Augmented Chironomia for Presenting Data to Remote Audiences

Brian D. Hall
University of Michigan
Ann Arbor, Michigan, USA
briandh@umich.edu

Lyn Bartram
Simon Fraser University
Surrey, British Columbia, Canada
lyn@sfu.ca

Matthew Brehmer
Tableau Research
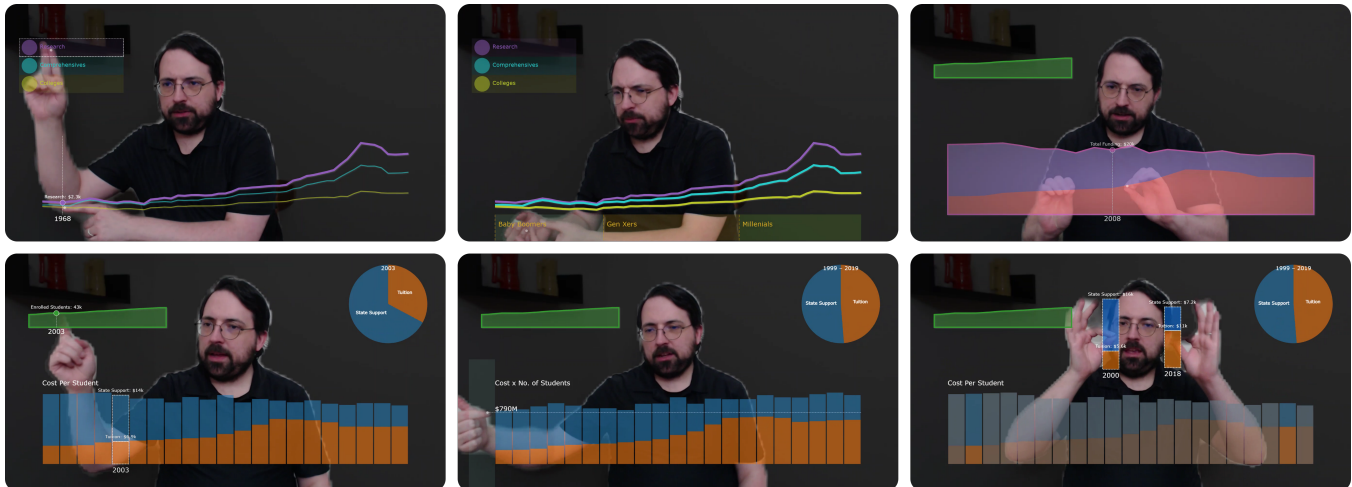Seattle, Washington, USA
mbrehmer@tableau.com

Figure 1: Frames from a video presentation in which a presenter manipulates and highlights elements in chart overlays.

ABSTRACT

To facilitate engaging and nuanced conversations around data, we contribute a touchless approach to interacting directly with visualization in remote presentations. We combine dynamic charts overlaid on a presenter's webcam feed with continuous bimanual hand tracking, demonstrating interactions that highlight and manipulate chart elements appearing in the foreground. These interactions are simultaneously functional and deictic, and some allow for the addition of "rhetorical flourish", or expressive movement used when speaking about quantities, categories, and time intervals. We evaluated our approach in two studies with professionals who routinely deliver and attend presentations about data. The first study considered the presenter perspective, where 12 participants delivered presentations to a remote audience using a presentation environment incorporating our approach. The second study considered the audience experience of 17 participants who attended presentations supported by our environment. Finally, we reflect on observations from these studies and discuss related implications for engaging remote audiences in conversations about data.

CCS CONCEPTS

• **Human-centered computing** → **Visualization**; *Gestural input*; *Mixed / augmented reality*.

KEYWORDS

Visualization, augmented reality, pointing, gesture, rhetoric, video

1 INTRODUCTION

Chironomia is the ancient art of manual rhetoric [4]: it refers to expressive hand movements that a speaker can employ when communicating to an audience. In this paper, we consider the combination of this art with the rhetorical power of data visualization [41]. We demonstrate an augmentation of webcam video with interactive visualization overlays in a speaker's foreground (Figure 1). We further augment a speaker's hand movement, granting them functional control of the overlays without precluding the potential for illustrative and affective body language. Together, these augmentations allow for nuanced and engaging conversations about data with remote audiences.

Our work is primarily motivated by the ever-growing needs of people within organizations who lead discussions and inform decisions that are grounded in data [13, 25]. These activities manifest in synchronous meetings and presentations among and between

colleagues, stakeholders, executives, and customers. In these settings, participants support their discussions and arguments with data, using charts and graphs to express quantities, proportions, categories, and time intervals. These visual aids are typically presented as a series of static slides featuring a small palette of familiar chart types [13]. Typically, people paste chart screenshots on to slides and add annotations or shape masks that draw attention to specific values or visual patterns; these attentional cues can only be revealed or removed in a pre-planned sequence during presentation delivery. If the underlying screenshots need to be updated or replaced, the attentional cues must also be manually updated to match the new content. Alternatively, a presenter could prepare multiple screenshots of a chart in various states, so as to reveal, emphasize, and hide specific elements, but preparing, organizing, and staging these assets is tedious. Part of the problem we address in this work is the inability to interact with such content in unplanned ways, such as in response to audience questions.

The other part of the problem is the increased prevalence of remote and hybrid work environments since the onset of the COVID-19 pandemic. Given this shift, presentations are now predominantly supported by the screen-sharing functionality of teleconference applications. This impoverished presentation medium reduces the capacity for presenters to draw their audience's attention to specific data points and fully illustrate their interpretations of the data. Absent the affective and illustrative cues that a presenter can provide when co-located with an audience, they cease to become 'sage on the stage'; instead, they are relegated to a secondary thumbnail video frame or just a disembodied voice. As a result, audiences become disengaged and opportunities for discussion are lost.

The approach we take in this paper was inspired in part by compelling public-facing presentations of data that take advantage of multiple communication modalities: the oration and body language of an engaging presenter coupled with dynamic data visualization. In particular, we draw inspiration from presentations delivered by the late public health expert Hans Rosling (e.g., [63–66]). In his 2010 BBC Documentary *200 Countries, 200 Years, 4 Minutes* [65], an animated scatterplot was composited over a recorded video of him speaking and gesticulating, giving the impression that he was controlling the chart with his body. In this paper, we demonstrate that it is now possible to deliver such presentations to a synchronous audience by compositing interactive charts over conventional webcam video. This approach allows a more attentive audience to interrupt the presenter with questions, prompting unplanned interactions with the data.

Contributions. We contribute a functional realization of a presentation environment in which live webcam video is overlaid with interactive visualization controlled by a presenter's hand movements. We also contribute demonstrations of several categories of touchless interaction that are simultaneously deictic and functional, without foreclosing the potential for affective gesture and rhetorical flourish. Finally, we contribute reflections and observations from two formative studies of our approach with 29 professionals who routinely attend and deliver presentations about data, considering both presenter and audience perspectives.

2 BACKGROUND AND RELATED WORK

We build upon several bodies of knowledge: classifications of hand movement in multimodal communication, gestural and spatial interaction, and the communicative use of data visualization.

2.1 Talking with the Hands

The non-verbal communication modalities of bodily movement, posture, gaze, and facial expression are interwoven with speech in social interaction [76, 79]. The communicative use of body language has been documented well into antiquity [4], with hand gesture in particular being viewed as integral to interpersonal communication [51]. According to a co-speech perspective [26, 82], gestures amplify and extend the production and perception of meaning, and they contribute to discourse coherence when speakers supplement and extend the interpretation of prior gestures and speech [46].

An embodied cognition perspective [49] posits that hand gestures are critical for comprehending and reasoning about complex content, such as when teaching mathematical concepts [2, 7, 30] or when establishing a shared understanding in engineering and design [16, 75]. Complex content also arises in persuasive presentations given by entrepreneurs and investors, and presenters who use their hands to depict and symbolize business ideas tend to have more positive entrepreneurial outcomes relative to those who present via the modalities of speech and text alone [20, 23]. There is consistency in hand gesture when speaking about quantities in particular; a study of television news archives [85] showed that people tend to compress or expand their fingers or hands when speaking about small and large values, respectively, or to move their hands laterally, vertically, or distally to indicate a progression of values. Given that many types of charts shown in presentations depict not only quantities but also time intervals and categories, our work seeks to better understand the role of expressive hand movements performed in the presence of these visual aids.

In considering prior categorizations of communicative hand movement [49, 76], we see two recurring categories as being particularly relevant to scenarios where people speak about data:

Illustrative gestures supplement what is being said, with three subcategories of interest: ***deictic*** gestures draw attention to artifacts visible to both speaker and audience through pointing, enumerating, and framing; ***iconic*** gestures illustrate distances, sizes, and shapes; and ***metaphoric*** gestures communicate abstractions [20], such as moving the hand clockwise to signify the passage of time.

Affective gestures convey emotion and emphasis; these *performative* gestures include the tightening of a fist or throwing one's hands up in the air to convey uncertainty or dismay. Consider how a *beat* of the finger, hand, or arm can mark an important point in a speech, while a repeated beat can variably convey urgency, importance, or steadiness, depending on the tempo. From the ancient Greeks to modern public speaking coaches, many have considered the richness of this category; Cicero used the term *chironomia* [4] to describe gesticulation in the service of rhetorical delivery, argumentation, and persuasion [73]. It is therefore no surprise that political speeches employing affective hand gestures are perceived as being more compelling, persuasive, and engaging [15].

Given the persuasive [58] and rhetorical [41, 45] potential of data visualization, our research considers scenarios where an audience

can simultaneously see visual representations of data with an orator performing illustrative and affective hand gestures, movements that may also take on a *functional* role, which we discuss next.

