Exploring Narrative Gestures on Digital Surfaces

Mehrnaz Mostafapour, Mark Hancock
Department of Management Sciences
University of Waterloo, Canada
first.last@uwaterloo.ca

ABSTRACT
A significant amount of research on digital tables has traditionally investigated the use of hands and fingers to control 2D and 3D artifacts, has even investigated people's expectations when interacting with these devices. However, people often use their hands and body to communicate and express ideas to others. In this work, we explore narrative gestures on a digital table for the purpose of telling stories. We present the results of an observational study of people illustrating stories on a digital table with virtual figurines, and in both a physical sandbox and water with physical figurines. Our results show that the narrative gestures people use to tell stories with objects are highly varied and, in some cases, fundamentally different from the gestures designers and researchers have suggested for controlling digital content. In contrast to smooth, pre-determined drags for movement and rotation, people use jiggling, repeated lifting, and bimanual actions to express rich, simultaneous, and independent actions by multiple characters in a story. Based on these results, we suggest that future storytelling designs consider the importance of touch actions for narration, in-place manipulations, the (possibly non-linear) path of a drag, allowing expression through manipulations, and two-handed simultaneous manipulation of multiple objects.

INTRODUCTION
Storytelling is an expressive form of art that can empower the expression of thoughts, beliefs, and emotions through narrative. The idea of storytelling often evokes thoughts of common media such as books, movies, and video games, but people tell anecdotes to one another every day around dinner tables, campfires, and water coolers. When narrating a story, people often make gestures with their hands, arms, and body to enhance the story, to build suspense, to exaggerate emotion, or to simply better engage the audience.

Digital tables [4,9,16] are a promising medium through which these anecdotes could be told, as the audience and storyteller can gather around the table, much like they would at a dinner table or campfire, and they support the ability to perform gestures that are immediately observable to the audience. The digital surface can then be used as a supportive medium that the storyteller can adopt to further enhance the storytelling experience; that is, their storytelling gestures could be made to have a greater impact on their story. A narrator can use on-screen objects to set the scene of a story, to draw paths or elements relevant to the plot of the story, or to improve a description of a story's characters.

While there has been a significant amount of research exploring the use of gestures on digital tables [23,24,36,37], and the use of one's fingers and hands to move and rotate artifacts [10,21,30], the understanding of a "gesture" in the literature has largely focussed on the control of on-screen content. Nonetheless, this research attempts to develop an understanding of people's behaviour [30] and simple gestures that people expect to use on digital surfaces [23,36]. Hinrichs and Carpendale [15] have highlighted the need to support the expressive power of gestures; however, much is still not known about the use of digital tables to support the creative processes involved in storytelling. What kinds of gestures do people perform to convey meaning in a story? Do gestures used to control on-screen artifacts interfere with storytelling gestures? How does the digital medium differ from the physical in its storytelling potential?

Some research already explores the use of digital tables for storytelling [10,13,39]. These prototypes allow a person to create a story by manipulating on-screen 3D artifacts [10]. These designs show promise for the support of this creative process, and integrate natural gestures and physical interaction techniques to improve the storytelling experience. However, this research has not yet explored whether and how these prototypes are used to tell stories and the gestures people use to convey these narratives.

In this paper, we present an observational study of people creating a story on a digital table, in a physical sandbox, and in water. We focussed our observations on the physical gestures used to tell a story as people narrate, create dialogue, and move characters. Our findings suggest that narrative gestures are inherently different than regular gestures; people use them to convey meaning to the audience. Beyond simple movement and rotations, storytellers animate characters in a variety of ways to convey meaning, and combine these narrative gestures with two hands in ways typically not expected from traditional movement and rotation interaction techniques. Moreover, the way people combine movements from multiple hands has meaning that may interfere with the common use of multiple hands and fingers to move and rotate digital artifacts in 3D.

ITS'14, November 16–19, 2014, Dresden, Germany.

ACM 978-1-4503-2587-5/14/11…$15.00
http://dx.doi.org/10.1145/2669485.2669510

RELATED WORK

In this section we first review the-state-of-the-art surface-based storytelling prototypes. We then review previous research on non-object mediated narrative gestures and the research on exploring hand gestures used to navigate on a digital surface. We then review the research comparing the interaction between physical and digital objects.

Storytelling on Multi-touch Surfaces

Storytelling is a powerful means of communication that empowers people to express their thoughts, beliefs, and emotions through narrative [19,35]. Given the importance of storytelling, many research prototypes already support storytelling on digital surfaces [1,10,13,29,39]. These prototypes support a variety of storytelling features, such as the ability to move and rotate photos [1,13,39] or 3D toys [10], to draw on the background [2,6,29], to record the story being told [30], and the ability to add one's own content, for example images from an existing photo collection [39] or captured from the surrounding physical environment [1]. These storytelling prototypes and applications have enabled collaboration [1,13,39], and supported therapy for children [10] or persons with aphasia [39]. While the purpose of these applications is to enable users to create a story, there is little research on how narrators manipulate objects to enact character actions and story events. In this paper, we study the way people benefit from hand gestures to manipulate on-screen objects to illustrate story events, with the intent of informing the design of such applications.

Narrative Gestures

Narrative gestures are usually used in conjunction with speech to illustrate an event or to communicate meaning [20,22]. Using hand gestures not only helps to engage the audience, but also serves as a tool for narrators to better focus and think [6,7]. These gestures are categorized as: *iconics*, *metaphorics*, *deictics*, *beats*, and *butterworths* [1,20,22]. Iconics are used to resemble and illustrate a concrete object or event, for instance, a gesture that depicts the act of hitting may be synchronous with the utterance, "she hit him on the shoulder". Metaphorics are similar to iconics except that they are used to depict abstract concepts, such as an upward hand gesture accompanied by the utterance "his IQ is very high". Deictics are hand movements used to point to a particular element, for instance, a pointing gesture to a door while speaking about that door. Beats are used to punctuate and give emphasis to discourse, for instance, a very quick and steady hand movement accompanying the utterance "that's it". Butterworths correspond to speech failure [1,22], for example, hand movements while trying to recall something. While narrative gestures have been studied in the context of open-handed non-object mediated communicative gestures that accompany speech [1,22], our work extends this literature by investigating storytelling that involves manipulating objects (e.g., story characters), in particular when illustrating story events through manipulating tangible and on-screen objects.

