Exploring the Design Space of Gestural Interaction with Active Tokens through User-Defined Gestures

Consuelo Valdes[1], Diana Eastman[1], Casey Grote[1], Shantanu Thatte[3],
Orit Shaer[1], Ali Mazalek[2], Brygg Ullmer[3,4], Miriam K. Konkel[5]

Dept. of Computer Science[1]
Wellesley College
Wellesley, MA, USA

Synaesthetic Media Lab[2]
Digital Media & GVU Center
Georgia Institute of Technology
Atlanta, GA, USA

School of EECS[3], Center for
Computation and Technology[4],
and Dept. of Biological Sciences[5]
Louisiana State University
Baton Rouge, LA, USA

oshaer@wellesley.edu, mazalek@gatech.edu, ullmer@lsu.edu, konkel@lsu.edu

ABSTRACT

Multi-touch and tangible interfaces provide unique opportunities for enhancing learning and discovery with big data. However, existing interaction techniques have limitations when manipulating large data sets. Our goal is to define novel interaction techniques for multi-touch and tangible interfaces, which support the construction of complex queries for big data. In this paper, we present results from a study, which investigates the use of gestural interaction with active tokens for manipulating large data sets. In particular, we studied user expectations of a hybrid tangible and gestural language engaging this space. Our main results include a vocabulary of user-defined gestures for interaction with active tokens, which extends beyond familiar multi-touch gestures; characterization of the design space of gestural interaction with active tokens; and insight into participants' mental models, including common metaphors. We also present implications for the design of multi-touch and tangible interfaces with active tokens.

Author Keywords

Gestures; tabletop; queries; multi-display environments; cross-device interaction; physical tokens.

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces---input devices and strategies, interaction styles

INTRODUCTION

Multi-touch and tangible interfaces are a growing area of human-computer interaction (HCI) research. Recent studies have focused on their use in entertainment, formal and informal learning environments, as well as collaborative search. Multi-touch and tangible interfaces provide form factors that foster collaboration through visibility of actions

and egalitarian input [16, 23] and support distributed cognition by affording spreading, piling, and organizing information artifacts [3, 4]. Such affordances are well suited for sensemaking [37]. Research also indicates that tabletop interfaces support active reading [29] and promote reflection [44]. Considering these qualities of tabletop interaction, as well as other factors such as improving display quality, increasing availability, and falling prices of commercial hardware platforms, tabletop interfaces provide unique opportunities for learning and discovery with large and abstract data sets.
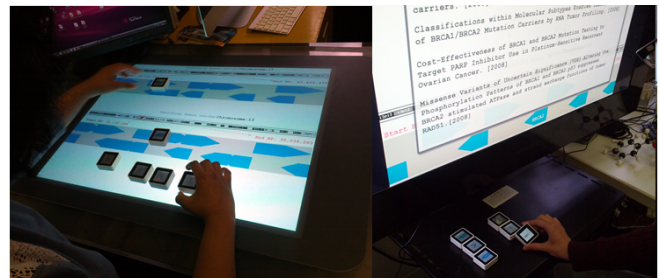

Figure 1: Horizontal experimental setup (left) - first generation Microsoft Surface and first generation Sifteo cubes. Vertical experimental setup (right) – second generation Microsoft Surface and first generation Sifteo cubes.

However, designing multi-touch and tangible interfaces that support learning and discovery in data-intense areas requires going beyond the application of existing interaction techniques [43]. While direct touch is currently a standard input method for multi-touch and tangible interfaces, in data-intense applications, representations are typically small [9]. Finger size and occlusion make it difficult to interact directly with small targets through touch [9, 54, 55]. Furthermore, in data-intense applications, WIMP-style control elements provided by various toolkits, such as scrollbars, sliders, and text fields, may often be either too small for effective and accurate touch interaction, consume expensive screen real-estate [5, 9], or lead to cognitive dissonance with the broader interaction paradigm.

Several studies have considered novel multi-touch interaction techniques for data driven multi-touch and tangible applications; e.g. [5, 9, 17, 55]. However, while providing advantage over touch interaction with WIMP-style controls, multi-touch gestures often suffer from low discoverability and lack of persistence [9]. In this paper, we investigate an alternative approach: exploring large data sets on multi-touch and tangible interfaces using tangible interaction with active tokens. Our goal is to identify novel interaction techniques for multi-touch and tangible interfaces, which support the construction of complex queries for large data sets.

Our work draws and expands on Tangible Query Interfaces (TQI) [50], which introduced tangible interaction techniques for querying large databases. TQI utilized systems of tokens embodying query parameters. These were manipulated, interpreted, and graphically augmented on a series of physical constraints. Technological advances and mass-market offerings such as Sifteo Cubes [1] have opened up the possibilities to revisit this approach and consider novel tangible and gesture-based interaction techniques for the manipulation of large data sets that use active tokens [60].

Active tokens are programmable physical objects with integrated display, sensing, or actuation technologies (e.g., [1, 6, 21, 24, 60]). Thus, they can be reconfigured over time, allowing users to dynamically modify their associations with datasets or controls. The use of active tokens could support the exploration of large data sets by expanding the design space of token and constraints interaction [42, 49], which leverages physical expression of digital syntax. Combining interactive multi-touch surfaces with active tokens could facilitate the presentation and manipulation of big data while preserving the benefits of tangible interaction [41] such as support for two-handed interaction, co-located collaboration, and strong affordances.

We focus on a sub-class of active tokens that can be manipulated using gestures independently from global constraints. Gestural interaction with active tokens blurs the boundaries between tangible and gestural interaction, between manipulations and gestures [20]; and fits within the area defined as "tangible gesture interaction" [52]. Active tokens also enable the expansion of tangible interaction with multi-touch and tangible interfaces beyond interaction on the surface into less explored areas such as tangible interaction *on-bezel* and *in-air,* hovering above or in front the surface.