2.2 Gestural Interfaces

HCI research spanning multiple decades has focused extensively on the technical and *functional* aspects of gestural interaction, such as how to control an interface using gesture (e.g., see Bolt (1980) [8]). While touch-based gestures common to mobile computing are now ubiquitous [19] (e.g., pinch, swipe, tap and hold), the lexicon of touchless or mid-air interaction is still evolving [1], particularly since the introduction of commercially-available visual sensing interfaces such as the Microsoft Kinect and the Leap Motion Controller. Despite the growing body of research dedicated to this interaction modality, it largely reflects a focus on recognition and usability [57, 81]. This research highlights issues of how common gestural techniques can both enhance and interfere with the contexts in which they are used. For example, O'Hara et al. [57] demonstrated a touchless interface for surgeons to control the display of medical imagery while performing surgery; while surgery is a collaborative context, the functional gestures considered were intended primarily for system control rather than for communicating with other members of the surgical team. This example underscores the need for a unified framework that encompasses both the *functional* and *communicative* power of touchless interaction with the hands.

Gesture-controlled presentations. According to Harrison [37], speakers giving presentations supported by visual aids such as slides produce complex multimodal 'ensembles' comprised of speech, body language, and images, where one element in an ensemble may emphasize, reinforce, or restate the meanings of other elements. When presenting visuals to a co-located audience, Fourney et al. [31] found that people employ various deictic gestures with either hand to variably emphasize different content groupings, from 'everything' to specific visual elements.

We are aware of several precedents for presentation systems incorporating touchless interaction with the hands [6, 24, 31, 43, 48, 50, 69], including sensor-based [6] and vision-based systems [31, 36, 43, 48, 50]. Collectively, these projects document appropriate algorithms to effectively capture a set of gestures that control a progression of visual aids [6, 36, 48] as well as considerations of interaction learnability and memorability [24]. While those who used these systems generally found their interactive experiences to be novel and engaging, the scope of interaction was limiting [6, 24, 31], in that system designers prioritized slide navigation over interacting with slide content [31]. Additional issues included the capture of unintentional movements [24, 31] and an increased cognitive load when the repertoire of interactions was large and unconstrained [69].

We specifically call attention to the recent work by Saquib et al. [69], who demonstrated a Kinect-based system for authoring creative video storytelling performances. Using their system, a performer could manage video navigation as well as point at and modify the appearance of animated graphical elements in the foreground via mid-air gesture. However, the interactions and associated behavior of linked visual elements had to be defined when authoring; while those who used the system explored a variety of interactions when authoring, they executed fewer interactions at performance time, and these interactions had to be performed in a predictable way.

In contrast to previous systems, our approach incorporates a standard webcam for tracking hand movements, and we focus on the small yet ubiquitous palette of chart types that are common across presentations of data [13]. Moreover, we demonstrate categories of interaction that can be performed flexibly across chart types during a presentation, and we prioritize the presentation delivery and audience experiences over the authoring experience.

Gestural interaction with data. Prior research has also considered the potential of touchless interaction for exploratory data analysis. For instance, in virtual reality environments, systems like ImAxes [22] allow people to move and coordinate three-dimensional chart components within a virtual environment. Elsewhere, we have seen mobile augmented reality applications allow for interaction with virtual objects with the dominant hand while the non-dominant hand holds the mobile device [42, 60]. Altogether, this body of work elicits the question of what happens when an audience is watching someone interact with data in these ways and the extent to which hand movement is interpreted as being communicative.

2.3 The Communicative Use of Visualization

Apart from data analysis, people also visualize data to communicate with others: it can help them tell stories about data [40], support an argument [45], or persuade an audience [58] to make decisions that are grounded in data [25]. Narrative visualization [72] can assume a variety of forms and rhetorical structures [41], though much of the prior research in this area has focused on forms typically associated with web-based journalism, where dynamic visualization manifests in magazine-style articles [21, 53], reader-controlled slide presentations [10, 70], or recorded data videos [3]. Each of these forms entail asynchronous consumption by individual viewers, and while they may be able to interact with the content, there is little capacity for them to interact with content authors or other viewers [52]. In contrast, we focus on synchronous multimodal communication involving visualization delivered by a presenter.

Meanwhile, the prevalence of visualization in live television broadcasts has increased in recent years, a medium that has been largely ignored by the research community [28]. From weather reporting to coverage of the COVID-19 pandemic or the 2020 US Federal Election, correspondents have made use of large high-resolution touchscreen displays or those controlled by a handheld tablet. We have also seen more elaborate presentations that place correspondents in an augmented reality environment produced with specialized cameras and studios (e.g., Vizrt [80]). While these are live presentations about data, we focus on scenarios where a direct engagement with the audience is possible, and we assume no specialized equipment aside from a standard webcam.

Conversations about data with a live, yet remote audience. We concentrate on teleconference-based presentation settings where the speaker and audience are able to interact in real-time. While prior research has demonstrated purpose-built tools for using dynamic data visualization in presentations for a co-located audience (such as SketchStory [47] or SandDance [27]), there is a dearth of analogous purpose-built tools for live presentations about data to remote audiences [86]. This absence is sorely felt in informal

presentations taking place within (distributed) organizations [13], where interruptions and impromptu discussion about the specifics of the data are commonplace. Currently, presenters resort to screen sharing slides or visual analysis tools, such as analytical notebooks or business intelligence dashboards. Slides impart an inappropriate level of formality and linearity, precluding spontaneous shifts of audience attention, while visual analysis tools introduce visually distracting interfaces, reflecting their intended use in individual interactive analysis. Our work is a response to this gap in the tool landscape for presenting data.

3 PROCESS AND IMPLEMENTATION

Our goal was to develop a better experience for presenters and remote audiences alike, encouraging a level of audience engagement that is lost in remote communication, one that fosters nuanced conversations about data. This goal shaped our first design imperative: to support a presenter's direct interaction with the *content* of the presentation and to look beyond purely functional view navigation and passive highlighting. Our second design imperative was to demonstrate a presentation environment that would incur little to no cost or special equipment while enabling presenters' existing communication skills, namely their fluency with respect to performing illustrative and affective gestures.

Iterative design. We iteratively developed our approach to presenting data. Our first proof-of-concept [11] made use of greenscreen-compatible presentation tools (OBS [56] and mmhmm [54]) along with a pose recognition model that we created using Teachable Machine [33] to trigger transformations to a chart composited behind a presenter. However, from an audience perspective, the poses seemed stilted, and echoing findings from Fourney et al. [31], it was undesirable for the presenter to occlude the content. From a presenter perspective, a choreographed performance of poses was fatiguing. Moreover, a set of discrete poses was insufficient for drawing attention to data; for instance, we could not use them to select or highlight specific chart elements.

Our second proof-of-concept was inspired by a news anchor sitting at a desk, with 'over-the-shoulder' graphics composited on top of their video. However, rather than composite charts with opaque backgrounds over webcam video, we lowered the opacity of chart elements and removed their backgrounds, so as to allow a presenter to point to chart elements from behind, reminiscent of the Lucid Touch interface [83]. In addition to semi-transparent over-the-shoulder charts, we incorporated continuous hand and finger tracking through the use of MediaPipe [34]. However, like our earlier greenscreen prototype, interacting repeatedly with over-the-shoulder charts was physically awkward and fatiguing.

An integrated browser-based presentation environment. We retained the continuous hand tracking from our second prototype, integrating it into a browser-based presentation environment developed as a Svelte [78] application. We rendered the overlay charts rendered as SVG elements, and we made use of D3.js [9] scale transformations for both element placement and interaction handling.

Following an observation that most presentations about data taking place within organizations incorporate a handful of simple chart types [13], our presentation environment is currently supports variants of bar, line, area, and pie charts. However, our environment can be easily extended to accommodate any SVG-based chart.

Realizing the limitations of over-the-shoulder charts and taking inspiration from Rosling's 2010 documentary [65], our environment allows for chart overlays to fill the entire video frame. This flexibility gives presenters the freedom to sit stationary or to stand and walk around the frame; it also allows them to keep their hands in more comfortable positions when speaking. Additionally, as presenters' surroundings and lighting arrangements will vary, we optionally apply background segmentation to darken the presenter's surroundings (Figure 1) as well as a grayscale filter to the presenter's video, so as to place additional emphasis on the overlays.

Given the stateless nature of the MediaPipe API [34], our implementation adds a time-based event protocol for gesture detection with dwell and timeout durations ranging between 0.1 to 1.0s; these settings reduce jitter as well as false positive and false negative classifications. For graphical performance, we avoid overprocessing the video stream, thereby reducing presenter-side latency to imperceptible levels. Specifically, we only send an image frame to MediaPipe for landmark detection if a previous frame is not being processed, and we perform gesture detection only once per result received from MediaPipe.

Prior to giving a presentation with this environment, a presenter needs to specify *scenes*, where each scene can contain one or more chart overlays connected to local data sources. For each overlay, the presenter can enable its ability to respond to interaction and specify its visibility, dimensions, and positions within the scene.

We reiterate that our focus in this paper is not the presentation authoring experience; while the current scene and overlay specification is JSON-based, future authoring experiences could emulate drag-and-drop dashboard creation in business intelligence tools.

