

Exploring Mid-Air Hand Interaction in Data Visualization

Zona Kostic, Catherine Dumas, Sarah Pratt, and Johanna Beyer

Abstract—Interacting with data visualizations without an instrument or touch surface is typically characterized by the use of mid-air hand gestures. While mid-air expressions can be quite intuitive for interacting with digital content at a distance, they frequently lack precision and necessitate a different way of expressing users' data-related intentions. In this work, we aim to identify new designs for mid-air hand gesture manipulations that can facilitate instrument-free, touch-free, and embedded interactions with visualizations, while utilizing the three-dimensional (3D) interaction space that mid-air gestures afford. We explore mid-air hand gestures for data visualization by searching for *natural* means to interact with content. We employ three studies—an Elicitation Study, a User Study, and an Expert Study, to provide insight into the users' mental models, explore the design space, and suggest considerations for future mid-air hand gesture design. In addition to forming strong associations with physical manipulations, we discovered that mid-air hand gestures can: promote space-multiplexed interaction, which allows for a greater degree of expression; play a functional role in visual cognition and comprehension; and enhance creativity and engagement. We further highlight the challenges that designers in this field may face to help set the stage for developing effective gestures for a wide range of touchless interactions with visualizations.

Index Terms—mid-air hand gestures, touchless interaction, design space, elicitation study, user study, expert study, data visualization

I. INTRODUCTION

In recent years, information visualization has become ubiquitous in representing data as a visual means. Visualization is now accessible to a large number of users and in a variety of platforms, such as ambient screens [1], wearables [2], or immersive environments [3]. These diverse audiences and spaces necessitate alternative methods of interacting with visualization. Nevertheless, many interaction modalities are frequently neglected in visualization, thereby limiting the ability to discover user intents [4]. To investigate engagement in information visualization for a general audience, it is necessary to evaluate different inputs used to perform operations and manipulations. Thus far, visualization has focused on many interaction modalities such as, instrumental interaction [5], proxemics [6], touch [7], and speech [8], and Lee et al. [9] provide a detailed overview of the potential of novel modalities in data visualization. Prior to the emergence of immersive environments and their use for data visualization, however,

Zona Kostic and Johanna Beyer are with John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, 02138. E-mail: {zonakostic, jbeyer}@seas.harvard.edu

Catherine Dumas and Sara Pratt are with Simmons University School of Library and Information Science, and State University of New York, Albany. E-mail: {catherine.dumas2, sarah.pratt}@simmons.edu

Manuscript received July 17, 2023.

the significance of mid-air hand gestures was frequently overlooked due to their limited precision and reliability. The advantages and disadvantages of hand gesture interaction are revealed by a comprehensive examination of techniques employed for immersive environments [10]. Despite technological advancements, there is no standard for ensuring that a set of mid-air hand gestures is appropriate for a domain. Moreover, to provide a consistent user experience, gesture sets should be transferable between domains, and Hosseini et al.'s [11] study seeks to identify cross-domain consensus of hand gestures in HCI. However, their study does not focus on hand interactions with data visualization. The discoverability of mid-air hand gestures for a set of visualizations and their applicability and usability have not been the primary focus of any visualization or HCI studies. To the best of our knowledge, the appropriateness of this mode of interaction has not been thoroughly researched.

Compared to conventional interaction modalities, mid-air gestures are mostly utilized in immersive environments, such as interactions with situated visualizations [12], or interactions with public displays [13]. In addition to the possibility of expressing physical manipulations while interacting with immersive content, gestures can also be used in a variety of other contexts. For example, the *Gesture Watch* [14] enables users to perform mid-air gestures above the watch due to the limited size of its interaction surface. With this, mid-air gestures expand the interaction space, particularly for transitions and changes that do not require high precision (e.g., guides, strokes, and manipulations). In situations involving interactive applications, where touch and speech are inadequate due to occlusion issues ("fat finger"), ambient noise, or the lack of suitable instruments, mid-air hand gestures are a viable touch-free interaction option [15], [16], [17]. This also presents the opportunity for touchless interaction, where certain tasks can be performed more effectively from a distance [18], allowing users to move their arms in more flexible and comfortable positions [19]. Besides their versatility, a significant advantage of mid-air gestures is their perceived naturalness and intuitiveness [20], [21].

In this paper, we examine the role of mid-air hand gestures as a touchless interaction modality with data visualization, paying particular attention to previously identified challenges such as recognition, input resolution, and comfort [9], [23], [24]. We conducted three studies: an **Elicitation Study (ES)**, a **User Study (SS)**, and an **Expert Study (XS)** (see Fig. 1). Details on our study goals and outcomes, and overall paper structure can be seen in Fig. 1. In the ES, we focus on the principles from Wobbrock et al. [25], which

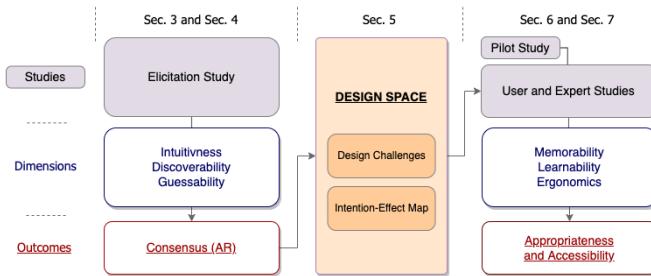


Fig. 1. Our research methodology. From the elicitation study (left column), we derive and explore the design space for mid-air gestures (middle column). We evaluate the gestures in subsequent user and expert studies (third column). Each column lists **Dimensions** of appropriateness [22] for gesture selection related to studies, as well as the **Outcomes** of studies.

are based on the guessability technique. We look to reduce the *gulf of execution* [26], a discrepancy between user intention and user execution. Next, we explore the *design space* of mid-air hand interactions with visualizations. By exploring this design space, we can balance multiple criteria when designing a gesture vocabulary. In the SS, we assess the *degree of appropriateness* [22] of the gesture vocabulary designed for a particular scenario. In the XS, we use an “ease-of-use” measure [27] to rate the *ergonomics* of the designed gestures. We follow the intuitive and ergonomic principles from Nielsen et al. [28] in the SS and XS. We seek to address the following research questions with all of our studies:

- What is the design space for mid-air hand interactions with data visualizations? (Elicitation Study)
- What is the learnability of designed gestures, and how does this affect the memorability for new users? (User Study)
- What is the trade-off between memorability and comfort for designed gestures? (Expert Study)

The main contribution of the paper attributes to the exploration of a design space of mid-air hand gestures for interacting with data visualizations. Following the exploration, we identify and evaluate a vocabulary of mid-air hand gestures in a user and expert study. Furthermore, we present a comprehensive discussion of the experiences and insights gained during the various stages of design, implementation, and evaluation. Our studies found that gesture suggestions are influenced by the visual encoding used to represent the data. However, the classification of gestures based on physical actions is consistent across visualizations for the same operation. Filtering time, for example, elicits semaphoric dynamic gestures, such as repeated hand flicks, consistently throughout different visual encodings, such as barcharts and scatterplots. In addition, legacy bias and physical actions have a greater impact on elicited gestures than visual representation, which directly affects the learnability and memorability of the gestures for new users. Users prefer gestures with strong physical associations and are able to use parallel inputs and hybrid configurations despite minor ergonomic challenges. To conclude, we provide considerations for mid-air hand interaction design and stress on design tensions and challenges that designers could face in the future.

II. RELATED WORK

Hand gestures serve multiple communication and comprehension functions [29]. An attempt to visualize an object, such as an interface component like a button, is frequently accompanied by gestures that express the object’s functionality (pressing or clicking) in addition to outlining its shape (round or square) [30]. Hands are frequently referred to as a *natural* mode of interaction because this is how we interact with the physical world and its objects [31]. In popular culture, artists have drawn inspiration from the portrayal of actions using *in-air* or *mid-air* hand gestures for art performances and various science fiction films; this is similar to the manipulations depicted in Steven Spielberg’s 2002 film *Minority Report*. Hand gesticulation, such as touching the surface, manipulating the object, or communicating via in-air expressions, enables the congruent representation of diverse concepts. It establishes a clear association between the operation and the task.

Congruent interactions have been explored in a variety of data visualization contexts, particularly to engage with the content on screens [32]. The space has been extended to a hand gestural interaction that activates kinetic manipulations and supports multiple inputs in parallel [7]. Mid-air hand interaction appears to be preferable in different scenarios to other interaction modalities. For example, Huang et al. [33] support the notion that mid-air hand gestures are more effective than a mouse for complex immersive graphs. Hegde et al. [34] posit that mid-air gestures, rather than voice or gaze, are preferred modes of interaction because of the stronger associations users can have through a more natural manipulation and reduced effort when interacting with the environment. Even before the invention of gesture recognition with head-mounted displays, complex interactions with virtual tools were facilitated by fingertip positioning and gesticulating in the air [35]. Mid-air input has proven effective for large amplitude slide gestures in immersive environments and large screens, but also on small screens, as a way to extend the interaction space or to complement other interaction modalities [36], [37], [16]. The location of a gesture can serve as a reminder of its function, and the motion pattern of a gesture does not need to be relearned if it resembles a shape associated with the action [38].

A. Hand Gestures and Data Visualization

The expression of data-related intentions with data visualizations generally adheres to interactions that resemble communication with physical objects. For example, Thompson et al. [39] and Schmidt et al. [40] discuss multi-touch gestures that allow hands to manipulate the visualization through natural movements, enhancing the experience of “virtual tangibility” [7]. Rzeszotarski and Kittur [41], in particular, see the use of “physics-based affordances” for direct dialogue with data as an easy method for users to manipulate, analyze, and interact with visualizations. Wall displays are particularly appropriate for hand gesture interfaces, such as gesticulating with tokens [42] or instruments

[43], but also using mid-air hand inputs [44]. Jansen [45] explores various types of data physicalizations that users can interact with directly using their hands; Taher et al. [46] later suggest that this can be accomplished with or without direct touch, such as hovering one's hand over a dynamic data representation.

Mid-air hand gestures, when combined with touch [47], pen [48], voice [49], or virtual pointer [50] effectively complement other modalities and play a significant role in a 3D interaction space. Visual feedback stimulates intuitive mid-air gesticulation and serves as an effective compensation for the lack of kinesthetic feedback in mid-air interaction [18]. In addition to translating button presses, augmented reality interfaces require the user to grab elements, pinch to zoom, and rotate knobs [51]. Despite the many potential applications for user interactions through mid-air hand gestures, this input method is still only used for immersive visualizations.

B. Hand Gesture Design

Designing a gesture system requires consideration of numerous factors in addition to technological constraints. Nielsen et al. [28] define the core of the human-based approach when developing gestures by recognizing the following characteristics: Gestures should be easy to perform and remember; intuitive; metaphorically and iconically logical towards functionality; and ergonomic, or not physically stressing when used frequently. Understanding users' mental models and cognitive biases aids in recognizing the associations that motivate users to take specific actions [52]. For example, hand gestures that resemble zooming into content (e.g., "pinch-to-zoom") are performed the same way regardless of the size of the content on the screen [53], [54], [39]; when interacting with remote visualizations, clicking or selecting is typically performed with a "pinching" or "tapping" in the air [15], [20], [55]; zoom and pan operations are frequently performed together and follow "grab&drag" movements, expressed using various interaction modalities [56], [57], [58]. There is cross-domain consistency for gestures as well as similar themes across the references [11]. Furthermore, mid-air hand gestures provide the possibility to expand mental models and interaction space into parallel, multi-fidelity, and hybrid hand gesture inputs [59], [60], [61], [12].

Gesture Studies. The existing literature on gesture studies and vocabulary design is highlighting iterative, user-focused, factor-centric methodologies [62]. *Elicitation studies* have yielded an impressive amount of knowledge; Villarreal-Narvaez et al. [27] offer a comprehensive review of the literature on user agreements, design, participants, and measures in elicitation studies. The goal of a gesture elicitation study is to determine the *first guess* [63], which categorizes gestures based on specific design criteria, utilizing the aforementioned dimensions of appropriateness [22] for gesture selection. This is achieved by demonstrating the result of a gesture and then requesting that users perform the action that caused it [64]. The criteria of *intuitive*

or natural or interacting with little or no instruction is often impacted by *legacy bias*, where gesture proposals are biased by the users' experiences with prior interfaces and technology. Several approaches exist for minimizing the legacy bias, such as *priming*, *partners*, *production* [65], or *increased production* [66]. Additional techniques, such as priming participants with a *frame*, are a useful design methodology for retaining participants within a particular scenario, thereby removing the reference to previous technology [67]. However, interaction designers can also use legacy bias to their advantage [68] when users intentionally suggest movements based on prior knowledge or experiences. Legacy bias can aid users' comfort and satisfaction, especially when designing novel interfaces. According to the Danilescu and Piorkowski's [66] synthesis of prior research, *culturally shared metaphors* may result in higher agreement scores, improved discoverability, and enhanced learnability. It is crucial to comprehend the equilibrium between leveraging prior experiences and proposing novel designs.

Another approach to eliciting gestures are Wizard of Oz studies, which replicate an automated system by incorporating a human operator, with the objective of minimizing the gulf of execution. In the absence of implementation, this form of research is employed to evaluate expensive concepts and define the scope of the problem. Nevertheless, this particular concept might be impractical for research involving remote participants due to the potential issues of maintaining consistency both across and within participants [69], as such challenges could be further exacerbated by delays in communication.