While technological advances facilitate the recognition and tracking of on-surface, in-air, and on-bezel interactions with active tokens, the gesture language for such interactions has yet to be explored in a task–driven scenario. What kinds of gestures should be employed to manipulate and explore large data sets? What gestures might typically or best be performed on-surface, on-bezel, and in-air? In users' minds, what are the important characteristics of such gestures? Are there guessable gestures for navigating a data set and for constructing complex queries? How consistently do different users employ gestures for the same commands?

On the one hand, we regard the design of a hybrid tangible and gestural language as a highly complex task, which pushes the skills of even expert tangible and gestural interaction designers. On the other, gestural interaction with smart phones and tablets is now routine for many millions of people. To understand and respect the expectations of users, we see it important to explore a cross-section of users' first anticipations of new systems in this space.

In this paper, we present findings from a user study with 19 users. The study investigated user expectations of a hybrid tangible and gestural language for multi-touch and tangible interaction with active tokens. Our findings include a vocabulary of user-defined gestures for active tokens, and insights into the mental models and experiences that shape users' expectations. Based on our findings we characterize the design space of gestural interaction with active tokens and discuss implications for the design of multi-touch and tangible interfaces with active tokens.

RELATED WORK

We begin with a survey of existing work on multi-touch and tangible interfaces for data exploration.

Multi-touch and Tangible Interfaces for Big Data

With the increasing importance of and access to large data sets by diverse prospective users [31], several multi-touch and tangible interfaces have been created to support learning and understanding in data intense areas in informal and formal settings. For example, Block et al. created the DeepTree exhibit for science museums [5], a tabletop interactive visualization of the Tree of Life that facilitates collaborative exploration of phylogenetic data while offering guided discovery. Involve [15] is a precursor of DeepTree that utilized Vornoi maps to visualize phylogenetic data. Both applications allow users to zoom, pan, and find a particular element in a visualization of the Life on Earth database. Thudt et al. created the Bohemian Bookshelf [46], a multi-touch interface, which supports serendipitous discoveries in the context of digital book collections through five interlinked visualizations. These multi-touch interfaces support collaborative and playful interaction with large data sets, which is appropriate for informal learning settings, but limit exploration through designer-defined guided discovery scenarios and visual representations.

G-nome Surfer [45] and GreenTouch [51] are tabletop multi-touch applications for collaborative exploration of genomic and phenology databases. Both applications target college-level inquiry-based learning and support open-ended data investigation. However, their expressivity is limited as users cannot define and set query operators directly. Schkolne et al. [39] developed an immersive tangible interface for the design of DNA; the interface does not support search and comparison. Morris et al. [30] surveyed the design space of collaborative tabletop search applications, examining several dimensions including domain, collaboration style, and search input. While application domains for search interfaces varied from personal data to large data-

bases, search input techniques have relied on touch, keyboard, and mouse input. Facet-Streams [19] is a recent tabletop search interface that explores the design space of tangible-tabletop interaction. The interface allows users to construct expressive queries using passive tokens that are augmented with content and functionality when placed upon the tabletop surface through the display of a surrounding circular menu. Similarly to the TQI [50], Facet-Streams utilizes passive tokens that could only be used on top of a tabletop interactive surface. While simplifying the construction of complex queries, this system also has some limitations such as physical clutter, separation of query formulation and result browsing, and lack of persistency for query parameters. Our focus is on the use of *active tokens* that, when combined with multi-touch and tangible interfaces, could be manipulated on-surface, on-bezel, and in-air. Such interaction remains unexplored within search interfaces.

Tangible Queries
Several tangible interfaces have explored the use of tokens for query formulation. Navigational Blocks [6] is an early tangible user interface for querying a database where each block represents one query parameter, its six faces represent the possible values. Depending on whether two parameter values would yield information, the blocks attract or repel each other, providing users with actuated feedback. The number of blocks and their fixed values limited the search space to a predefined number of queries. DataTiles [36] combines tiles – graspable windows for digital information – on top of a horizontal display. Each tile represents particular information or control (e.g. parameter, storage, portal). While DataTiles provides a highly expressive physical language, interaction is constrained to a horizontal surface.

Ullmer et al. [50] introduced two tangible interfaces for query formulation that use tokens to represent query parameters. One employed "parameter wheels" for fixed query parameters; the other, "parameter bars" that can be dynamically assigned to represent various parameters. In both interfaces, tokens can represent either discrete or continuous values, and are manipulated and interpreted on a series of physical constraints (e.g. racks). Such token and constraint interaction, while expressive, makes the tokens less portable and limits possibilities for collaboration. In later evolutions, cartouches [48] build on DataTiles and TQI toward generalizing representation of digital information across diverse systems. Examples include systems of tokens and constraints that could be combined with systems ranging from interactive walls to information kiosks. We extend this work by investigating the design space of active tokens, which as active physical objects could be manipulated using gestures independently from mechanical constraints.

While most tangible query interfaces utilize horizontal arrangements [6, 19, 36, 50], Stackables [21] explores the use of tangible objects that are arranged vertically. Each stackable token represents a query parameter; multiple tokens are stacked to express a query. Stackables utilize active tokens that could be reprogrammed and support interaction independently from global constraints. Multiple tokens can be combined with a logical AND or negation through vertical stacking. Interaction with a single stackable token is limited to using its sliders and button. Stackables provides a step in the direction of using active tokens for search. Our focus is on identifying possibilities for rich gestural interaction with active tokens combined with multi-touch and tangible interfaces. We seek to explore gestural interaction that extends beyond the surface into less explored areas such as on-bezel and in-air, hovering above or in front of the surface.