What you see is what I see. Initially, we had intended for presenters to perform both communicative hand movements as well as those dedicated to scene management, such as adding, removing, positioning, and resizing overlays. However, scene management would require either a set of unique functional gestures or an on-screen mode-switching widget. Both approaches would be distracting for audiences, particularly as neither serve a direct communicative purpose. Furthermore, as with interacting with over-the-shoulder overlays, scene management gestures can be tiring, echoing reports of actors becoming easily fatigued while filming similar touchless interactions in science fiction films [55]. Alternatively, we considered a secondary presenter view, such as a control panel not visible to the audience, but this would run the risk of splitting the presenter's attention, requiring them to coordinate interaction across two displays. As a result, the presenter and audience both see the same content, and we relegated scene navigation in our environment to off-screen keyboard shortcuts that trigger animated transitions. Chart overlays can enter or exit through translation or fade transitions, or if two consecutive scenes contain the same overlay, it can smoothly translate and scale if required. Combining a *'what you see is what I see'* approach [74] with keyboard-based scene navigation allowed us to focus on a presenter's communicative movement, which we discuss in the next section.
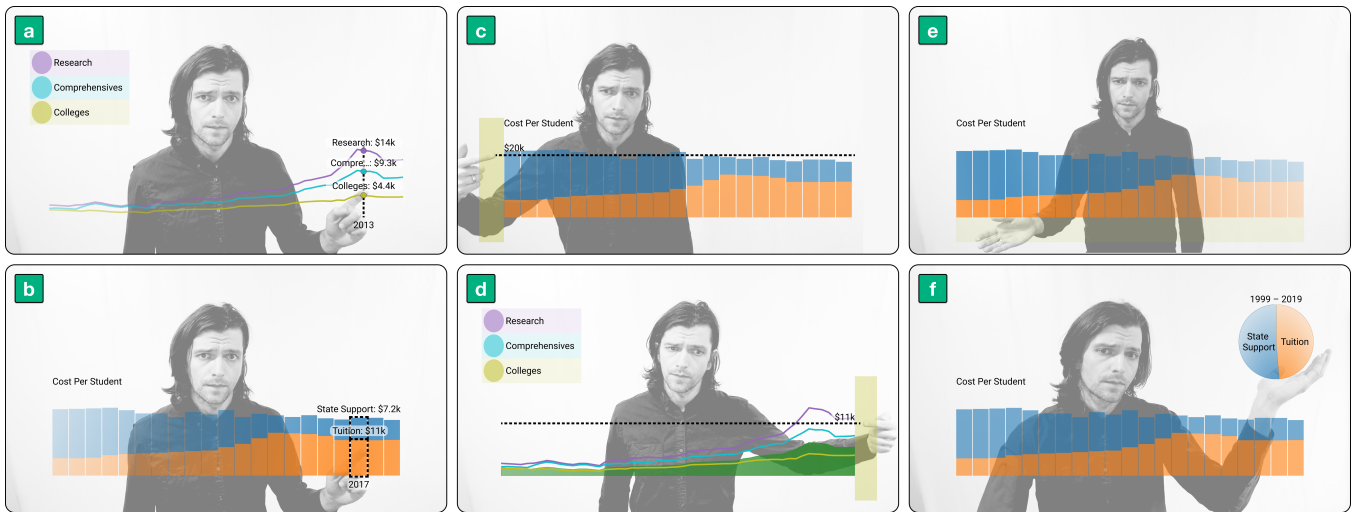
Figure 2: Three interactions applicable across chart types: (a—b) pointing with the index finger highlights the nearest value(s); (c—d) pointing with the index finger in either side margin triggers a reference line and value annotation; (e—f) an opacity gradient radiates from the palm's centroid to coarsely emphasize content near the hand.
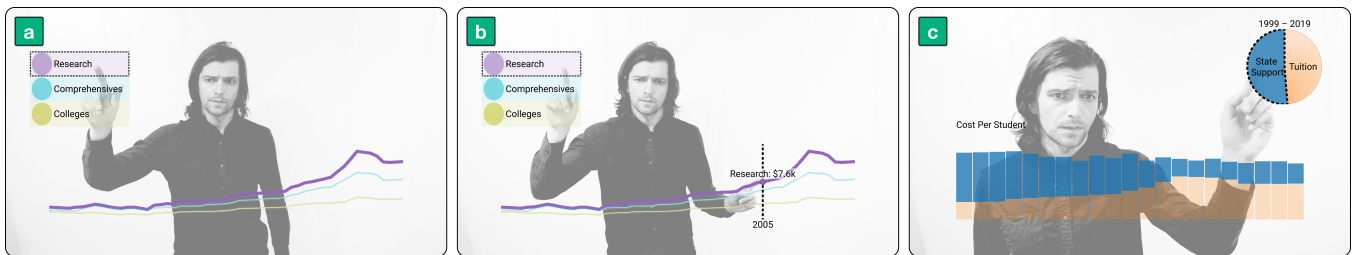

Figure 3: Highlighting across overlays: (a) pointing to a legend swatch emphasizes the corresponding series in an adjacent line chart; (b) if pointing at a legend swatch, only values corresponding to that series will appear as annotations (compare to Figure 2a); (c) pointing to a wedge in a pie chart emphasizes marks having that category in an adjacent bar chart.

Figure 4: Coordinating and manipulating overlays: linked highlighting across charts (a) can be superseded with (b) bimanual pointing; (c) pinching and dragging a cloned copy of the green student enrollment area chart; (d) releasing the cloned area chart over the cost-per-student bar chart (which shares the same temporal domain) multiplies corresponding values, yielding total cost values over time; (e—f) pinching generates ephemeral and draggable cloned elements for spatially-adjacent comparisons.

4 INTERACTING WITH CHART OVERLAYS

We now describe the categories of chart overlay interactions supported by our environment. In addition to their depictions in Figures 2 through 6, the supplemental video includes demonstrations of interactions from each category. Overall, these touchless interactions are simultaneously deictic and functional, designed with the goal of avoiding conflict with affective movement. Most are agnostic to chart type, while some are unique to the type of data being shown or the geometry of a particular visual encoding. While interactions can be performed with a single chart overlay, we also show their effects with multiple adjacent or layered overlays. Many allow for the addition of *rhetorical flourish*, or expressive movement that reinforces what is being said to advance a narrative about data.

Our environment tracks both of the presenters' hands, specifically the tip of the index finger, the tip of the thumb, and the palm centroid of both hands. The environment can optionally highlight index finger and thumb control points, and these markers will change color when the presenter is pinching. It can also distinguish between right and left hands, and the two hands can perform different interactions concurrently. When both hands are within the same chart overlay, the right hand is designated as dominant and given precedence for single-handed interactions; however, it is possible to configure the environment for left-handed individuals.

4.1 Deictic Highlighting

Figure 2 demonstrates three categories of deictic highlighting applicable across chart types, drawing inspiration from previous tools for emphasizing chart elements such as SmartCues [77] and ChartAccent [62]. These interactions are continuous and ephemeral, updating based on the position of the fingertip or palm, with highlights disappearing once the hand leaves the bounds of an overlay.

Pointing to reveal values. Figure 2a—b illustrates overt highlighting via pointing. This form of highlighting directs attention to specific values in a chart, so we paired this interaction with tooltip-like value annotations. For rectilinear charts with continuous scales along the horizontal axis (such as line and bar charts), we map the position of the index finger to the nearest bisected value in the underlying data; we display those values as text labels along with a vertical reference line and an emboldened mark stroke. For circular charts (such as a pie chart), we use a trigonometric function to determine the segment nearest the index finger (Figure 3c).

Pointing in the margin. Figure 2c—d shows the highlighting of an axis value spanning the entire chart. Whenever an index fingertip is found within the side margins of a rectilinear chart, we draw a horizontal reference line emanating from the fingertip. This form of highlighting is both deictic and iconic, in that the raising of a hand tends to signify a larger quantity.

Illuminating from the palm. Figure 2e—f shows a more subtle form of highlighting, where a radial opacity gradient emanates from the centroid of the palm. This form of highlighting is intended to draw the audience's attention to a general region within a chart. This gradient can be activated whenever the presenter's palm appears within a chart or its margins; it is not activated when the palm is kept outside the margin (2c—d). A presenter can therefore selectively combine this coarse highlighting with finer highlighting

by intentionally placing the palm relative to the index finger. If the palms of both hands are visible within a single chart, the gradient will emanate from the dominant hand.

4.2 Interacting with Multiple Overlays

The display of multiple adjacent or superimposed overlays offers an opportunity to coordinate interaction between them. While there are many chart coordination design patterns to consider [17], we predominantly focus on linked highlighting and selection.

Highlighting categories. Figure 3 illustrates linked highlighting triggered by pointing at a categorical legend overlay (3a), whereupon the corresponding series in the adjacent line chart is emphasized and other series are de-emphasized. At this point, the right hand can point at features in the line chart (3b) while continuing to point at a particular legend swatch with the left hand; doing so will suppress the value annotations of other categories (compare to Figure 2a). Finally, linked highlighting need not be driven from a dedicated categorical legend: Figure 3c illustrates how pointing at a category in a pie chart emphasizes the corresponding category in an adjacent bar chart.

Highlighting values along a common domain. Linked highlighting can also be powerful when multiple chart components exhibit a common value domain [61]. For instance, the green area chart in Figure 4a—b shares a temporal domain with the stacked bar and pie charts (4a) and the orange area chart (4b). The size and peripheral location of the green area chart and pie chart suggest that they are of secondary importance, providing context to the more visually prominent chart below it. Accordingly, highlighting a point in time in the stacked bar chart can trigger corresponding highlights in the peripheral charts. However, we allow presenters to override this linked highlighting (4b) by simultaneously pointing at different positions along the domains of two charts.

4.3 Pinch and Drag to Transform and Compare

Figure 4c—f illustrates the selecting and repositioning of individual chart elements. While the direct manipulation and dragging of individual chart elements may evoke the activities of chart construction [71] and exploratory data analysis [68], this category of interaction can also serve a communicative purpose.

Illustrating a value transformation. Consider how to explain a uniform transformation applied to a set of values, such as multiplying the GDP per capita of countries by their populations to derive total GDP values. While this should be explained in a presenter's oration, this explanation could also be reinforced by interaction. We illustrate such a transformation in Figure 4c—d, in which the presenter pinches to select the green area chart corresponding to student enrollment over time, generating a temporary copy (4c). In this example, the bar chart of individual student costs spans the same temporal domain, and in dragging and releasing the copied enrollment values over the bar chart, the copy is destroyed and the bar chart updates to reflect total cost values over time (4d).