Gestures on Multi-Touch Surfaces

Many studies have investigated gestures on multi-touch surfaces to understand and develop natural [23,24,36], ergonomic [25,26], or novel interaction techniques [10,32,37]. This research primarily focuses on understanding and designing interaction techniques for digital surfaces to perform common tasks (e.g., movement, scaling, etc.); however, these studies have not considered the context of illustrating a story or describing an object. The focus of our work is not to design specific interaction techniques, such as the number and combination of fingers used to interact with the objects [11,25,38], nor to identify people's expectations about what command a gesture should invoke [5,24], nor to develop multi-user gestural interaction,23], but to study the nature of interactions and physical actions used to perform story events and character actions in the context of narration. We thus observe physical and digital interactions of the storytelling process, using an exploratory approach similar to Hinrichs and Carpendale's [15].

Physical vs. Digital Interactions

Comparing the method of interaction with objects in a digital 2D and physical 3D space helps to incorporate methods of physical interaction into the design of multi-touch devices [33]. For instance, Scott et al. [31] studied collaborative interactions in the physical world to provide their territoriality framework that can be applied in the design of collaborative applications on digital tables. Terrenghi et al. [33] studied the nature of interactions in 3D and 2D by asking participants to sort pictures and complete a puzzle in a physical environment and on a digital table. Using a similar method, North et al. [28] compared gestures used to "manipulate many small objects, in three different interaction paradigms: physical, multi-touch, and mouse interaction," [28, p. 5] to understand the similarities and differences between the interactions used in these environments. These examples used lab studies, where they asked participants to complete tasks in different environments. In our work, we use a similar approach that compares how people manipulate objects in physical environments and on the digital table in the context of narration and storytelling.

OBSERVING STORYTELLING GESTURES

To investigate the use of narrative gestures, we focused our attention on the act of *illustrating a story*—a specific instance of storytelling where the narrator enacts character movements and story events by manipulating figurines—as in this act, people frequently demonstrate a variety of emotion and draw the audience in in a variety of ways. We thus focused our attention on the following research questions:

- *How does a narrator make use of gestures to illustrate story events?*
- *How do these gestures differ when the story is told in a digital space, rather than a physical one?*
- *How do the digital or physical artifacts in the story affect the gestures performed?*

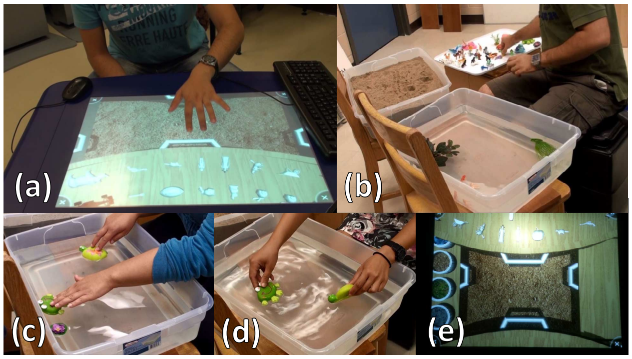

Figure 1. The setup of the study in (a) the digital condition, (b) the sand and water conditions. Participants would sometimes (c) touch objects or (d) grasp objects. A screenshot of the digital sandtray (e) had a figurine, paint, and resize drawer.

To provide a basis for developing interaction techniques in digital mediums, HCI researchers frequently study physical interactions [28,31,33,34]. Thus, to develop a better understanding of narrative gestures, we observed participants illustrating a story to the experimenter in one digital and two physical conditions. As a basis for the digital storytelling, we used an application designed to support storytelling—a digital sandtray [11,19]. We had participants tell a story in this digital medium, and the original physical setup of a sandbox. Because the digital medium only allows interaction in 2D, while physical sandboxes allow movement on and above the surface, we included another physical condition where interaction could occur in a 2D plane, water. Water allows for both forms of physical interaction; participants were able to either touch and push, or grasp and move objects to manipulate them (Figure 1c & 1d).

Participants

Twenty-nine university students aged 19 to 45 (*Mdn*=26) participated in our study (11 female). Six (21%) did not own and had never used a multi-touch device. Twenty-three (79%) owned a multi-touch device, such as a smartphone or tablet, twenty of whom (87%) for more than a year. Only two (7%) had worked with a digital table before.

Apparatus

Participants were asked to illustrate a story in one digital and two physical environments. In all conditions, we modelled the environment after sandtray therapy, a type of art therapy that provides clients with a tray of sand and a shelf full of figurines with which to tell a story. Clients use these figurines to create and tell a story to the therapist [19]. We chose this setting for several reasons: (1) this type of storytelling is already used in the practice of therapy, and so our results can directly inform this current practice, (2) an existing digital tabletop display application, was available and modelled directly after this physical practice [11], (3) this form of storytelling had already been refined by therapists to quickly engage the client in storytelling, and have the story take on personal meaning.

In the digital condition, participants created and told a story on a SMART Table, a rear-projected 92 cm × 74 cm multi-touch table with a resolution of 1024 × 768 pixels and a height of 64 cm. The software used for the study on the digital table was Hancock et al.'s [11] sandtray application (Figure 1a & 1e). This multi-touch sandtray application was built in Java and includes three drawers: a characters drawer that includes a set of figurines, a paint drawer that enables drawing on the background, and a resize drawer that allows resizing of figurines (Figure 1e).

