A Design Studio Course for Gesture Aware Design of Tangible Interaction

Mary Lou Maher  
Software and Information Systems  
UNC Charlotte  
m.maher@uncc.edu

Kazjon Grace  
Software and Information Systems  
UNC Charlotte  
k.grace@uncc.edu

Berto Gonzalez  
Software and Information Systems  
UNC Charlotte  
agonza32@uncc.edu

ABSTRACT
Tangible computing incorporates gesture, grasping, and the use of physical objects as integral parts of interaction design. This paper describes a theory-driven design studio course in which students were asked to develop a new tangible application grounded in the psychology of gesture and thought and informed by protocol studies and critiques of the role of gesture in existing tangible applications. The following tangible computing devices and platforms were included in the studio: the Sifteo Siftables, the Samsung Pixelsense multi-touch table, multi-touch tablets and displays, the Orbotix Sphero, the Microsoft Kinect, and the Leap Motion. The students designed new applications for the Siftables and the Samsung Pixelsense tabletop computers. From the students’ studies and designs, we have developed a framework for studying and designing gesture in tangible computing and a set of design principles for effective use of gesture in interaction design.

Author Keywords
Gesture; tangible interaction; design principles; studio teaching

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human Factors; Design; Measurement.

INTRODUCTION
Tangible computing incorporates gesture, grasping, and the use of physical objects as integral parts of interaction design. We define gesture in tangible computing to encompass both the use of body or hand movement to interact with digital information as well as other bodily or manual movements accompanying the specific actions that are interpreted by the tangible application. TUI’s encourage gesture, those supported by the input devices as well as other related bodily movements. The design and subsequent use of gesture in the context of tangible computing can be based on a consideration of how gesture and thought are related. There are several studies in the field of psychology that show that gesture affects cognition, and this effect can be positive or negative. For example, Cook, Yip, and Goldin-Meadow [3] have results that show that gesture can make memories last; while Cook, Mitchell and Goldin-Meadow [2] show that in addition to aiding the persistence of specific memories, gestures can facilitate task and concept learning. McNeil [14] shows how gesture is related to thought and speech production in the gesturer rather than being only a form of communication; Wu and Coulson [19] show that gestures prime concepts that are related to similar sensorimotor events in listeners, and Krauss [8] shows that restricting gesture can cause verbal disfluency. These findings suggest that gesture has a significant effect on the cognition of both performer and observer. The intent of our design studio was to explore the relationships between the psychological line of inquiry and the experience design of tangible computing. Our approach emphasizes the significance of a theoretical grounding, as argued for in Marshall [13] in the emerging discipline of tangible user experience design.

The following section is organized around a design studio course taught Jan-May 2013 in which students used the results and methods of psychological studies of gesture and thought as the basis for studying and designing tangible applications. We conclude the paper with a framework for studying how gesture is used in tangible computing and some design principles for tangible computing that consider the role of gesture and thought in user experience design.

DESIGN STUDIO SCHEDULE AND PROJECTS
The design studio schedule was organized into three parts: (1) research on gesture, thought, and tangible interaction; (2) user studies of gesture, action, and interaction design; and (3) design and implementation of a new application for an existing tangible computing device. Three projects were the principle focus of studio sessions: as a studio style learning environment there were no lectures and each class comprised an interactive activity that progressed one of their projects.

Project 1: User studies of tangible computing
Students selected a tangible technology and application to use as the focus for this user study. Students studied one or more users whilst they completed a task with the application. Students observed the interaction and collected video data. Protocol analysis was used to analyze how users’ gestures and actions contributed towards goal achievement.
Project 2. Design critique of a selected tangible computing platform or application
After completing their protocol analysis of a tangible application, students were asked to perform a design critique of a second application. The students were asked to describe the interaction design in terms of the different functions that can be achieved with the application, as well as the mapping from gestures to actions on the device. They were asked to critique the design with respect to how gesture and action are synergistic. Pedagogically these critiques served to develop students’ reflective capacity in the design domain of tangible interaction, but they also served as an exploration of the cognitive influences in present tangible practice.