Comparisons on demand. Another typical scenario is comparing values that are not spatially adjacent. While such comparisons could be planned in advance, we expect unplanned comparisons, such as those prompted by audience questions [13]. Although these
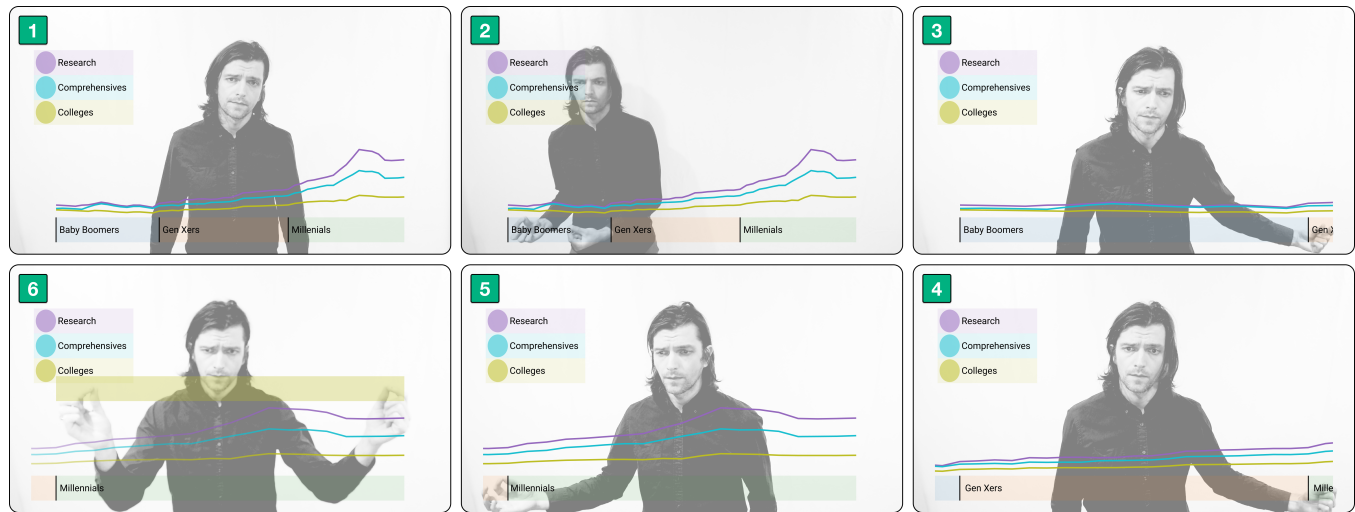
Figure 5: Manipulating the temporal scale of a chart where time appears on the X axis (clockwise starting from #1): bimanual pinching in the bottom margin to zoom in on the timeline (2), pinching in a bottom corner margin to pan the timeline in either direction (3—5), and bimanual pinching in the top margin to zoom out on the timeline (6). The associated flourishes include *expanding* when zooming in, *flicking* or *swiping* when panning, and *compressing* when zooming out.

Figure 6: Pinching reveals a second band in the stacked area chart (a, b), where a possible flourish includes *snapping* one's fingers; bimanual pinching ephemerally aggregates bands in the stacked area chart into a single band until the both pinches release, as shown (c), a presenter can *crimp* the boundary between the two bands as a flourish.

unplanned comparisons could be performed by executing a filter command, this may result in a loss of context for the audience. We demonstrate an alternative way to compare elements in Figure 4e—f, where each hand can pinch within the bounds of an element to create a cloned copy, one that is tethered to the position of the fingertips. As long as the pinches are maintained, these cloned elements can be freely repositioned (4e), so as to afford a side-by-side comparison of values. Meanwhile, the original chart elements are de-emphasized apart from a stroke to indicate the source of the cloned elements. Releasing the pinch destroys the cloned copies and restores the source chart to a normal opacity level (4f).

4.4 Supporting Rhetorical Flourish

Some of the most prominent visual transformations that a presenter can apply to a chart include manipulations of scale and the addition or removal of chart content. While animation can be a powerful cue to draw viewers' attention to these transformations [39], a presenter's body language can provide complementary illustrative and affective cues. However, the salience and deliberateness of a presenter's movement should be commensurate with the extent of visual transformation taking place, and with the transformation's relative

importance with respect to the overall narrative. Accordingly, several of the interactions supported by our presentation environment are quite deliberate relative to deictic highlighting. Moreover, we give presenters the freedom to adjust the visual salience of these interactions; we describe this as the ability to add a *rhetorical flourish* when performing the interactions.

Express changes of temporal scale. We demonstrate several opportunities for flourish when manipulating the temporal scale of a chart in Figure 5, beginning with 5.1 and proceeding clockwise. This particular line chart of higher education tuition spans more than 50 years, with regions along the horizontal margin indicating when three generations attended college. To zoom in on the Baby Boomer generation, for instance, the presenter pinches the corresponding span along the horizontal margin with both hands (5.2), which is a fairly deliberate and salient gesture. Optionally, the presenter can pull their hands apart after pinching; this *expansion* flourish does not affect the recognition of the interaction, but it visually reinforces that a zooming in is occurring. The salience of this flourish is left to the discretion of the presenter: they could perform a small lateral expansion, or a more emphatic expansion heralding a more important reveal of information. After zooming in, pinching the

corner of the margin with one hand will trigger a panning of the timeline (5.3—5.5). We allow for a degree of flourish here too, in that the presenter can *flick* or *swipe* their hand in the direction the pan, once again complementing the visual transformation taking place. Lastly, bimanual pinching in the top margin or with a hand in both side margins will zoom out to restore the full timeline (5.6). Mirroring the flourish for zooming in, pairing these pinches with a *compression* gesture can draw additional attention to the change in temporal scale.

Revealing new chart elements. Progressively unveiling highly salient chart elements offers another opportunity for flourish. For instance, pinching within the bounds of an area chart can reveal an initially hidden second band (Figure 6a—b). Pinching within the contours of the hidden band will draw more attention to the band as it appears, and pairing this pinch with an optional and audible finger *snap* may draw even more attention. Highly salient visual changes may also be ephemeral: for instance, bimanual pinching within the stacked area chart reveals a total value band superimposed over the individual bands (6c), and this band remains in place as long as one hand is pinching within the bounds of the chart, leaving the other hand free to point out specific total values. A possible flourish here is one in which the presenter assumes a bimanual *crimping* pose along the boundary between the bands in the area chart, simultaneously satisfying the requirements of a bimanual pinch while visually reinforcing the aggregation of values.

5 EVALUATION

We evaluated our approach to presenting data in two independent studies: one considering the perspective of the presenter and another considering the perspective of the remote audience.

5.1 Procedure

Irrespective of whether a participant was participating in the presenter study or audience study, we began each session by asking them about their current experiences with respect to attending and delivering remote presentations about data. We then introduced the remote presentation scenario, which we described as an informal presentation among colleagues, one where the audience could interrupt the presenter and ask questions. The topic of the presentation pertained to higher education costs, for which we used data published by the Economic Opportunity Institute [29]. At this point, the procedure diverged depending on the study, which we describe below. We closed both session types with a reflection on the participant's experience, in which we asked them to extrapolate from our approach to consider other chart types and configurations, as well as other possible interactions that might support a presentation about data. In both studies, sessions lasted between 45 and 60 minutes, which we recorded and transcribed.

Presenter-oriented study. The aim of the presenter-oriented study was to collect feedback on the utility, usability, and learnability of the presentation environment. To familiarize presenter participants with the charts and their underlying data prior to their sessions, we directed them to a Tableau Public workbook [12] containing a series of charts (featured in the supplemental video). To ensure a consistent experience for participants, we observed them as they interacted with our presentation environment in a

small meeting room in a corporate office setting (Figure 7). One researcher was present in the room with the participant, while another researcher assumed the role of a remote audience, with their video feed shown on one display. We displayed our presentation environment on another display and also shared it via a videoconferencing application. Finally, we pointed a Logitech 1080p webcam at the participant, positioning it to capture them from the waist upwards and to allow for either a seated or standing presentation delivery, depending on their preference. After toggling our environment's option to display fingertip control point markers, we then progressed through eight scenes of a presentation, with each scene featuring different combinations of chart overlays and associated interactions. For each scene, we invited participants to discover and practice the scene's associated interactions, and to think aloud as they interacted; we also provided them with a set of printed interaction reference sheets (adapted from Figures 2 — 6). When participants struggled to perform a particular interaction, we provided verbal hints. Once familiar with the interactions in each scene, the remote researcher posed a question about the data, so as to serve as a presentation prompt, such as *"can you describe how tuition costs changed for millennials during the 2010s?".*

Figure 7: In the presenter evaluation, a webcam (mounted on a tripod placed at the center of the table) captures a participant's hand motions while chart overlays on their video appear in the right monitor; a researcher assuming the role of the remote audience appears in the left monitor.

Audience-oriented study. The aim of the audience-oriented study was to better understand how presentations delivered using our environment might engage audiences. In each session, one researcher took on the role of a presenter, while another researcher assumed the role of a meeting host or moderator. To better understand how a remote audience would experience a presentation about data delivered using our environment, we prepared two 5-minute presentations adapted from the same content used in the presenter-oriented study; we include abbreviated versions of these presentations in the supplemental video. We additionally prepared alternate versions of these two presentations that were more representative of current presentation practices [13]; these retained the same speaking points but we delivered them by screen-sharing the Tableau Public workbook [12] used in the presenter-oriented study. Unlike a slide presentation, a workbook allowed us to present content in unplanned ways, such as in response to audience questions.

The presenter then delivered two of the four presentations; to mitigate a potential novelty effect, half of the audience participants

saw the alternate screen-share presentation followed by one delivered using our presentation environment, while the other half experienced the reverse order. Midway through each presentation, the moderator posed a clarifying question about the data in cases where the participant had not already interrupted or asked a question, so as to remind participants of the live presentation scenario. After each presentation, we asked participants to summarize the content of the presentation and whether they had any content-related questions for the presenter. As the interviews were open-ended, we found it useful with some participants to briefly demonstrate what the alternate presentation delivered via screen-share would have looked like had it been delivered using our presentation environment, so as to discuss specific points of contrast.