This prototype supports the illustrating of a rich narrative by enabling narrators to move and rotate objects both in a 2D and 3D space. Narrators can move objects on the surface with one point of contact, and rotate them in a 2D space through two points of contact. These same two points can be used to lift or lower an object by spreading them apart or pinching them together. In order to rotate objects in a 3D space, narrators need to have two fixed points of contact and use a third touch to rotate the object in the space along any desired axis. For a more complete description of the these interaction technique, see [10,11].

In both sand and water, participants completed the story in two 90 cm×70 cm trays with a depth of 16 cm. The top edge of both trays was adjusted to be the height of the SMART Table. The seat was adjusted so participants could reach all available areas on the digital table and in the trays. A set of different toys was provided next to the trays (Figure 1b).

Note that a rabbit, a turtle, and a tree figurine were provided in all three environments, but the other physical and digital figurines were not similar. Two groups of toys were provided in the physical conditions, including a group of 6 decorative items and a group of 44 animal figurines. Some of the toys were specifically made for water (bath toys), so that they would float. Participants could choose any of the toys for both the water and sand conditions, regardless of whether they were bath toys (i.e., intended for use in water). Also 161 figurines were provided in the digital table from which 6 figurines were different types of tree and flower figurines. The figurines included some animals, some fictional characters (e.g., a Pegasus), some furniture figurines (e.g., a couch), and some transportation vehicles (e.g., an airplane).

Conditions

Our primary factor in the design of this study was the story telling *environment*, with three levels: *sand*, *water*, and *digital* table. We used a within-participants design where each participant was asked to tell a story in all three media. We included a secondary between-participants factor where half the participants were asked to stick to the script of the original story (fixed) and the other half were allowed to deviate in theme and plot (free-form); however, in practice, participants tended to ignore this request, with many in the fixed condition deviating frequently and many in the free-form condition sticking to the original plot. Thus, we did not consider this secondary factor in our analysis.

Task & Data Collection

All of the participants were given a short summary of the famous children's story "The Tortoise and the Hare" to read. In the given story, the rabbit was a boastful character that was challenged by the tortoise to a race. The rabbit lost the race because he decided to sleep along the way. Additional setting information (e.g., place, time, etc.) was not described to participants. Each participant started by reading the story, and then proceeded to illustrate his/her story in all three environments. The experimenter played the role of an audience member as each participant told his/her story, with the experimenter sitting in front of him/her and actively listening (i.e., displaying emotional responses, such as smiling/laughing at funny moments, making eye contact, and otherwise responding to the narrative), without physically interfering with the surface and objects. Note that all reactions were genuine, and no script or acting was used. While we recognize these reactions may have influenced participant behaviour, we believe it created a more realistic setting, and the absence of these reactions would have been more detrimental to our results (e.g., not laughing at a joke).

Each session was videotaped and participants completed a post-study questionnaire, which included demographic questions and asked participants to explain their comfort level while manipulating objects in different environments.

Data Elimination

Note that the results from five participants were eliminated from the data and all the presented analysis is done based on the data gathered from the 24 remaining participants. Four participants were eliminated because they only narrated stories in two environments, as they did not have enough time to complete the whole study in the allotted hour. The results of a fifth participant were eliminated because she was not easily able to work with the digital table, so the experimenter had to interfere. After elimination, the order of presentation of the storytelling environment was still balanced (4 participants per order), with the exceptions of the order sand, digital, water (5) and water, sand, digital (3).

GESTURE CLASSIFICATION

Character actions and story events are different from story to story; one story might be about characters who are climbing mountains while another story might be about characters who are sitting in a room and talking to each other. Therefore, in order to analyze how participants exploit possible actions (e.g., lifting, rotation, and dragging) to manipulate objects and enact character actions, we selected a set of story events and character actions that were common among all the stories told. We thus followed a two-pass video analysis strategy suggested by Jacucci et al. [18].

Video Analysis and Gesture Classification

In the first pass, we watched all the study sessions and identified the actions commonly used by all participants:

Dialogue: In all stories, there was always at least one conversation between two or more characters in which a story character was talking. For instance, when a narrator said, "The turtle said, 'Let's see who wins.'"

Narration: In all stories, there was always at least one point at which the narrator was explaining what was going on. In these moments, the narrator usually described a scene, a story event, or a character's thoughts, feelings, etc. For instance, when a narrator said, "The animals decided to come and watch the race," or, "The rabbit was very angry."

Character movements: In all stories, there was at least one character who moved from one location to another.

Throughout this section, we use the term "narrate" to describe verbal utterances and "illustrate" to describe any narrative act that involves visual cues or physical action (e.g., object movement).

In the second pass, we used a grounded theory approach to elucidate and identify different categories of narrative gestures. However, our method of coding was similar to McNeill's [22]. We looked at the utterance and simultaneous hand gesture to see when a particular gesture was used. Throughout this section, we discuss when our gestures could be categorized using McNeil's terminology (iconics, metaphorics, deictics, beats, and butterworths [22]), but chose a grounded theory approach, since these gestures were not intended for gestures with objects, specifically.

Gestures for Dialogue and Narration

We observed that while participants were illustrating *dialogue* or *narration* they performed the following gestures:

Touch/Hold: Participants sometimes touched a character on the digital table, or touched/held a character in their hand in the physical conditions. This included any touches more than 2 seconds. We found that people sometimes touch/hold an object when talking about it. These types of gestures can be considered as deictic gestures that are used to point to an element. However, in this case narrators actually touched the object instead of just pointing to it. We also observed that participants touched or held objects when they were thinking about what to say or when they wanted to come up with a creative storyline to tell. In this case narrators usually touched an object even if that object was not related to what they were talking about. This could be due to the reason that touching/holding an object could help narrators to focus on what they were saying. This type of touch/hold gesture could be considered as butterworths that are used as an effort to recall a word or a sentence.