In addition to comprehending users' mental models, gesture design should consider technological constraints. For example, technological issues such as accuracy [70] and input resolution [9] are observed with mid-air interactions. The direction of translation, hand posture, and translation distance are all subject to change [71]. Mid-air hand interaction with spatial representations requires a greater degree of freedom of manipulation, which can be effectively carried out with one's hands. Still, it is subject to certain ergonomic restrictions [59], [70]. To overcome the lack of uniformity and low levels of directness, elicited gestures must be tested with users in various environments, such as varying lighting or screen sizes on personal computers. A well-executed *user study* can aid designers in comprehending the evolution of technique improvements over time and strike a balance between elicited and pre-designed gestures. In addition, it is necessary to evaluate the ergonomics of designed gestures and understand the trade-off between the intuitiveness and accessibility of the users' suggestions [72]. In our paper, we conducted an elicitation study, which was subsequently complemented by user and expert studies, in order to achieve an optimal integration of design and implementation.

Gesture Classifications. A large number of gesture taxonomies exist, such as Karam and schraefel [73], Wobbrock et al. [64], or Hoffmann et al. [74]. More recently, Stachl [44] suggested a mid-air gesture taxonomy which adds *static*

pose to the previously defined taxonomies. To define the nature of gestures, Aigner et al. [63] analyzed and grouped hand gestures into different **gesture types**. Their classification combines three gesture taxonomies and focuses on the functional level of gestures, regardless of context. They demonstrate a preference in performing specific physical actions in respect to users' intentions. For example, an intention to select an object is conveyed through pantomimic gestures, such as grabbing, grasping, or tapping the air. Classifying physical actions and gesture types rather than attempting to create a dictionary of gestures for a given set of functions is another trade-off between elicited and designed gestures. In addition, previous research by Pham et al. [57] on augmented reality gestures, Seyed et al. [75] on gestures for multi-display environments, and Chan et al. [76] on single hand microgestures explore **gesture themes**. Participants select comparable gestures for fundamentally related tasks, such as various hand expressions of swiping (using one finger, two fingers, or the entire hand) to select a value within a continuous range. Grouping gestures by types and themes focuses on classifying expressions based on physical actions, that refer to similar mental models and facilitate a more accurate interpretation of the elicited suggestion. We identify and extract gesture types and themes from our elicitation study and include them in the design space dimensions.

Summary. Although the importance of mid-air hand gesture interaction in data visualization is acknowledged, little progress has been made in comprehending its design and applicability for a set of visualization- and data-centric operations. To fill this gap, in our work we conduct an elicitation study to identify and categorize preferred actions for each gesture effect on a data visualization. To explore the design space of mid-air gestures in data visualization, we build upon prior work to classify physical actions into gesture types [63]. Additionally, using the concept of gesture themes, we conceptualize and interpret gestures for more complex tasks, such as actions requiring parameterization. Based on the findings of our elicitation study, we implement a set of gestures and evaluate them in subsequent user and expert studies.

III. ELICITATION STUDY (ES)

This section describes our Elicitation Study (ES) design, implementation, and findings. In our ES procedure, we first show the effect of an interaction on a data visualization (known as a *referent*) to each participant and, next, ask them to suggest a gesture intended to produce that effect (examples of ES referents are presented with Figure 2). For instance, we ask participants to propose a gesture for selecting individual bars on a bar chart. To control for legacy bias in the ES while providing ample space for users' expression of opinion, we implemented a combination of *production* and *priming* techniques [65]. We were also curious to find out whether or not people were able to draw on previous experiences to learn or discover mid-air gestures [77]. For example, we asked users to compare their suggestions with already established mid-air techniques,

TABLE I

WE MAP HIGH-LEVEL INTERACTION GOALS [78] (INTENTIONS), TO *how* TASKS ARE EXECUTED IN TERMS OF METHODS [79]. THEN, WE IDENTIFY LOW-LEVEL OPERATIONS THAT CORRESPOND TO THESE METHODS. FOR EACH OPERATION, WE CREATE ANIMATIONS OF INTERACTIONS WITH DATA VISUALIZATIONS, REFERRED TO AS REFERENTS, TO PRESENT TO THE ES PARTICIPANTS.

Intention	Method	Operation	Referents
Elaborate	Annotate	<u>Hover</u>	
Select	Select	<u>Select</u>	
Abstract	Navigate	<u>Zoom</u>	
Explore	Navigate	<u>Pan</u> (drag view)	
Explore	Filter (view)	<u>Slide</u>	
Explore	Arrange	<u>Sort</u>	
Reconfigure	Aggregate	<u>Reconfigure</u>	
Reconfigure	Navigate	<u>Rotate</u>	
Encode	Change	<u>Change represent.</u>	
Connect	Select	<u>Lasso</u>	
Connect	Filter (subset)	<u>Filter category</u>	
Filter	Filter (value)	<u>Filter data</u>	

such as comparing the suggested gesture with the "pinch" gesture widely used in immersive spaces (*production*). We further asked participants to keep their hands above the keyboard and focus on showing their hands in front of the camera, which is the spatial range where the Leap Motion can sense the gestures (*priming*).

We ran all our studies (ES, SS, XS) remotely via Zoom¹ with participants' laptops. There are multiple reasons for this strategy. First, *time constraints and user volume*—it would be hard to have 52 participants complete the study in a single lab given public health and social distancing guidelines; Second, *subjects' attention*—we aim to restrict attention to the novelty of the user interface and to eliminate the novelty of surroundings, technology settings or visualizations; And finally, *reducing hand fatigue*—the comfort of one's home or office invites more natural postures (e.g., elbow resting on a table).

A. ES Design

There are many possible interaction operations on data visualizations. In our ES, we started with high-level user intentions based on Yi et al. [78] (such as marking all data points with the same color) and mapped those user intentions to low-level methods and operations as defined by Brehmer and Munzner [79] (e.g., categorical filtering).

Intention-to-Operation. There are numerous taxonomies in data visualization; Dimara and Stasko's [80] work on decision making provides a comprehensive review of visualization task classification systems. We start with users' data-related intentions based on Yi et al.'s [78] high-level categorization. For instance, if a user is interested in "exploring" a set of data points, we can look for hand suggestions they intend to employ. However, their categorization does not establish a clear relationship between intentions, low-level operations, and visualizations. Brehmer & Munzner's

¹<https://zoom.us/>.

[79] typology bridges the gap between the low-level and high-level tasks focusing on the previous work (including Yi et al.’s classification). To relate intents and interaction techniques, we utilized the *how* portion of the typology (methods that focus on modifying or altering existing data points) that focuses on low-level operations from each group of methods. Table I shows the final *Intention-to-Operation* map.

Referents. The selection of visual encodings was the first step in determining the referents used in the ES. Each referent is a data visualization with an animation that demonstrates a desired effect, such as selecting a data point. Due to the lack of a clear taxonomy for visual representations, we focused on different visual encodings as defined by Munzner [81], and selected a subset of encodings which are most familiar to a general audience. In addition, we analyzed the choices of representations employed in related studies on interaction design with data visualizations (for example, Saket et al. [82], [83], or Kondo and Collins [84]). The final set of referents explore the same operations across different visual representations (e.g., scatterplot and barchart for select operations), as shown in Table I. We wanted to investigate whether different visualizations would elicit different gestures for the same operation (also noted by Willett et al. [85]). We designed referents without graphical widgets for user interface components to motivate subjects to focus on *embedded* interactions [86], and address the lack of discoverability in contemporary user interfaces [87]. The referents used in the ES and their corresponding questions are listed in Table III.

B. ES Participants and Procedure

The ES included 12 participants. All participants were between 29–60 years old; 8 identified as female; 4 as male; all participants had at least a college degree, 9 had Ph.D. degrees; all were residing in the United States at the time of the study; and none of the participants were HCI or data visualization experts. We recruited participants from the research team’s professional network via email.

We first introduced participants to a series of referents (i.e., visual encodings with an animated operation). Next, we prompted participants to suggest and perform a mid-air gesture to carry out the effect shown. Table I lists all the referents, whereas Fig. 2 illustrates some of the animations in action and the suggestions of a single participant. We presented the referents in increasing difficulty to the participants to encourage them to develop their own mental models and to not overwhelm users. Referents range from simple actions with common visualizations to parameterized actions and transformations not typically encountered by the general public. Therefore, all participants completed the tasks in the same order. We asked participants to think aloud and show each gesture in view of their camera and recorded all Zoom sessions for further review. At the conclusion of the sessions, our team shared a post-study survey with participants via email. Participants suggested unimanual and bimanual gestures, and we observed different types of biases.

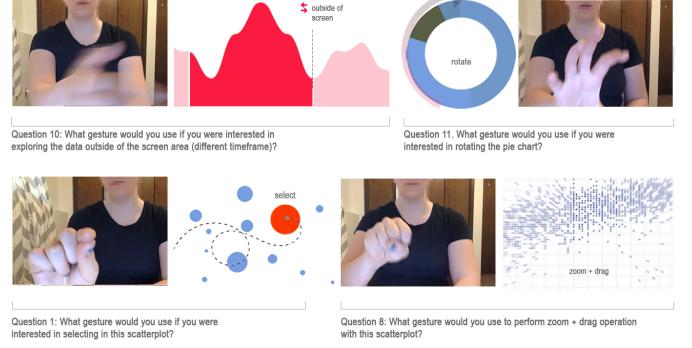


Fig. 2. Examples of elicitation study questions, referents, and suggestions from a single participant. A complete list of questions can be found in the [supplementary materials, Table 1](#).

C. ES Measures

We use the measure Agreement Rate— $AR(r)$ —to calculate gesture agreements across 12 participants. Initially introduced by Wobbrock et al. [25] and later redefined by Vatavu and Wobbrock [88], $AR(r)$ is a quantitative measure that assesses the congruence between the proposed gestures of several participants. The following formula defines AR :

$$AR(r) = \frac{\sum_i \sum_{i \neq j} [\delta(p_i, p_j) \leq \epsilon]}{N(N-1)}$$

where r represents the referent, N is the number of participants, p_i and p_j are the proposals of participants ($1 \leq i, j \leq N$), and $[\delta(p_i, p_j) \leq \epsilon]$ represents Kronecker’s function that evaluates to 1 when the inner expression is true and to 0 when false. $AR(r)$ values range between 0 and 1, from total disagreement to absolute agreement. We used the tolerance $\epsilon=0$ following the previous research on in-air and on-body hand gestures [89]. To indicate that two gestures are equivalent, handedness, number of fingers, and hand translations must be identical. Vatavu and Wobbrock [90] delivered a toolkit to assist practitioners to compute agreement rates.

In addition to agreement rates, we categorize gestures by type (Table II) based on the Aigner et al. [63] classification. The classification of gestures based on their type facilitates the categorization of gestures into themes for individual or multiple operations.

Three researchers independently analyzed the videos, compared notes, and used a majority vote to classify the gestures into distinct categories. There were no disagreements among researchers. According to Vatavu and Wobbrock [90], and later used by Hosseini et al. [11], the agreement rate level is to be interpreted as low agreement ($AR(r)<0.1$), medium agreement ($0.1<AR(r)<0.3$), high agreement ($0.3<AR(r)<0.5$), and very high agreement ($AR(r)>0.5$).

IV. ELICITATION STUDY FINDINGS

As participants frequently described gestures aloud, there were no disagreements regarding the classification of a gesture’s effect, type, or theme within our research team.

TABLE II

THE CLASSIFICATION OF GESTURE TYPES BY AIGNER ET AL. [63] USED IN OUR ES. TO IDENTIFY PATTERNS OF GESTURE TYPES FOR EACH REFERENT, GESTURE SUGGESTIONS IN THE ES ARE CLASSIFIED INTO THE FOLLOWING GROUPS.

Gesture type	Description
<i>Pointing</i>	Indicate objects and directions
<i>Pantomimic</i>	Imitation, multiple low-level gestures
<i>Semaphoric Static</i>	Static hand postures (e.g., “stop” sign)
<i>Semaphoric Dynamic</i>	Repeatedly flicking or waving
<i>Semaphoric Strokes</i>	Single, stroke-like movements
<i>Iconic Static</i>	Spontaneous static hand postures
<i>Iconic Dynamic</i>	Motions describing paths or shapes
<i>Manipulation</i>	Motions with subsequent reactions

Table III shows the calculated agreement rates and Fig 3 shows details on the gesture types participants used.

Close to fifty percent of displayed referents have a high degree of agreement ($AR(r) \geq 0.5$). For example, participants proposed similar “index [finger] up” movements to hover over data points. We found that for the same gesture effect, the translation direction of the proposed gestures varied, and we observed that participants’ hands often followed the shape of visual representations (↗, ↘). When participants were asked to propose a gesture to reconfigure a bubble chart ↗ into clusters ↗, nine participants performed a bimanual gesture that mimicked the animated action they observed in the visualization. When participants were asked to propose gestures that filter by category in a scatterplot ↗, we observed a heavy influence of the visualization design on the elicited gestures. We observed, that even for the same operation, the recommended **gestures vary across different visualization types**. For example, when data points appeared layered on top of one another (z -axis), participants moved their hands forward and backward to filter data points. When working with the pie chart ↙, participants’ gestures mimicked that of turning a door knob or turning a volume dial, sometimes rotating the whole hand from the wrist and other times using small finger strokes to rotate. Interestingly, gestures between pie and donut chart ↙ differed, indicating that participants were influenced by the visual encoding as well as the white space in a visualization.