5.2 Participants

We recruited 29 professionals who attend, prepare, and deliver customer-facing presentations about data at a multinational software company, as well from its' subsidiary and customer organizations. Our participants varied in terms of years of experience (from several months to over two decades) and job title (e.g., solution engineer, product manager, account executive). While some participants were accustomed to presenting data to remote or mixed audiences prior to the COVID-19 pandemic, all of them have been presenting primarily to remote audiences since the pandemic's onset. The majority of our participants reported attending multiple presentations a week involving some discussion about data, and all of our participants reported giving such presentations themselves, with cadences ranging from quarterly to daily. For the presenter study (N=12), we limited our recruitment to those who were willing and able to meet with us in person at an office that required all employees to observe strict COVID-19 safety protocols. For the audience study (N=17), participants joined remotely from locations spanning North America, South America, and Europe.

5.3 Observations & Participant Reflections

We performed a thematic analysis of our observations and transcribed quotes from the session recordings. This process yielded several themes and tags, which we list in the supplemental material. Throughout this section, participant attributions are study-specific, with P# referring to a presenter study participant and A# referring to an audience study participant.

Reflecting on utility. When we asked presenter study participants to describe the potential value of adopting our approach for their own presentations. P3 hypothesized that this would *"probably lead to more discussion than a normal presentation,"* adding that audiences are *"definitely not gonna forget,"* suggesting that seeing the presenter interact with data in this way will elicit more audience engagement, which will in turn result in a more memorable presentation experience. As to what might drive this focused attention, P8 suggested that the format would force his audience *"to look at the data closer, just by nature of me putting the effort into doing all this…it would put an organic impetus on the audience…it's like a circus, I'm juggling things — you got to watch me, right?"* This focused audience attention could be particularly important when communicating causal relationships in data: P5 will start *"with an overview and break it down into 'here's what's causing that change'",*

and he saw our approach as being potentially *"really helpful to explain and understand data at that level"*.

A potential to engage. From the audience perspective, participants were able to directly contrast our approach with the convention of presenting data via screen-sharing. In describing our approach, A5 said: *"the storytelling is more engaging…it connected more."* The overlay presentation also suspended A3's tendency to critique a presenter's delivery, likening it to instances where an audience unfamiliar with interactive business intelligence dashboards might similarly suspend their critical tendencies if they were accustomed to seeing presentations delivered primarily via slideware. Also notable is A10's criticism of the screen-share presentation as lacking annotation and axis labelling, which we attribute to his tendency to associate business intelligence workbooks with data analysis, as opposed to synchronous presentation. Following the overlay presentation, he did not offer the same criticism despite a similar absence of labelling: *"I'm actually listening more, I'm paying attention more because I'm more engaged by your physical body, …I like it better now that it's clean and not have all those numbers that I was looking for before…in this version, I'm able to just listen to your voice and focus to the story…my heart is more patient with you."*

For some audience members, their engagement was palpable: *"I love this thing…that gets my heart going a little bit"* [A12]. A4 summarized the value of this potential to increase audience engagement: *"it's drawing attention because it's so engaging…I actually want to understand what's going on behind the data, and that's where we want to be, because ultimately, data is for driving business decisions."* Given this guided attention and a realization that the presenter can interact directly with the content, A17 foresaw audiences asking more specific questions: *"if I said: 'hey, what is that peak over there?'…you could just put your finger."*

A potential to distract. Audience study participants were split as to how and when our approach to presenting data might distract audiences. On the one hand, A10 found to the experience to be neither confusing or distracting, and A4 described how *"it pulls you in a good way where the technology doesn't become a distraction"*. On the other hand, participants A16 found the presenter's body language distracting, while A8 worried that audiences would derail presentations by interrupting presenters with requests to perform more interactions.

Falling between these perspectives, A1 wavered on the potential of the interactions to distract audiences: *"it sometimes feels natural and it sometimes feels stilted"*, while both A3 and A14 remarking on how their perspective shifted as the presentation proceeded, with A3 stating: *"initially I wrote down 'gimmick', because I felt like the effect of what you were doing was more distracting than the data I was getting, but that very quickly subsided…I feel like I'm being guided through a story."*

To assess if audiences were distracted by our approach, we collected and contrasted their presentation content summaries, along with any content-related interruptions and questions. Ultimately, we did not observe any pronounced differences in the quality of content summaries for a screen-share presentation and those for a presentation supported by interactive overlays, however the latter did elicit more substantive content-related interruptions and questions. Despite encouraging participants to assume informality

and to interrupt the presenter with content-related questions, the majority of the interruptions made during the overlay-supported presentations pertained to *how* the content was being presented.

When not to use our approach. There are teleconference scenarios involving visualization where our approach might not be applicable. Both P4 and P11 described giving demonstrations of the functionality of data analysis software, where audiences already have an understanding of relevant analytical concepts. P5 also described customer service meetings wherein he would present a dense dashboard to customers, meetings where there is already a shared context and an intent to resolve an issue, rather than an intent to reveal a narrative based in data.

Natural and personal. Most of our presenter study participants commented on how the experience of presenting data with interactive chart overlays made them feel relative to a conventional screen-sharing approach. P6 stated that *"it feels more natural"* while also alluding to the agency or ownership that he felt with respect to the presentation content: *"it's more obvious that I'm the one presenting this information… [the audience] can ask me the question, and not just ask a chart."* Similarly, P8 *"felt more control over the data… you feel like you're really telling the story with your whole body."* Continuing this theme, P5 described how the interface gave him more independence with respect to the language he would use when speaking to an audience: *"I felt like I had more freedom,"* specifically referring to the use of tedious spatial cues such as *"in the bottom left of this visualization."* Beyond personal agency, P1 described how the experience *"felt more like teaching,"* highlighting the potential to apply our approach in situations where remote audiences require data literacy and graphicacy [5] training. Finally, P6 commented that our approach adds a human element to the seemingly precise domain of data: *"it adds an emotional and personal element to something that's usually pretty right and wrong,* [where] *there's one logical answer."*

In contrast to the comments about how our approach might make presentations about data feel more personal, P9 offered a different perspective based on his familiarity with livestreaming practices on YouTube and Twitch: overlay-supported presentations *"would feel more professional"* and *"performative"* than screen-shared slide presentations, and this would make them more engaging. On the other hand, P10 felt *"like you're playing more with the data,"* suggesting a more casual or informal presentation experience.

Flourishes and interaction variations. The majority of audience study participants made references to specific interactions performed by the presenter in presentations supported with interactive overlays. In particular, pinching to clone and drag chart components for value transformation or comparison (Figure 4c—4f) was mentioned by six participants, with A13 remarking how difficult it would be to perform these interactions using conventional presentation approaches.

We also heard comments pertaining to how some of our interactions allowed for flourish or embellishment without unintended effects on the content. A9, A12, and A15 each singled out the *snap* to reveal a band in the stacked area chart (Figure 6a—6b), despite acknowledging that this reveal was triggered by a simple pinch. A12 additionally praised *swipe* flourish when panning the timeline

of the line chart (Figure 5.3—5.5) and a *bloom* gesture that the presenter performed while the stacked area chart transitioned into a stacked bar chart (see supplemental video), describing these as *"jazz hand maneuvers… the theatrical performing part of my brain did really enjoy that."* From the presenter perspective, P8 specifically called out the allowance for flourish when panning the timeline of a chart: *"it's like on the iPhone… give it a little wrist and it goes… you feel like you're more connected with it."*

One unanticipated reflection on the interactions came from A9, who routinely attends and leads presentations with his customers in the public sector, though he specifically spoke about his customers in law enforcement and the armed forces. He mentioned how in these domains, mannerisms such as single-digit pointing are discouraged in interpersonal communication, as they have been found to elicit a heightened stress response. As our deictic interactions can accommodate open-handed pointing, this alleviated his concerns regarding the viability of our approach in these domains.

Reflecting on usability. Many presenter study participants initially assumed awkward or uncomfortable poses, such as stiffly pointing straight at the camera or raising their elbow to be level with their hand. However, within minutes, they assumed more comfortable poses. Retrospectively, several participants appreciated being able to interact with content using their hands, describing it as *'comfortable'* [P7] and *"more effective than speaking to something and using your mouse to point out stuff in a presentation"* [P3].

On the other hand, P11 seemed less comfortable: *"I'm using my whole body, which is uncomfortable… maybe just because it's different, not because it's bad."* P12 also expressed concern that *"some of the natural hand gestures that we're using for speaking could be confusing for this technology,"* which certainly resonated with us given our intention to avoid any such collisions.

Specific interactions also elicited some concerns, such as P6's need for precision pointing: *"the advantage of having a cursor or track pad is that it's a little easier to know exactly where you're going and take your hand off the mouse… if you want something to stop, you can't really cut your hand off and make the visual pause,"* highlighting a need to disable and enable hand-tracking on demand. Aside from pointing, the pinch-based interactions were also an occasional source of frustration; P3 voiced a concern about testing her audience's patience should she require multiple attempts at executing a pinch, and while P7 stated that *"not many things were challenging,"* he found it difficult to know *"how to keep my hand away from the camera so that it feeds it properly."* Lastly, pinching proved especially difficult for P12, whose dark nail polish appeared to thwart hand pose recognition.