Project 3. Tangible design project
After students had analyzed and critiqued existing applications from a theoretical perspective, the final component of the studio was to develop an application for one of the tangible platforms. The students formed groups where each member of the group had one of the following roles: interaction designer, software developer, or system designer. The students were asked to design and implement a new application for the tangible device and write a report on the role of gesture in the interaction design of their applications.

STUDIES & CRITIQUES OF GESTURAL INTERACTION
The students were introduced to several applications on the following tangible devices: Sifteo Siftables, Samsung Pixelsense tabletop computer, various multi-touch tablets and displays, Orbotix Sphero, Microsoft Kinect, and Leap Motion. Below is a sample of their studies and critiques.

Sifteo Siftables
The Siftable cubes [12] are small, handheld devices with integrated displays that communicate with each other and a computer acting as a base-station. Each cube is equipped with a 3-axis accelerometer and a near-field sensor to detect the presence of other cubes.

One student study examined the interaction techniques used with the “Word Play” application available for the Sifteo cubes. In this application, each cube has a letter and the user (or group of users) arranges the cubes into words. Different letters are available by pressing on the cube’s display. When a valid word is formed, the cubes respond with a sound and a change in the display. The students used a think aloud session to capture the verbalized thought processes of the users. The students found that the actions/gestures that occurred most frequently were “slide” and “bump” even though the application provided support for 12 gestures/actions on the cubes (see Figure 1). In contrast, a student studying a different game on the Siftables, “Smörgåsbord”, found a more even distribution of gestures (illustrated in Figure 2). This dichotomy, which was observed in other applications across a number of platforms, indicates two contrasting approaches to the design of tangible interfaces: encouraging the primacy of one or two gestures that can be combined to achieve goals, or composing a set of different gestures that each map to different tasks in the application’s interaction lexicon.

Dell Multi-touch All-in-One
Multi-touch workstations are one interesting direction in which tangible interaction is increasingly being incorporated into mainstream computing activities. An
example is the Dell One Touch range, which incorporates a multi-touch display into all-in-one desktop workstations.

A student study of the Dell One Touch investigated the role of touch in a visual information-seeking context, and how users switch between the familiar desktop interaction of mouse and keyboard interaction and the relatively unfamiliar gestural modality. In this study subjects were asked to plan a hypothetical vacation to a destination they had not visited using a combination of the Bing Travel App and a web browser. The task was structured so that users had to discover places to visit from amongst mostly pictorial representations, like those seen in Figure 3, then research those sites and plan their itinerary. The users were recorded performing these tasks and those recordings were segmented by subtask (e.g.: browsing image search results, browsing text search results, reading detailed text) as seen in Figure 4 and then coded using protocol analysis.

The results of this student study suggested that in a mixed-modality touch and keyboard/mouse interaction users would tend to alternate between the two modalities. However, after a task was encountered that required mouse activity it was likely that users would stick with the traditional interaction modality even when returning to performing tasks for which touch interaction was supported. This indicates that the cost – whether physical or cognitive – of switching between modalities is high enough to discourage rapid alternation. This observation parallels the mention of potential conflicts between multiple interface modalities in Price’s [16] framework for tangible interaction. It was also observed, although not to a statistically significant level, that users tended to rest their hand on the mouse when reading blocks of text, even when they had been previously navigating by touch, perhaps because this body position was an unconscious default. These observations raise questions about the sufficiency of novelty and engagement in encouraging users to interact tangibly, instead highlighting the role of familiarity and consistency in tangible interface design.

**Sphero**

The Orbotix Sphero is both a robotic ball that is remotely controlled from a multi-touch tablet, and a motion-based gestural input device for said tablet. The device can be placed on the floor and controlled with tilting or swiping gestures on the tablet, or held in the hand and rotated in three axes to control a tablet application.