Jiggle: Participants sometimes touched/held an object while doing small up-down or right-left motions (Figure 3). These events were coded as jiggle actions, and not as touch/hold actions. Jiggling was mostly used to resemble talking, dancing, or emotions such as anger, happiness, etc. For instance, a participant jiggled the rabbit and accompanied it with the utterance "The rabbit said no way!" In these cases, jiggling

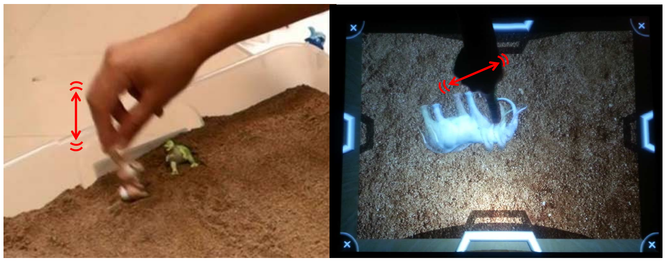

Figure 2. Jiggling

could be considered an iconic gesture through which narrators want to represent a particular meaning or event. However, jiggling could also be considered as a beat gesture, since it is used to emphasize a particular concept. Meaning that by jiggling, participants not only represented talking, but also emphasized *how* a character talks, (e.g., when the character was angry they jiggled faster).

Tap: Participants sometimes touched/held an object for less than 2 seconds. As our smallest unit to measure the time was a second, any touch gesture up to two seconds was considered as a tap. Tap could be considered a deictic gesture, since it was mostly used to point to particular elements of the scene (e.g. a scenic element or a story character).

Above surface hand gestures: Participants sometimes narrated the story while performing hand gestures above the surface (i.e. digital table, sand or water). These gestures include all five of McNeil's gestures [22].

As an example of how we coded dialogue one participant stated, "the rabbit said, 'I am much better at running,'" as she jiggled the rabbit to enact talking. We classified this gesture as a *jiggle* in the *dialogue* category. We followed the same method for coding narration.

Gestures for Character Movements

We also observed that *character movements* were performed through four types of gestures which could be considered iconic or beat gestures [1,20,22]:

Dragging: Moving an object while always in contact with the surface. This gesture could be considered iconic, as the narrator only wants to illustrate movement.

Dragging while jiggling: Moving an object forward while jiggling it. This gesture could be considered as either iconic or as a beat gesture, as the narrator not only illustrates that a character is moving but also emphasizes how it moves.

Lift and drag: In the physical conditions, participants sometimes picked an object up and moved it to another location. The same action could be performed on the digital table through lifting and dragging an object. This type of gesture could again be considered as an iconic gesture, as it only depicts that a character moves from one location to another without any visual information about how it moves.

Repeated lift and drag: Participants moved an object through several small/repeated lift and drags (e.g., to show that the rabbit is jumping to get to the finish line). Repeated lift and drag could effectively illustrate "hopping". Similar to dragging while jiggling, this gesture could also be considered as both iconic and as a beat gesture.

As an example of how we coded character movements, one participant said "the turtle went slowly but surely" and did slow, deliberate, left and right motions while moving the turtle, which we classified as: *dragging while jiggling* in the *character movement* category.

Two-Handed Interactions

In addition to coding gestures used for dialogue, narration, and character movements, we noticed that many participants performed interesting combinations of gestures with both hands (Figure 3). While we observed asymmetric bimanual actions [8] (e.g., the bimanual interaction required to rotate objects in the digital sandtray), of particular note were the two-handed gestures driven by the narrative being told, such as simultaneous drags to represent two characters racing. We thus coded examples of simultaneous gestures on two figurines. To simplify this analysis, we considered all touch, jiggle, tap, and rotate gestures as *in-place* actions and all character movements, including drag, drag and jiggle, repeated lift and drag, and lift and drag, as *move* actions. Therefore, three simultaneous bimanual actions were observed: *move+move*, *in-place+move*, *in-place+in-place*. Note that these gestures cannot easily be classified using the common HCI terminology of "symmetric" and "asymmetric" [8], since actions were sometimes half-way between (e.g., characters running at different speeds, or one character interrupting the dialogue of another).

We counted the number of instances that each gesture occurred for each action (dialogue, narration, or character movement). Sustained gestures were counted in 10 second intervals: a drag gesture held for 23 seconds would be counted as 3 drags (two 10 second and one 3 second drag).

RESULTS

We separate our analyses according to the codes identified in the first pass of analysis. Specifically, we separately con-

Figure 3. Two-Handed combination

sider gestures used to perform *dialogue and narration*, *character movements*, and *two-handed interaction*.

While the sample size in our tests is 24 and the data may not be normally distributed, parametric tests have been shown to be robust to violations of these assumptions [27], hence we used Repeated Measures Analyses of Variance (RM-ANOVAs) with storytelling environment and gesture as primary factors. We also included action (dialogue vs. narration) as a factor when analyzing these data, as the gestures used to perform these actions were often similar. We note the exact test used in each subsection.

Our dependent measure was the number of instances of each gesture, either as a raw count or normalized by condition. Normalized results appear as percentages (%), and were calculated using the number of instances of all gestures within that condition (e.g., water, sand, or digital) as the denominator. When this denominator was zero (i.e., no gestures were performed in that condition), we represented this as 0%. The decision to normalize had the effect of focusing the analysis on the differences in frequency of gesture within each condition (e.g., which gestures were used to illustrate), rather than on the number of instances across conditions (e.g., how many gestures were performed in sand vs. water vs. digital). We chose normalized analysis when investigating story-centred actions (dialogue, narration, and character movements) and raw counts when investigating interaction-centred actions (two-handed interaction). Note that mean differences between storytelling environments or any factor other than gesture are nonsensical for normalized data, since the conditions add up to 100%. We thus consider only main effects and interactions involving gesture in our normalized analyses.

Gestures for Dialogue and Narration

We analyzed dialogue and narration with a 3 environment (digital, water, sand) × 2 action (dialogue vs. narration) × 4 gesture (jiggle, touch, tap, above-surface) RM-ANOVA.

