger movements [Papetti and Fröhlich, 2019]. One promising technique is used by physicists to allow participants to visceralize dark matter as cave formations [Trotta, 2018]. There is open-source technology that we believe will be useful in extending haptics to statistical tools [Teller et al., 2013, Forrest, 2015, Force Dimension, 2021].

MIT's Artificial Intelligence Laboratory has developed tools for virtual tactile sense communication. Their interface allows users to alter "shape, texture, temperature, weight and rigidity . . . through a device called the PHANToM haptic interface" [MIT News, 1988]. That particular haptic interface is currently produced by SensAble Technologies; Immersion produces a competing product called TouchSense.

2.4 Balance and kinesthetic sense

The Wii technology—motion sensors in the form of a small remote ("Wiimote") which allow a user's movements in three dimensions to be conveyed to a computer, and which also allow the computer to send feedback to the hands—has already been extended beyond sports video games (golf, baseball, and the like) to assist in rehabilitating stroke victims [Vassou, 2008] and in physical therapy sessions [Drummond, 2008]. At the moment the authors are not entirely sure how the Wii technology would be adapted to statistical analysis and data visualization, but we expect that there are some good possibilities and applications to be discovered or created.

2.5 Sound

During the Cold War, Speeth [1961] used audification techniques to discriminate between a bomb and an earthquake and had a 90% detection rate. He expected his methods to be quickly adopted but was unfortunately rejected and isolated, and his security clearance was retracted [Volmar, 2013].

Anyone who has heard an unusual rattling on their bike understands the usefulness of audification. Additionally, mechanics' ability to diagnose problems based upon sound demonstrates that refining listening skills can enhance our understanding of audified data. The authors also know of instances where people with low vision have been able to diagnose issues with wheel alignment in a car by touching the steering wheel (they were not driving).

Pioneering software engineer Margaret Hamilton told the following story [Creighton, 2016] of the SAGE project from around 1960:

> "When my program was running, it always sounded like the sound of waves on a beautiful seashore, so we all referred to it as the 'seashore program.' Until one night at 4 am in the morning when I got a call from one of the computer operators who said something terribly wrong has happened with your program. When I asked him how he knew, he said 'it no longer sounds like a seashore!' Now, we had a new way to debug, using sound!"

This episode indicates how our brains' ability to distinguish timbre can be used to detect changes in data.

Attempts to use sonification for data have been studied since Bly [1982], and there are many directions for improvement: in addition to being beautiful, information can also sound beautiful. Currently, some astronomers develop beautiful sonifications [Siliezar, 2021] but the laborious process is not yet automated or available in multiple programming languages. Scientists have built censuses of sonifications [Lenzi et al., 2021]. Existing open-source software can be used as a start-

ing point [Downey, 2014, Siegert and Williams, 2017, Ligges, 2018, Highcharts, 2021, Twotone, 2021]. Using musical skills, the authors plan to improve and simplify this process.

There are multiple facets to auditory display [Hermann et al., 2011], including exploratory data analysis, parameter mapping sonification, model-based sonification, parametric, and nonparametric methods.

In heliophysics, audification helped researchers identify "artificial, instrument-induced noise that was not previously observed by the scientist and also the identification of wave activity embedded within turbulent solar wind data" [Alexander et al., 2014]. Two months later the researchers continued use of this technique.

Positions of points on a scatterplot and lines on a graph could, with care, be indicated with sound. Compared with the senses of touch or body location, sound is less localized, so it would make sense to use pings not to tell where data are but rather when they are encountered. For example, if a robot arm (or, perhaps better still, the user's bare hand) is swept over a constrained space, data thus encountered on the virtual scatterplot could make the sound of rain, with individual drops or pings for isolated datapoints and a louder pitter or torrent for increasingly dense clusters. This method was used to audify COVID-19 deaths [Klirs, 2021]. The authors could also incorporate spatialized 3D sound technology in this area.

More generally, sound could be used at the user's discretion to reveal the values of additional variables that are not in the physical display. For example, the user could click a mouse at any time to learn additional information about any data point that he or she encounters in exploration, or the computer could be asked about information such as the number of nearby points, the integral of a function in a specified range, or other summaries of a graph.

2.6 Voice dialogue

Voice driven development is another technological tool that can increase accessibility [Nowogrodzki, 2018, Pearl, 2019]. There is software that assists those with repetitive strain injury that can no longer type with ease to "speak" with their computer [Caballar, 2021, Talon, 2021, Serenade, 2021]. The software allows users to verbally code in computing languages [Rudd, 2013]. We will build something more conversational where the computer recognizes keywords and fills in a template with variable names. The authors show a hypothetical example script in Section 4.2.

The DataBreeze interface currently ties together pen, touch, and speech-based interactions to explore data [Srinivasan et al., 2020]. This type of technology could be made available in statistical software.

Another way to interact with visualizations is natural language interfaces. Eviza [Setlur et al., 2016] is a Tableau prototype that allows users to interactively dialogue with an existing visualization. Users do not need to know Tableau; the phrase "show me wildfires in California" will

execute a query. DataTone uses voice recognition software to the address multiple ambiguities when querying a database [Gao et al., 2015]. For example, the above query does not specify the year or the size of the wildfire. DataTone will display boxes from which the user can select both year and size and updates the plot as the user makes their selections. FlowSense is a visualization tool that also has voice interaction and allows users to type queries into the online tool [Yu and Silva, 2020].

The authors do not expect voice interaction to supplant the usual menu or command-based systems. Oddly enough, however, many of the steps that will be required for the voice dialogue system—an ability to anticipate the user's requests and to present results concisely—are important in statistical software in general. It may be that the tools we develop here will prove useful generally as well as for those who cannot type for reasons relating to their dexterity.

2.7 Smell and taste

In 1960 Elizabeth Taylor starred in The Scent of Mystery, which involved not only audio and visual stimulus but novel technology called Smell-O-Vision; it was the first ever movie that stimulated olfactory senses beyond the smell of buttered popcorn. The company Scentevents participated in a remake [Scentevents, 2015]. Scratch and sniff cards have also been used for a similar purpose for the movie Polyester [Phelps, 2018]. Nakamoto [2016] offers an overview of current research in conveying information through smell and taste; the author admits that this research is in a primitive state. A chef from a three Michelin-starred restaurant was building an olfactory extension to the iPhone [Forbes, 2014]. A company in Vermont is also working to include scents into virtual reality headsets [Lefrak, 2021]. In addition to visual cues, stimulation of olfactory senses can be considered to explore data.