A student study of this technology investigated the relationship between the use of spatially metaphorical language and the use of the Sphero as a hand-held gestural input controller. Participants were recorded playing a scrolling shooter game called Sphero Exile, shown in Figure 5, with the coded period coming after a five minute free-play introduction. Participants were told to speak aloud as if they were playing the game in the presence of friends.

The Leap Motion Controller

The Leap Motion is a short-range touch-free interaction device allowing three-dimensional gestures in the air in front of a display. The device is designed with gestural fidelity in mind, tracking finger movements to within a tenth of a millimeter.

Students studied the Leap Motion as an alternative input device to the popular mobile app “Fruit Ninja”, in which players perform swipe gestures on falling fruits in order to score points. Gestures were coded based on the number of digits and hands involved, as well as whether the gesture was continuous or discrete. These codes were then compared to spatial and metaphorical language, as well as
success in targeting and “slicing” the on-screen fruit consistent with the goal of the game.

While the continuous one-fingered swipe was found to be the preferred interaction for accuracy, all gestures were found to be highly inaccurate at targeting fruit. Users’ utterances commonly referenced the difficulty, futility or exhausting nature of the task. The students hypothesized that this difficulty came not from the precision of the Leap – which as expected was very high – but from the difficulty of calibrating the Leap so that there was a one-to-one correspondence between the position of a user’s finger in the air and the pixel on-screen targeted by the resulting game cursor. User expectation was that the position of their gestures would correspond to the point on screen occluded by their fingers from their own vantage point, an assumption strongly based on real-world physical interaction. This expectation proved too difficult to unlearn for the Leap-enabled Fruit Ninja app. This observation highlights that precision sensing alone will not make a touch-free tangible user experience intuitive, engaging or “natural” unless the sensor conforms to human expectations grounded in a lifetime of sensorimotor interaction.

**Samsung SUR40 with Microsoft PixelSense**

The Microsoft PixelSense framework embeds vision-based multi-touch technology in a large LCD that can be mounted horizontally or vertically. The matrix of embedded IR emitters and reflectors is sensitive to fingers, objects placed on the surface, and hands or fingers hovering just above the surface, making the PixelSense both an on-screen and in-air gestural interface.

Several students analyzed and critiqued the design of two PixelSense applications. The Scatter Puzzle application is a simple demonstration app illustrating dragging and rotating behaviors using on-screen jigsaw puzzle pieces. and it was found to be highly intuitive and easy to learn. Interacting with the puzzle requires a small range of gestures and behaviors that are analogous to real-world manipulation of physical puzzle pieces on a tabletop. However, the extreme sensitivity of the PixelSense platform to hovering hands and near-touches did lead some users to accidentally load a new puzzle, highlighting the importance of consistency and user expectations even in very simple applications.

The PixelSense application “NUiverse”, an exploratory representation of the solar system shown in Figure 6, was also critiqued. This application provides a variety of tools, mostly in the form of radial menus, for exploring objects of interest within a 3D world. Student critique found this application to be intuitive once learnt, with a strong reliance on touch-and-drag to interact with both menus and objects. However, there were significant shortcomings in affordance and discoverability, with users expressing that several menus were “hidden” and difficult to invoke unless already known of. It was suggested that menus could additionally be activated from designated areas on the edge of the table surface to provide visual cues of interactability. These critiques highlight the critical role of affordances in shaping user mental models, particularly while users acquire competence with a new interface.

**Figure 6: The many menus of NUiverse, intuitive to expert users but with critical flaws in discoverability**

**FRAMEWORK FOR ANALYZING GESTURE IN INTERACTION DESIGN**

The studies reported by the students did not include a significant number of participants, so the reported observations are not statistically valid and should not be considered scientifically supported. In addition to educating students in the methods and principles of a psychologically grounded approach to tangible interface design, the purpose of the studies was to provide an initial exploration of the space as a precursor to more robust analysis. In this section, we draw from these student projects to propose a framework for analyzing gesture in interaction design.

Our framework for understanding how gesture is included in interaction design has three components: gesture, action, and function. Actions are the result of gestures, and one or more actions result in the performance of a function by the application.