There was a significant main effect of gesture ($F_{3,69}$=15.97, p<.001), shown in Figure 4, green. Pairwise comparisons revealed that tapping was used significantly less than every other type of narrative action (p<.001) and that above-surface gestures were used significantly less than touching (p=.01), and less than jiggling, but this difference was not significant (p=.06). There was no significant difference between touching and jiggling (p=1.00). Thus, while some dialogue and narrative actions were represented using above the surface gestures, participants tended to prefer contact with the objects (i.e., touch/jiggle). It may be that contact with the story objects helps the narrator focus on the story being told. However, the effect of gesture can be further explained by the interactions.

Two-Way Interaction: Environment and Gesture

We found a significant interaction between environment and gesture ($F_{6,138}$=5.3, p<.001, Figure 4, blue). We performed post-hoc pairwise comparisons grouped by gesture.

Jiggling: We found jiggling to be used significantly more in sand than in water (p=.02). There was no significant difference between water and the digital table in terms of jiggling (p=1.00). This finding may suggest that participants had difficulty accomplishing jiggling on the digital table and in water, or that sand lends itself better to this action.

Touching/Holding: We also found that touching was used significantly more in water than sand (p=.02). However there was no significant difference between water and digital (p=.47), nor between sand and digital (p=1.00). We suspect that participants felt the need to hold objects in place in the water condition to prevent them from floating away.

Tapping: There were no significant differences between all the environments in terms of tapping (p=1.00); participants did not tend to tap much in any environment, as per the main effect. This may suggest that participants preferred to touch objects for longer than 2 seconds while narrating.

Above-surface: Participants used more above-surface hand gestures in digital than both sand (p=.01) and water (p<.01). This finding indicates a possible hesitation by participants when using the digital table vs. physical media; they felt a need to indicate dialogue or narration, but resisted using another gesture (jiggle, touch, or tap). This may be due to the Midas Touch phenomenon [12,14]. Interestingly, this phenomenon may have partially extended to (gritty)

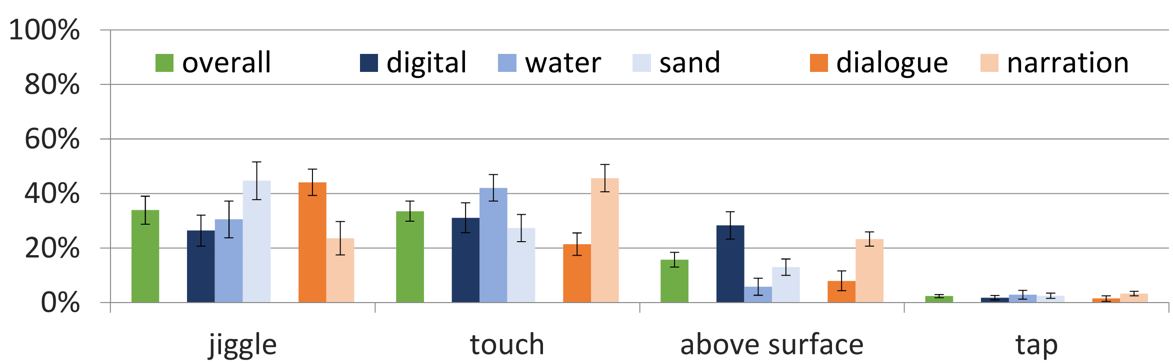

Figure 4. Main effect of gestures used for dialogue and narration (left, green), interaction between environment and gesture (middle, blue), and interaction between gesture and action (right, orange). Means are normalized (%), and show error bars (SE).

sand, though this difference may also be due to participants' apparent need to "hold" objects in water.

Two-Way Interaction: Action and Gesture

We also found a significant interaction between action and gesture ($F_{3,69}$=24.9, p<.001, Figure 4, orange). Above-surface and touch gestures were used significantly more (p<.001) when participants were narrating than when they were enacting dialogue. Conversely, participants jiggled objects significantly more to enact dialogue than when they were narrating a part of the story (p<.001). There was no significant difference between dialogue and narration in terms of tapping (p=.18). This finding suggests that participants used the more animated jiggling action to indicate dialogue, and preferred more touching and above surface actions when narrating a part of the story.

Three-Way Interaction: Environment, Action, Gesture

There was also a significant three-way interaction between environment, gesture, and action ($F_{6,138}$=4.4, p<.001), but we did not explore this interaction further.

Gestures for Character Movement

We analyzed character movement with a 3 condition (digital, water, sand) × 4 gesture (drag, drag & jiggle, repeated lift & drag, lift & drag) RM-ANOVA.

There was a significant main effect of gesture ($F_{3,69}$=36.31, p<.001, Figure 5, green). We found that lift and drag was used significantly less than both dragging and dragging while jiggling (p<.001). This result may suggest that participants preferred to "perform" object movement. That is, the journey from one location to another was as important as the start and end locations. Similarly, repeated lift and drag was used significantly less than dragging and dragging while jiggling (p<.001); however, this may be partially explained by the interaction between environment and gesture (see below), as this gesture was not easy to perform on the digital table. There was no significant difference between dragging and dragging while jiggling (p=1.00) nor between lift and drag and repeated lift and drag (p=.38).

Two-Way Interaction: Environment and Gesture

We also found a significant interaction between environment and gesture ($F_{6,138}$=23.32, p<.001, Figure 5, blue). Post-hoc pairwise comparisons were grouped by gesture.

Dragging: Dragging was used significantly more in digital than each of water (p<.002) and sand (p<.001). Dragging in water was also used significantly more than sand (p<.002). As participants could easily move objects on the surface of water, but not in sand. This finding also suggests that dragging was the most used action to move the object around the digital table. That is, participants tended to mostly drag objects to move them around the digital table.