Moderate agreement rates ($0.5 < AR(r) < 0.3$) display a wide variety of suggestions for the same referent. However, **the different gestures elicited for the same referent still indicate the same gesture type**. For example, participants suggested “tap” with index finger only, “tap” with the whole hand, or “pinch” (index and thumb) movements when asked what gesture they would use to select a data point on a scatterplot ↗, or a bar chart ↘; however, all suggestions fall into the same gesture type (pantomimic). To show changes over time in a scatterplot ↗, the visualization displayed data points moving left and right (linear), and ten participants proposed a gesture that followed the animation but using different hand symbols. Participants either used a single index finger, two fingers, or their whole hand to move left to right, indicating different hand postures, but the same gesture type.

When prompted to explore changes over time in a bar chart ↘, participants compared this to scrolling with

TABLE III

AGREEMENT RATES $AR(r)$ FOR TWELVE PARTICIPANTS ($N=12$), WHERE P_i IS THE NUMBER OF DIFFERENT PROPOSALS FOR A REFERENT, AND P IS THE COUNT OF THE MOST COMMON IDENTICAL GESTURE PROPOSAL. THE ORDER OF THE REFERENTS IS BASED ON THE QUESTION (Q) ASKED DURING THE STUDY. THE COLOR CODE DISPLAYS THE FOLLOWING AGREEMENT RATES: VERY HIGH ■■■, HIGH ■■, MEDIUM ■, LOW ■■■■■.

Referents(r)	P_i	P	$AR(r)$
Q1 Hover with scatterplot ↗	3	8	0.5
	3	6	0.3
Q2 Hover with bar chart ↘	2	9	0.6
	3	6	0.3
Q3 Lasso with scatterplot ↗	4	7	0.4
Q4 Reconfiguration ↗, ↘, ↙, ↛, ↜	4	9	0.6
Q5.1 Filter time with scatterplot ↗	3	10	0.7
Q5.2 Filter categories with scatterplot ↗	2	8	0.4
Q6.1 Filter time with barchart ↘, ↘, ↘	4	6	0.3
Q6.2 Filter categories with barchart ↘, ↘	3	8	0.4
Q7 Zoom into clusters ↗	6	4	0.1
Q8 Zoom&drag with scatterplot ↗	4	6	0.3
Q9 Zoom&drag with bubble chart ↗, ↙	4	5	0.2
Q10 Slide with area chart ↘	2	7	0.5
Q11. Rotate with pie chart ↙	2	9	0.6
Q12. Rotate with donut chart ↙	2	7	0.5
Q13. Sort with vertical barchart ↘	7	3	0.1
Q14. Sort with horizontal barchart ↘	4	5	0.2
Q15. Animated transition ↘, ↘, ↙	9	3	0.1

a mouse wheel or scrollbar widget, **demonstrating legacy bias**. When prompted to explore zoom&drag² in a scatterplot ↗, participants followed “grab&drag” movements, but also “pinch-zoom” used with touch screens. This was also presented with referents that scored low agreement rates (Table III, $AR(r) < 0.3$). Despite a great deal of variation in their recommendations, participants suggested the gestures performed with other modalities. This includes filtering by category in a bar chart ↘/↖ as participants fixated on interacting with the legend rather than on the *embedded* interaction with the visualization. Moderators prompted participants to consider the legend as static. Still, four participants could not imagine a gesture for filtering in the bar chart when the option of the legend was not available to them. Referents like sort with horizontal ↘ and vertical ↘ bar chart as well as animated transition ↘, ↙ (Table III, Q12, Q13, and Q14), suggested different variations of drawing and swiping in the air, similar to the gestures demonstrated on touch screens.

The post-study survey at the end of the ES focused on participants’ subjective experiences with the novel interaction modality. Participants reported that they could see themselves interacting using bare hands with various digital resources, **engaging with data visualizations on home systems**, especially home management systems, as well as entertainment and communication. Opinions regarding mid-air gesture interaction, in general, were somewhat divided between extreme enthusiasm and extreme skepticism, with many participants remarking that it felt artificial and that they would need to become accustomed to the system and its features first.

²Instead of zoom&pan, we are using the zoom&drag operation which is inspired by the actual “grab&drag” gesture [56].

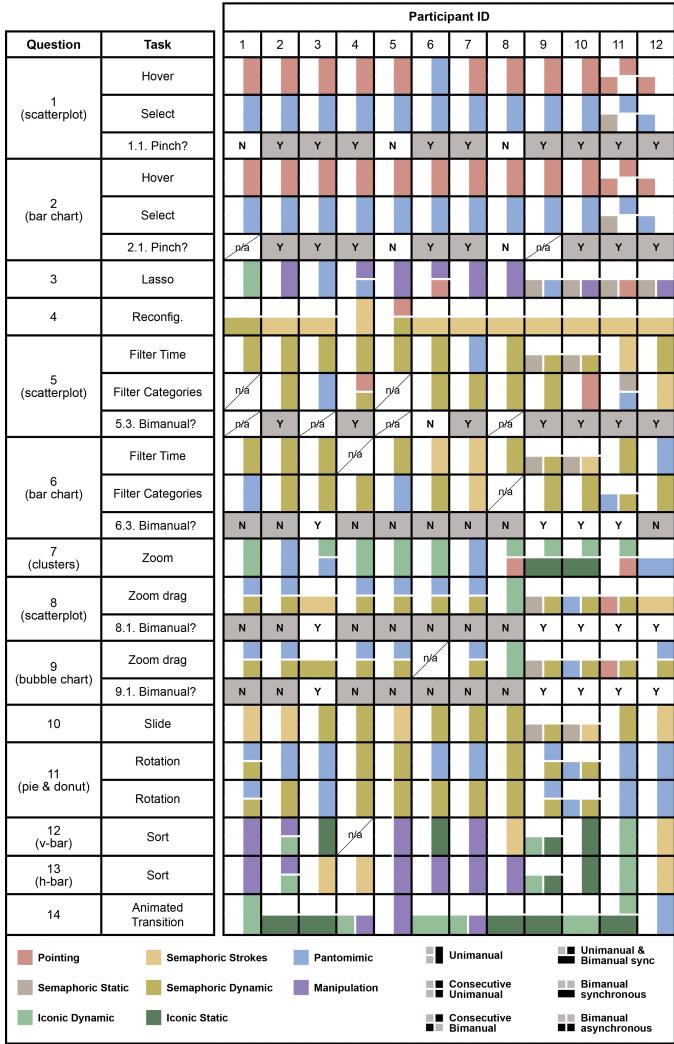


Fig. 3. Intention-effect map. The color represents the gesture type and different patterns show handedness. For more information on gesture suggestions, see [supplementary materials, Figure 4](#)

A. ES Discussion

Legacy bias. Regardless of the gesture type (e.g., pointing gestures for hover operation), gesture theme (e.g., swiping gestures across multiple operations), or agreement rate $AR(r)$, elicited gestures in our study were heavily influenced by the participant's bias. For example, when prompted to suggest a gesture for referents that show hover and select operations, participants proposed gestures that mimicked the use of a handheld mouse or touch screens; however, participants were also able to adapt to suggestions (e.g., using the “pinch” gesture for selection). Participants were sometimes distracted by design choices, like the presence of a legend within the various visualizations, and often defaulted to interacting with the legend to manipulate the visualization. Additionally, participants suggested the full-hand pinch gesture for zoom and a swiping motion for any scrolling gesture, which mimics interactions with touch screens. Bias was also present due to natural physical or cultural and social gestures, such as putting one hand up parallel to the screen to signal “stop.” Certain referents (e.g.,

reconfigure

Awareness of scale and white space. Participants displayed an awareness of the use of space and white space in the visualizations. Participants often changed gestures for the same effect depending on the size of data points, the amount of white space in the visualization, and the level of specificity required by the prompt. This spatial awareness translated to an awareness of the three-dimensional interaction space and bimanual parallel inputs.

Interface limitations. Similarly, participants were influenced by the simulated sequence of the animated visualizations; one participant suggested that since the visualization already had an inherent logic that displayed an action in a certain way, this limited both their imagination to suggest gestures and the functionality of the visualization itself. During the ES, participants demonstrated an understanding that systems are programmed and therefore limited to only responding to or performing certain functions or actions.

Mental models. Participants were often concerned with the application itself, what the system would allow them to do, and how it would react and respond. Across all participants, moderators noticed a desire for more context and imposed limitations. Moderators observed that participants often developed their own gesture vocabulary as they moved through the visualizations and task sets. Participants were influenced by their own suggestions from an earlier visualization, either to re-use a gesture to carry out a similar function or not to use a gesture because it had already been exhausted. Participants discussed the issue of disambiguation, desiring a defined set of gestures to control the system and also that gestures are distinct from each other. This was evident in observing participants' self-designed gesture set as they progressed through the study.

B. ES Limitations

For a complete comprehension of the ES and its results, we must also discuss its limitations. Overall, participants in the ES were not advised whether or not mid-air hand gestures should be consistent across all required interaction scenarios. This, however, was omitted on purpose because we wanted to determine whether or not users would be inspired to suggest a single gesture for multiple operations or if they have some sort of comprehension of the need for disambiguation. The ES has limitations in the following aspects:

Recruitment. The recruited subjects were professors and graduate students in academia. Despite the fact that none of them have prior experience with HCI or VIS, it could be beneficial to extend the ES beyond an academic audience.

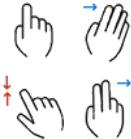
Visualization and cognition. We designed visualizations together with a simulated action or animation to elicit gestures. However, participants' creativity might have been limited by the animation shown to them. In order to clearly present operations in the elicitation study, some of the indicators were rendered together with the effect

(e.g., highlight data to show hover and render a cursor to represent the center of a manipulation by hand gestures). These indicators and feedback have strong associations with traditional WIMP interfaces and thus might explain the high consensus on elicited gestures for those visualizations.

Social aspects. Besides technology bias, elicited gestures might be subjected to cultural biases as well. All participants of our elicitation study were residing in the United States. Thus, elicited gestures may have been tied to social or cultural norms (e.g., one hand up for “stop”); we did not request demographic information regarding cultural associations, heritage, race, or ethnicity from participants. Because of this limitation, the cultural influence of persons who have lived outside of the United States (e.g., arms crossed in “X” pattern for “stop” in Japan) was not considered or captured by our elicitation study.

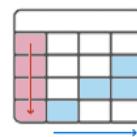
V. DESIGN SPACE EXPLORATION

Based on the results of the elicitation study, we aim to define a **gesture vocabulary** for our user study by exploring the design space of mid-air gestures. In particular, we have identified the following dimensions: *gesture type*, *gesture theme*, *space-multiplexing*, and *visual representation*. Note, that these dimensions can overlap and are not meant to be exhaustive. Rather, we aim to provide gesture designers with a common vocabulary to balance different aspects in their gesture designs according to the requirements of applications.

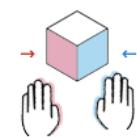


Gesture Types. The classification of gestures into gesture types [63] supports the identification of underlying user intentions and tasks. Not only can we draw strong correlations between gesture types and the agreement rates of elicited gestures (for more details, see [supplementary materials, Figure 4](#)), but we can also search for instances in which users perform gestures of the same type while exhibiting posture differences for the same operation. One such instance is reconfigure, in which users imitate the separation of bubbles into clusters and demonstrate strong physical associations to the performed action despite variations in translation direction. Regardless of the $AR(r)$ in all tasks, gestures typically fall into the same type category; this categorization is independent of visual representation and serves as a first-level design guideline for gestures suitable for the intended operations. For example, *semaphoric strokes* are hand flicks characterized by single stroke-like movements; their nature, which permits repetitions, suggests interaction with a broader space. Strokes are observed when users intend to bring the data outside of the screen, such as interacting with area charts. On the other hand, *semaphoric dynamic* gestures operate in a more constrained space (e.g., slide). In contrast to dynamic gestures, the actual range of movement in strokes does not convey information regarding the action. *Pantomimic* gestures are used to demonstrate a particular task, such as selecting a single data point, rotating charts, or zooming, regardless of data size. *Manipulation* gestures are used for operations where the effect “follows” the gesture,

such as sorting operations. Nevertheless, when the hand gesture is aligned with the effect (e.g., animated transition), participants prefer dynamic gestures that are typically used to describe the shape or the transition. *Static* gestures, both *semaphoric* or *iconic*, are observed only when combined with other dynamic gestures across different referents, and are used as a way to manifest the action “trigger” (e.g., lasso). Participants understood that there must be an initial action (similar to a long click on the mouse to initiate “drawing” in the air) and suggested several ways of the same type to imitate this action.