Active Token Interaction Languages
The use of active tokens has been explored in various contexts. For example, the Tangible Video Editor [60] employed active tokens to represent video clips. To guide users' interaction, video clip tokens were embedded in cases that resembled jigsaw puzzle pieces. Video clips were associated by connecting the puzzle pieces on a surface or in-air. However, while the affordances and physical constraints provided a clear association pattern, they also restricted the use of the active tokens. Siftables [24], a precursor of Sifteo Cubes [1], explored interaction gestures based on reality-based [18] metaphors such as grouping cubes to tag them with a common element; shaking to express yes or no; and "sugar pack snap" to clear cube contents. Sifteo [1] implements various gestures including shaking, flipping, tilting, neighboring, and touching. However, these gesture are expert-designed and were not evaluated with users. Six-Forty by Four-Eighty [7] is an art installation that uses a set of gestures for manipulating "tile" displays. SynFlo [59] is an interactive installation for introducing Synthetic Biology concepts to non-scientists. It utilizes Sifteo cubes to simulate a wet-lab bench and evoke gestures such as pouring and shaking. The installation incorporates a Microsoft PixelSense, which serves as a test bed for a virtual experiment.

Though these systems present explorations into the design space of gesture-based interaction with active tokens, they neither tackle the big data context nor evaluate the gesture set in a task-driven scenario.

User-Elicitation Studies
User-elicitation studies are a specific type of participatory design methodology that involves end-users in the design of gesture-sets [12, 28, 54]. Such studies have been used to design gesture interfaces of various types including multi-touch gestures on small and large surfaces [2, 10, 11, 22, 25, 32, 57] and multi-modal interactions [28]. There is evidence that user-defined gesture sets are more complete than those defined solely by experts [2, 32, 35, 57].

In a user-elicitation study, users are shown referents (an action's effects) and are asked to demonstrate the interactions that result in a given referent [57]. In this work, we draw upon the user-elicitation methodology to identify user expectations and suggestions for active-token interaction with multi-touch and tangible interfaces for big data explo-

ration. Our intention is not to present a complete user-defined gesture language, but rather to identify and understand user expectations when transitioning from familiar multi-touch interaction into a new space. We examine when users use on-bezel and in-air interactions, and how users associate active tokens with data displayed on top of multi-touch and tangible interfaces. We identify common gestures for search and query formulation tasks and discuss users' perceptions of the benefits and drawbacks of each.

ELICITING USER EXPECTATION THROUGH USER-DEFINED GESTURES

To explore user expectations of a hybrid tangible and gestural interaction language with active tokens, we designed a user elicitation study that prompts users to suggest gestures for various tasks within a particular task scenario (i.e. model world) of personal genomics. Personal genomics engages the study of individual (typically human) genomes.

We chose personal genomics as an application area for several reasons. First, user interfaces for personal genomics could prospectively benefit from the novel HCI techniques we investigate here [43]. Second, we were motivated by the increasing public interest and media coverage of personal genomics, broadening its relevance to both experts and lay people. Third, the genes chosen for the study task, BRCA1 and BRCA2, received extensive news coverage reflecting actress Jolie's medical decision around the time of the study. While personal genomics was chosen as a context for this study, the tasks investigated were selected from taxonomy of search interactions and could be generalized to various contexts. The chosen application area did not pose observed difficulties to our participants (college students). Participants typically did not have domain knowledge beyond general familiarity with basic genetic concepts.

Users where asked to perform eight distinct commands for exploring the human genome, with selection based on a taxonomy of search interactions [38]. Four of the commands were repeated throughout the task. These included select an item from a list; display information from the cube onto the surface; modify and define a range; associate information to a cube; create a query; zoom in; make a query more specific; replace part of a query, and create a compound query. Table 2 lists the eight commands for which participants were asked to choose a gesture. The order of the commands was determined by the task, which mirrors a real-world scenario of genomic investigation.

To avoid restricting users' gesture interactions, we presented users with a prompt asking them to perform a gesture for a particular command (e.g., select an item from a list). Only upon confirming the gesture, we displayed the effect of that command. To further avoid bias [10], no elements specific to a particular platform (e.g., Microsoft PixelSense or Sifteo) were presented.

The system did not recognize users' gestures, but did log sensor information, touch events, and neighboring events generated by the cubes. Participants used the think-aloud protocol. Video from these sessions was recorded. Participants also supplied subjective preference ratings. By using a think-aloud protocol, we obtained rich qualitative data that illuminates users' mental models. The logged data and questionnaires supplied quantitative measures regarding gesture performance, timing, and user preferences. The results provide a detailed description of user-defined gestures for the manipulation of large data sets, which combine active physical tokens with multi-touch interaction, and insight into the mental models accompanying them.

Study Design

We conducted a within-subjects design with 19 users. They were asked to complete a query-building task in two conditions: 1) using six active tokens combined with a vertical multi-touch interface; 2) using six active tokens combined with a horizontal multi-touch interface. The task was identical across both conditions, with the order of conditions counterbalanced to account for learning. A within-subjects design was employed to solicit gestures for both form factors (vertical and horizontal), and to explore the impact of a particular form factor on user-defined gestures. We decided to compare horizontal and vertical form factors. This was motivated by observations from pilot studies where we noticed that cubes tend to stay on the surface, occluding the workspace when a horizontal surface is used. Our goal was to test for differences between these two form factors.

Device	Mean (SD)
Smart phones	4.11 (0.83)
Tablets	3.44 (1.04)
Accelerometers	2.06 (1.21)
Sifteo cubes	1.11 (0.32)
Xbox Kinect	1.17 (0.71)

Table 1: Technology experience demographic questionnaire.

Experimental Task

To elicit task-driven gestures for querying large data sets using an interactive surface with active tokens, we created a casual genome browser application (see Figure 1). The application enables users to search the human genome for a particular gene and find related scientific publications; we prompted users to search for BRCA1 and BRCA2 genes.

The task began with the selection of a chromosome on a cube (i.e., active physical token). Results were displayed on the surface. Next, participants zoomed in to a gene of interest, and saved this back to a cube. The task was then repeated for another gene on another chromosome. Next, the participant was guided through a series of steps to formulate a publication query by retrieving publications on a particular gene, defining the publication date parameter, adding a Boolean operator and another gene to the query, and saving the results to a cube. Following the guided query formula-

tion, participants were asked to formulate a new compound query on their own and save the results.