2.8 Combining multiple senses

In addition to technological advances that deliver data differently, there are also creative methods. An example of this is the Boston Coastlines Future Past in which artists organized a walking tour of the projected coastline after sea levels rise from global warming [Sutton, 2016]. A Sort of Joy [Museum of Modern Art, 2015] is another example of delivering data differently. In this short play, artists with work at the museum are organized by first name. Instead of a bar chart to show which first names are most common, actors spoke the name and allowed time to pass that was related to the height of the bar. Additionally, instead of coloring bar charts based upon gender, male voices verbalized names of male artists and female voices names of female artists. Another example involves data physicalization of water pollution created collaboratively with community partners in Chelsea, Massachusetts [Perovich et al., 2018]. Some artists are asking the question,

8

"What does data theatre look like?" [Brea, 2021]. The authors believe there are scientists who may be interested in these questions.

A more concrete example of combining senses could be the exploration of level curves of a price function (or, for a more homely example, isotherms in a weather map) using the robot arm. As the user pushes the arm across the curves, the system could say aloud what is happening ("Crossing the twenty-degree line from below ... crossing the thirty-degree line from below ... ").

3   Perceptual differences

Each sense has a unique interaction with the brain. Most people have favorite songs that they repeat, but it is unusual to reread a book or rewatch a movie. Music cognitionists agree that music is a repeating sense [Margulis, 2013]. Music also elicits emotion as a matter of course, unlike visual imagery which tends to elicit emotional reaction only when tied to particular content. Another advantage of a sound is that it may not consume all the attention of the brain the same way that movies, books, or visuals do. Background music is a common term that describes this. In visual senses, focus can impact upon our ability to perceive obvious things even if our eyes see them. One example of this is the invisible-gorilla study [Chabris, 2010].

The authors believe that the processing methods of the brain should impact upon how the senses are used. Compared to sound and sight, a haptic tool making use of the muscular-resistance sense is more conversational. The tool reacts to user input by giving more information which results in a looping effect similar to speaking with someone.

Unlike some of the other senses, smell has a strong discrete aspect and does not always naturally fall on a continuum; hence the authors suspect it could be used in the way that we use different symbols or colors to distinguish points on a graph, or as an alarm (if the sound channel has already been taken for more useful purposes).

Taste, like sound, has an emotional valence that responds positively to repetition. Researchers have shown that childhood exposure to high salt and high fat diets "exacerbate[s] a number of illnesses and conditions in adulthood" [Henney et al., 2010, Gidding et al., 2009, Ekelund et al., 2007, Barker, 2004, Mennella, 2014] which suggests that taste is a sense that is calibrated in early childhood. It is not clear to us how taste would be used in data perception, but to the extent this is done, we would want to make use of this property.

Touch and the sense of body location are inherently spatial but in a different way than the spatial organization of vision. Another complexity is the relation between sensory perception in time. For senses other than sound, observation in time can require active focus on short-term memory. For example, the dynamic scatterplots of Rosling [2015] have been extremely popular as a way of visualizing the way that differences between countries change over time, but it becomes difficult to "read" these graphs with a goal of understanding details: there are too many patterns

to focus on at once, and this leads to a frustrating pattern of repeated viewing and attempts at memorization. The goal is not just to see things as they are happening; we also need to store the patterns, and deviations from these patterns, in an accessible and inviting way.

Designing tools that engage non-visual senses in a statistical workflow can be an overwhelming task. The authors suggest that scientists think carefully about the perceptual advantages of the sense they would like to incorporate and work to discern how the advantages can improve their workflow. Braille is an accessible form of writing for blind and low vision people, but American Sign Language is a completely different language dissimilar to English. Both are necessary. Vivifications of plots for the blind are important, but the authors additionally want to develop a whole language of interactions with computers that will improve statistical workflows. Our goal is beyond replacing visualizations.

4   Some possible applications to statistics

The authors would like to use vivification technology to improve statistical practice, both for people with low vision and for analysts more generally who can use non-visual senses in order to increase their understanding of data and models. In this section we speculate on different ways in which technological developments can be turned into tools for data analysis.

4.1   Sensory scatterplots and imaging

Remember the toy that is made up of pins held loosely in a grid of holes? You push up with your hand on one side, and then a three-dimensional image of your hand in pinheads appears on the other side. One could use an electromechanical version of this toy to graph surfaces, curves, and scatterplots in a way that could be felt by the user's hand. Linked up to a computer, the pinboard could also convey a series of snapshots or changes over time, thus potentially displaying four dimensions in a way that is not possible using standard visualization tools. Whether or not the pinboard becomes useful as a supplementary tool for sighted users, the authors anticipate that it could be an excellent way to convey complex graphical information for people with low vision. The American Printing House for the Blind confirmed it is currently developing such a device, and a more expensive European model is currently available.

When 3D printing a plot, variables could be mapped to different textures. With X-ray data, NASA maps silicon to a smooth texture, iron is clumpy, and argon and neon are sandpaper with differing grit [Arcand, 2021]. If each variable is mapped to a different texture, the authors would recommend 3D printing the path of each coefficient alone, and the axes with tick marks, and a grid alone. The full plot could be built by stacking all the pieces together which is similar in process to building a ggplot.

4.2 Voice dialogue for real-time data analysis

Voice-driven development combined with natural language interfaces has the potential to benefit data scientists and statisticians. Here is a potential script:

> User: Regress earnings on height.
> Computer: Earnings equals sixteen thousand plus four hundred times centered height plus error.
> User: Error?
> Computer: Mean zero, standard deviation eighteen thousand.
> User: Centered height?
> Computer: Height in inches minus sixty-five.
> User: One more significant digit.
> Computer: Sixty-five point two.
> User: Standard errors of the regression.
> Computer: Intercept: sixteen thousand, standard error two thousand. Coefficient on centered height: four hundred, standard error one hundred fifty.
> User: Include sex as a predictor.
> Computer: Earnings equals four thousand plus eleven thousand if male plus two hundred times centered height plus error.
> User: Interact sex and height.
> . . .

The idea is that interactive dialogue maps to statistical workflow in a way that is not captured by usual coding or menu-based approaches. Thus, this could be a tool that would augment the toolkit of sighted data analysts as well as providing a replacement for visualizations for those with vision impairments.