Gesture Themes. By mapping gestures to a user’s conceptual model, gesture themes assist in identifying commonalities between gestures on a higher, conceptual level. We discovered two types of gesture themes, *vertical* and *horizontal*. Vertical gesture themes are concepts that present similar gesture suggestions for each referent, and they can be useful for detecting patterns when agreement rates are low. For example, when we ask participants to suggest a gesture for the referent that demonstrates changes over time (filter data), a variety of hand postures are evoked, including “index [finger] up”, “pinch”, and “whole hand”; however, each mimics the use of a slider user interface component with moving the hand “left-right” to indicate the desired change. The horizontal concept is exemplified by a consistent behavior across multiple referents. It is useful for identifying identical hand postures and movements across various operations. For example, users “swipe” with their hands when they wish to filter time, filter view, or rearrange data. Movements such as “swipe” are spatially unrestricted and comprise a series of strokes to bring the data outside the screen. In addition, participants have the same general idea (concept or theme) regardless of the visual representation, which can aid in identifying gesture design patterns applicable to various scenarios. Examples are illustrated in the [supplementary materials, Table 4](#).



Space-multiplexing. With space-multiplexed input, each hand gesture is assigned its own space and channel. Due to legacy bias, the initial suggestions of ES participants contained a few original ideas for parallel inputs (e.g., inputs for multiple operations). Nonetheless, we observed the use of two consecutive gestures in a number of instances, as well as bimanual interactions, indicating that the interaction space of gestures can be expanded to accommodate the simultaneous execution of multiple operations. In the zoom&pan task, for instance, participants opened their hands to perform zoom and then made a second gesture to pan. Accessing the “layers” of data in categorical filtering with scatterplots, involved consecutive spatial movements, which were sometimes performed with both hands accessing multiple categories at the same time. Parallel manipulation reduces the number of distinct postures for a given set of tasks, and the possibility of expanding the interaction space to three dimensions enables hybrid gesture design.



Visual representation. The results of the ES indicate that the visual encoding of a referent has a significant impact on the gestures participants chose. As participants progressed through the tasks, users typically re-used some gestures to perform comparable functions across several tasks. Nevertheless, when the referent visualization was altered, the user responded and modified their gesture. In addition, participants concentrated on manipulating only “visible” areas of visualizations, and once the visual representation changed, they thought their previous gesture was no longer applicable. For rotation tasks, which occurred later in the ES, the visualization itself (donut chart) influenced the participant’s gestures as they had the sense that they could not manipulate the white space within the visualization itself. In addition to exploring gesture types and themes, physical associations impact gesture suggestions. This idea is consistent with transitions embedded in visual representation. In fact, the visual representation triggers these associations, and we can translate these scenarios into appropriate designs. For example, as a method of manipulating orientation, rotation gestures could be applied to any representation provided that it was properly associated with it (e.g., rotate to change orientation, to change direction, or to swap axes). The connection between the physical association and the user’s mental model will effectively align with the user’s data-related intention.

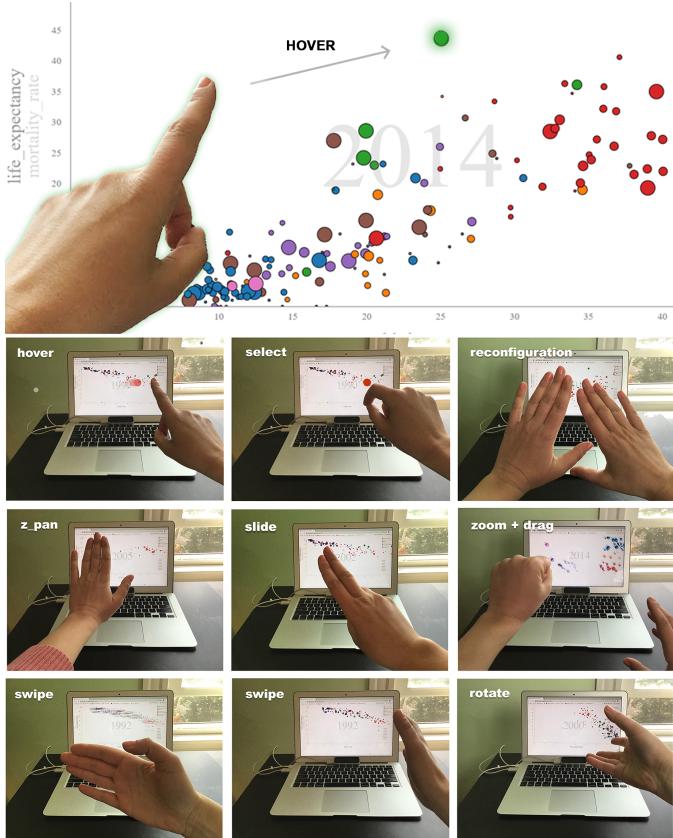


Fig. 4. The screenshot of the web application running the *Gapminder* visualization with embedded mid-air hand gesture interaction. The presented setup also depicts the equipment used for the user study.

VI. VISUALIZATION AND GESTURE VOCABULARY DESIGN

Based on the ES, we implemented different mid-air gestures and evaluated them in a user and expert study. To place emphasis on the interaction modality, we chose a visualization and dataset that are well-known to a general audience.

Therefore, we used a United Nations dataset, which is also used by the popular *Gapminder Trendalyzer*³, to populate the visualization. The visualization (referred to as the *Gapminder* visualization throughout the article) encodes four variables of the data in a scatterplot (see Fig. 4, top). Each data point represents a country, the *x*-axis displays *child birth rate*, the *y*-axis toggles between *life expectancy* and *child mortality rate*, and the size of the data point shows the country’s *population size*. Furthermore, countries are grouped by region, which is represented by color. In addition, users can explore changes over time, filter data by category, and swap axes. The visualization design accounts for embedded interaction without user interface components.

Next, we discuss the vocabulary design for our subsequent studies. We designed gestures with several considerations in mind: *AR(r)*, gesture types and themes, the ability to transition from time-multiplexed to space-multiplexed inputs, and the effect of visual representation. We implemented the following eight gestures: HOVER, SELECT, Z-PAN, SWIPE, SLIDE, ROTATE, RECONFIGURE, ZOOM&DRAG (Figure 4). The majority of gestures (5 out of 8) are designed as elicited, with respect to gesture types and themes. In consideration of technological limitations and disambiguation, we redesigned ZOOM&DRAG (combining two gestures into one), SLIDE (using two fingers instead of one), and Z-PAN (using the whole hand instead of one finger) to improve the memorability and learnability of the vocabulary, as well as to comprehend the capability for diverse inputs. Utilizing the 3D interaction space maximizes the concept of hand translation direction (e.g., left-right for SLIDE; forward-backward for Z-PAN). To link gestures to physical operations, we looked for the user’s intentions expressed in the ES (e.g., for SLIDE, the perception of time is linear, so we use the *y*-axis; for Z-PAN, the *z*-axis is used to access the *layers* of the data participants often perceived with multi-category scatterplots). To avoid confusion, the HOVER gesture employs one-finger manipulations, whereas the SLIDE gesture employs two fingers. ZOOM&DRAG enables parallel inputs and leaves room for additional operations to be executed. ROTATE is designed as elicited (turning the door nob movement), but it is applied to a new scenario (swap the axes); thus, it adheres to the principle of linking the physical association and the user’s intent.

VII. USER AND EXPERT STUDIES (SS & XS)

We conducted a user study (SS) to explore the intuitiveness of our designed gestures. In an expert study (XS), we evaluated the ergonomics of our gestures. We asked

³<https://www.gapminder.org/tools>.

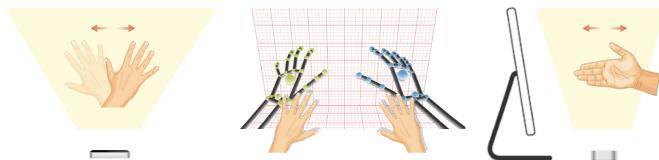


Fig. 5. Tutorial for spatial understanding: Participants practice finding the “sweet spot” where the controller can accurately track their hands.



Fig. 6. Tutorial tasks are designed with physical associations in mind and feature distinct visual cues compared to SS tasks.

participants to analyze hand postures and movements and collected feedback on how to reduce hand fatigue.

A. SS Design and Implementation

Apparatus. We provided Leap Motion Controllers and setup instructions to all participants for the duration of the study (devices were mailed to each participant’s home.) The setup included a laptop (iMac, Macbook, etc.) and the Leap Motion Controller, connected to the laptop via a USB cable. We developed a web application using *JavaScript* language and Leap Motion SDK⁴. For detecting gestures, we utilized *Frame Objects* and associated attributes to understand the positions of hands and fingers. Visualizations are implemented in D3⁵. We ran our study as a web application optimized for a screen resolution of $2,560 \times 1,600$ pixels. The controller’s position had to be individually adjusted by each participant. The laptop screen sizes varied among participants.

SS Participants. We conducted the SS with 20 participants. All participants were between the ages of 18–49 years; 16 identified as female; 3 as male; 1 identified as non-binary; all participants have at least a college degree; all resided in the United States at the time of the study; and none of the participants were domain experts. We recruited participants from members of the research team’s university via email.

Tutorial. Participants first completed a tutorial where we introduced them to the gesture vocabulary and how to position their hands in relation to the sensor (Fig. 5). The tutorial further introduced simple visual cues (Fig. 6). It encouraged users to feel and learn gestures and make logical connections between a gesture’s action and its effect, without introducing any SS-specific tasks. After completion of the tutorial, we tested participants on their ability to perform a different variant of a task from the SS. After successful completion, participants moved on to the SS tasks.

⁴<https://developer.leapmotion.com/>.

⁵<https://d3js.org/>.

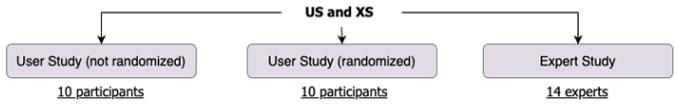


Fig. 7. Participants and experts in the SS and XS studies.

SS Procedure. To increase the validity of our study, we randomized the tasks with some users and kept the same order with another group (10 users in each group as shown in Fig. 7). This helped determine whether the order of the tasks in the tutorial had an effect on the recall of the gestures. (We found no difference in recall between the two groups.) We recorded all SS sessions (including the tutorial and task sets) in Zoom and asked participants to think aloud. We further asked participants to show their hands for each gesture in view of their camera for later review. After completion of the tutorial, we asked participants to complete nine tasks in the user study. For each task, we showed participants the *Gapminder* visualization and asked them to perform a certain user interaction (e.g., filtering, panning) using the mid-air gesture vocabulary introduced in the tutorial. In addition, we asked participants to execute a number of bimanual interactions with parallel inputs. The complete list of tasks is presented with Figure 8. After completing the task set in the SS, we sent a link to a post-study survey to all participants. We asked participants about their experience interacting with the data visualization via gestures (including both Likert scale and open-ended questions) and some demographic questions.

Pilot study. Before running the user study, we conducted a pilot study with six participants who helped us finalize the moderator script and tutorial, and troubleshoot the application. We redesigned the system to improve input reliability and required remote pilot participants to complete the task in their own environment. No changes were made to gestures following the pilot study, and no gender bias was observed among participants. However, pilot participants requested some way to know whether their hands were within the sensor’s range or “sweet spot”, and we tested the system with *virtual hands*. As we progressed with pilot users, we found that participants could easily overcome initial uncertainty regarding new technology and that *virtual hands* occluded visualizations and negatively impacted recall (the participants focused solely on the virtual appearance of their hands on the screen). Therefore, neither the tutorial nor the subsequent studies utilized *virtual hands*. In addition, the pilot study aided moderators in comprehending the need to keep their hands away from the screen and to express the tasks without gesticulation.

B. XS Design and Implementation

XS participants. We conducted our expert study with 14 experts to evaluate the ergonomics of the eight gestures used in the SS. Two co-authors reached out to their domain expert colleagues, including researchers and PhD students. All XS participants are HCI domain experts (7 postdoctoral fellows, 7 faculty), and eight self-identified as having experience with hand gestures. We solicited participants via

Gesture	Participant ID Recall (Y/N)																				Questions
	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	
Hover/ Select	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Q1. Select a data point for a country from Sub-Saharan Africa (red). (this question assumes users will perform hover, before making a selection).
Slide	Y	Y	Y	Y	N	Y	N	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Q2. With that country selected, now we want to explore data in another decade, the 2000s. What gesture would you use to explore this data?
Swipe	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Q3. Now we are interested in changing the category along the y-axis from life_expectancy to mortality_rate. What gesture would you use to make this change?
Slide	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Q4. With this new data set on the y-axis, navigate to a year in the 2000s. Would it be possible to alter the data along the y-axis and simultaneously navigate through multiple periods?
Z-Pan	Y	N	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Q5. What region has the lowest birth rate in the year you've selected? (this is the first question regarding "z-pan" gesture)
Reconfigure	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Q6. We want to reorganize data into groups by region. What gesture would you use to reorganize the data in the visualization? What do you see?
Zoom&Drag	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Q7. In this view, some data points are outside of the visualization. We want to explore the data specifically in Europe and Central Asia. What gesture would you use to see all the data for that region (blue)?
Rotate	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Q8. In this visualization, we are able to identify certain trends based on the data that you see here. We want to see if there are other trends by changing the orientation of the visualization. What gesture would you use to swap the axes to do this?
Z-Pan	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Q9. We see data for Europe and Central Asia, but we want to see data for Sub-Saharan Africa. What gesture would you use to filter the data by that region? Can we filter the data by region while changing the visualization's orientation?