Dragging while jiggling: This action was used significantly more in sand than on the digital table (p<.001), and water

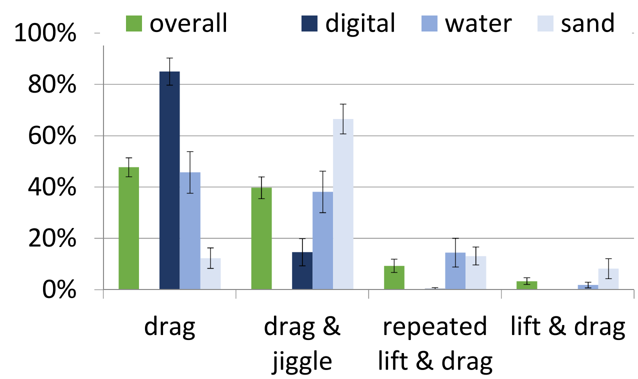

Figure 5. Main effect of gesture (left, green) and gesture × environment interaction (right, blue) for character movement.

(p<.02). It was also used significantly more in water than on the digital table (p<.05). Participants could move objects without jiggling them to show that they are moving; however, they jiggled objects (i.e. animated movements) while moving them. And this happened significantly more in the physical conditions than on the digital table.

Repeated lift and drag: This action was used significantly less on the digital table than in the sand (p<.003) and in the water (p<.05). There was no significant difference between sand and water (p=1).

Lift and drag: There were no significant differences between the environments in terms of this type of movement, and it was not used much overall. This finding might suggest that participants preferred to be in contact with the objects and enact how they move while moving them from one point to another point, instead of just picking them up and putting them down in another location.

Two-Handed Interactions

Participants used two-handed manipulation while either manipulating one object or two objects at a time.

Manipulation of one object: Manipulation of one object, was always performed using only one hand in physical conditions; however, participants sometimes used two hands to rotate or move one object on the digital table. In order to perform 2D and 3D rotations in the sandtray application, participants were required to have respectively two or three points of contact with the surface [10]. Therefore we observed a variety of bimanual interactions to rotate an object on the digital table. We also sometimes observed that participants moved (i.e. drag, drag while jiggling, lift and drag) an object on the digital table using two hands (usually one finger from each hand)

Manipulation of two objects: While in some instances, participants used bimanual actions to interact with one object, we observed that sometimes they used two-handed coordination to simultaneously interact with two different on-screen objects. To the best of our knowledge little work has been done on studying two-handed coordination while simultaneously manipulating more than one object, we focused on investigating these types of actions to first understand what types of two-handed coordination are generally used while working with two different objects and second to explore how frequently narrators tend to use these two-handed manipulations in the different environments.

We ran a 3 environment (digital, water, and sand) × 3 combination (in-place+in-place, in-place+move, and move+move) RM-ANOVA on the number of times that participants used any combination. There was a significant main effect of environment ($F_{2,46}$=28.08, p<.001). The number of times that participants used two-handed coordination in water (M=2.3, SE=0.3) was significantly more than each of sand (M=0.9, SE=0.2) and digital (M=0.3, SE=0.1, p<.001). This result could be due to the fact that participants often hesitated to leave objects on the water to float, and instead kept them in their hands. The number of times participants used two-handed coordination in the sand was also significantly more than the digital condition (p<.009). These results suggest that participants used significantly more two-handed interaction in the physical environments than on the digital table. This may be because certain interactions on the digital table (e.g., rotation) required two or three points of contact with the object, and consequently participants were not able to manipulate other objects at the same time.

While bimanual interaction techniques are a common approach for manipulation of objects on a digital surface [11,21,30], this finding might suggest that there are some down sides, as they could constrain the animated manipulation of more than one object at a time. We recommend a more in-depth study to understand both the benefits and drawbacks that bimanual interaction might cause in different contexts. In both the physical conditions of our study, we observed that participants were able to rotate the rabbit and move the turtle figurine at the same time to show two simultaneous events in the story. However, this representation could not be easily performed in the digital condition as the participants needed to use three points of contact to rotate the rabbit, so they used two fingers of one hand and one finger from the other to rotate it. Consequently, they could not manipulate any other object at the same time.

We also found a significant main effect of two-handed combination ($F_{2,46}$= 10.616, p<0.001, Figure 6, green). We found that *in-place+in-place* were used together significantly more than both *in-place+move* (p=.005) and *move+move* (p=.013). However, there was no significant difference between *in-place-move* and *move+move* (p=.243). The increase in *in-place+in-place* is best explained through the interaction between environment and combination, as this effect was likely dominated by the water condition.

Two-Way Interaction: Combination and Environment

There was a significant interaction between environment and combination ($F_{8,184}$=7.35, p<.001, Figure 6, blue). Post-hoc pairwise comparisons were grouped by combination.

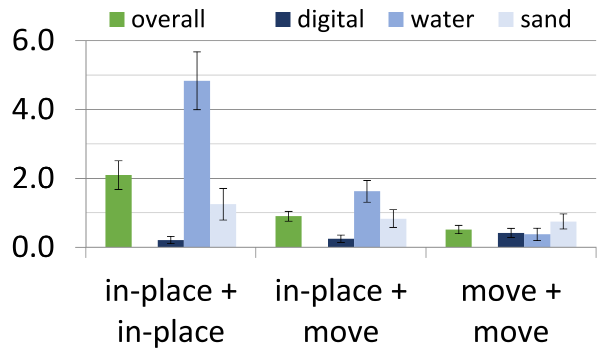

Figure 6. Mean counts and standard error (SE) of simultaneous two-handed actions, showing a main effect (left, green) and environment × combination interaction (right, blue).

In-place+in-place: This combination was used significantly more in water than both sand and digital (p<.001). The difference between sand and digital was not significant (p=.06). Note that this combination was used far more in water (M=4.8, SE=0.8) than in any other combination and environment (M < 1.7). This may be because participants did not appreciate figures “floating away”, and so performed many in-place actions in the water with two hands where they just held the characters.