The authors suspect that a voice-recognition program where users can regress on chosen variables, which then speaks out coefficients and errors would greatly improve statistical processes. The user may then ask for regression with interaction terms. This process could be used with linear regression, logistic regression, and include multilevel models with different voices that could then speak coefficients and errors.

One of the main commercial applications of a computer voice dialogue is for it to act as a human or even to fool users into thinking it's human (as for example when it's the back-end for an online tool to resolve customer complaints). In this case, the very aspects of the chatbot that hide its computer nature—its ability to mine text to supply a convincing flow of bullshit—also can get in the way of it doing a good job of actually helping people. This suggests that chatbots would be more useful if they more transparently revealed their computer nature and didn't try to

emulate people, in the same way that its better for a chessbot to just be itself and not try to fake us into thinking we're playing against a human.

4.3   Voice recognition for exploring matching in clinical trials

Sturdevant et al. [2021] present a tool for finding balance in a two-armed clinical trial at randomization. The authors provide a Shiny web application [Chang et al., 2015] where researchers can practice randomization conditional on a set of matches. To update the matches, researchers must scroll down to the variable, update the weight, then scroll to the top again to view possible randomizations. Scrolling disrupts the exploration process and locating the desired variable is tedious. One solution to this is voice recognition. After turning on voice recognition, researchers could use the keywords to update weight, then the variable name and the weight to interact with the computer. The user could then click on a button at the top of the "Matching Variables Info" tab that tells the application to update the matches and produce a plot.

4.4   Using musical sounds to convey the progress of iterative algorithms

Instead of using sound as a replacement for visuals, we can augment our interactions with computers with sound. For example, modern data-fitting algorithms tend to be iterative, with an initial transient period as the algorithm gets close to the solution and then a slower convergence to a point (for an optimization algorithm) or a distribution (for a Monte Carlo algorithm). In either case, the progress of the algorithm could be expressed using musical sound, with a change in pitch during the transient period, settling to a single note or distribution of notes. More can be done along these lines: savvy users know to run multiple chains of algorithms such as variational inference or Markov chain Monte Carlo using different starting points. The spaces in parameter space being occupied by different chains can be represented by different keys and timbres, so that the progress of mixing of initially-separate chains can sound something like different musical instruments being tuned.

We can also use pitch, volume, rhythm, timbre, and the progressions of musical expectation to convey what is happening within the fits and starts of an iterative algorithm. For example, Hamiltonian Monte Carlo can get stuck when its steps are too small, too large, or misaligned to the local geometry of the space being explored; within an algorithm such as the no-U-turn sampler, these problems can appear as divergences or max treedepths. Instead of expressing these as discrete warnings on the screen, these could be conveyed, for example, through pitch, with well-behaving trajectories sounding like calm music, gradually turning into annoying high-pitched buzzing when the algorithm is getting stuck. This sonic summary could be tied to a dynamic

visualization so that the user could then take a look at where in the parameter space this is happening.

A related application is the training of neural networks, where progress is tracked using improvement of some objective function during the training process, and, again, sound can be used to convey progress of the algorithm from different starting points, with unusual or unexpected sound patterns indicating a lack of smooth progress. The sonic of musical output can serve to reassure that the algorithm is progressing well or alert users of problems, and also to signal when and where to perform further exploration, perhaps using visualizations, to diagnose and fix problems.

4.5   Exploring fitted curves and likelihood surfaces haptically

Imagine a robot arm that can be moved within a 30×40 cm box. Such an arm could be used by a blind user to "feel" a curve (for example, a nonlinear regression or spline, or a mathematical function such the logarithm or the normal distribution curve), as follows: the arm would start at one end of the curve and the user could grip it and move it along, with the motion physically constrained so that the arm would trace the curve.

In displaying several surfaces—for example, level curves indicating indifference curves—the arm could start on one of the curves and be programmed to stay on that curve, unless the user pushes hard, in which case there would be resistance during which the arm moves between curves. It would then lock into the next curve, which the user could again trace until he or she pushes hard enough to get the arm unstuck again.

More generally, the robot arm could be used for exploring three-dimensional functions such as physical potentials, likelihood functions, and probability densities. From any point in the two-dimensional box, a "gravitational force" would pull the arm toward a local minimum (or, for a likelihood or probability density, the maximum) of the function. Then with moderate effort the user could move the arm around and, by feeling the resistance, get a sense of gradients, minima, and constraints.

Various displays could be implemented as three-dimensional functions and then explored using the robot arm. For example, a scatterplot can be converted to a kernel density estimate. When moving across sparse data, the arm would feel a slight pull or stickiness as it moved past individual data points; in data-dense parts of the scatterplot, the massed points would pull the density down while at the same time allowing the user to feel individual data as small local tugs on the arm. For this purpose, it might make sense to use a kernel function such as a Student-t so that each point has a sharp local peak within a broader area where it contributes to the density estimate.

Alternatively, walking with each hand on a physical curve to feel the differences between them is another possibility. The differences in heights between the two hands and the comparisons of the sensations in the arm and attached muscles would allow users to feel the relationships between

the curves.

4.6   Parallel processing

The authors envision a program that would integrate with cellphones and send messages when simulations have finished. From here, researchers could update arguments and rerun models through voice-recognition similar to sending a text message. We anticipate that some aspects of modeling can be done hands-free and remotely from a cell phone.

4.7   Teaching

Histograms are a standard data visualization tool. Teaching this plot involves discussing an appropriate bin width. With current technology learners plot multiple plots to discern an appropriate range for this argument. A better teaching tool might be developed using a haptic glove where the distance between the thumb and forefinger align with the bin width. The bin width of the histogram would increase simultaneously with the distance between the fingers. This mimics the growth of the bin and allows for somatic understanding of what the argument does that is impossible to achieve by solely changing the argument in a function. This tool would engage students who struggle to understand this and deepen the understanding of others.

An extension of the histogram is the density plot. Like the histogram, drawing this plot involves the bin width argument for which the authors suggest a glove. In addition to the bin width argument there is a kernel argument that could be voice activated. If the output is a smooth curve—say, from a Gaussian kernel—sonifying the resultant density function could deepen the understanding of the learners and engage learners who might not understand. When the bin width is too small the pitch of the sonification would vary widely; when it is too big it would be overly simple.