Fig. 8. The efficacy rates of gesture recall in the SS. Additional questions regarding bimanual manipulation are **highlighted**.

private Slack workspaces dedicated to an HCI group of scholars (from the Georgia Institute of Technology, Harvard University, and UC Berkeley).

XS procedure. The Expert Study consisted of a questionnaire comprised of eight sections, one for each of the eight gestures. We first asked the experts about their general familiarity with gesture and interaction design. Each of the eight sections included a brief video of a user performing gestures on our web application and additional reference images illustrating hand positions from various angles (see Fig. 9). We asked experts to rate the following: discomfort level of hand position relative to their body, muscle strain level resulting from the hand position, and wrist movement. The remaining questions were open-ended and encouraged participant commentary. We list detailed survey question in the [supplementary materials, Table 8 and Table 9](#).

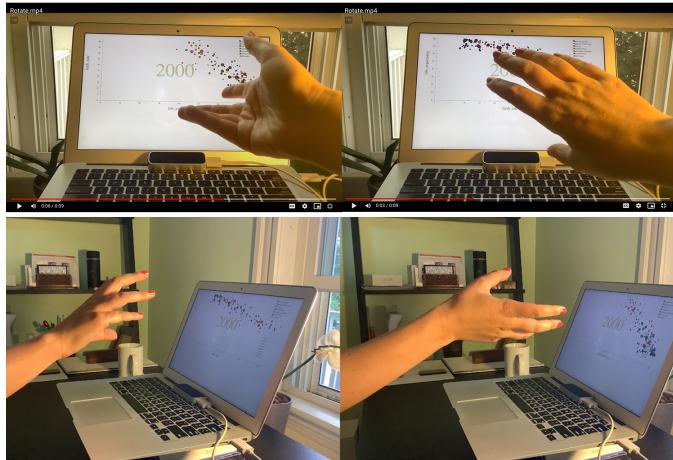


Fig. 9. In addition to videos, we provided experts with images depicting the same hand posture from different angles.

C. SS & XS Measures

We evaluate the effectiveness of our designed gestures based on learnability, memorability, and ergonomics (see Table IV). We analyze *recall* of a novel task variation [91], *bimanual manipulation* comprehension, and *ergonomics* characteristics evaluated by subject-matter experts.

After the SS and XS studies, the research team watched the video recordings, calculated the participants' success rates for each task, and observed cognition. Three researchers watched the videos independently and used majority voting for reaching consensus. The classification of whether a participant correctly recalled a gesture was based on: (1) the user took longer than 3 seconds to recall, (2) the user answered with "I don't know" or "not sure", or (3) the user demonstrated a wrong gesture. All research team members agreed on the recall results. Additionally, the research team also took note of participants' demonstrating bimanual manipulation and their verbal comments.

TABLE IV
USER AND EXPERT STUDIES MEASURES.

Study	Dimensions	Measures
SS	Learnability Memorability	Mapping Recall
XS	Ergonomics Characteristics	Hand Tensions

VIII. USER AND EXPERT STUDIES FINDINGS

In this section, we report on the findings of SS and XS studies. Detailed results are shown in [supplementary materials \(Fig. 11, 12, and 13\)](#).

A. Learnability and Memorability

All SS participants completed the tasks in under 15 minutes. The recall success rates, as shown in Fig. 8, are depicted with "Y" if the participant recalled the gesture correctly on the first attempt and "N" if they did not. Almost all gestures have high recall rates, except for the Z-PAN gesture (5 out of 20 users were unable to recall the gesture). The randomization of the task set had no effect on hand gesture recall. Two tests show no statistical significance between the two groups: Independent Group t-test on overall recall ($p\text{-value}=0.7$), and MANOVA test run on the recall success without HOVER and SELECT gestures ($Pr(>F)=0.9$). On multiple occasions, regardless of whether the question was explicitly asked or not, bimanual manipulation comprehension was observed; when explicitly asked (e.g., Q7 and Q9, Fig. 8), all users were able to perform

parallel inputs immediately. It is also interesting to observe the results for each user individually; some users had an excellent memory for all gestures, whereas others struggled with the majority of the vocabulary, primarily due to the sensor's input range (e.g., p7 did not recall a gesture in 6 out of 9 tasks).

We observed various demonstrations of visual understanding with different gestures from the vocabulary. This was noted in consideration of both the intuitiveness and logic (metaphorical and/or iconical) of a given gesture. We observed **increased visual cognition when gestures were related to or mimicked interactions in familiar technology interfaces or the physical world** (e.g., ROTATE). Users commented on an increased understanding of such operations, especially the combination of multiple gestures. For example, parallel operations, such as categorical filtering and axis swapping, make it possible to compare extreme values across multiple groups simultaneously. The presence of embedded interactions in hand inputs allows the user's attention to remain undivided between the visual representation and interface components, thereby facilitating the observation of diverse patterns.

Participants felt that interacting with the system using mid-air gestures was easy, and that the gestures were easy to use. Participants commented on the intuitiveness of the application, saying that they “*thought the gestures were intuitive and did not have a problem using them*”. Other participants stated “*that it feels like a more organic way of interacting with the data*” despite the fact that none of the participants had used this application before. Another participant mentioned that gestures were “*intuitive and matched how you would move data if it were in front of you on a paper*”. Another said that they liked it better than using a mouse and keyboard because “*you can do a lot more with visualizations of data very quickly*”. Participants noted that they enjoyed using the mid-air gestures to navigate the data visualization and said would recommend the system to their friends. Based on the survey responses, moderator observations, and participant comments, all participants found interacting with the gestures and the data visualizations enjoyable and exhibited playfulness.

B. Ergonomics

Feedback from the XS study indicates that from an accessibility standpoint, all gestures are ergonomically sound (see Fig. 10). Gesture ergonomic ratings ranged from 1 (low/bad ergonomics) to 5 (high/good ergonomics): 11 (out of 14) experts rated the *posture* of all gestures at four or higher; 10 experts rated the *muscle strain* of all gestures at four or higher; 11 experts rated the *wrist movement* of all gestures at four or higher. One expert disclosed that they suffer from a chronic illness that can restrict hand use. Their responses and remarks reflect this. In addition to the overall results, participants identified the following characteristics that could improve inclusion and accessibility. Here is how they compare with the efficacy rates of gesture recall:

HOVER. The exceptional memorability of this gesture may conflict with minor ergonomic concerns. The gesture is more challenging for extended periods of time, when using an ergonomic desk setup, or if mobility in the wrist or elbow is limited. Muscle strain could also become an issue if an action requires precision.

SELECT. Despite its high learnability, this gesture can be improved ergonomically, as any prolonged hold would be uncomfortable. As a result, users may unintentionally move beyond the system's range, resulting in a loss of accuracy. One participant suggested that users with functional deficits would perhaps benefit from this gesture.

SLIDE. The gesture is very comfortable and natural, but the tendon strain in the wrist from holding up two fingers is higher than when holding up just one finger.

SWIPE. No recommendations for improvements were suggested as long as the system captures the motion of the hand (movement from wrist) or forearm (from the elbow). Users with carpal tunnel, however, would find this gesture uncomfortable.

ZOOM&DRAG. The complexity of the gesture impacts memorability and muscle strain. Users with functional deficits would perhaps benefit from the “pinch” gesture instead. If the gesture was combined with a ROTATE gesture, ergonomics could be improved for some users by the system perceiving the rotation of the forearm (rather than the rotation of the wrist) to interact with the visualization.

Z-PAN. This gesture's posture, muscle strain, and wrist movements are rated highly, despite the fact that this gesture has the lowest recall score in the vocabulary. Experts noted that users might experience elbow discomfort if they are too close to the remote display.

RECONFIGURE. This gesture was rated as easy to perform, learn, and remember; however, bimanual gestures are limited for users that are one-handed. An alternative gesture would need to be designed as an inclusive and accessible option.

ROTATE. The gesture is simple to learn and easy to remember, with a high recall rate, but it presents some ergonomic challenges, particularly regarding posture and wrist movement. Users with functional deficits might benefit from a “pinch” gesture instead of a “grab”. One participant commented that the gesture feels “very normal”.

C. SS & XS Discussion

Physical association and legacy bias increase memorability. We attribute great recall for HOVER and SELECT to legacy bias due to the existence of similar gestures for other interfaces (e.g., touch screens). Due to the physical gesture mimicking a real-world action, participants also had perfect recall when gestures were mapped to physical actions, such as ROTATE (e.g., turning a door knob). Multiple repetitions of the same gestures throughout the tasks may contribute to the ability to comprehend some bimanual and parallel manipulations, particularly the final one (Fig. 8, Q9 that focuses on using ROTATE and Z-PAN at the same time). Nevertheless, given that the questions represent a novel task

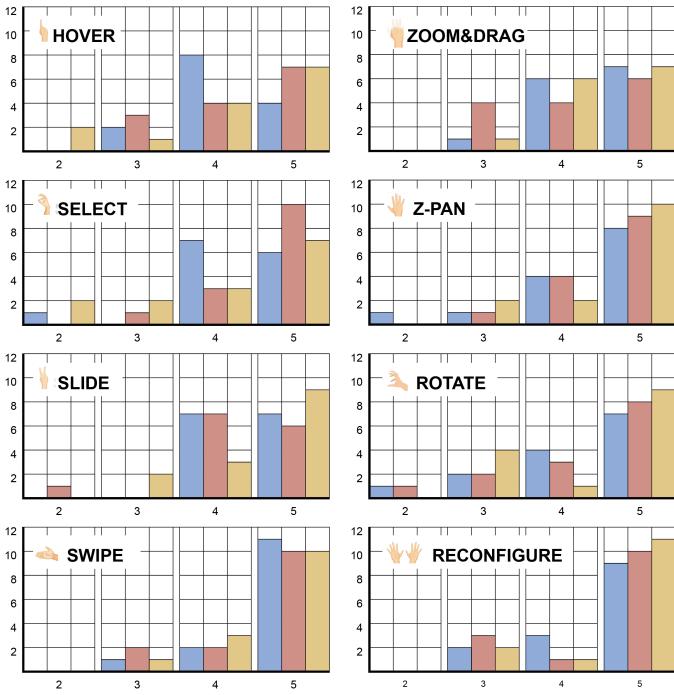


Fig. 10. The average ergonomics ratings in XS for the following attributes: Posture ■, Muscle Strain ■, and Wrist Movement ■.

variation, we are confident that the map between the intention and gesture provides a strong bond and, consequently, strong recall results. Visual cognition was enhanced when gestures related to or mimicked interactions with familiar technology interfaces or the physical world. For example, the SLIDE gesture allows participants to explore changes in data over time. In a traditional WIMP interface, this interaction is often performed using a slider widget and metaphorically maps to the classic interpretation of time being linear. Embedded interaction (directly manipulating the representation) enables the user to focus on the task. At the same time, the direction and speed of the movement can help reveal patterns more quickly as both gesture and data moving manifest the temporal change. This also applies to the RECONFIGURE as an easy switch from the default orientation to a map view. When we explore Z-PAN, the gesture is tied directly to the nature of the visual representation (overlapping data points). Participants observed the “layers” of data within the visualization while discovering patterns of specific regions. The ZOOM&DRAG gesture prompts users to “grab” data to move it into view. The operation and outcome mimic the physical act of grabbing an object and bringing it to a space in better view, enabling faster overview–detail comprehension. The ROTATE gesture allows users to change the default orientation of a given visualization. The operation itself (rotating the visualization instead of swapping the axes demonstrated with transition) resembled the physical action of the gesture, which further reinforced the connection between the operation and outcome.

Subjective feedback indicates high learnability and engagement. In a post-study survey, participants were neutral as to whether or not most individuals would be able to

learn this method of interaction quickly. However, when prompted to reflect on their own experiences learning the gesture vocabulary, participants suggested it was very easy to learn. Participants felt more engaged with the data, having a sense of controlling the visualization “physically”. One participant commented that understanding the 3D nature of the visualization was easier to understand because they were interacting with it through physical gestures. Participants exhibited excitement when interacting with the system and using the novel interaction techniques like Z-PAN, RECONFIGURE, and especially bimanual interaction in both the tutorial and task set. Participants commented on feeling “powerful” and some referenced the movie Minority Report. These comments echoed those of the elicitation study, where participants also connected their experience with mid-air gestures to modern cultural access points.

Technology constraints. We asked participants to reflect on what they disliked about using mid-air gestures to interact with the data visualization application. The problems identified are with the system, not the gestures, specifically the frustration the users felt when their hands were not detected by the sensor, primarily during the tutorial phase. We noticed the same inconsistencies between user’s intentions and the lack of systems responses during all of our studies. Most of the problems relate to the issue of *exit error* described by Tuddenham et al. [92] for multi-touch and tangible user interfaces. It refers to a system’s difficulty in determining when people switch from non-interacting to action-causing motions. We asked participants to show their hands for each gesture in view of their camera, and we recognize that this prompt may have contributed to participants having their hands outside of the sensor’s detection range. Detection and lag time also contributed to an observed decrease in user confidence. Lag time was more prevalent for gestures similar to other gestures in the vocabulary, like HOVER and SLIDE, confirming the need for disambiguation in gesture design.