Speculating about learnability. All but one of the presenter study participants commented on the process of learning how to effectively interact with chart overlays. An overall sense of optimism was notable, such as when P4 stated *"it's a learning process; that doesn't bother me… I actually really enjoy figuring this thing out and watching it respond."* However, as P9 points out, our approach could be *"a lot to figure out on the fly,",* with P12 suggesting that presenters might benefit from a *"significant amount of training before"* giving their first presentation to an audience. Beyond individual discovery of the interactions, P3 adds: *"I think watching someone else do it first is probably very critical for training."* Despite a perceived learning

curve, we were encouraged by P5 and P8 describing our interactions as *'intuitive'*, and as with any new approach, P6 stated: *"it probably just takes some getting used to."*

Audience participants also speculated on learnability, expressing a similar blend of optimism and concern with respect to an expected learning curve, particularly if we contrast A5 imagining a time when *"we all learn these skills — I presume that's not a challenge"* with A16 suggesting that the presenter had extensively practiced the presentation beforehand.

5.4 Eliciting Ideas to Extend our Approach

When asked about what is uniquely challenging or frustrating about presenting data to remote audiences, two themes stood out: the ability to engage an audience whose attention is divided and difficulties with presenting either complex visuals or complex data. We now consider the ideas we elicited from participants that speak to these themes as a basis for extending our approach.

Engaging an audience and managing attention. While many audience study participants reported feeling engaged and focused on a dynamic presenter, the interaction with overlays was ephemeral, making it difficult for audiences to attend to highlighted values and their related insights. Meanwhile, presenter study participants P2, P7, and P9 independently expressed a desire to pin value annotations in place; *"I instinctively wanted to push it hard and stick there"* [P2], which suggests a potential for gesture recognition at various depths from the camera, and recalls a similar interaction demonstrated by Gong et al. in their HoloBoard presentation system [32]. While a purely functional *pin* gesture like this may seem natural for presenters, we are unsure of how it might be interpreted by audiences. This potential mismatch calls for a deeper exploration of *pinning* and *unpinning* interactions for data marks and annotations from the perspectives of both presenter and audience.

We also foresee a need for unplanned annotations to be added during a presentation. Such functionality could encourage audiences to volunteer their own insights on the data; a presenter could capture and externalize these insights, perhaps through speech-to-text annotation [P6], binding these annotations to data marks based on hand proximity or via linked highlighting (Section 4.2). Alternatively, A6 suggested that a presenter could pass interactive privileges to an audience member: *"when you picked up one of those [bars], you could have said 'do you want to try and move it?'… And then I could have moved it to show you what I'm talking about."*

Presenting complex data. While bar, line, area, and pie charts may satisfy the needs of many presentations of data [13], 16 of our participants urged us to consider a wider palette of data and chart types, which could in turn expand our repertoire of interactions. Specifically, multiple participants mentioned scatterplots, both symbol and choropleth maps, richly-formatted text tables, treemaps, and representations of distributions such as boxplots. However, we appreciate P12's sense of caution: *"it's going to be difficult… especially when you get into very detailed charts,"* suggesting the need for a different approach to interaction for charts with a greater number or density of elements.

Beyond a wider palette of supported charts, A4 and A16 urged us to consider non-linear narratives and the ability to break out of a planned sequence or arrangement of charts in response to audience

engagement. Whether such scene management and content transformation functionality are activated by introducing a new set of non-distracting gestures or through off-screen presenter controls is a question worthy of further research and design.

6 DISCUSSION

We now reflect on the potential of our approach, proposing design implications for presentation tools that reinforce and extend previous research. In Section 3, we stated our goal of improving the data presentation experience for both presenter and audience, with two design imperatives: (1) to support direct interaction with visualization content rather than interaction with the containing scene or slide; and (2) to realize a presentation environment that would require no special equipment and leverage an existing fluency in communicative body language encompassing illustrative and affective hand gestures. In developing our environment, the interplay between these imperatives made us aware of how introducing novel functional interactions can be an imposition on presenters as well as a potential source of distraction for audiences. As a result, we focused on identifying interactions that were simultaneously functional and deictic, interactions that could be embellished with rhetorical flourishes.
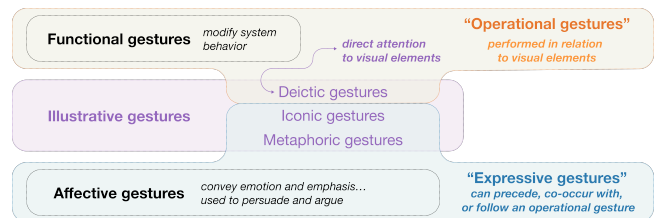

Figure 8: Our pragmatic reclassification of functional, illustrative, and affective gestures for presenting data, which distinguishes between operational gestures performed in relation to specific visual elements and expressive gestures that can be performed independent of any visual aid.

By reflecting on these design imperatives, we now have a better understanding of what hand gestures are useful when talking about data with interactive chart overlays. In particular, we can revisit the categorization of hand gestures described in Section 2 from a pragmatic perspective (Figure 8). Deictic and functional gestures can be considered to be *operational* in that they mediate the experience of attending to dynamic visual artifacts such as interactive chart overlays. In contrast, iconic, metaphorical, and affective gestures facilitate *expressive* human-to-human communication; they are an essential component of rhetoric, and they can be performed independent of any visual aid. With this new understanding, we propose three design implications for presentation tools that support touchless interaction with data:

Recognize operational gestures. We contend that the most fruitful potential path lies in recognizing operational gestures performed in the context of dynamic charts, and this recognition should be independent of any affective interpretation. As a comparison, consider the recent presentation authoring system described by Saquib et al.[69], one capable of recognizing individual presenters' idiosyncratic mix of illustrative and affective gesture. When speaking about

familiar data abstractions and supported by a limited palette of associated charts, a presenter should be able to reliably perform a minimum set of operational gestures, each with an allowable tolerance in terms of how it is recognized.

Operational gestures should be familiar and leverage cultural norms, such as deictic pointing gestures in interpersonal conversation or functional gestures learned from using ubiquitous touchscreen devices [19], such as swiping to pan along a continuous dimension. Familiarity with operational gestures could also explain participants' optimism with respect to learnability; although many participants expected that learning how to use a presentation environment integrating our approach would take time and practice, we also heard our operational gestures described as *'intuitive'*.

Assign gestures according to frequency and salience. As intuitive as some operational gestures may be, presenters likely need time to learn how to adapt these familiar gestures to a new medium, so system designers must be judicious when mapping operational gestures to interface functions. Consider the functional pinch gesture: easy to perform on a touchscreen, but as we saw in the presenter study, it can be harder to perform with webcam-based hand tracking. System response should also be commensurate with the deliberateness and expected frequency of the operational gesture. For instance, single-handed pointing at elements in a chart can happen quite often during a presentation; in response, our environment reveals unobtrusive and ephemeral value annotations. In contrast, pinching to select a chart element is a more deliberate and infrequent operational gesture, and our system responses are accordingly more overt (e.g., zooming, panning, element cloning).

Do no harm: avoid conflict with expressive gestures. A challenge and opportunity for future presentation systems is to encourage the richness of expressive body language without compromising system behaviour. In other words, the expressive gestures that a system ignores are just as important as the operational gestures that it acts upon. While expressive gesture can be performed independently of an operational gesture, we demonstrated how an expressive gesture can immediately precede, follow, or take place concurrently with an operational gesture. For instance, in Section 4.4, we described *expand*, *swipe*, and *compress* gestures as flourishes that follow an operational pinch gesture; the pinch triggered the system response of zooming or panning along the temporal domain of a line chart, while the flourishes helped to express the concept of a continuum of time, complementing the visual transition without triggering their own system response. Consider that these iconic gestures could have supported verbal statements about different time intervals without the line chart even being shown. Similarly, we described the affective *snap* gesture as a rhetorical flourish following a pinch that triggered the reveal of a new chart element; here the snap both visually and audibly emphasized the importance of the newly revealed content. Distinguishing expressive and operational gestures will undoubtedly present challenges, though a system could make inferences by collecting additional movement data, such as the acceleration or beat frequency of a fingertip or hand, as well as additional position data, such as the location of the hand relative to the body.

Limitations and future work. In our studies, we limited recruitment to those associated with a multinational software company,

and given the office setting of the presenter study, COVID-19 safety protocols prohibited us from inviting non-employee participants. We hesitated to conduct the presenter study remotely, as we could not guarantee a consistent environment free from distractions in presenters' homes or remote co-working spaces. On the other hand, our meeting room setting (Figure 7) may have been potentially unrepresentative with respect to the relative positioning of the presenter, the webcam, and the displays. It was also challenging for participants to remain in character during the presentation scenarios. In the presenter study, we reminded participants of the scenario with prompts related to the presentation content, while in the audience study, we noted several presentation interruptions focusing not on the content but on the mode of delivery and the technology in use. Whether audiences will engage the presenter with meaningful content-focused discussion and retain key messages should be examined in a longitudinal deployment study of a presentation system integrating our approach.

We identified several opportunities to extend our approach, though two of the recurring categories of ideas elicited by our study participants stand out.

First, we can identify additional ways to capture attention and increase engagement. This includes allowing the audience to interact with and annotate the content during a presentation [18, 84], or provide a shared awareness of what audience members are looking or pointing at [35, 38]. Alternatively, we could explore more flexible and expressive chart annotation options for presenters by recognizing handheld peripherals such as pens or pointing devices, drawing inspiration from Perlin et al.'s Chalktalk [59].

Second, we recognize that a palette of basic palette of charts will satisfy many presentation scenarios [13], but we must nevertheless consider how to interact with more detailed or unfamiliar communication-oriented charts [44]. This could involve incrementally constructing a detailed chart through a series of operational gestures, or transforming [14, 67] a more familiar chart into a new configuration, such as through bending, stretching, rolling, or tearing chart components.

7 CONCLUSION

The capacity to perform communicative and collaborative knowledge work at distance has increased in recent years, particularly since the onset of the COVID-19 pandemic. At the same time, organizations both large and small have acknowledged the importance of making decisions and adopting policies that are grounded in data. Despite a need to discuss data with remote stakeholders, colleagues, and customers, existing tools for presenting data to these audiences fall short: presentations fail to engage audiences and the multimodal expressivity of presenters goes unseen.