In both of the above examples 3D printing would also deepen understanding. For the histogram, a 3D dot plot of the values that can be layered onto multiple 3D histograms with different bin widths would add an extra dimension. Density curves could also be 3D plotted.

An even more fundamental example is the mean. The mean is also the center of mass for a distribution which could be built with 3D printing. Learners would have a deeper understanding of the importance of this computation.

Correlations could be explored using a virtual reality (VR) headset. A scatterplot where the axes can be shifted from the origin to the means of x and y, then the squares for the numerator drawn with different colors for negative and positive values. Seeing multiple data sets used to iterate the numerator would allow learners to deeply understand the reasoning behind the formula in a way that is difficult.

Iterative algorithms like the expectation-maximization algorithm are similar to walking up a lighthouse. We could demonstrate this with a VR headset. The observed data log-likelihood after convergence would be visible at the top. The expectation step can be seen on one side of the lighthouse and maximization on the opposite. For people with low vision, the distribution could be 3D printed. This reinforces the dependence of the current distribution on prior iterations in a way that is difficult without these somatic cues. This could be extended to explore different initial values by changing the location of the initial step. Additionally, the number of iterations can vary between viewings, or controlled through voice-recognition software.

The authors take a broad view of learners. Data scientists often need to explain concepts to stake holders and we anticipate that these tools could be used outside of a traditional classroom.

4.8   Using pitch, volume, and timbre to convey statistical information

The past two years have been altered by COVID-19. Several months into the pandemic there were a number to tests available to people but understanding sensitivity, the true positive rate, and specificity, the true negative rate, could be communicated multiple ways besides stating the definition. Sonification could be explored. An octave is familiar to most ears. To sonify testing data we could start with a note and use the octave increase for 100% sensitivity, and a decrease of an octave for 100% specificity. The actual sensitivity and specificity notes would be prorated to the actual value; we wouldn't expect the actual octave since tests with 100% sensitivity or specificity are uncommon. In addition, we could use a glissando (glide) sound starting at the original then rising to the sensitivity level, return to the original tone, then glissando down to the specificity. As different tests take different amounts of time to process, the speed of the glissando could be related to the processing time, with quicker tests having shorter "songs" and vice versa. To extend this, multiple tests could be considered with dependence in the calculations of sensitivity and specificity and fixed speed of the glissando.

Another opportunity to sonify COVID relates to the differences in rates of change in different countries. Vanderplas et al. [2021] explores pandemic graphs, and one plot, "COVID-19: How many days it took to reach 500 deaths?", could be sonified using glissandos. Each group of 500 deaths could be a one-octave glissando with duration related to the time taken for a country to reach that iteration. Timbre should be chosen carefully to reflect the seriousness of the topic.

As the size of data increases, variable selection has increased in importance. Lasso and ridge regression are both methods that can be used to fit multiple models each with different penalties (lasso is the sum of the squares of the coefficients, ridge is the sum of the absolute values of the coefficients). There are a number of ways to vivify variable selection. The resulting plot from ridge or lasso regression of coefficient estimates vs. the regularization parameter λ could be sonified, 3D printed, or seen through a VR headset. For sonification, each variable could be mapped to a

different timbre dependent on the type of variable. For one-variable data, astronomers visualize x-rays (high-energy photons) with blueish purple which is the highest frequency of visible color. To map to timbre, the xylophone is the instrument with the highest timbre so it is mapped to the highest-energy photons. Following this trend, violins is used for near-infrared light, and soft piano for infrared light [Arcand, 2021]. A similar process could be used for multiple variables; synesthesia might be helpful.

As λ becomes large, all coefficients go to zero due to the penalty. For this reason, the authors suggest reading the plot from right to left instead of the normal left to right. All variables/timbres/instruments start on the same tone which represents 0 and a high value of λ, as time passes, the pitches adjust with the values of the coefficients as the value of λ decreases. The sonification ends with the coefficients from the least squares fit sonified when λ is 0. Users should have the ability to choose the variables they would like sonified to simplify the sound if they desire.

In terms of a VR headset, the authors again recommend reading the plot from right to left and enlarging the plot a considerable amount. Visualizing each coefficient path as a tunnel wherein the user can travel from a large λ to a λ of 0 could increase understanding. One day it might be possible to experience it like an amusement park ride.

5   Discussion

5.1   First steps

One can start by applying existing technology to tasks of statistical data analysis, For example, it should be possible to build a digital signal processing Downey [2014] package in R to enable works like those done at System Sounds to be automated in R [Siliezar, 2021, System Sounds, 2017]. Sonification is a natural extension of data visualization for low-vision and blind people. The next step would be to develop data vivifications that include both sight and sound. Some preliminary work has been completed in this field [Storms and Zyda, 2000, Hermann et al., 2011] but more is needed.

Simultaneous to this, it would make sense to develop a network of people within the low-vision and blind community to support this process. To achieve the goal of making data more accessible quickly, listening to those who have been continually disenfranchised is imperative. Technical developments should be done in concert with this group to ensure that the tools being built, and the order in which they are being built, are optimized to meet the needs of the community. This would build upon community-based research initiatives where participants guide researchers towards a shared goal.

5.2 Using theoretical ideas to evaluate and understand particular data vivifications

Chernoff reduced the multiple complexities of faces to a few characteristics like length of the nose and curvature of the mouth. To extend his method, Chernoff [1973] suggested "adding ears, hair, [and] facial lines." The abstract of the paper uses the term "cartoon," and humans have evolved to read people not cartoons. Chernoff did not understand the perceptual properties of faces before trying to create them [Kosara, 2007]. This method simplified the human Umwelt of a face.

It is clear to us that Arabic numerals are much easier to manipulate than Khipu. Khipu were not calculators so extensions of this method to basic arithmetic was not possible [O'Connor and Robertson, 2001]. This method was too complex to be extended.

To summarize, Chernoff faces simplify the human Umwelt without understanding them, and Khipu is not simple enough to extend. Twenty years from now, data vivification methods will likely include methods that are simple enough to be extended, as well as complicated enough to represent the human Umwelt. This is not to say that Chernoff's work is valueless; we often need to try out ideas and see what does not work in order to better understand how to do better.

An exception to this is the famous image by Charles Minard of Napoleon's armies diminishing as they marched through Russia. This mode of display is not extendable or applicable in many other settings—it takes advantage of a particular data structure that allows multiplexing of time and two spatial dimensions—but many experts hold it in high regard as a model of data storytelling.