**Gesture**: Gestures are the set of physical and bodily manipulations and motions that can be used while interacting with the application. This may be direct interaction such as point or touch, grasp, shake, or require no contact at all such as waving a hand across a sensor. A set of gestures supported by the application shows what the designer expected of the user, but our investigations indicate that not only do users augment the list of supported gestures with unsupported gestures (similar to the “offline interactions” described in Fernaeus et al.[6]), they also tend to use only a subset of the supported gestures.

**Actions**: The actions available within an application characterize the set of ways the user can change the application’s state. Actions are distinct from the input required to trigger them, and are in turn distinct from the functions carried out by the application as a result of them.
**Functions**: These are the tasks that a user can carry out using an application. This is usually described as a hierarchy of functions, with the root of the hierarchy being the most abstract function, like “play a game”, with nodes beneath that root taking on more specific functions within that application; “play a level”, “defeat an opponent”, or “fire a shot”. These functions are triggered by interface actions by the user, and those actions are performed with gestures.

As an example of applying this framework, consider the case of a user swiping their finger to view additional information about a video. The gesture of swiping downwards to view details is distinct from the user action of opening the detail view, which is in turn distinct from the application function that presents those details in the form of a window. These distinctions enable a designer to specify the functions available to users (e.g. how should the application support users who want to view more information?), the actions that enable those functions (e.g. when and how can the user access the detailed view?) and the gestures that enact those functions (e.g. what kind of swipe should be performed on an element to view its details?). This distinction is particularly relevant in tangible computing as the language of available gestures is much richer than a traditional keyboard and mouse interface, in which pointing, clicking, and typing are the only available gestures and thus interaction design is more concerned with mapping actions to functions. In tangible computing an interaction designer must consider both the mapping of gestures to actions and the mapping of actions to functions in order to create a consistent, engaging and effective user experience.

This framework gave the students a basis for characterizing and critiquing the interaction design of tangible and gesture-based interfaces. Some of the observations made with this framework include: when one or two gestures dominate and are the required body movements for multiple actions; when the same gesture is used for more than one action and potentially introduces confusion or inconsistency; or when the relationship between function and gesture, via action, violates a metaphorical expectation for a user (e.g. waving a hand upward to close a window). Further studies and comparisons of tangible interaction design using this framework could provide a comparative analysis that goes beyond the usability of interface actions to perform functions. Such studies could also investigate how the design of tangible interaction is or isn’t consistent with a cognitive understanding of gesture’s impact on thought.

**DESIGN PROJECTS**

The design project component of the studio was focused on the design and implementation of a new application for a device that enables touch and/or graspable object interaction. The intention was for students to experience designing and implementing all three stages of our gesture, action and function framework and the mappings between them. The students were also asked to build upon their user studies and design critiques to design their new application, while also mindful of the role of gesture in thinking and how that translates into interaction design. The students used a combination of user-centered design methods and agile computing systems design. They were asked to select two user centered design methods from Martin and Hanington [11] and utilize them in their design process. A description of the three resulting student-designed tangible applications and their consideration of the cognitive impacts of gesturing follows.

**What Comes Next?**

What Comes Next is a game for the Sifteo Cubes in which players assemble sequences or compounds of objects represented by words, pictures and sounds. Each cube represents an object, idea or component, with media items (words, images, sounds, mathematical symbols, etc) relating to that idea being presented on the cube and cycled through by pressing on the display. The central metaphor of the game is that placing two cubes physically adjacent to each other corresponds to combining the two concepts displayed on them. Additionally, it was hypothesized (after [18]) that users would perform a variety of orientation and alignment gestures as a way to offload the cognitive load of searching for combinations.

The system was designed to work with any kind of sequence or compound that could be displayed on the cubes, but the domain of compound words was explored in the most depth. Six simple nouns (i.e. dog, mail, stick, box, house, and fish) were displayed at the same time with one word on each of the six cubes, and users were rewarded to find as many compounds (e.g. “doghouse”, “mailbox”) as they could, and then cycle the application to another set of words. Figure 7 shows testing of the What Comes Next application as compared to a tangible but non-interactive paper prototype.