This task was designed to examine eight commands for the manipulation of large data sets and query formulation. We selected these commands based on existing taxonomies for search interactions [38]. Users manipulated both discrete and continuous query parameters. Table 2 shows the eight commands with corresponding user task description.

Procedure

Participants were greeted, briefed about the goal of the study, and given a standard consent form. Users were then introduced to the Sifteo cubes with the official demo video [1] followed by five minutes of "free-play." Our decision to expose users to the Sifteo instruction video resulted from our pilot study, which showed that Sifteo gestures suffered from low discoverability (partly attributable to limited public exposure, esp. outside of a gaming context). The video provides users with additional possibilities for action beyond gestures users are familiar with based on their experience using phone and tablet interfaces. Playing with Sifteo cubes prior to the experimental task allowed users to explore gesture possibilities independently of the task context.

Participants were introduced to and walked through the process using a simple demo task. Following introduction, the interactive surface and active physical tokens displayed still images, serving as a starting point for the task. Users were read a prompt describing a subtask (e.g. select chromosome 13 from the cube and display it on the surface) and asked to perform a gesture that would accomplish this subtask. Prompts were read from a standard script for consistency. Users were instructed to think aloud and confirm their gesture for each subtask verbally. The surface application did not respond in real-time to any touch or gesture, and the Sifteo application provided only basic visual feedback: offsetting the image depending on the tilt, rotating the image when a cube was rotated, and magnifying the image if the screen was clicked. Upon confirming their gesture, participants were shown two 5-point scales and asked to rate the gesture goodness and ease. The effect of the gesture would then be presented on both the surface and the Sifteo cubes: a still image illustrating the state of the application following the user's action.

The complete experimental task was comprised of 11 outlined sub-tasks that tested eight query commands. Upon completion of the experimental task, users were asked to repeat the task on a different condition. Condition order was counterbalanced to account for learning. For each session, log and video data were collected. An observer tracked the gestures generated by participants while a second observer transcribed user comments.

Participants

19 participants (15 female) ranging in age from 18 to 25 with an average age of 20, volunteered for the study. All participants were compensated for their time with a five-

dollar gift card. Before the study, users completed a questionnaire about their exposure to multi-touch and gesture-based interfaces. Results are shown in Table 1.
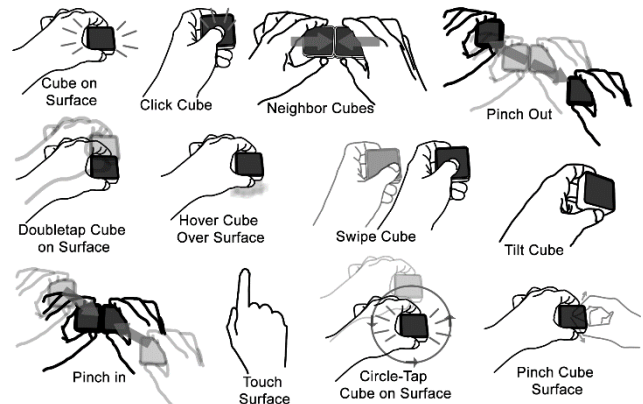

Figure 2: User-generated gesture vocabulary.

Apparatus

The study was conducted using six Sifteo Cubes that served as active physical tokens. A first generation Microsoft Surface measuring 30 inches at 1024 by 768 resolution was used as a horizontal surface. A Microsoft PixelSense measuring 40 inches at 1920 by 1080 resolution was used as a vertical surface. Our decision to use different platforms (Microsoft Surface vs. Pixelsense) was based on the availability and ease of setup of the devices. The surface interfaces were resized accordingly for each resolution to maintain a comparable user experience. We used Sifteo Cubes to accelerate realization of an initial prototype, as early commercial reprogrammable examples of active tokens. However, our results could be generalized to active tokens with similar functionality using other commercially available technologies (e.g. small Android devices or a future iOS iPod Nano). Figure 1 shows the experimental setup.

The Sifteo application gave users some feedback for interaction such as offsetting the image depending on the tilt, rotating the image if two are neighbored, or magnifying the image if the screen is clicked; but no menu structure was presented so as not to affect the users' mental model of information association on the cubes. It also collected accelerometer data, touch events, and neighboring events for gesture analysis. The Sifteo application was implemented using Sifteo 1.0 (written in C#); the surface application was identical for both conditions. It displayed still images as feedback for user gestures. It was built on the Surface 1.0 (for Horizontal condition) and Surface 2.0 SDK (for Vertical condition) and was written in C#.

DATA ANALYSIS AND RESULTS

In this section, we present qualitative and quantitative results from our gesture elicitation study.

User-Defined Gestures

Overall, 19 participants proposed a total of 560 gestures, for eight commands, of which 21 were distinct (considering

distinctness on a per-command basis). Figure 2 and Table 2 show proposed gestures per command for each condition with a consensus threshold of two (i.e., with repetition by at least two participants).

Command	Gesture	H	V
Select an item from a list	Swipe cube screen	45%	40%
	Tilt cube	39%	40%
	Cube on Surface	5%	9%
Display information from the cube onto the surface	Cube on Surface	71%	72%
	GUI-WIMP	21%	17%
	Hover cube over surface	0%	4%
	Click cube	6%	3%
Modify and define a range	Cube on Surface	26%	29%
	Pinch 2 cubes out	9%	17%
	Tilt cube	3%	14%
	Pinch 2 cubes in	14%	0%
	Pinch on cube	9%	17%
	Touch surface	6%	9%
	Hover cube over surface	6%	9%
	Click cube	9%	6%
Associate information to a cube	Cube on Surface	35%	38%
	Click cube	28%	28%
	GUI-WIMP	7%	10%
	Double tap cube on surface	10%	7%
	Hover cube over surface	7%	6%
	Circle tap cube on surface	4%	4%
	Touch surface	3%	4%
Create a query	Cube on Surface	23%	33%
	Click cube	27%	22%
	GUI-WIMP	18%	6%
	Tilt cube	5%	17%
	Touch surface	9%	17%
Zoom in - make a query more specific	Cube on Surface	14%	31%
	Tilt cube	7%	19%
	Neighbor cubes	21%	13%
	Click cube	21%	13%
Shift - Replace part of a query	Neighbor cubes	43%	31%
	Cube on Surface	14%	38%
Create a compound query	Neighbor cubes	25%	29%
	Cube on Surface	25%	14%
	Click cube	25%	5%

Table 2: Gestures generated per command by condition: H (horizontal) and V (vertical). GUI-WIMP refers to gestures such as tapping a button or using a slider on the cube.