Lack of diversity in technical fields has had negative consequences on what data are collected and on the development of algorithms [Angwin et al., 2016, Buolamwini and Gebru, 2018, Eubanks, 2018, Sjoding et al., 2020]. This goes beyond racism and sexism; the community also lacks members with physical constraints such as restricted sight and hearing. This impacts all aspects of the data pipeline, including perhaps a problem of integrating communication and visualization with other aspects of statistics. Through delivering data differently, the authors hope to foster a more inclusive environment throughout science.

5.3 Engaging the open-source software community

A large community exists around the open-source statistical programming language R. The pre-eminent conference, useR! 2021, has multiple tutorials available in languages other than English. In addition, the R Consortium sponsors over 92 R User Groups in over 38 countries with over 60,000 members. R Forwards is another community whose mission is "leading the R community forwards in widening the participation of women and other under-represented groups." R Ladies is group that works towards diversity in the R community; at publication, there are over 75,000 members in 57 countries. The community is focused on including those from diverse backgrounds

and experiences. One focus of The Minority in R community is accessibility. Several members are blind and others have extensive backgrounds in accessibility and are presenting and teaching the topic.

Although the RStudio interactive development environment is not yet accessible, using R through emacs is accessible for those with low vision [Godfrey and Loots, 2014]. Our second author confirms that, despite being relatively unfamiliar with R, he readily embraced it after publishing a comparative analysis of R, SAS, SPSS, and Minitab for people with low vision [Godfrey and Loots, 2014].

It is important to engage this community about delivering data differently as part of a long conversation that will not be limited by today's thinking.

5.4   Replacements and improvements

We tell our children, "Don't try to be someone else. Just be the best 'you' you can be." The same goes for the methods discussed in the present article:

- Sonifications in the form of music can be emotionally engaging and can also be processed in the background, unlike visualizations, which require conscious attention to convey their information.

- The haptic sense is inherently interactive, as it is through pushing that you feel the resistance in your muscles.

- Through the sense of touch you can perceive three dimensions directly with no need for projections.

- An effective voice dialogue has no need to pass a Turing test; it is more effective when it is transparent enough that the user can work directly with its algorithmic nature.

The data-exploration tools we would like to develop can serve as replacements for visual tools for blind users, but they also have their own unique advantages. Thus, when considering vivification of data or statistical models using sound or touch or the muscular resistance sense, let's try to go directly from the statistical or scientific problem to the sense in question, rather than first "translating" to a visual display and then "translating" to the non-visual mode.

5.5   Looking forward

Technology can expand the human Umwelt [Hutson, 2018]. The scientific term for this is sensory augmentation. Skin is a particularly good at sensing different patterns and pressures, and there has been research on building sound-to-touch interfaces that work without lipreading [Novich

and Eagleman, 2015]. The future is likely to include interactions with computers far beyond our imagination as we learn to develop new skills. They will likely not be limited to keyboard and mouse; indeed, those technologies may go the way of the rotary phone and typewriter.

Computer scientists will benefit from the support of consultants with advanced somatic skills to imagine new methods of interaction. The seminal work The Body Keeps the Score discusses the consequences of trauma upon the brain [van der Kolk, 2014]. Few people make it through life without times of loss of safety; imagine the results of such intrusions on your body. Scientists believe that if the body feels safe the brain heals [van der Kolk and Fry, 2021], and prioritizing feelings of safety could be one way forward. Methods that assess how technology impacts upon the human body will support this work.

In addition, consider the expansion of of nascent technologies, and technologies not yet developed, to include statistical methodologies. Currently, work on visualization and vivifications focuses on data display, but this is only part of statistical or data science workflow. A first step going forward would be to identify a set of example problems in applied statistics and explore implementations of different sorts of vivifications. Once some examples have been made available, it should be possible to think of how to integrate a range of vivification tools into data science workflow. Given that it took many years to incorporate visualization into applied statistics, we do not anticipate this will be easy, but we think the way forward is to take advantage from all the available technology (some of which has been reviewed in the present article) and think about how to apply them to various challenges of real-world data analysis.

We believe that future delivery of data will stimulate more of the human Umwelt then current delivery methods. Research suggests that high quality audio increases perception when coupled with high quality visualizations [Storms and Zyda, 2000, Hermann et al., 2011]. This has the potential to increase accessibility for those with low vision and increase perceptiveness for all. We would like to see this happen faster and would like those with statistical training to take leadership roles.

Scientific misunderstandings abound even among the well educated [Schneps et al., 1989]. Indeed, misconceptions are an important part of the learning process as students integrate new knowledge [Sadler, 1998]. One way to address this knowledge is data vivification where students have the opportunity to integrate new knowledge using multiple aspects of their Umwelt [Schneps et al., 2014].

Vaccines only work if communities take them. With the rise of mistrust in science, scientists must carefully consider the limitations of the impact of their work if communities do not believe them. Research suggests that delivering data differently may mitigate some of these issues. We encourage scientists to learn about and utilize these approaches in reaching out to their communities.

In the words of Matsouaka [2021], "the most important skill for a translational scientist [is] to

be able to work in a collaborative environment with people from other fields. As a biostatistician, having this skill is not a choice. You have to work with doctors and scientists to solve problems. You must communicate across disciplinary boundaries, earn trust and learn to speak your colleagues' specialized language. Over time, you discover what is important to your collaborators and understand why they work a certain way. For example, the cardiologists I [Roland Matsouaka] work with not only have training in medicine but are informed by constant interaction with patients. Their perspective reminds me that my job is not just to write complicated equations—I also need to translate those equations into a common language that people can understand." We see our work as extending the work of translational science within scientific practice and among the general public.

5.6   Win conditions

We see two success criteria. Most directly, we would like for blind and low vision people can more easily access, experience, and work with. Beyond that, we would like to develop vivifications that make statistical workflow more effective for sighted people by making use of the unique benefits that other senses have to offer. This should not be controversial; consider, for example, the familiar observation that learning and memory are enhanced when multiple senses are involved, as when we can remember a particular formula by visualizing where it was on the blackboard when it was presented in a lecture, or when a long train of thought is triggered by an unexpected smell or a sudden change in temperature. Stilgoe [1998] discusses the way that even familiar landscapes convey unexpected information and stimulate creative thought, if we train ourselves to focus on details. Vivifications should be a way to do this. The authors would like to develop tools that will continue to be useful and be developed, not to solely go viral and then be forgotten. The future of our current practice of machine learning, data science, and artificial intelligence is what we would like to develop and explore and what we are addressing in this article, and one challenge going forward is to consider how to measure the success of this endeavor.