In this paper, we described an approach that aims to help restore the multimodal richness and nuance of co-located presentations about data. We demonstrated a presentation environment that composites presenter webcam video with interactive visualization overlays that respond to a presenter's hand movements. We identified a set of interactions that can support presentations about data, using them with different types of charts to draw an audience's attention to differences in quantity, proportion, and interval. We evaluated our presentation environment in two studies, examining

the perspectives of 12 presenters and 17 audience members with respect to utility, usability, and the capacity to engage or distract audiences. Ultimately, we remain optimistic about the potential of multimodal presentation tools for presenting data to remote audience.1 However, as tool builders expand the vocabulary of operational gestures applicable across chart types, we urge them to avoid potential conflicts with expressive gestures.

REFERENCES

[1] Roland Aigner, Daniel Wigdor, Hrvoje Benko, Michael Haller, Alexandra Ion, Shengdong Zhao, and Jeffrey Tzu Kwan Valino Koh. 2012. Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI. Microsoft Research Technical Report MSR-TR-2012-111 https://tinyurl.com/msrgesture2012.

[2] Mary Aldugom, Kimberly Fenn, and Susan Wagner Cook. 2020. Gesture during math instruction specifically benefits learners with high visuospatial working memory capacity. *Cognitive Research: Principles and Implications* 5, 1 (2020). https://doi.org/10.1186/s41235-020-00215-8

[3] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Christophe Hurter, and Pourang Irani. 2015. Understanding data videos: Looking at narrative visualization through the cinematography lens. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. https://doi.org/10.1145/2702123.2702431

[4] Gilbert Austin. 1806. *Chironomia; Or, a Treatise on Rhetorical Delivery.* T. Cadell and W. Davies.

[5] William GV Balchin and Alice M Coleman. 1966. Graphicacy should be the fourth ace in the pack. *Cartographica: The International Journal for Geographic Information and Geovisualization* 3, 1 (1966). https://doi.org/10.3138/C7Q0-MM01-6161-7315

[6] Thomas Baudel and Michel Beaudouin-Lafon. 1993. Charade: remote control of objects using free-hand gestures. *Commun. ACM* 36, 7 (1993). https://doi.org/10.1145/159544.159562

[7] Raymond Bjuland, Maria Luiza Cestari, and Hans Erik Borgersen. 2008. The interplay between gesture and discourse as mediating devices in collaborative mathematical reasoning: A Multimodal approach. *Mathematical Thinking and Learning* 10, 3 (2008). https://doi.org/10.1080/10986060802216169

[8] Richard A Bolt. 1980. "Put-that-there" Voice and gesture at the graphics interface. In *Proceedings of the ACM Conference on Computer graphics and Interactive Techniques.* https://doi.org/10.1145/800250.807503

[9] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D^3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 17, 12 (2011). https://doi.org/10.1109/TVCG.2011.185

[10] Jeremy Boy, Francoise Detienne, and Jean-Daniel Fekete. 2015. Storytelling in information visualizations: Does it engage users to explore data?. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. https://doi.org/10.1145/2702123.2702452

[11] Matthew Brehmer. 2021. The Information in Our Hands. Information+ 2021 conference presentation. https://vimeo.com/592591860.

[12] Matthew Brehmer. 2022. Higher Education Costs in Washington (1965 - 2019). Tableau Public workbook. https://tinyurl.com/augchiroworkbook.

[13] Matthew Brehmer and Robert Kosara. 2022. From jam session to recital: Synchronous communication and collaboration around data in organizations. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VIS)* 28, 1 (2022). https://doi.org/10.1109/TVCG.2021.3114760

[14] John Brosz, Miguel A Nacenta, Richard Pusch, Sheelagh Carpendale, and Christophe Hurter. 2013. Transmogrification: Causal manipulation of visualizations. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. https://doi.org/10.1145/2501988.2502046

[15] Peter Bull. 1986. The use of hand gesture in political speeches: Some case studies. *Journal of Language and Social Psychology* 5, 2 (1986). https://doi.org/10.1177/0261927X8652002

[16] Philip Cash and Anja Maier. 2016. Prototyping with your hands: The many roles of gesture in the communication of design concepts. *Journal of Engineering Design* 27, 1-3 (2016). https://doi.org/10.1080/09544828.2015.1126702

[17] Xi Chen, Wei Zeng, Yanna Lin, Hayder Mahdi Ai-Maneea, Jonathan Roberts, and Remco Chang. 2021. Composition and configuration patterns in multiple-view visualizations. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 27, 2 (2021). https://doi.org/10.1109/TVCG.2020.3030338

[18] John Joon Young Chung, Hijung Valentina Shin, Haijun Xia, Li-yi Wei, and Rubaiat Habib Kazi. 2021. Beyond show of hands: Engage viewers via expressive and scalable visual communication in live stream. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. https://doi.org/10.1145/3411764.3445419

[19] Mauricio Cirelli and Ricardo Nakamura. 2014. A survey on multi-touch gesture recognition and multi-touch frameworks. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS)*. https://doi.org/10.1145/2669485.2669509

[20] Jean S. Clarke, Joep P. Cornelissen, and Mark P. Healey. 2019. Actions speak louder than words: How figurative language and gesturing in entrepreneurial pitches influences investment judgments. *Academy of Management Journal* 62 (2019). https://doi.org/10.5465/amj.2016.1008

[21] Matthew Conlen and Jeffrey Heer. 2018. Idyll: A markup language for authoring and publishing interactive articles on the web. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. https://doi.org/10.1145/3242587.3242600

[22] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H Thomas, and Kim Marriott. 2017. ImAxes: Immersive axes as embodied affordances for interactive multivariate data visualisation. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. https://doi.org/10.1145/3126594.3126613

[23] Joep P. Cornelissen, Jean S. Clarke, and Alan Cienki. 2012. Sensegiving in entrepreneurial contexts: The use of metaphors in speech and gesture to gain and sustain support for novel business ventures. *International Small Business Journal* 30, 3 (2012). https://doi.org/10.1177/0266242610364427

[24] Stefania Cuccurullo, Rita Francese, Sharefa Murad, Ignazio Passero, and Maurizio Tucci. 2012. A gestural approach to presentation exploiting motion capture metaphors. In *Proceedings of the ACM International Working Conference on Advanced Visual Interfaces (AVI)*. https://doi.org/10.1145/2254556.2254584

[25] Evanthia Dimara, Harry Zhang, Melanie Tory, and Steven Franconeri. 2022. The unmet data visualization needs of decision makers within organizations. *To appear in IEEE Transactions on Visualization and Computer Graphics (TVCG)* 28, TBA (2022). https://doi.org/10.1109/TVCG.2022.3074023

[26] Marion Dohen and Benjamin Roustan. 2017. Co-production of speech and pointing gestures in clear and perturbed interactive tasks: multimodal designation strategies. In *Proceedings of the Annual Conference of the International Speech Communication Association (Interspeech)*. https://hal.inserm.fr/GIPSA-PCMD/hal-02367749v1.

[27] Steven Drucker and Roland Fernandez. 2015. A unifying framework for animated and interactive unit visualizations. Microsoft Research Technical Report MSR-TR-2015-65 https://tinyurl.com/sanddance2015.

[28] Steven Drucker, Samuel Huron, Robert Kosara, Jonathan Schwabish, and Nicholas Diakopoulos. 2018. Communicating data to an audience. In *Data-Driven Storytelling*, Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale (Eds.). CRC Press. https://tinyurl.com/drucker2018.

[29] Economic Opportunity Institute. 2021. Education. Tableau Public workbook. https://tinyurl.com/EOIeducation.

[30] Laurie Edwards. 2005. Gesture and mathematical talk: Remembering and problem solving. In *Proceedings of the American Educational Research Association Conference.*

[31] Adam Fourney, Michael Terry, and Richard Mann. 2010. Gesturing in the wild: Understanding the effects and implications of gesture-based interaction for dynamic presentations. In *Proceedings of the Human-Computer Interaction Conference (HCI)*. https://doi.org/10.14236/ewic/HCI2010.29

[32] Jiangtao Gong, Teng Han, Siling Guo, Jiannan Li, Siyu Zha, Liuxin Zhang, Feng Tian, Qianying Wang, and Yong Rui. 2021. HoloBoard: A large-format immersive teaching board based on pseudo holographics. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. https://doi.org/10.1145/3472749.3474761

[33] Google. 2019. Teachable Machine. https://teachablemachine.withgoogle.com.

[34] Google. 2020. MediaPipe. https://google.github.io/mediapipe/.

[35] Jens Emil Grønbæk, Banu Saatçi, Carla F Griggio, and Clemens Nylandsted Klokmose. 2021. MirrorBlender: Supporting hybrid meetings with a malleable video-conferencing system. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. https://doi.org/10.1145/3411764.3445698

[36] Maisevli Harika, Ary Setijadi P, Hilwadi Hindersah, and Bong-Kee Sin. 2016. Finger-pointing gesture analysis for slide presentation. *Journal of Korea Multimedia Society* 19, 8 (2016). https://doi.org/10.9717/kmms.2016.19.8.1225

[37] Simon Harrison. 2021. Showing as sense-making in oral presentations: The speech-gesture-slide interplay in TED talks by Professor Brian Cox. *Journal of English for Academic Purposes* 53 (2021). https://doi.org/10.1016/j.jeap.2021.101002

[38] Zhenyi He, Keru Wang, Brandon Yushan Feng, Ruofei Du, and Ken Perlin. 2021. GazeChat: Enhancing virtual conferences with gaze-aware 3D photos. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. https://doi.org/10.1145/3472749.3474785

[39] Jeffrey Heer and George G. Robertson. 2007. Animated transitions in statistical data graphics. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 13, 6 (2007). https://doi.org/10.1109/TVCG.2007.70539

[40] Nathalie Henry Riche, Christophe Hurter, Nicholas Diakopoulos, and Sheelagh Carpendale (Eds.). 2018. *Data-Driven Storytelling.* CRC Press.