In-place+move: This combination was used significantly less in the digital condition than both sand (p=.05) and water (p=.001). This may again be because participants wanted to keep at least one character “still” in the water. There was no significant difference between sand and water (p=.243).

*Move+m*ove: There was no significant difference between environments in *move+move* combination (p>.40).

Crossed Hands

We observed eight incidents on the digital table in which participants crossed their hands while moving two objects simultaneously (Figure 7), three times in water and never in the sand. When characters would cross paths in the sand, participants would exchange figures between hands; however, they hesitated to change hands in the other environments. In the water, changing hands could have resulted in objects floating on the water for few seconds and that could have been undesirable in the story. On the digital table, participants may have been concerned that the objects would “drop” or lose their orientation once they let go.

Limitations

The focus of our study was to find out what types of narrative gestures are used on a digital surface and how they are different from the gestures used in physical environments. While this study had some limitations, we believe they do not alter the main contributions of the paper. This study was done one on one, as opposed to with a larger audience. However, our findings still show differences from traditional gestures, even with only one audience member. Nonetheless, future work could analyze larger audiences. Even

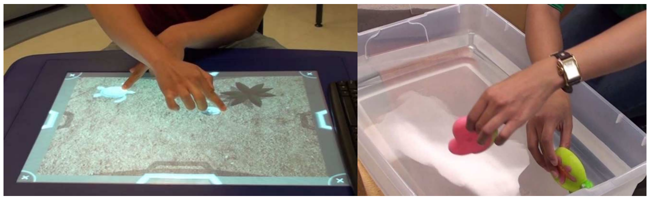

Figure 7. Participants crossed their hands on eight separate occasions while telling their stories.

though some digital interactions, such as repeated lift and drag, were sometimes hard to perform, which may be seen as a study limitation, participants still demonstrated similar gestures in physical environments that we feel indicate a need for new 3D object interaction techniques.

DISCUSSION

Our results show that the object-mediated narrative gestures in our study had some similarities to open-handed narrative gestures [1,20,22] and were mostly used to convey meaning by, for instance, touching a character to show that it is being described or by jiggling while moving to enact how characters move. We also observed that, most of the time, participants preferred to touch and hold or to manipulate objects while telling a story, instead of only pointing to them from above the surface. This observation might suggest that participants preferred to be in contact with objects and be engaged with the process physically. We suggest the following design considerations based on our observations:

Consider touch to narrate actions. In most digital surface applications, input is often handled through directly touching the screen. In a similar vein to others who have noted a potential Midas Touch problem [12,14,17,40], our study shows that narrators sometimes touch the screen to mention an object or to focus on what they are saying, and may not consider the act a command. Considering these unique touches is important in designing narrative applications.

Consider in-place manipulations of on-screen objects. Designers should also be aware of narrative in-place actions while designing interactive applications. For instance, in Microsoft Windows (7 and 8), when a person jiggles a window, other windows on the desktop minimize. This type of interaction may interfere with the narrative.

Consider the path to get there. While in some circumstances, it may be more desirable to consider more efficient and automatic interaction techniques (e.g., to avoid fatigue), we found that designing assistive or automatic interaction may not be always a good design decision, as narrators may prefer to be engaged with the narration process and movements related to story events. For instance, designing voice commands for movements or a move interaction that involves tapping an object and then tapping its target destination may seem more optimal, but may not be desirable for a narrative application, as narrators may prefer to move the objects around with their hands to engage in the process.

Support expressive and animated actions through manipulation techniques. We also observed that, while participants tended to animate character actions and story events, animated movements were employed more in physical environments than on the digital table. This finding might suggest that manipulating objects on the digital table was not as easy as it was in the physical conditions. Therefore, designers should exploit new technology to enable narrators to be more expressive and animated in their movements.

Consider two-handed, simultaneous manipulation of multiple on-screen objects. We also found that two-handed interactions were often used to simultaneously manipulate two objects. These types of actions were used significantly more in physical environments than on the digital table. This could be due to the fact that certain actions (2D and 3D rotation) were mostly performed by participants while using two hands and that could prevent them from manipulating any other object at the same time. However, this constraint did not exist in physical environments. Therefore, while a bimanual interaction technique could be suitable for movement and rotation in many circumstances, it might not always be suitable for narrative or expressive applications.

CONCLUSION

In this paper, we reported on a large and detailed observational study to explore how people make use of gestures to tell a story on a digital surface, a physical sandbox, and in water. We showed that these expressive gestures are fundamentally different than the movement and rotation gestures common on a digital table, and that people use two hands to richly and creatively express meaning in a story.

ACKNOWLEDGEMENTS

We thank the Natural Sciences and Engineering Council of Canada (NSERC), NSERC's Digital Surface Software Application Network (Surfnet), and the Graphics Animation & New Media (GRAND) NCE for funding.

REFERENCES

1. Cassell, J., 2001, Nudge Nudge Wink Wink: Elements of Face-to-Face Conversation for Embodied Conversational Agents. In Embodied conversational agents. MIT Press, Cambridge, MA, USA, 1-27
2. Daeman, E., Dadlani, P., Du, J., Li, Y., Erik-Paker, P., Martens, J., & De Ruyter, B. (2007). Designing a free style, indirect, and interactive storytelling application for people with aphasia. *INTERACT'07*, 221–234.
3. Decortis, F. and Rizzo, A. (2002). New Active Tools for Supporting Narrative Structures. *Personal and Ubiquitous Computing*, Vol. 6, 416-429
4. Dietz, P., & Leigh, D. (2001). DiamondTouch: a multi-user touch technology. In *Proc. UIST*, 219-226.
5. Epps, J., Lichman, S. and Wu, M. (2006). A study of hand shape use in tabletop gesture interaction. In *Ext. Abstracts CHI*, 748-753.