References

R. L. Alexander, S. O'Modhrain, D. A. Roberts, J. A. Gilbert, and T. H. Zurbuchen. The bird's ear view of space physics: Audification as a tool for the spectral analysis of time series data. *Journal of Geophysical Research: Space Physics*, 119(7):5259–5271, 2014. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JA020025.

American Printing House. Tactile graphics kit, 2021. URL https://www.aph.org/product/tactile-graphics-kit/.

J. Angwin, J. Larson, S. Mattu, and L. Kirchner. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. URL `https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing`.

K. Arcand. Exploring our universe through sight & sound (or touch!). YouTube video, 2021. URL `https://www.youtube.com/watch?v=Q_r7KKdH38E`.

D. J. P. Barker. Developmental origins of adult disease. *Journal of the American College of Nutrition*, 23(90006), 2004.

J. Bertin. Graphique et mathématique : Généralisation du traitement graphique de l'information. *Annales. Histoire, Sciences Sociales*, 24(1):70–101, 1969.

S. Bly. Presenting information in sound. In *Proceedings of the 1982 Conference on Human Factors in Computing Systems*, CHI '82, page 371–375, New York, 1982. Association for Computing Machinery. URL `https://doi.org/10.1145/800049.801814`.

Braillo. Tactile graphics, 2020. URL `https://braillo.com/tactile-graphics/`.

A. Brea. What does data theatre look like? Lessons from our first workshop. Data Culture blog, Northeastern University, 2021. URL `https://medium.com/data-culture-group/what-does-data-theatre-look-like-lessons-from-our-first-workshops-287f32e9bc50`.

G. Brokaw. *A History of the Khipu.* Cambridge University Press, 2010.

J. Buolamwini and T. Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81:77–91, 2018. URL `http://proceedings.mlr.press/v81/buolamwini18a.html`.

P. W. S. Butcher, N. W. John, and P. D. Ritsos. Vria: A web-based framework for creating immersive analytics experiences. *IEEE Transactions on Visualization and Computer Graphics*, 27(7):3213–3225, 2021.

R. D. Caballar. Programming by voice may be the next frontier in software development. IEEE Spectrum, 22 Mar., 2021. URL `https://spectrum.ieee.org/computing/software/programming-by-voice-may-be-the-next-frontier-in-software-development`.

C. F. Chabris. *The Invisible Gorilla and Other Ways Our Intuitions Deceive Us.* Crown, New York, 2010.

W. Chang, J. Cheng, J. Allaire, C. Sievert, B. Schloerke, Y. Xie, J. Allen, J. McPherson, A. Dipert, B. Borges, RStudio, jQuery Foundation, jQuery contributors, jQuery UI contributors, M. Otto, J. Thornton, B. contributors, I. Twitter, P. N. Khan, V. T. D. Lembree, S. Chakravarthula, C. O'Connor, I. PayPal, S. Petre, A. Rowls, D. Gandy, B. Reavis, S. Bejaoui, D. Ineshin, S. Samhuri, S. Limited, J. Fraser, J. Gruber, I. Sagalaev, and R. C. Team. Shiny, 2015. URL `https://CRAN.R-project.org/package=shiny`.

H. Chernoff. The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 68(342):361–368, 1973.

J. Creighton. Margaret Hamilton: The untold story of the woman who took us to the moon. Futurism, 20 July, 2016. URL `https://futurism.com/margaret-hamilton-the-untold-story-of-the-woman-who-took-us-to-the-moon`.

A. B. Downey. Think DSP: Digital signal processing in Python, 2014. URL `https://greenteapress.com/thinkdsp/thinkdsp.pdf`.

P. Dragicevic and Y. Jansen. Gallery of physical visualizations, 2021. URL `http://dataphys.org/list/gallery/`.

M. Drummond. Wii-hab: Video games are becoming the new trend in rehabilitation hospitals, 2008. URL `https://web.archive.org/web/20080328144429/http://disabilities.suite101.com/article.cfm/wiihab`.

D. Eagleman. Sensory substitution, 2021. URL `https://eagleman.com/science/sensory-substitution/`.

U. Ekelund, K. K. Ong, Y. Linné, M. Neovius, S. Brage, D. B. Dunger, N. J. Wareham, and S. Rössner. Association of weight gain in infancy and early childhood with metabolic risk in young adults. *Journal of Clinical Endocrinology and Metabolism*, 92(1):98–103, 2007.

B. Ens, S. Goodwin, A. Prouzeau, F. Anderson, F. Y. Wang, S. Gratzl, Z. Lucarelli, B. Moyle, J. Smiley, and T. Dwyer. Uplift: A tangible and immersive tabletop system for casual collaborative visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 27(2): 1193–1203, 2020.

V. Eubanks. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin's Press, 2018.

P. Forbes. Mugaritz is developing smartphone smellovision. Eater, 3 Feb, 2014. URL `https://www.eater.com/2014/2/3/6286937/mugaritz-is-developing-smartphone-smellovision`.

Force Dimension. What is haptics?, 2021. URL `https://www.forcedimension.com/`.

N. D. Forrest. Protohaptic: An interactive haptic prototyping tool. `https://code.google.com/archive/p/protohaptic/`, 2015.

M. Friendly and D. Denis. The early origins and development of the scatterplot. *Journal of the History of the Behavioral Sciences*, 41:103–130, 2005.

T. Gao, M. Dontcheva, E. Adar, Z. Liu, and K. G. Karahalios. Datatone: Managing ambiguity in natural language interfaces for data visualization. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, page 489–500, New York, 2015. Association for Computing Machinery. URL `https://doi.org/10.1145/2807442.2807478`.

A. Gelman and A. Unwin. Infovis and statistical graphics: Different goals, different looks (with discussion). *Journal of Computational and Graphical Statistics*, 22(1):2–28, 2013.

S. S. Gidding, A. H. Lichtenstein, M. S. Faith, A. Karpyn, J. A. Mennella, B. Popkin, J. Rowe, L. Van Horn, and L. Whitsel. Implementing American Heart Association pediatric and adult nutrition guidelines. *Circulation*, 119(8):1161–1175, 2009.