Appropriateness and accessibility. The discussion surrounding gesture ergonomics is an initial attempt to evaluate the accessibility of designed gesture vocabulary with experts. In addition, we wanted to investigate how experts’ opinions align with those of the elicited suggestions, particularly if any expert ratings contradict the suggested design dimensions. Although not tested in our research, the use of intention-to-operation mapping facilitates the implementation of the same types of gestures for multiple interaction techniques (touch, mid-air, or wearables). This mapping suggests the possibility of achieving uniformity of gestures across multiple modalities for interacting with visualizations. Notably, the positive ratings and high recall rates provided by users further support the viability of the mapping, as most gestures are designed with a specific physical association in mind and align with the predominant gesture types observed during the elicitation study. However, it is worth noting that several recommendations put forth by experts propose the use of the same gesture for different operations, which contradicts the requirement for disambiguation; this is also in line with some of the findings

from the elicitation study, particularly when considering user mental models (exhausting the same gesture). In many instances, however, experts suggested that the ergonomics of gestures could be improved by permitting users to hold their hands more loosely to avoid muscle strain. None of the recommendations suggest redesigns that alter the gesture type (for instance, suggestions that lead from dynamic to static gestures) or the need to decompose hybrid inputs into time-multiplexed inputs.

D. SS & XS Limitations

Next, we discuss the limitations of our SS and XS studies.

System setup. We ran all the studies remotely, and participants had to install and troubleshoot their setup in advance. Although there are some advantages to this configuration, it is not entirely clear whether these local conditions had a direct relationship with performance and recall. The functionality of the Leap Motion Controller and application are not supported by PCs. Thus our study was limited to participants using Apple machines (iMac, Macbook, etc.).

Recruitment and context. Two participants in the SS study had prior experience with gesture technology in virtual reality headsets, but none had used hand gestures, and we did not find any differences between them and novices. Domain experts in the XS study performed gestures after watching recordings. We are unaware if trying the system or watching experts in person would bring different insights and conclusions from these experts.

Visualization and cognition. Even though we utilized the transfer component by using different visual cues for tutorial and SS study tasks, we only investigated a scatterplot visualization and did not compare gestures using different visual representations. Furthermore, did not consider the significance or potential effect of color, palette, or intensity design choices in the elicitation or user studies. Since the average length of sessions was 15 minutes, we were unable to investigate the recency effect.

Accessibility. The design of the visualization used in the SS revealed an accessibility concern as one participant was colorblind; we did not consider this concern during recruitment. During the user study sessions, the moderator adjusted prompts to accommodate for this.

IX. MID-AIR HAND GESTURE DESIGN CONSIDERATIONS

All three studies (ES, SS, and XS) improved our understanding of mid-air hand gestures used to manipulate data visualizations. These findings, along with the above-mentioned conclusions from prior research, help clarify the relationship between gestures and intended operations and the desire for rules and disambiguation. This section discusses a set of considerations and strategies for resolving future problems that designers may face.

Appropriateness. The gesture vocabulary used in an application must be tailored to the specific purpose of the application. When designing a gesture vocabulary for a general audience to interact with casual data visualizations or public displays, designers should rely on legacy bias. We call

this a *legacy consensus* as a measure for designing a gesture set that users can easily discover and perform naturally and instantly. For all other applications that necessitate novel gesture designs, designers can rely on the *intention-effect* map, since gesture types are associated with operations regardless of the visualization.

Furthermore, the visual encoding of data, such as the size of data points and the direction of data translation, play crucial roles. According to the ES, scatterplots with numerous small data points lead to single-finger gestures. In contrast, bubble charts with larger data points inspire palm-up gestures for the same operation. This can also be translated to scenarios with different screen sizes. Furthermore, animated visualizations present the challenge of translation direction, which is dependent on the data and cannot be predicted in advance. In this case, designers should focus on physical associations to overcome the effect of the translation direction (such as the perception of time being linear). In addition, we recommend that designers encode hand gestures into physical associations, especially for abstract operations. Particularly when the system's reaction occurs after the execution of the gesture, users might form stronger links and even infer novel, unknown parts of the system by performing familiar movements.

Disambiguation. One of the strengths of WIMP interfaces is the ability to disambiguate, whereas gestures are fundamentally less discrete. Users must demonstrate distinct hand movements, which, when enforced on the user's end, has a negative effect on comfort and playfulness. Hand gesture designs must be clearly differentiated from one another, regardless of visual representation, to promote system effectiveness, understanding, and utility. This does not imply that the gesture styles must be distinct, and designers can experiment with various gesture types for different operations. The disambiguation must be present in the direction of translation and distance of designed gestures. These parameters are often observed in individual repetitions in the elicitation study (variability in hand postures directly correlates with memorability). This way, gestures are distinct enough for the user to comprehend the entire vocabulary and for the sensor to differentiate between hand movements more quickly, reducing systems' lag time and enhancing users' confidence.

Input Mapping. In practice, the number of operations to perform on a visualization can be quite high, resulting in the development of extensive gesture sets, which directly contradicts the challenge of disambiguation. A 3D interaction space can aid in designing diverse inputs, focusing on the changes in direction translation while keeping the same hand posture. Reducing the number of different postures is also possible by using hybrid techniques. Looking for inputs to be combined into a single technique is challenging. However, the possibility for parallel inputs (combining two distinct operations) may serve as a good starting point. Although hybrid techniques (e.g., *DiveZoom* or *TerraceZoom*[12]) are not the primary focus of our research, we are confident that this space provides an additional opportunity for the development of innovative mid-air interaction

designs that can outperform existing modalities.

Accessibility. It is important to recognize user group needs in relation to ergonomics. The direction of the gesture itself should hint at the direction of translation or manipulation, but it must be comfortable and discourage fatigue. For desktop settings, gestures should allow users to remain in a static position. Repetitions should be avoided if different muscle groups are required to perform the gesture. If gestures are not ergonomic, it may be necessary to make compromises to promote memorability and reduce fatigue. The gesture vocabulary should maintain spatial consistency that is comfortable for the average user and within reach for both unimanual and bimanual gestures.

Embodiment. The dimension of embodiment describes the degree to which a person perceives technology to be an extension of themselves [9]. As a perspective on the relationship between users and systems, embodied interaction [93] implies understanding in both directions: the system's *understanding* of a user's intent and the user's comprehension of the system as a "ready-to-hand" tool. Every technology has its limitations, which do not need to be explained explicitly to users but must be considered when designing interaction systems. To face challenges with input reliability, designers of mid-air gestures can incorporate embodiment via virtual hands, increasing the user's connection to the system and their confidence. However, the embodiment should correspond visually to the application and not obscure the visualization with opaque objects. In our pilot study, virtual hands distracted users, as they often paid more attention to the hands than the visualization. In immersive representations, particularly in VR, the variety of possibilities is much greater; the appearance of virtual hands does not interfere with the content since the content is in close proximity to the user and is often much bigger than the screens of home users. For desktop applications, incorporating visual responses and highlighting different parts of a visualization with different iconic gestures is a great substitute for virtual hands. For instance, if the system detects a specific gesture, it could highlight the visualization elements that can be interacted with. Visual feedback can subtly alert users when their postures and gestures are becoming ambiguous, thereby guiding their interactions. Ultimately, the user's connection with the system and the system's visual response to a gesture will enhance memorization, cognition, and self-confidence.

X. CONCLUSIONS AND FUTURE WORK

This paper explores the design considerations for mid-air hand gestures used to interact with data visualization. We conducted three studies—an Elicitation, User, and Expert Study—to understand user mental models, explore the design space of mid-air gestures, and provide recommendations for future work. There is the potential to rely on users' prior experiences and build upon gesture concepts previously explored or learned in other interfaces or interacting with the physical environment. We can leverage user knowledge and legacy bias to make novel systems more intuitive. If prior knowledge cannot be effectively utilized,

design space dimensions explored in this paper can be used to inform the appropriate design for mid-air hand interaction with data visualizations.

We consider our research as a foundation for future work on more specific scenarios. In that respect, future avenues could include collaboratively eliciting gestures by investigating how users perform and interpret a wide variety of operations and strategies through a mass performance study and collective learning. Although we were impressed by the ease of application of the equipment and the way in which Zoom users handled online usability studies, conducting our research in a controlled lab with larger displays may elicit different hand gestures. In order to design gestures that are ergonomic-aware [51], future research should examine the tradeoffs between comfort and precision of mid-air hand interactions. Furthermore, we want to investigate gestures for more complex operations and compare them to the findings of this study (e.g., if the gestures differ, are they still the same type?). We are also interested in using *aliasing* [94] or having multiple *gesture synonyms* for a single operation and seeing how that could affect memorability. Focusing on how designed gestures *feel* in the presence of haptic technology (e.g., [95]) is an additional dimension that could help mitigate minor accessibility and memorability challenges. Haptic feedback is a complementary modality to mid-air interaction that is used to confirm system response and increase confidence, particularly among visually impaired individuals.

REFERENCES

- [1] Z. Pousman and J. Stasko, "A taxonomy of ambient information systems: Four patterns of design," in *Proc. of the Working Conference on Advanced Visual Interfaces*. Venezia, Italy: ACM, May 2006, p. 67–74. [Online]. Available: <https://doi.org/10.1145/1133265.1133277>
- [2] T. Blascheck, L. Besançon, A. Bezerianos, B. Lee, and P. Isenberg, "Glanceable visualization: Studies of data comparison performance on smartwatches," *IEEE Trans. on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 630–640, Aug. 2019. [Online]. Available: <https://doi.org/10.1109/TVCG.2018.2865142>
- [3] A. Fonnet and Y. Prié, "Survey of immersive analytics," *IEEE Trans. on Visualization and Computer Graphics*, vol. 27, no. 3, pp. 2101–2122, Jul. 2021. [Online]. Available: <https://doi.org/10.1109/TVCG.2019.2929033>
- [4] E. Dimara and C. Perin, "What is interaction for data visualization?" *IEEE Trans. on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 119–129, Aug. 2020. [Online]. Available: <https://doi.org/10.1109/TVCG.2019.2934283>
- [5] Y. Jansen and P. Dragicevic, "An interaction model for visualizations beyond the desktop," *IEEE Trans. on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2396–2405, Oct. 2013. [Online]. Available: <https://doi.org/10.1109/TVCG.2013.134>
- [6] M. R. Jakobsen, Y. Sahlemariam Haile, S. Knudsen, and K. Hornbæk, "Information visualization and proxemics: Design opportunities and empirical findings," *IEEE Trans. on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2386–2395, 2013. [Online]. Available: <https://doi.org/10.1109/TVCG.2013.166>
- [7] D. Baur, B. Lee, and S. Carpendale, "Touchwave: Kinetic multi-touch manipulation for hierarchical stacked graphs," in *Proc. of the ACM International Conference on Interactive Tabletops and Surfaces*, Cambridge, MA, USA, Nov. 2012, p. 255–264. [Online]. Available: <https://doi.org/10.1145/2396636.2396675>
- [8] A. Saktheeswaran, A. Srinivasan, and J. Stasko, "Touch? Speech? or Touch and Speech? investigating multimodal interaction for visual network exploration and analysis," *IEEE Trans. on Visualization and Computer Graphics*, vol. 26, no. 6, pp. 2168–2179, Jan. 2020. [Online]. Available: <https://doi.org/10.1109/TVCG.2020.2970512>