[41] Jessica Hullman and Nick Diakopoulos. 2011. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer*

Graphics (Proceedings of InfoVis) 17, 12 (2011). https://doi.org/10.1109/TVCG.2011.255

[42] Paul Issartel, Florimond Guéniat, and Mehdi Ammi. 2014. Slicing techniques for handheld augmented reality. In *Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI)*. https://doi.org/10.1109/3DUI.2014.6798839

[43] Minju Kim and Kwangyun Wohn. 2018. HoloBox: Augmented visualization and presentation with spatially integrated presenter. *Interacting with Computers* 30, 3 (2018). https://doi.org/10.1093/iwc/iwy007

[44] Robert Kosara. 2016. Presentation-oriented visualization techniques. *IEEE Computer Graphics and Applications (CG&A)* 36, 1 (2016). https://doi.org/10.1109/MCG.2016.2

[45] Robert Kosara. 2017. An argument structure for data stories. In *Short Paper Proceedings of the Eurographics / IEEE VGTC Conference on Visualization (EuroVis)*. https://doi.org/10.2312/eurovisshort.20171129

[46] Alex Lascarides and Matthew Stone. 2009. Discourse coherence and gesture interpretation. *Gesture* 9, 2 (2009). https://doi.org/10.1075/gest.9.2.01las

[47] Bongshin Lee, Rubaiat Habib Kazi, and Greg Smith. 2013. SketchStory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 19, 12 (2013). https://doi.org/10.1109/TVCG.2013.191

[48] Hyeon-Kyu Lee and Jin-Hyung Kim. 1999. An HMM-based threshold model approach for gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21, 10 (1999). https://doi.org/10.1109/34.799904

[49] David Matsumoto, Mark G. Frank, and Hyi Sung Hwang. 2013. *Nonverbal communication: Science and applications: Science and Applications*. SAGE.

[50] Fabrice Matulic, Lars Engeln, Christoph Träger, and Raimund Dachselt. 2016. Embodied interactions for novel immersive presentational experiences. In *Extended Abstract Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. https://doi.org/10.1145/2851581.2892501

[51] Bernhard Maurer, Alina Krischkowsky, and Manfred Tscheligi. 2017. Exploring gaze and hand gestures for non-verbal in-game communication. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play*. https://doi.org/10.1145/3130859.3131296

[52] Brian James McInnis, Lu Sun, Jungwon Shin, and Steven P Dow. 2020. Rare, but valuable: Understanding data-centered talk in news website comment sections. *Proceedings of the ACM on Human-Computer Interaction (CSCW)* 4 (2020). https://doi.org/10.1145/3415245

[53] Sean McKenna, Nathalie Henry Riche, Bongshin Lee, Jeremy Boy, and Miriah Meyer. 2017. Visual narrative flow: Exploring factors shaping data visualization story reading experiences. *Computer Graphics Forum (Proceedings of EuroVis)* 36, 3 (2017). https://doi.org/10.1111/cgf.13195

[54] mmhmm 2022. mmhmm. https://mmhmm.app.

[55] Christopher Noessel and Nathan Shedroff. 2012. *Make It So: Interaction Design Lessons from Science Fiction*. Rosenfeld Media.

[56] OBS 2022. OBS Studio. https://obsproject.com.

[57] Kenton O'Hara, Gerardo Gonzalez, Graeme Penney, Abigail Sellen, Robert Corish, Helena Mentis, Andreas Varnavas, Antonio Criminisi, Mark Rouncefield, Neville Dastur, and Tom Carrell. 2014. Interactional order and constructed ways of seeing with touchless imaging systems in surgery. *Computer Supported Cooperative Work (CSCW): The Journal of Collaborative Computing and Work Practices* 23, 3 (2014). https://doi.org/10.1007/s10606-014-9203-4

[58] Anshul Vikram Pandey, Anjali Manivannan, Oded Nov, Margaret Satterthwaite, and Enrico Bertini. 2014. The persuasive power of data visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 20, 12 (2014). https://doi.org/10.1109/TVCG.2014.2346419

[59] Ken Perlin, Zhenyi He, and Karl Rosenberg. 2018. Chalktalk: A visualization and communication language – As a tool in the domain of computer science education. arXiv preprint https://arxiv.org/abs/1809.07166.

[60] Jing Qian, Jiaju Ma, Xiangyu Li, Benjamin Attal, Haoming Lai, James Tompkin, John F Hughes, and Jeff Huang. 2019. Portal-ble: Intuitive free-hand manipulation in unbounded smartphone-based augmented reality. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. https://doi.org/10.1145/3332165.3347904

[61] Zening Qu and Jessica Hullman. 2017. Keeping multiple views consistent: Constraints, validations, and exceptions in visualization authoring. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 24, 1 (2017). https://doi.org/10.1109/TVCG.2017.2744198

[62] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer, and Eun Kyoung Choe. 2017. ChartAccent: Annotation for data-driven storytelling. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*. https://doi.org/10.1109/PACIFICVIS.2017.8031599

[63] Hans Rosling. 2006. Debunking myths about the "Third World". TED (Technology, Entertainment, Design) conference presentation. https://gapminder.org/videos.

[64] Hans Rosling. 2007. The seemingly impossible is possible. TED (Technology, Entertainment, Design) conference presentation. https://gapminder.org/videos.

[65] Hans Rosling. 2010. 200 Countries, 200 Years, 4 Minutes. BBC Four. https://youtu.be/jbkSRLYSojo.

[66] Hans Rosling. 2013. The River of Myths. https://youtu.be/lYpX4l2UeZg.

[67] Puripant Ruchikachorn and Klaus Mueller. 2015. Learning visualizations by analogy: Promoting visual literacy through visualization morphing. *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 21, 9 (2015). https://doi.org/10.1109/TVCG.2015.2413786

[68] Bahador Saket, Hannah Kim, Eli T Brown, and Alex Endert. 2017. Visualization by demonstration: An interaction paradigm for visual data exploration. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 23, 1 (2017). https://doi.org/10.1109/TVCG.2016.2598839

[69] Nazmus Saquib, Rubaiat Habib Kazi, Li-Yi Wei, and Wilmot Li. 2019. Interactive body-driven graphics for augmented video performance. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. https://doi.org/10.1145/3290605.3300852

[70] Arvind Satyanarayan and Jeffrey Heer. 2014. Authoring narrative visualizations with Ellipsis.... *Computer Graphics Forum (Proceedings of EuroVis)* 33, 3 (2014). https://doi.org/10.1111/cgf.12392

[71] Arvind Satyanarayan, Bongshin Lee, Donghao Ren, Jeffrey Heer, John Stasko, John Thompson, Matthew Brehmer, and Zhicheng Liu. 2019. Critical reflections on visualization authoring systems. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 26, 1 (2019). https://doi.org/10.1109/TVCG.2019.2934281

[72] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 16, 6 (2010). https://doi.org/10.1109/TVCG.2010.179

[73] Steven Smith. 2019. The digital chironomia. *Enculturation*. https://enculturation.net/the-digital-chironomia.

[74] Mark Stefik, Daniel G Bobrow, Gregg Foster, Stan Lanning, and Deborah Tatar. 1987. WYSIWIS revised: Early experiences with multiuser interfaces. *ACM Transactions on Information Systems (TOIS)* 5, 2 (1987). https://doi.org/10.1145/27636.28056

[75] Jürgen Streeck. 2009. *Gesturecraft: The Manu-Facture of Meaning*. John Benjamins Publishing.

[76] Michael Studdert-Kennedy. 1994. *Hand and Mind: What Gestures Reveal About Thought*. By David McNeill. *Language and Speech* 37, 2 (1994). https://doi.org/10.1177/002383099403700208 Review article.

[77] Hariharan Subramonyam and Eytan Adar. 2018. SmartCues: A multitouch query approach for details-on-demand through dynamically computed overlays. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)* 25, 1 (2018). https://doi.org/10.1109/TVCG.2018.2865231

[78] Svelte 2022. Svelte. https://svelte.dev/.

[79] Gustav Verhulsdonck and Jacquelyn Ford Morie. 2009. Virtual chironomia: Developing standards for non-verbal communication in virtual worlds. *Journal For Virtual Worlds Research* 2, 3 (2009). https://doi.org/10.4101/jvwr.v2i3.657

[80] Vizrt 2021. Vizrt Extended Reality. https://vizrt.com/flexible-access/extended-reality.

[81] Tijana Vuletic, Alex Duffy, Laura Hay, Chris McTeague, Gerard Campbell, and Madeleine Grealy. 2019. Systematic literature review of hand gestures used in human computer interaction interfaces. *International Journal of Human-Computer Studies* 129 (2019). https://doi.org/10.1016/j.ijhcs.2019.03.011

[82] Petra Wagner, Zofia Malisz, and Stefan Kopp. 2014. Gesture and speech in interaction: An overview. *Speech Communication* 57 (2014). https://doi.org/10.1016/j.specom.2013.09.008

[83] Daniel Wigdor, Clifton Forlines, Patrick Baudisch, John Barnwell, and Chia Shen. 2007. Lucid Touch: A see-through mobile device. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. https://doi.org/10.1145/1294211.1294259

[84] Wesley Willett, Jeffrey Heer, and Maneesh Agrawala. 2012. Strategies for crowd-sourcing social data analysis. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*. https://doi.org/10.1145/2207676.2207709

[85] Bodo Winter, Marcus Perlman, and Teenie Matlock. 2013. Using space to talk and gesture about numbers: Evidence from the TV News Archive. *Gesture* 13, 3 (2013). https://doi.org/10.1075/gest.13.3.06win

[86] Zhenpeng Zhao and Niklas Elmqvist. 2022. The stories we tell about data: Media types for data-driven storytelling. arXiv preprint. https://arxiv.org/abs/2202.00047.