A. J. R. Godfrey and M. T. Loots. Statistical software (R, SAS, SPSS, and Minitab) for blind students and practitioners. *Journal of Statistical Software*, 58(1):1–25, 2014. URL `https://www.jstatsoft.org/v058/s01`.

D. Hajas, D. Pittera, A. Nasce, O. Georgiou, and M. Obrist. Mid-air haptic rendering of 2D geometric shapes with a dynamic tactile pointer. *IEEE Transactions on Haptics*, 13(1), 2020. URL `http://sro.sussex.ac.uk/id/eprint/89337/`.

M. Hansen. Data-driven aesthetics. *New York Times*, 2013. URL `https://bits.blogs.nytimes.com/2013/06/19/data-driven-aesthetics/`.

M. Hansen and B. Rubin. Listening Post, Whitney Museum of American Art, New York, 2003.

J. E. Henney, C. L. Taylor, and C. S. Boon, editors. *Strategies to Reduce Sodium Intake in the United States*. National Academies Press, Washington, D.C, 2010.

T. Hermann, A. Hunt, and J. G. Neuhoff, editors. *The Sonification Handbook*. Logos Verlag, Berlin, 2011.

Highcharts. Accessibility module, 2021. URL `https://www.highcharts.com/docs/accessibility/accessibility-module`.

M. Hutson. The sensory revolution. *Psychology Today*, May 2018. URL `https://www.psychologytoday.com/us/articles/201805/the-sensory-revolution`.

G. Höst, K. Palmerius, and K. Schönborn. Nano for the public: An exploranation perspective. *IEEE Computer Graphics and Applications*, 40(2):32–42, 2020.

Y. Jansen, P. Dragicevic, P. Isenberg, J. Alexander, A. Karnik, J. Kildal, S. Subramanian, and K. Hornbæk. Opportunities and challenges for data physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, page 3227–3236, New York, 2015. Association for Computing Machinery. URL `https://doi.org/10.1145/2702123.2702180`.

C. Klirs. What does 425,000 covid deaths sound like? Twitter, 2021. URL `https://threadreaderapp.com/thread/1356267828874256387.html`.

R. Kosara. A critique of Chernoff faces. Eager Eyes blog, 25 Feb, 2007. URL `https://eagereyes.org/criticism/chernoff-faces`.

B. Lee, X. Hu, M. Cordeil, A. Prouzeau, B. Jenny, and T. Dwyer. Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *IEEE Transactions on Visualization & Computer Graphics*, 27(2):1171–1181, 2021.

M. Lefrak. Vermont tech firm believes to experience the metaverse, you have to smell it too, 2021. URL `https://www.npr.org/2022/03/16/1086832763/vermont-tech-firm-believes-to-experience-the-metaverse-you-have-to-smell-it-too`. Accessed: 2022-03-16.

S. Lenzi, P. Ciuccarelli, H. Liu, and Y. Hua. Data sonification archive, 2021. URL `https://sonification.design/`.

B. Li, Z. Wang, J. Liu, and G. Fang. New blind graphic printing technology based on new braille printer. In *2011 International Conference on Computer Science and Service System (CSSS)*, pages 1403–1406, 2011.

U. Ligges. tuneR, 2018. URL `https://cran.r-project.org/web/packages/tuneR/tuneR.pdf`.

E. H. Margulis. *On Repeat: How Music Plays the Mind*. Oxford University Press, New York, 2013.

R. A. Matsouaka. Using abstract math to solve real-world crises, 2021. URL `https://ncats.nih.gov/ctsa/profiles/roland-matsouaka?fbclid=IwAR3Qxalc2Db4poPtbiiiFPQr-_AVkg5YpBO9qURaz06C2ay3MKbtgxPsKCA`.

J. A. Mennella. Ontogeny of taste preferences: Basic biology and implications for health. *American Journal of Clinical Nutrition*, 99(3):704S–711S, 2014.

Mimu. Glover, 2021. URL `https://mimugloves.com/glover/`.

MIT News. Device turns computer into tactile tool, lets user touch virtual objects, 1988. URL `https://news.mit.edu/1998/phantom`.

T. Morgan-Wall. From R to reality: 3D printing maps with Rayshader, 2018. URL `https://www.tylermw.com/3d-printing-rayshader/`.

D. Murdoch. rgl, 2021. URL `https://cran.r-project.org/web/packages/rgl/vignettes/rgl.html`.

Museum of Modern Art. A sort of joy (thousands of exhausted things), 2015. URL `https://www.moma.org/calendar/events/1005`.

T. Nakamoto. *Essentials of Machine Olfaction and Taste*. Wiley, Hoboken, 2016.

S. D. Novich and D. M. Eagleman. Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput. *Experimental Brain Research*, 233(10):2777–2788, 2015.

A. Nowogrodzki. Speaking in code: How to program by voice. *Nature*, 559(7712):141–142, 2018.

J. J. O'Connor and E. F. Robertson. Mathematics of the Incas. MacTutor, 2001. URL `https://mathshistory.st-andrews.ac.uk/HistTopics/Inca_mathematics/`.

S. Papetti and M. Fröhlich. The TouchBox: Open-source haptic device for finger-based interaction. In *International Workshop on Haptic and Audio Interaction Design - HAID2019*, Lille, France, 2019. URL `https://hal.archives-ouvertes.fr/hal-02045943`.

C. Pearl. Using voice interfaces to make products more inclusive. Harvard Business Review, 16 May, 2019. URL `https://hbr.org/2019/05/using-voice-interfaces-to-make-products-more-inclusive`.

L. J. Perovich, S. Wylie, and R. Bongiovanni. Pokémon Go, pH, and projectors: Applying transformation design and participatory action research to an environmental justice collaboration in Chelsea, MA. *Cogent Arts & Humanities*, 5(1):1483874, 2018. URL `https://doi.org/10.1080/23311983.2018.1483874`.

B. Phelps. Odorama - John Waters' Polyester, 2018. URL `https://interactivemediaarchive.wordpress.com/odorama-john-waters-polyester/`.

Precision Content. On rocket ships and horses' asses: Where did your content standards come from? Futurism, 20 July, 2016. URL `https://www.precisioncontent.com/blog/on-rocket-ships-and-horses-behinds-where-did-your-content-standards-come-from/`.