While Table 2 shows popular gestures, it is not conflict free. The same gestures were observed to be reused for similar commands. For example, conflicting gestures were suggested for "displaying information from a cube onto the surface" and "associating information to a cube".

Agreement Performance and Preference

We used the formula introduced by Wobbrock et al. [57] for calculating an agreement score for each of the eight commands listed in Table 2. We also calculated and compared the means of the goodness and ease scores obtained from user preference questionnaires. Table 3 shows agreement, goodness, and ease scores for each of the eight commands used in the elicitation study per condition. We did not find significant differences between conditions in terms of agreement, goodness, and ease of gestures. Overall, we found that when simple gestures were mapped to simple commands they had more agreement and higher goodness.

Commands	Agreement		Good Mean SD		Ease Mean SD	
	H	V	H	V	H	V
Select an item from a list	0.36	0.33	4.26 *1.05*	4.00 *1.00*	4.58 *0.84*	4.32 *1.11*
Display information from the cube onto the surface	0.54	0.56	4.47 *1.07*	4.16 *1.21*	4.63 *0.96*	4.37 *1.07*
Modify and define a range	0.12	0.16	4.00 *0.88*	3.55 *0.96*	4.13 *1.01*	3.97 *0.82*
Associate information to a cube	0.23	0.24	4.34 *1.05*	4.29 *1.10*	4.58 *1.04*	4.42 *1.02*
Create a query	0.18	0.21	4.37 *0.83*	3.53 *1.31*	4.37 *1.01*	3.89 *1.05*
Zoom in - make a query more specific	0.15	0.19	3.79 *0.98*	3.58 *1.35*	4.00 *0.88*	3.58 *1.22*
Shift - Replace part of a query	0.20	0.17	3.84 *1.12*	3.58 *1.31*	4.05 *1.03*	3.74 *0.87*
Create a compound query	0.17	0.17	3.53 *1.17*	3.58 *1.12*	3.79 *1.13*	3.47 *1.12*

Table 3: Gesture agreement, goodness, and ease by condition: H (horizontal) and V (vertical).

Gesture classification

In order to better understand the design space of active tokens with multi-touch and tangible interfaces, we manually classified each gesture along three dimensions: space, flow, and cardinality. While there are numerous gesture taxonomies in the literature, our focus is on characterizing dimensions that are mostly relevant for the manipulation of active physical tokens within multi-touch and tangible interfaces.

Interaction *space* describes where a gesture was performed: *on-surface* – upon the interactive workspace, *on-bezel* – in the area surrounding the interactive surface, or *in-air* – above or in-front the interactive surface. Some gestures have a hybrid interaction space; for example initiated in-air and concluded on-surface. The dimension of *flow* is adopted from [57]; within this dimension *discrete* refers to a gesture where a computationally mediated response is expected

after the user acts, and *continuous* refers to a gesture where response is expected while the user acts. To accomplish more complex tasks, some gestures had combined flow; for example, one hand was sliding along the screen while another was clicking a cube screen.

Cardinality indicates the number of hands and tokens involved in a gesture. We regard *atomic* gestures as performed by one hand on a single token, and *compound* gestures as comprising a sequence of atomic gestures and involving one or two hands interacting with multiple tokens. Parallel gestures are bimanual and consist of two atomic gestures performed at the same time. Guiard's Kinematic Chain theory [13] provides a theoretical basis for the understanding and design of two-handed interactions, also several HCI studies have explored cooperation of the hands (e.g. Hinckley [14]). We intend to continue to explore the design space of two-handed interaction with active tokens and multi-touch interfaces.

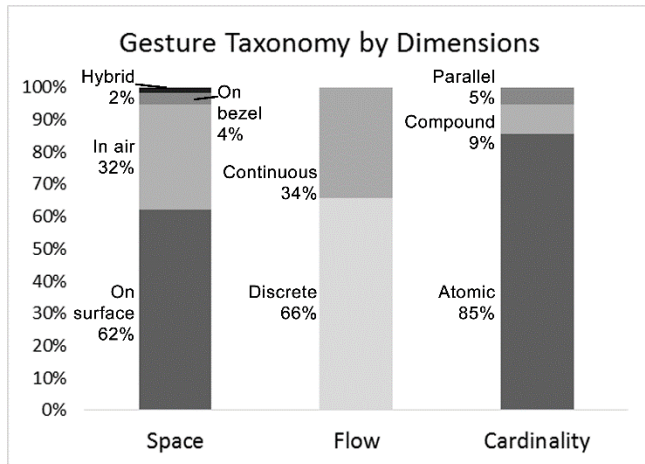Figure 3 shows percentage of gestures in each category per dimension.

Figure 3: Gesture Taxonomy presented as percentage per dimension

Mental Model Observations
Common Metaphors – the 'think aloud' protocol provided us with insight into participants' mental models. We identified several common metaphors that were suggested by multiple participants. Users often used the active tokens (i.e. cubes) as containers, dragging information from the display surface into the cube, displaying information from the cube by placing it on the surface, or tilting the cube above the surface. Several users treated the cubes as physical knobs for information presented on the surface. For example, one user described: "put [the cube] on the BRCA1 gene and then rotate the cube to bring up different data sources…" Some participants used active tokens as tools for indirect touch, such as a pen or a marker; for example: "I will circle 2007 with the drawing edge of the cube." The use of naïve physics metaphors to describe the way the active tokens (i.e. cubes) operate was also common. For ex-

ample, in the words of users: "Depending on [the angle of] the tilt, the speed will be faster or slower…"; "Tilt it again so that the lower number would gravitationally fall…"

Impact of Experience with Existing Interfaces – we observed that gestures suggested by users were heavily influenced by participants' experiences with traditional GUIs and with multi-touch interfaces. For example, one popular gesture was to swipe a finger on top of a cube's display, other common gestures included pinching-in and out with two fingers on the face of a cube. Users also suggested that a keypad or keyboard appear on the cube, as well as various buttons such as 'ok' and 'save.' One user described, "Instinctively, I just want to zoom in like that [pinches out] and it is probably because I have a Mac." Another user said, "Try to swipe the cube like on an iPad or on a phone." These observations are aligned with findings from previous studies indicating that users' creativity in suggesting gestures for new form-factors are impacted by their experiences with existing interface technologies [10, 28, 57].