**Figure 7**: The What Comes Next tangible application (left) and the paper prototype to which it was compared (right).

The designers of What Comes Next intended it to be used in a classroom environment, and after discussions with educational stakeholders the application was integrated with a desktop computer for the purposes of displaying user progress. This output of what compounds had been constructed could be displayed privately to an educator at a terminal or publically to the class via a projector.
Participants in both the paper and interactive conditions were able to quickly and efficiently search the space of possible combinations and generate compounds and sequences. Alignment, orientation and movement gestures which were not recognized by the system were frequently observed, although no difference was discernible between the paper and interactive versions in this regard. Users of the Siftable application self-reported substantially greater engagement, possibly due to the interactivity of cycling through the multiple representations of each idea. Usability investigations revealed that some participants were able to discover the “shake” gesture (which exchanged a word for another in the database) without prompting, but others, particularly younger (<8) participants, required verbal prompting. This indicates that the shake-to-change-words gesture, while metaphorically consistent, could potentially benefit from some additional affordances to aid discoverability. Use of a “shake” gesture to change or cancel is gaining some mainstream acceptance (e.g. the undo typing feature of the Apple iPhone), but these results indicate it is not yet familiar enough to be considered intuitive without instruction, cues or prompting.

One first-grade student was extremely excited about having discovered a word, “fishstick”, to which the cubes would not respond, indicating that he had discovered a “new” word the game’s designers had not considered. The enthusiasm that this discovery provoked suggests that tangible interfaces for the creative construction of new compounds and sequences have significant educative potential, a potential domain raised in Marshall [13].

The studies with What Comes Next also revealed a social component to gestural interaction. When working individually children would lay out all six cubes in front of them and make many epistemic movements while searching. By contrast, when playing in pairs children tended to be significantly more possessive, picking up or moving a pair of cubes towards them to indicate “ownership” while they were searching for compounds. In pairs the number of shake-to-change gestures increased dramatically, and the number of compounds found per word displayed dropped rapidly – both clearly a result of the lack of sharing between participants. This effect, similar to the “Tucking Away” and other territorial gestures described in Olsen et. al. [15] indicates the complexity of tangibly interactive collaboration, as participants chose (perhaps subconsciously) to ensure their sole possession of a few cubes over the potential benefits of considering and arranging the set of six cubes as a collaborative whole. Designing to promote collaboration in tangible interfaces will need to address this complex interaction between individual engagement and collective collaboration.

**Blockade**

Blockade, shown in Figure 8, is a game application for the Microsoft PixelSense that combines multi-touch tabletop interaction with tangible pieces. Players take turns moving their 3D-printed player pieces around the board, with the PixelSense recognizing the fiducial tags on the base of each piece. Players attempt to strategically organize their pieces into formations that block opponents from passing, which are the eponymous “blockades”. These barriers force any opponent who passes them to pay up, in a manner reminiscent of highway robbery. Players must count out the tax to be paid to an opponent whose blockade they encounter, and that opponent then uses a special “money” piece to take the resulting money. The objective is to reach a monetary threshold by successfully blockading and stealing from other players. The game was designed as a quasi-educational game for ages 7-12.

Figure 8: The Blockade app on the PixelSense, with insets of (top to bottom) the player pieces, money pieces, and dice.

Blockade exhibits a number of design considerations that reflect cognitive influences on tangible interaction. User responses to having graspable, tangible pieces on a touch-screen interface were extremely positive, with those users who found the game most engaging tending to specifically mention the integration of physical board-game style pieces with a digital “board”.

The function of money transfer from the victim of a blockade to its perpetrator was the focus of the designers’ integration of the cognitive impact of gesture. When a player hands over money they are required to count it out in real-world denominations and the application does not enforce that the money exchanged matches the required penalty. In this way the cognitive load of counting out the amounts is placed upon the player, who must move the correct amount from their personal “bank” to a staging area where it can be reviewed by the other player(s). The counting player’s movement of each note is accomplished by a drag-and-drop gesture on the multi-touch table, facilitating the practice of mathematical skills by tangible interaction. Players have an incentive to count the correct number so as to not overpay their attacker, and they cannot rely on tertiary parties to intervene on their behalf unless doing so serves their own needs.