P. Reipschläger, T. Flemisch, and R. Dachselt. Personal augmented reality for information visualization on large interactive displays. 2020. URL `https://virtual.ieeevis.org/paper_f-info-1243.html`.

H. Rosling. Gapminder: A fact-based worldview, 2015. URL `http://www.gapminder.org/`.

T. Rudd. Using Python to code by voice. YouTube video, 2013. URL `https://www.youtube.com/watch?v=8SkdfdXWYaI`.

P. M. Sadler. Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, 35(3):265–296, 1998.

Scentevents. Scent of Mystery with Smell-O-Vision powered by Scentevents, 2015. URL `https://www.scentevents.com/`.

M. H. Schneps, P. M. Sadler, S. Woll, and L. Crouse. A private universe. Astronomical Society of the Pacific, 1989. URL `https://vimeo.com/113349804`.

M. H. Schneps, J. Ruel, G. Sonnert, M. Dussault, M. Griffin, and P. M. Sadler. Conceptualizing astronomical scale: Virtual simulations on handheld tablet computers reverse misconceptions. *Computers & Education*, 70:269–280, 2014. URL `https://www.sciencedirect.com/science/article/pii/S0360131513002534`.

T. A. Sebeok. *Contributions to the Doctrine of Signs*. Peter de Ridder Press, Lisse, Netherlands, 1976.

J. Seo. Tactiler, 2020. URL `https://github.com/jooyoungseo/tactileR`.

Serenade. Code with voice, 2021. URL `https://serenade.ai/`.

V. Setlur, S. E. Battersby, M. Tory, R. Gossweiler, and A. X. Chang. Eviza: A natural language interface for visual analysis. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, pages 365–377, New York, 2016. Association for Computing Machinery. URL `https://doi.org/10.1145/2984511.2984588`.

S. Siegert and R. Williams. Sonify: Data sonification - turning data into sound, 2017. URL `https://cran.r-project.org/web/packages/sonify/index.html`.

J. Siliezar. Music of the spheres, 2021. URL `https://news.harvard.edu/gazette/story/2021/01/harvard-scientist-turns-space-images-into-music/`.

M. W. Sjoding, R. P. Dickson, T. J. Iwashyna, S. E. Gay, and T. S. Valley. Racial bias in pulse oximetry measurement. *New England Journal of Medicine*, 383(25):2477–2478, 2020.

S. D. Speeth. Seismometer sounds. *Journal of the Acoustical Society of America*, 33:909–916, 1961.

A. Srinivasan, B. Lee, and J. T. Stasko. Interweaving multimodal interaction with flexible unit visualizations for data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 2020.

J. R. Stilgoe. *Outside Lies Magic: Regaining History and Awareness in Everyday Places*. Walker, New York, 1998.

R. L. Storms and M. J. Zyda. Interactions in perceived quality of auditory-visual displays. *Presence*, 9(6):557–580, 2000.

S. G. Sturdevant, S. S. Huang, R. Platt, and K. Kleinman. Matching in cluster randomized trials using the Goldilocks approach. *Contemporary Clinical Trials Communications*, 22:100746, 2021. URL `https://www.sciencedirect.com/science/article/pii/S245186542100048X`.

A. Sutton. Boston coastline: Future past, 2016. URL `https://www.andisutton.net/boston-coastlines-future-past`.

System Sounds. SYSTEM Sounds. `https://www.system-sounds.com/`, 2017.

Talon. Powerful hands-free input, 2021. URL `https://talonvoice.com/`.

S. Teller, J. Glass, C. Livermore, R. Miller, A. Oliva, N. Roy, and A. Torralba. MIT Fifth Sense Project: Providing the functions of sight to blind people, 2013. URL `http://projects.csail.mit.edu/bocelli/index.html`.

R. Trotta. Dark matter: A multi-sensory experience. YouTube video, 2018. URL `https://www.youtube.com/watch?v=zKRsjGqz5Ls`.

E. R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Conn., 1983.

E. R. Tufte. *Beautiful Evidence*. Graphics Press, Cheshire, Conn., 2006.

Twotone. Welcome to TwoTone, 2021. URL `https://twotone.io/`.

M. Tønnessen, R. Magnus, and C. Brentari. The biosemiotic glossary project: Umwelt. *Biosemiotics*, 9(1):129–149, 2016.

Ultraleap. Digital worlds that seem human, 2021. URL `https://www.ultraleap.com/`.

A. Unwin, M. Theus, and H. Hofmann. *Graphics of Large Datasets: Visualizing a Million.* Springer-Verlag, 2006.

B. van der Kolk and B. Fry. Healing trauma & how the body keeps the score. How To Academy Mindset, 2021. URL `https://www.youtube.com/watch?v=d_YApSkqsxM`.

B. A. van der Kolk. *The Body Keeps the Score: Brain, Mind, and Body in the Healing of Trauma.* Viking, 2014.

S. Vanderplas, E. Robinson, and R. Howard. Pandemics, graphics, and perception of log scales. R-Ladies DC, 2021. URL `http://srvanderplas.github.io/Presentations/2021-NUMBATS/#4h`.

A.-M. Vassou. Motion sensor technology could help stroke victims, 2008. URL `https://web.archive.org/web/20090528043355/http://www.computeractive.co.uk/computeractive/news/2226200/motion-sensor-technology-help`.

Viewplus, 2016. URL `https://viewplus.com/tag/tactile-graphics/`.

A. Volmar. Listening to the Cold War: The nuclear test ban negotiations, seismology, and psychoacoustics, 1958–1963. *Osiris*, 28(1):80–102, 2013.

H. Wickham. *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag, 2016. URL `https://ggplot2.tidyverse.org`.

L. Wilkinson. *The Grammar of Graphics, second edition.* Springer-Verlag, 2005.

L. Wilkinson. nViZn. `https://www.cs.uic.edu/~wilkinson/nViZn/nvizn.html`, 2010. Accessed: 2022-03-24.

Y. Yang, T. Dwyer, K. Marriott, B. Jenny, and S. Goodwin. Tilt map: Interactive transitions between choropleth map, prism map and bar chart in immersive environments. *IEEE Transactions on Visualization & Computer Graphics*, 2021.

B. Yu and C. T. Silva. Flowsense: A natural language interface for visual data exploration within a dataflow system. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1–11, 2020.