Display orientation – while we did not find significant differences between horizontal and vertical conditions in terms of agreement, ease, and goodness (i.e. fit) (see Table 3), several users expressed strong preference toward the use of active tokens with a horizontal surface. Reasons include hand fatigue, effort, and persistence. For example, one user said, "I thought it was easier on the horizontal. I felt like with the vertical, I couldn't put the cubes down." Another user mentioned, "I like the table better than the PixelSense. It felt weird tapping a cube to an upright screen. You can drag things around and leave them there. More intuitive." A different user described, "The horizontal surface feels more personal, it's like you own space. When it moved to the vertical surface, I was more reluctant to use more than one cube at a time because that requires more effort ..."

Possibilities for action –The minimal visual feedback (and the absence of any other form of feedback) combined with the simple and clean design of the cubes provides only minimal real and perceived affordances. Several users commented on the difficulty of coming up with new ideas for interaction. In the words of one user, "It's hard to come up with a way to use this cube"; and in the words of another user, "I really wouldn't use these cubes at all. I don't find them intuitive." One explanation to these views is that the form of the Sifteo cubes did not provide users with sufficient clues regarding possibilities for action.

DISCUSSION
In this section, we describe the implications of our results for the design of multi-touch and tangible interfaces with active tokens. While we elicited user-defined gestures for active physical tokens within a particular model world of personal genomic investigation, the design considerations we discuss below could apply to many domains; For example, Twitter data mining, or the exploration of medical records. Following, we describe implications for design, which are based on our findings, and intended to inform expert

designers in the creation of novel interactions combining active tokens with multi-touch and tangible interfaces.

Implications for design

Several implications for the design of novel interactions that combine active tokens with interactive surfaces have emerged based on our findings:

Interaction beyond the surface – roughly 30% of gestures in both conditions were performed in-air, around 5% of gestures performed on the bezel, while 2% were hybrid (e.g. initiated off the surface and concluded on-surface). This suggests interaction above, in-front, or next-to an interactive surface might profitably be considered by designers along multi-touch interaction. Designing interactions beyond the surface could help designers to overcome challenges common to data-intense multi-touch and tangible interfaces such as finger size, occlusion, and clutter. In our study, participants employed in-air and on-bezel gestures for selecting and setting query operators, modifying and defining ranges, and formulating queries.

Continuous interaction – roughly 40% of gestures operating active tokens were continuous. Among these, hovering above the surface, rotating a cube, and scrolling a cube were all common interactions. This suggests that users might value immediate response and continuous control over information. The use of parallel and compound gestures, combined with on-bezel and hybrid interactions (see Figure 3) suggests that system designers might also consider using active tokens in control panel-like structures (as in [47]). Such structures could enable users to gain fine control of the displayed information while avoiding occlusion and fat finger challenges.

Gesture reuse – we found that participants occasionally reused gestures for multiple commands. For example, the gesture of placing a cube on the surface was suggested for each of the commands; however, the target on the surface was different for each command. For example, placing a cube on a chromosome start region was suggested to bind modifying functionality to the cube. Alternately, placing a cube on the gene BRCA1 was suggested to cause the gene to be saved to that cube. Gesture reuse and consistency are important for increasing learnability and memorability. Multi-touch gestures often suffer from low memorability, which could be improved using active tokens combined with contextual constraints to disambiguate gestures. The notion of interpreting an action based on a combination of token and constraints is discussed in the TUI literature [42, 49] and should be applied when using active tokens with multi-touch and tangible interfaces.

Affordances, metaphors, and constraints – we observed that the design of the cubes combined with the lack of feedback in our study did not provide users with sufficient affordances. Affordances denote the possibilities for action; inviting and allowing specific actions [33, 53]. Norman distinguishes perceived affordances, which are only visually

conveyed and rely on the interpretation of images, from real physical affordances [33]. The power of tangible interaction lies in providing both real and perceived affordances [41]. When integrating active physical tokens with multi-touch and tangible interfaces, designers should conduct a careful investigation of tokens' affordances, considering both perceived (i.e. visual) and physical affordances. Variations in size, shape, and material of a token as simple as a cube can affect the ways in which users handle it. We found that with the lack of sufficient affordance users' expectations were deeply influenced by traditional GUIs and multi-touch interfaces. For example, some users used the cubes as means for indirect touch. Users also suggested gestures that draw upon reality-based [18] metaphors such as naïve physics, containers, and knobs. Constraints (physical or digital), which restrict and provide context to possible actions, work in tandem with affordances [8]. Designers could employ both physical and digital constraints (e.g. visual targets) to guide users through sequences of action [8, 42, 49].

Limitations and Future Work

Our study has several limitations that point toward future work. First, we aimed to gain insight into users' behavior and preferences without the bias caused by computationally-mediated feedback. Hence, we removed the dialogue between the user and the system; instead creating a monologue, in which users' behavior is always accepted. This approach has several weaknesses; such as the absence of guiding cues, which makes it difficult for users to imagine possible gestures, and the lack of opportunities for users to reflect back on their gestures and change the interactions they proposed. In the words of one user: "It was really weird trying to build something from the ground up because I'm more used to relying on the feedback of whatever I'm interacting with to know what my next move is."