- [9] B. Lee, P. Isenberg, N. H. Riche, and S. Carpendale, "Beyond mouse and keyboard: Expanding design considerations for information visualization interactions," *IEEE Trans. on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2689–2698, Oct. 2012. [Online]. Available: <https://doi.org/10.1109/TVCG.2012.204>
- [10] B. Spittle, M. Frutos-Pascual, C. Creed, and I. Williams, "A review of interaction techniques for immersive environments," *IEEE Trans. on Visualization and Computer Graphics*, pp. 1–1, May 2022. [Online]. Available: <https://doi.org/10.1109/TVCG.2022.3174805>
- [11] M. Hosseini, T. Ihmels, Z. Chen, M. Koelle, H. Müller, and S. Boll, "Towards a consensus gesture set: A survey of mid-air gestures in HCI for maximized agreement across domains," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany, Apr. 2023, pp. 311:1–311:24. [Online]. Available: <https://doi.org/10.1145/3544548.3581420>
- [12] K. A. Satriadi, B. Ens, M. Cordeil, B. Jenny, T. Czauderna, and W. Willett, "Augmented reality map navigation with freehand gestures," in *IEEE Conference on Virtual Reality and 3D User Interfaces*, Mar. 2019, pp. 593–603. [Online]. Available: <https://doi.org/10.1109/VR.2019.8798340>
- [13] S. Claes and A. V. Moere, "The role of tangible interaction in exploring information on public visualization displays," in *Proc. of the 4th International Symposium on Pervasive Displays*, Saarbruecken, Germany: ACM, Jun. 2015, p. 201–207. [Online]. Available: <https://doi.org/10.1145/2757710.2757733>
- [14] J. Kim, J. He, K. Lyons, and T. Starner, "The gesture watch: A wireless contact-free gesture based wrist interface," in *11th IEEE International Symposium on Wearable Computers*, Oct. 2007, pp. 15–22. [Online]. Available: <https://doi.org/10.1109/ISWC.2007.4373770>
- [15] S. S. Arefin Shimon, C. Lutton, Z. Xu, S. Morrison-Smith, C. Boucher, and J. Ruiz, "Exploring non-touchscreen gestures for smartwatches," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, San Jose, CA, USA, Mar. 2016, p. 3822–3833. [Online]. Available: <https://doi.org/10.1145/2858036.2858385>
- [16] J. Han, S. Ahn, K. Park, and G. Lee, "Designing touch gestures using the space around the smartwatch as continuous input space," in *Proc. of the ACM International Conference on Interactive Surfaces and Spaces*, Brighton, United Kingdom, Oct. 2017, p. 210–219. [Online]. Available: <https://doi.org/10.1145/3132272.3134134>
- [17] G. Ren, C. Li, E. O'Neill, and P. Willis, "3D freehand gestural navigation for interactive public displays," *IEEE Computer Graphics and Applications*, vol. 33, no. 2, pp. 47–55, Jan. 2013. [Online]. Available: <https://doi.org/10.1109/MCG.2013.15>
- [18] D. Vogel and R. Balakrishnan, "Distant freehand pointing and clicking on very large, high resolution displays," in *Proc. of the 18th Annual ACM Symposium on User Interface Software and Technology*, Seattle, WA, USA, Oct. 2005, p. 33–42. [Online]. Available: <https://doi.org/10.1145/1095034.1095041>
- [19] E. Brasier, O. Chapuis, N. Ferey, J. Vezien, and C. Appert, "Arpads: Mid-air indirect input for augmented reality," in *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Nov. 2020, pp. 332–343. [Online]. Available: <https://doi.org/ISMAR50242.2020.00060>
- [20] A. T. Cabreira and F. Hwang, "An analysis of mid-air gestures used across three platforms," in *Proc. of the British ACM HCI Conference*, Lincoln, United Kingdom, Jul. 2015, p. 257–258. [Online]. Available: <https://doi.org/10.1145/2783446.2783599>
- [21] S. A. Grandhi, G. Joue, and I. Mittelberg, "Understanding naturalness and intuitiveness in gesture production: Insights for touchless gestural interfaces," in *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, Canada, May 2011, p. 821–824. [Online]. Available: <https://doi.org/10.1145/1978942.1979061>
- [22] P. Vogiatzidakis and P. Koutsabasis, "Gesture elicitation studies for mid-air interaction: A review," *Multimodal Technologies and Interaction, MDPI*, vol. 2, no. 4, pp. 65:1–65:21, Sep. 2018. [Online]. Available: <https://www.mdpi.com/2414-4088/2/4/65>
- [23] W. Lu, Z. Tong, and J. Chu, "Dynamic hand gesture recognition with leap motion controller," *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1188–1192, Apr. 2016. [Online]. Available: <https://doi.org/10.1109/LSP.2016.2590470>
- [24] R. Veras, G. Singh, F. Farhadi-Niaki, R. Udhani, P. P. Patekar, W. Zhou, P. Irani, and W. Li, "Elbow-anchored interaction: Designing restful mid-air input," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan, May 2021, pp. 737:1–737:15. [Online]. Available: <https://doi.org/10.1145/3411764.3445546>
- [25] J. O. Wobbrock, H. H. Aung, B. Rothrock, and B. A. Myers, "Maximizing the guessability of symbolic input," in *ACM CHI Extended Abstracts on Human Factors in Computing Systems*, Portland, OR, USA, Apr. 2005, p. 1869–1872. [Online]. Available: <https://doi.org/10.1145/1056808.1057043>
- [26] D. A. Norman and S. W. Draper, *User Centered System Design; New Perspectives on Human-Computer Interaction*. USA: L. Erlbaum Associates Inc., Jan. 1986, isbn: 0898597811. [Online]. Available: <https://dl.acm.org/doi/book/10.5555/576915>
- [27] S. Villarreal-Narvaez, J. Vanderdonckt, R.-D. Vatavu, and J. O. Wobbrock, "A systematic review of gesture elicitation studies: What can we learn from 216 studies?" in *Proc. of the ACM Designing Interactive Systems Conference*, Eindhoven, Netherlands, Jul. 2020, p. 855–872. [Online]. Available: <https://doi.org/10.1145/3357236.3395511>
- [28] M. Nielsen, M. Stoerring, T. Moeslund, and E. Granum, "A procedure for developing intuitive and ergonomic gesture interfaces for HCI," in *In Proc. of the Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop*. Genova, Italy: Springer, Jan. 2003, pp. 409–420. [Online]. Available: https://doi.org/10.1007/978-3-540-24598-8_38
- [29] D. McNeill, *Hand and mind:what gestures reveal about thought*. Chicago, IL, USA: University of Chicago Press, 1992, isbn: 978-0226561349. [Online]. Available: <https://www.jstor.org/stable/1576015?origin=crossref>
- [30] E. Morsella and R. M. Krauss, "The role of gestures in spatial working memory and speech," in *The American Journal of Psychology*. Chicago, IL, USA: University of Illinois Press, Sep. 2004, vol. 117, no. 3, pp. 411–424. [Online]. Available: <http://www.jstor.org/stable/4149008>
- [31] T. Vuletic, A. Duffy, L. Hay, C. McTeague, G. Campbell, and M. Grealy, "Systematic literature review of hand gestures used in human computer interaction interfaces," in *International Journal of Human-Computer Studies*. Elsevier, Sep. 2019, vol. 129, no. 1, pp. 74–94. [Online]. Available: <https://doi.org/10.1016/j.ijhcs.2019.03.011>
- [32] C. Perin, "Direct manipulation for information visualization," Ph.D. dissertation, Université Paris-Sud, France, Nov. 2014. [Online]. Available: <https://inria.hal.science/tel-01096366/document>
- [33] Y.-J. Huang, T. Fujiwara, Y.-X. Lin, W.-C. Lin, and K.-L. Ma, "A gesture system for graph visualization in virtual reality environments," in *IEEE PacificVis Pacific Visualization Symposium*, Apr. 2017, pp. 41–45. [Online]. Available: <https://doi.org/10.1109/PACIFICVIS.2017.8031577>
- [34] S. Hegde, R. Perla, R. Hebbalaguppe, and E. Hassan, "Gestar: Real time gesture interaction for ar with egocentric view," in *IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, Sep. 2016, pp. 262–267. [Online]. Available: <https://doi.org/10.1109/ISMAR-Adjunct.2016.0090>
- [35] T. Lee and T. Hollerer, "Hybrid feature tracking and user interaction for markerless augmented reality," in *IEEE Virtual Reality Conference*, Apr. 2008, pp. 145–152. [Online]. Available: <https://doi.org/10.1109/VR.2008.4480766>
- [36] C. Arslan, Y. Rekik, and L. Grisoni, "E-pad: Large display pointing in a continuous interaction space around a mobile device," in *Proc. on Designing Interactive Systems Conference*. San Diego, CA, USA: ACM, Jun. 2019, p. 1101–1108. [Online]. Available: <https://doi.org/10.1145/3322276.3322284>
- [37] X. A. Chen, J. Schwarz, C. Harrison, J. Mankoff, and S. E. Hudson, "Air+touch: Interweaving touch and in-air gestures," in *Proc. of the ACM UIST Symposium on User Interface Software and Technology*, Honolulu, HI, USA, Oct. 2014, p. 519–525. [Online]. Available: <https://doi.org/10.1145/2642918.2647392>
- [38] T. Isenberg and M. Hancock, "Gestures vs. Postures: 'Gestural' Touch Interaction in 3D Environments," in *Proc. of the CHI Workshop on "The 3rd Dimension of CHI: Touching and Designing 3D User Interfaces"*, Austin, TX, USA, May 2012, pp. 53–61, hal: 00781237. [Online]. Available: <https://inria.hal.science/hal-00781237>
- [39] J. Thompson, A. Srinivasan, and J. Stasko, "Tangraph: Interactive exploration of network visualizations using single hand, multi-touch gestures," in *Proc. of the ACM International Conference on Advanced Visual Interfaces*, Castiglione della Pescaia, Italy, May 2018, pp. 43:1–43:5. [Online]. Available: <https://doi.org/10.1145/3206505.3206519>
- [40] S. Schmidt, M. A. Nacenta, R. Dachsel, and S. Carpendale, "A set of multi-touch graph interaction techniques," in *ACM International Conference on Interactive Tabletops and Surfaces*, Saarbrücken, Germany, Nov. 2010, p. 113–116. [Online]. Available: <https://doi.org/10.1145/1936652.1936673>
- [41] J. M. Rzeszotarski and A. Kittur, "Kinética: Naturalistic multi-touch data visualization," in *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Toronto, Canada, Apr. 2014, p. 897–906. [Online]. Available: <https://doi.org/10.1145/2556288.2557231>