6. Goldin-Meadow, S. (1999). The role of gesture in communication and thinking.*Trends in cognitive sciences*, *3*(11), 419-429.

7. Goldin-Meadow, S. (2005). *Hearing gesture: How our hands help us think*. Harvard University Press.

8. Guiard. Y. (1987). Asymmetric divisoin of labor in human-skilled bimanual action: the kinematic chain as a model. *J. Motor Behavior*, 485–517.

9. Han, J. Y. (2005). Low-cost multi-touch sensing through frustrated total internal reflection. In *Proc UIST*, 115-118.

10. Hancock, M., Carpendale, S., and Cockburn, A. (2007). Shallow-depth 3D interaction: Design and evaluation of one-, two- and three-touch techniques. In *Proc. CHI*, 1147–1156.

11. Hancock M, Cate T, Carpendale S, Isenberg T. (2010). Supporting sandtray therapy on an interactive tabletop. In *Proc. CHI*, 2133–2142.

12. Hansen, J.P., Tørning, K., Johansen, A.S., Itoh, K., and Aoki, H. Gaze typing compared with input by head and hand. In Proc. *ETRA*, ACM Press (2004), 131–138.

13. Helmes, J., Cao, X., Lindley, S.E., and Sellen, A. (2009). Developing the story: designing an interactive storytelling application. In *Proc. Tabletop*, 49-52.

14. Hinckley, K., Yatani, K., Pahud, M., Coddington, N., Rodenhouse, J., Wilson, A., Benko, H., and Buxton, B. (2010). Pen + touch = new tools. In *Proc. UIST*, 27-36.

15. Hinrichs, U., and Carpendale, S. (2011). Gestures in the wild: studying multi-touch gesture sequences on interactive tabletop exhibits. In *Proc. CHI*, 3023–3032.

16. Hodges, S., Izadi, S., Butler, A., Rrustemi, A., & Buxton, B. (2007). ThinSight: versatile multi-touch sensing for thin form-factor displays. In *Proc. UIST*, 259-268.

17. Jacob, R.J.K. (1993). Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in human-computer interaction* 4, 151-190.

18. Jacucci, G., Morrison, A., Richard, G. T., Kleimola, J., peltonen, P., Parisi, L., and Laitinen, T. (2010). Worlds of information: designing for engagement at a public multi-touch display. In *Proc. CHI*, 2267–2276.

19. Kalff, D. M. (2003). *Sandplay: A psychotherapeutic approach to the psyche*.

20. Kendon A. (2004). Gesture: Visible action as utterance. Cambridge U Press.

21. Martinet, A., Casiez, G., and Grisoni, L. (2012). Integrality and separability of multitouch interaction techniques in 3D manipulation tasks. *Trans. Vis. & CG* 18(3), 369–380.

22. McNeill, D. (1992). Hand and mind: What gestures reveal about thought. University of Chicago Press.

23. Morris, M.R., Huang, A., Paepcke, A. and Winograd, T. (2006). Cooperative gestures: Multi-user gestural interactions for co-located groupware. In *Proc. CHI*, 1201-1210.

24. Morris, M.R., Wobbrock, J.O., and Wilson, A.D. (2010). Understanding users' preferences for surface gestures. In *Proc. GI*, 261–268.

25. Moscovich, T. (2007). *Principles and Applications of Multi-touch Interaction*. Brown University.

26. Nielsen. M, Storring. M, Moeslund. T, and Granum, E. (2004). A procedure for developing intuitive and ergonomic gesture interfaces for HCI. *LNCS*, 2915, 409-420.

27. Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. Advances in health sciences education, 15(5), 625-632.

28. North. C, Dwyer. T, Lee. B, Fisher. D, Isenberg. P, Robertson. G, and Inkpen-Quinn. K. (2009). Understanding multi-touch manipulation for surface computing. In *Proc. HCI Int.*, 236–249.

29. Polkinghorne, D. (1988). *Narrative knowing and the human sciences*. Suny Press.

30. Reisman, J. L., Davidson, P. L., and Han, J. Y. (2009). A screen-space formulation for 2D and 3D direct manipulation. In *Proc. UIST*, 69–78.

31. Scott, S.D., Carpendale, M.S.T., and Inkpen, K. (2004). Territoriality in Collaborative Tabletop Workspaces. In *Proc. CSCW*, 294-303.

32. Shen, C., Ryall, K., Forlines, C., Esenther, A., Vernier, F., Everitt, K., Wu, M., Wigdor, D., Morris, M.R., Hancock, M., Tse, E. (2006). Informing the design of direct-touch tabletops. *IEEE CG&A*, 26(5), 36-46.

33. Terrenghi, L., Kirk, D., Sellen, A., Izadi, S. (2007). Affordances for Manipulation of Physical versus Digital Media on Interactive Surfaces.In *Proc. CHI*, 1157-1166.

34. Underkoffler, J., and Ishii, H. (1999). Urp: A luminous-tangible workbench for urban planning and design. In *Proc. CHI*, 386-393.

35. White, M., & Epston, D. (1990). *Narrative means to therapeutic ends*. WW Norton & Company.

36. Wobbrock, J.O., Morris, M.R. and Wilson, A.D. (2009). User-defined gestures for surface computing. In *Proc. CHI*, 1083-1092.

37. Wu, M. and Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proc. UIST*, 193-202.

38. Wu, M., Shen, C., Ryall, K., Forlines, C. and Balakrishnan, R. (2006). Gesture registration, relaxation, and reuse for multi-point direct-touch surfaces. In *Proc. Tabletop*, 185-192.

39. Zancanaro, M., Cappelletti, A., & Stock, O. (2003). StoryTable: Computer supported collaborative storytelling. In *Proc. UIST Extended Abstracts*.

40. Zhai, S., Morimoto, C., and Ihde, S. (1999). Manual and gaze input cascaded (MAGIC) pointing. In *Proc CHI*, 246-253.