Also, the use of active physical tokens offers opportunities for supporting collaborative work with big data around an interactive surface or between distant surfaces; however, in this study we did not investigate a collaborative scenario. Rather, we focused on the engagement of a single user. It is possible that a collaborative scenario would highlight additional classes of gestures – e.g. proximal (but non-contiguous or coordinated across distant surfaces). These issues are worthy of investigation, but are beyond the scope of the current work. In the future, we intend to draw on our findings to design, implement, and evaluate a multi-touch and tangible interface for querying genomic data.

CONCLUSION

In this paper, we conducted a study with 19 users, which investigated gestural interaction with active-tokens. Our work provides four main contributions: 1) a vocabulary of user-defined gestures for interaction with active tokens; this gesture set extends familiar multi-touch gestures and includes gestures on-surface, in-air, and on-bezel; 2) characterization of the design space of gestural interaction with active tokens; 3) insight into participants' mental models

including common metaphors, sources of influence, and a strong preference of horizontal (vs. vertical) surfaces; and 4) implications for the design of multi-touch and tangible interfaces with active tokens for exploring large data sets. We intend to apply our findings in the design of tangible interface for the exploration of large genomic data sets.

ACKNOWLEDGMENTS

REFERENCES

1. (n.d.). Retrieved from Sifteo: https://www.sifteo.com/

2. Anthony, L., Brown, Q., Nias, J., Tate, B., & Mohan, S. (2012). Interaction and recognition challenges in interpreting children's touch and gesture input on mobile devices. In Proc.of the ITS, ACM.

3. Antle, A. N., & Wise, A. F. (2013). Getting down to details: Using theories of cognition and learning to inform tangible user interface design. Interacting with Computers, 25(1).

4. Antle, A. N., Droumeva, M., & Ha, D. (2009). Hands on what?: comparing children's mouse-based and tangible-based interaction. In Proc. of IDC, ACM.

5. Block, F., Horn, M. S., Phillips, B. C., Diamond, J., Evans, E. M., & Shen, C. (2012). The DeepTree Exhibit: Visualizing the tree of life to facilitate informal learning.Visualization and Computer Graphics, IEEE, 18(12).

6. Camarata, K., Do, E. Y. L., Johnson, B. R., & Gross, M. D. (2002). Navigational blocks: navigating information space with tangible media. In Proc. of IUI, ACM.

7. Coelho, M., Zigelbaum, J., & Kopin, J. (2011). Six-forty by four-eighty: the post-industrial design of computational materials. In Proc. of TEI. ACM.

8. Djajadiningrat, T., Overbeeke, K., & Wensveen, S. (2002). But how, Donald, tell us how?: on the creation of meaning in interaction design through feedforward and inherent feedback. In Proc. of DIS, ACM.

9. Drucker, S., Fisher, D., Sadana, R., Herron, J., & Schraefel, M. C. (2013). TouchVix: A Case Study Comparing Two Interfaces for Data Analytics on Tablets. In Proc. of CHI, ACM.

10. Epps, J., Lichman, S., & Wu, M. (2006). A study of hand shape use in tabletop gesture interaction. In Proc. of CHI extended abstracts, ACM.

11. Findlater, L., Lee, B., & Wobbrock, J. (2012). Beyond QWERTY: augmenting touch screen keyboards with multi-touch gestures for non-alphanumeric input. In Proc. of CHI, ACM.

12. Good, M. D., Whiteside, J. A., Wixon, D. R., & Jones, S. J. (1984). Building a user-derived interface. Communications of the ACM, 27(10), 1032-1043.

13. Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. Journal of motor behavior, 19, 486-517.

14. Hinckley, K., Pausch, R., Proffitt, D., Patten, J., & Kassell, N. (1997). Cooperative bimanual action. In Proc. of CHI, ACM.

15. Horn, M. S., Tobiasz, M., & Shen, C. (2009). Visualizing biodiversity with voronoi treemaps. In Voronoi Diagrams, IEEE.

16. Hornecker, E., Marshall, P., Dalton, N. S., & Rogers, Y. (2008). Collaboration and interference: awareness with mice or touch input. In Proc. of CSCW, ACM.

17. Isenberg, P., Isenberg, T., Hesselmann, T., Lee, B., Von Zadow, U., & Tang, A. (2013). Data visualization on interactive surfaces: A research agenda. Computer Graphics and Applications, IEEE, 33(2), 16-24.

18. Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces. In Proc. of CHI, ACM.

19. Jetter, H. C., Gerken, J., Zöllner, M., Reiterer, H., & Milic-Frayling, N. (2011). Materializing the query with facet-streams: a hybrid surface for collaborative search on tabletops. In Proc. of CHI, ACM.

20. Jetter, H.-C., Reiterer, H., and Geyer, F. (2013). Blended Interaction: understanding natural human–computer interaction in post-WIMP interactive spaces. Personal and Ubiquitous Computing.

21. Klum, S., Isenberg, P., Langner, R., Fekete, J. D., & Dachselt, R. (2012). Stackables: combining tangibles for faceted browsing. In Proc. of AVI. ACM.

22. Lahey, B., Girouard, A., Burleson, W., Vertegaal, R., and Canada, K.L.(2011). PaperPhone : Understanding the Use of Bend Gestures in Mobile Devices with Flexible Electronic Paper Displays. In Proc. of CHI, ACM.

23. Marshall, P., Hornecker, E., Morris, R., Sheep Dalton, N., & Rogers, Y. (2008). When the fingers do the talking: A study of group participation with varying constraints to a tabletop interface. In Proc. of ITS, IEEE.

24. Merrill, D., Kalanithi, J., & Maes, P. (2007). Siftables: towards sensor network user interfaces. In Proc. of TEI, ACM.

25. Micire, M., Desai, M., Courtemanche, A., Tsui, K. M., & Yanco, H. A. (2009). Analysis of natural gestures for controlling robot teams on multi-touch tabletop surfaces. In Proc. of ITS, ACM.