- [42] E. Courtoix, C. Appert, and O. Chapuis, "Walltokens: Surface tangibles for vertical displays," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, Yokohama, Japan, May 2021, p. 421:1–421:13. [Online]. Available: <https://doi.org/10.1145/3411764.3445404>
- [43] M. Cordeil, A. Cunningham, T. Dwyer, B. H. Thomas, and K. Marriott, "Imaxes: Immersive axes as embodied affordances for interactive multivariate data visualisation," in *Proc. of the ACM UIST Symposium on User Interface Software and Technology*, Québec City, Canada, Oct. 2017, p. 71–83. [Online]. Available: <https://doi.org/10.1145/3126594.3126613>
- [44] T. Stachl, "Elicitation and evaluation of mid-air hand gestures for global earth observation data presented on large public displays," Master's thesis, Technische Universität Wien, Jul. 2021. [Online]. Available: <https://doi.org/10.34726/hss.2021.90363>
- [45] Y. Jansen, "Physical and tangible information visualization," Ph.D. dissertation, Université Paris-Sud, France, Mar. 2014. [Online]. Available: https://theses.hal.science/file/index/docid/981521/filename/VA2_JANSEN_YVONNE_10032014.pdf
- [46] F. Taher, Y. Jansen, J. Woodruff, J. Hardy, K. Hornbæk, and J. Alexander, "Investigating the use of a dynamic physical bar chart for data exploration and presentation," *IEEE Trans. on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 451–460, Aug. 2017. [Online]. Available: <https://doi.org/10.1109/TVCG.2016.2598498>
- [47] A. Balali Moghaddam, J. Svendsen, M. Tory, and A. Branzan Albu, "Integrating touch and near touch interactions for information visualizations," in *CHI Extended Abstracts on Human Factors in Computing Systems*. NY, USA: ACM, May 2011, p. 2347–2352. [Online]. Available: <https://doi.org/10.1145/1979742.1979934>
- [48] I. Aslan, T. Schmidt, J. Woehrle, L. Vogel, and E. André, "Pen + mid-air gestures: Eliciting contextual gestures," in *Proc. of the 20th International Conference on Multimodal Interaction*. Boulder, CO, USA: ACM, Oct. 2018, p. 135–144. [Online]. Available: <https://doi.org/10.1145/3242969.3242979>
- [49] K. Sabir, C. Stoltz, B. Tabor, and S. I. O'Donoghue, "The molecular control toolkit: Controlling 3D molecular graphics via gesture and voice," in *IEEE Symposium on Biological Data Visualization*, Nov. 2013, pp. 49–56. [Online]. Available: <https://doi.org/10.1109/BioVis.2013.6664346>
- [50] J. Wagner, W. Stuerzlinger, and L. Nedel, "Comparing and combining virtual hand and virtual ray pointer interactions for data manipulation in immersive analytics," *IEEE Trans. on Visualization and Computer Graphics*, vol. 27, no. 5, pp. 2513–2523, May 2021. [Online]. Available: <https://doi.org/10.1109/TVCG.2021.3067759>
- [51] D. Schön, T. Kosch, F. Müller, M. Schmitz, S. Günther, L. Bommhardt, and M. Mühlhäuser, "Tailor twist: Assessing rotational mid-air interactions for augmented reality," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany, Apr. 2023, pp. 400:1–400:14. [Online]. Available: <https://doi.org/10.1145/3544548.3581461>
- [52] Z. Liu and J. Stasko, "Mental models, visual reasoning and interaction in information visualization: A top-down perspective," *IEEE Trans. on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 999–1008, 2010. [Online]. Available: <https://doi.org/10.1109/TVCG.2010.177>
- [53] J. Avery, M. Choi, D. Vogel, and E. Lank, "Pinch-to-zoom-plus: An enhanced pinch-to-zoom that reduces clutching and panning," in *Proc. of the 27th ACM Symposium on User Interface Software and Technology*, Honolulu, HI, USA, Oct. 2014, p. 595–604. [Online]. Available: <https://doi.org/10.1145/2642918.2647352>
- [54] M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay, "Mid-air pan-and-zoom on wall-sized displays," in *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems*, May 2011, p. 177–186. [Online]. Available: <https://doi.org/10.1145/1978942.1978969>
- [55] F. van de Camp, A. Schick, and R. Stiefelhagen, "How to click in mid-air," in *First International Conference on Distributed, Ambient, and Pervasive Interactions*. Berlin, Heidelberg: Springer, Jul. 2013, pp. 78–86. [Online]. Available: https://doi.org/10.1007/978-3-642-39351-8_9
- [56] C. R. Austin, B. Ens, K. A. Satriadi, and B. Jenny, "Elicitation study investigating hand and foot gesture interaction for immersive maps in augmented reality," in *Journal of Cartography and Geographic Information Science*. Taylor & Francis, Jan. 2020, vol. 47, no. 3, pp. 214–228. [Online]. Available: <https://doi.org/10.1080/15230406.2019.1696232>
- [57] T. Pham, J. Vermeulen, A. Tang, and L. MacDonald Vermeulen, "Scale impacts elicited gestures for manipulating holograms: Implications for ar gesture design," in *Proc. of Designing Interactive Systems Conference*. Hong Kong, China: ACM, Jun. 2018, p. 227–240. [Online]. Available: <https://doi.org/10.1145/3196709.3196719>
- [58] W. Xu, H.-N. Liang, Y. Chen, X. Li, and K. Yu, "Exploring visual techniques for boundary awareness during interaction in augmented reality head-mounted displays," in *Proc. of the IEEE Conference on Virtual Reality and 3D User Interfaces*, May 2020, pp. 204–211. [Online]. Available: <https://doi.org/10.1109/VR46266.2020.00039>
- [59] L. Besançon, A. Ynnerman, D. Keefe, L. Yu, and T. Isenberg, "The state of the art of spatial interfaces for 3D visualization," *Computer Graphics Forum*, vol. 40, no. 1, pp. 293–326, Feb. 2021. [Online]. Available: <https://doi.org/10.1111/cgf.14189>
- [60] F. Block and H. Gellersen, "Space-multiplexed input on mouse-extended notebooks," in *Proc. of the 6th ACM Nordic Conference on Human-Computer Interaction: Extending Boundaries*, Reykjavik, Iceland, Oct. 2010, p. 603–606. [Online]. Available: <https://doi.org/10.1145/1868914.1868984>
- [61] J. Lee, A. Olwal, H. Ishii, and C. Boulanger, "Spacetop: Integrating 2D and spatial 3D interactions in a see-through desktop environment," in *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Paris, France, Apr. 2013, p. 189–192. [Online]. Available: <https://doi.org/10.1145/2470654.2470680>
- [62] H. Xia, M. Glueck, M. Annett, M. Wang, and D. Wigdor, "Iteratively designing gesture vocabularies: A survey and analysis of best practices in the HCI literature," *ACM Trans. on Computer-Human Interaction*, vol. 29, no. 4, pp. 1–54, May 2022. [Online]. Available: <https://doi.org/10.1145/3503537>
- [63] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, S. Zhao, and J. Koh, "Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI," Microsoft Corporation, Redmond, WA, USA, Tech. Rep., Nov. 2012. [Online]. Available: <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/GesturesTR-20121107-RoA.pdf>
- [64] J. O. Wobbrock, M. R. Morris, and A. D. Wilson, "User-defined gestures for surface computing," in *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA, USA, Apr. 2009, p. 1083–1092. [Online]. Available: <https://doi.org/10.1145/1518701.1518866>
- [65] M. R. Morris, A. Danilescu, S. Drucker, D. Fisher, B. Lee, m. c. schraefel, and J. O. Wobbrock, "Reducing legacy bias in gesture elicitation studies," *ACM Interactions*, vol. 21, no. 3, p. 40–45, May 2014. [Online]. Available: <https://doi.org/10.1145/2591689>
- [66] A. Danilescu and D. Piorkowski, "Iterative design of gestures during elicitation: Understanding the role of increased production," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, New Orleans, LA, USA, Apr. 2022, p. 558:1–558:14. [Online]. Available: <https://doi.org/10.1145/3491102.3501962>
- [67] F. Cafaro, L. Lyons, and A. N. Antle, "Framed guessability: Improving the discoverability of gestures and body movements for full-body interaction," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, Montreal, Canada, Apr. 2018, p. 1–12. [Online]. Available: <https://doi.org/10.1145/3173574.3174167>
- [68] A. Köpsel and N. Bubalo, "Benefiting from legacy bias," *ACM Interactions*, vol. 22, no. 5, p. 44–47, Aug. 2015. [Online]. Available: <https://doi.org/10.1145/2803169>
- [69] J. Walny, B. Lee, P. Johns, N. Henry Riche, and S. Carpendale, "Understanding pen and touch interaction for data exploration on interactive whiteboards," *IEEE Trans. on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2779–2788, Oct. 2012. [Online]. Available: <https://doi.org/10.1109/TVCG.2012.275>
- [70] L. Gallo, "A study on the degrees of freedom in touchless interaction," in *SIGGRAPH Asia Technical Briefs*. Hong Kong, China: ACM, Nov. 2013, pp. 28:1–28:4. [Online]. Available: <https://doi.org/10.1145/2542355.2542390>
- [71] D. Freeman, R. Vennelakanti, and S. Madhvanath, "Freehand pose-based gestural interaction: Studies and implications for interface design," in *the Proceedings of the 4th IEEE International Conference on Intelligent Human Computer Interaction*, Mar. 2012, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/IHCI.2012.6481816>
- [72] S. Aromaa, N. Frangakis, D. Tedone, J. Viitaniemi, and I. Aaltonen, "Digital human models in human factors and ergonomics evaluation of gesture interfaces," *Proc. of the ACM on Human-Computer Interaction*, vol. 2, no. 6, pp. 1–14, Jun. 2018. [Online]. Available: <https://doi.org/10.1145/3229088>
- [73] M. Karam and m. c. schraefel, "A taxonomy of gestures in human computer interactions," University of Southampton, Tech. Rep., Aug. 2005. [Online]. Available: <https://eprints.soton.ac.uk/261149/>

- [74] F. Hoffmann, M.-I. Tyroller, F. Wende, and N. Henze, "User-defined interaction for smart homes: Voice, touch, or mid-air gestures?" in *Proc. of the 18th ACM International Conference on Mobile and Ubiquitous Multimedia*, Pisa, Italy, Nov. 2019, pp. 33:1–33:7. [Online]. Available: <https://doi.org/10.1145/3365610.3365624>
- [75] T. Seyed, C. Burns, M. Costa Sousa, F. Maurer, and A. Tang, "Eliciting usable gestures for multi-display environments," in *Proc. of the ACM International Conference on Interactive Tabletops and Surfaces*, Cambridge, MA, USA, Nov. 2012, p. 41–50. [Online]. Available: <https://doi.org/10.1145/2396636.2396643>
- [76] E. Chan, T. Seyed, W. Stuerzlinger, X.-D. Yang, and F. Maurer, "User elicitation on single-hand microgestures," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, San Jose, CA, USA, May 2016, p. 3403–3414. [Online]. Available: <https://doi.org/10.1145/2858036.2858589>
- [77] C. Ackad, A. Clayphan, M. Tomitsch, and J. Kay, "An in-the-wild study of learning mid-air gestures to browse hierarchical information at a large interactive public display," in *Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Osaka, Japan, Sep. 2015, p. 1227–1238. [Online]. Available: <https://doi.org/10.1145/2750858.2807532>
- [78] J. S. Yi, Y. a. Kang, J. Stasko, and J. A. Jacko, "Toward a deeper understanding of the role of interaction in information visualization," *IEEE Trans. on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1224–1231, Nov. 2007. [Online]. Available: <https://doi.org/10.1109/TVCG.2007.70515>
- [79] M. Brehmer and T. Munzner, "A multi-level typology of abstract visualization tasks," *IEEE Trans. on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2376–2385, Oct. 2013. [Online]. Available: <https://doi.org/10.1109/TVCG.2013.124>
- [80] E. Dimara and J. Stasko, "A critical reflection on visualization research: Where do decision making tasks hide?" *IEEE Trans. on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 1128 – 1138, Sep. 2021. [Online]. Available: <https://doi.org/10.1109/TVCG.2021.3114813>
- [81] T. Munzner, *Visualization Analysis and Design*. USA: A K Peters Visualization Series, CRC Press, Dec. 2014, isbn: 9781498759717. [Online]. Available: <https://www.cs.ubc.ca/~tmm/vadbook/>
- [82] B. Saket, S. Huron, C. Perin, and A. Endert, "Investigating direct manipulation of graphical encodings as a method for user interaction," *IEEE Trans. on Visualization and Computer Graphics*, vol. 26, pp. 482–491, Aug. 2019. [Online]. Available: <https://doi.org/10.1109/TVCG.2019.2934534>
- [83] B. Saket, H. Kim, E. T. Brown, and A. Endert, "Visualization by demonstration: An interaction paradigm for visual data exploration," *IEEE Trans. on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 331–340, Aug. 2017. [Online]. Available: <https://doi.org/10.1109/TVCG.2016.2598839>
- [84] B. Kondo and C. Collins, "Dimpvis: Exploring time-varying information visualizations by direct manipulation," *IEEE Trans. on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2003–2012, Dec. 2014. [Online]. Available: <https://doi.org/10.1109/TVCG.2014.2346250>
- [85] W. Willett, Q. Lan, and P. Isenberg, "Eliciting Multi-touch Selection Gestures for Interactive Data Graphics," in *EuroVis - Short Papers*. The Eurographics Association, Jun. 2014. [Online]. Available: <https://doi.org/10.2312/eurovisshort.20141161>
- [86] B. Saket, A. Srinivasan, E. D. Ragan, and A. Endert, "Evaluating interactive graphical encodings for data visualization," *IEEE Trans. on Visualization and Computer Graphics*, vol. 24, no. 3, pp. 1316–1330, Aug. 2018. [Online]. Available: <https://doi.org/10.1109/TVCG.2017.2680452>
- [87] D. A. Norman and J. Nielsen, "Gestural interfaces: A step backward in usability," *ACM Interactions*, vol. 17, no. 5, p. 46–49, Sep. 2010. [Online]. Available: <https://doi.org/10.1145/1836216.1836228>
- [88] R.-D. Vatavu and J. O. Wobbrock, "Clarifying agreement calculations and analysis for end-user elicitation studies," *ACM Trans. on Computer-Human Interaction*, vol. 29, no. 1, pp. 5:1–5:70, Jan. 2022. [Online]. Available: <https://doi.org/10.1145/3476101>
- [89] L.-B. Bilius, O.-C. Ungurean, and R.-D. Vatavu, "Understanding wheelchair users' preferences for on-body, in-air, and on-wheelchair gestures," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany, Apr. 2023, pp. 78:1–78:16. [Online]. Available: <https://doi.org/10.1145/3544548.3580929>
- [90] R.-D. Vatavu and J. O. Wobbrock, "Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit," in *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*, Seoul, Republic of Korea, Apr. 2015, p. 1325–1334. [Online]. Available: <https://doi.org/10.1145/2702123.2702223>
- [91] F. Anderson and W. F. Bischof, "Learning and performance with gesture guides," in *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Paris, France, Apr. 2013, p. 1109–1118. [Online]. Available: <https://dl.acm.org/doi/10.1145/2470654.2466143>
- [92] P. Tuddenham, D. Kirk, and S. Izadi, "Graspables revisited: Multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks," in *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems*, Atlanta, Georgia, USA, Apr. 2010, p. 2223–2232. [Online]. Available: <https://doi.org/10.1145/175326.1753662>
- [93] P. Dourish, *Where the Action Is: The Foundations of Embodied Interaction*. USA: The MIT Press, Oct. 2001, isbn: 0262041960. [Online]. Available: <https://direct.mit.edu/books/book/3875/Where-the-Action-IsThe-Foundations-of-Embodied>
- [94] A. S. Williams, J. Garcia, and F. Ortega, "Understanding multimodal user gesture and speech behavior for object manipulation in augmented reality using elicitation," *IEEE Trans. on Visualization and Computer Graphics*, vol. 26, no. 12, pp. 3479–3489, Sep. 2020. [Online]. Available: <https://doi.org/10.1109/TVCG.2020.3023566>
- [95] J. Jang and J. Park, "SPH fluid tactile rendering for ultrasonic mid-air haptics," *IEEE Trans. on Haptics*, vol. 13, no. 1, pp. 116–122, Jan. 2020. [Online]. Available: <https://doi.org/10.1109/TOH.2020.2966605>



Zona Kostic is a research fellow and lecturer at Harvard University. Her research lies at the intersection of data visualization, machine learning, and immersive technologies. Zona conducted research on immersive visualization and interfaces with the Visual Computing Group, Harvard SEAS. While pursuing her Ph.D., Zona published six books on computer graphics and design, conducted industry research on human-computer interfaces in virtual reality, and published research on immersive technologies for distance learning.



Catherine Dumas is an Assistant Professor and INF Director at the State University of New York at Albany (Assistant Professor, School of Library and Information at Simmons University at the time of the research). Catherine's research is inspired by data used for online collective action or digital activism, specifically electronic petitioning. Catherine is currently researching the utility and effectiveness of immersive technologies in education, as well as on the usability, collaboration, and interaction in large virtual reality environments.



Sarah Pratt is Associate University Archivist for Community Engagement at Harvard University (Assistant Professor of Practice at Simmons University's School of Library and Information Science at the time of the research). Sarah is currently serving as co-chair of the Society of American Archivists' Committee on Research, Data, and Assessment. Sarah received her MSLIS with a concentration in Archival Management from Simmons and has a post-Master's certificate in user experience research and design.



Johanna Beyer is a research scientist and lecturer at the Visual Computing Lab at Harvard University. Before joining Harvard, she was a postdoctoral fellow at the Geometric Modeling and Scientific Visualization Center at KAUST. She received her Ph.D. in computer science at the University of Technology Vienna, Austria, in 2009. Her research interests include the combination of abstract information visualization with scientific visualization for novel domain-specific applications and immersive visual analytics.