26. Microsoft. (n.d.). Microsoft PixelSense. Retrieved from Microsoft PixelSense: http://www.microsoft.com/en-us/pixelsense/default.aspx

27. Microsoft Surface. (n.d.). Retrieved from www.microsoft.com/surface

28. Morris, M. R. (2012). Web on the wall: insights from a multimodal interaction elicitation study. In Proc. of ITS, ACM.

29. Morris, M. R., Brush, A. B., & Meyers, B. R. (2007). Reading revisited: Evaluating the usability of digital display surfaces for active reading tasks. In Proc. of ITS,2007, IEEE.

30. Morris, M. R., Fisher, D., & Wigdor, D. (2010). Search on surfaces: Exploring the potential of interactive tabletops for collaborative search tasks. Information processing & management, 46(6), 703-717.

31. Nature Specials: Big Data. (2013). Retrieved from Nature: http://www.nature.com/news/specials/bigdata/index.html

32. Nielsen, M., Störring, M., Moeslund, T. B., & Granum, E. (2004). A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In Gesture-Based Communication in HCI. Springer Berlin Heidelberg.

33. Norman, D. A. (1999). Affordance, conventions, and design. Interactions, 6(3).

34. Patten, J., & Ishii, H. (2000). A comparison of spatial organization strategies in graphical and tangible user interfaces. In Proc. of DARE, ACM.

35. Pyryeskin, D., Hancock, M., & Hoey, J. (2012). Comparing elicited gestures to designer-created gestures for selection above a multi-touch surface. In Proc. of ITS, ACM.

36. Rekimoto, J., Ullmer, B., & Oba, H. (2001). DataTiles: a modular platform for mixed physical and graphical interactions. In Proc. of CHI, ACM.

37. Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). The cost structure of sensemaking. In Proc. of CHI, ACM.

38. Sacco, G. M., & Tzitzikas, Y. (2009). Dynamic taxonomies and faceted search (Vol. 25). Springer.

39. Schkolne, S., Ishii, H., Schroder, P. (2004). Immersive design of DNA molecules with a tangible interface. Visualization , IEEE, pp. 227- 234.

40. Seyed, T., Burns, C., Costa Sousa, M., Maurer, F., & Tang, A. (2012). Eliciting usable gestures for multi-display environments. In Proc. of ITS, ACM.

41. Shaer, O., & Hornecker, E. (2010). Tangible user interfaces: past, present, and future directions. Foundations and Trends in HCI, 3(1–2).

42. Shaer, O., Leland, N., Calvillo-Gamez, E. H., & Jacob, R. J. (2004). The TAC paradigm: specifying tangible user interfaces. Personal and Ubiquitous Computing, 8(5).

43. Shaer, O., Mazalek, A., Ullmer, B., & Konkel, M. (2013). From Big Data to Insights: Opportunities and Challenges for TEI in Genomics. In Proc. of TEI. ACM.

44. Shaer, O., Strait, M., Valdes, C., Feng, T., Lintz, M., & Wang, H. (2011). Enhancing genomic learning through tabletop interaction. In Proc. of CHI, ACM.

45. Shaer, O., Strait, M., Valdes, C., Wang, H., Feng, T., Lintz, M., ... & Liu, S. (2012). The design, development, and deployment of a tabletop interface for collaborative

exploration of genomic data. International Journal of Human-Computer Studies, 70(10), 746-764.

46. Thudt, A., Hinrichs, U., & Carpendale, S. (2012). The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In Proc. of CHI, ACM.

47. Ullmer, B., Dell, C., Gill, C., Toole, C., Wiley, C., Dever, Z., . . . Parker, R. (2011). Casier: Structures for Composing Tangibles and Complementary Interactors for User Across Diverse Systems. In Proc of TEI, ACM.

48. Ullmer, B., Dever, Z., Sankaran, R., Toole, C., Freeman, C., Cassady, B., . . . Parker. (2009). Cartouche: conventions for tangibles bridging diverse interactive systems. In Proc of TEI, ACM.

49. Ullmer, B., Ishii, H., & Jacob, R. J. (2005). Token+ constraint systems for tangible interaction with digital information. ACM TOCHI, 12(1), 81-118.

50. Ullmer, B., Ishii, H., & Jacob, R. J. (2003). Tangible query interfaces: Physically constrained tokens for manipulating database queries. In Proc. of Interact.

51. Valdes, C., Ferreirae, M., Feng, T., Wang, H., Tempel, K., Liu, S., & Shaer, O. (2012). A collaborative environment for engaging novices in scientific inquiry. In Proc. of ITS, ACM.

52. van den Hoven, E., Mazalek, A. (2011). Grasping gestures: Gesturing with physical artifacts. AI EDAM 25(3).

53. Vermeulen, J., Luyten, K., van den Hoven, E., and Coninx, C. (2013). Crossing the bridge over Norman's gulf of execution: revealing feedforward's true identity. In Proc. Of CHI, ACM.

54. Vogel, D., & Baudisch, P. (2007). Shift: a technique for operating pen-based interfaces using touch. In Proc of CHI, ACM.

55. Voida, S., Tobiasz, M., Stromer, J., Isenberg, P., & Carpendale, S. (2009). Getting practical with interactive tabletop displays: designing for dense data, fat fingers, diverse interactions, and face-to-face collaboration. In Proc. of the ITS, ACM.

56. Weiser, M. (1999). The computer for the 21 st century. ACM SIGMOBILE, 3(3), 3-11.

57. Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009). User-defined gestures for surface computing. In Proc. of CHI, ACM.

58. Wu, M., Shen, C., Ryall, K., Forlines, C., & Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In Proc. of HIHCS,

59. Xu, W., Chang, K., Francisco, N., Valdes, C., Kincaid, R., & Shaer, O. (2013). From wet lab bench to tangible virtual experiment: SynFlo. In Proc. of TEI. ACM.

60. Zigelbaum, J., Horn, M. S., Shaer, O., & Jacob, R. J. (2007). The tangible video editor: collaborative video editing with active tokens. In Proc. of TEI, ACM.