Once the victim has counted out the required money to the satisfaction of other players, the blockading player reaches over the table and places their money piece within the victim’s personal space, initiating the transfer of funds.
Thus the victim is required to physically count out their losses, present them to the table for agreement, and then sit passively while the attacking player reaches into their personal space to complete the exchange. In the terms of our framework, the function of money exchange is supported by the actions of money-counting by the blockaded player and of money-taking by the blockading player. Additionally, the gesture to trigger money-taking – reaching across the large table to place a piece within the opponent’s personal space – was metaphorically aligned with the concept of forcefully taking valuables from a victim. Participants in the user studies of Blockade consistently rated taking money as the most satisfying and engaging portion of the game.

Students did not compare the level of user engagement of this hybrid tangible/digital game with a traditional board-game setup, but users did rate the game as significantly more engaging than a traditional physical-only setup. This integration of the physicality of tangible pieces with the interactivity of digital components leverages the advantages of each paradigm.

DreamAScene

DreamAScene is an application for the Microsoft PixelSense that was designed to be an interactive display in a children’s library. In DreamAScene, up to six users stand around the PixelSense table and collaboratively design a collage of clip-art images while collectively narrating a story to accompany their scene. Users select a background image and then drag clip-art from bars at the edge of the table into the “scene” where they can be rotated, positioned and scaled. The application can be seen in use in Figure 9.

Figure 9: DreamAScene. A user is seen scaling an on-screen element with a two-fingered pinch gesture; arrows indicate where other users could join in.

With a focus on how the interface could provide affordances for promoting collaborative contribution to the scene, the students placed multiple clip-art menus around the exterior of the table. This prompted individuals to check whether each menu contained the same selection of images. The final design includes six small arrows placed around the edges of the table, labeled with the word “Join”, each of which can be dragged or tapped to bring up one of the clip-art menus. The presence of the arrows indicating that additional menus could be dragged out onto the table – even when some users were already interacting with it – encouraged additional participants to begin interacting with the table.

In contrast to the successful affordance of collaborative contribution provided by the “join” arrows, the designers found that the menu for background selection, invoked by a double-tap gesture performed on the background image, was not discoverable. This is a similar issue to that observed with the menus in the NUIverse application analyzed earlier. These observations provide further indication of the importance of “interactability” affordances for gestural interaction, similar to the challenge of conveying interactivity discussed in Grace et al. [7].

There was an overall tendency of DreamAScene users to comprehensively narrate the scene as they were developing it. There was, however, a strong negative correlation found between speech and gesture on the table. Participants typically alternated between dragging and arranging scene elements and describing them, contrasting with expectations that tangible interaction would increase speech. It is hypothesized that this negative relationship was the result of verbal disfluency caused by the cognitive effort necessary to precisely position elements on the highly sensitive PixelSense display. Further investigation is necessary to determine whether sensitivity correlates with the size of this disfluency effect, and whether the effect leads to reduced narrative ideation overall or just causes gesture and speech to occur consecutively rather than concurrently.

**DESIGN PRINCIPLES**

While the studies, critiques and projects produced during the Tangible Design Studio were not rigorously controlled, methodologically sound or their findings statistically significant, they led to a number of insights into the design of compelling tangible user experiences. We have synthesized a set of gestural interaction design principles for tangible computing out of the experience of designing and teaching this course.

- **Considering action and gesture separately** in order to distinguish the design of user actions from the design of user gestures.
- **Consistency among gestures**, both in terms of what gestures map to what kind of actions (activate a button, drag out an element) in different parts of an application, and what gestures map to what kinds of function (open, close, like, view more, etc).
- **Metaphorical alignment** between gestures and the kinds of actions and functions which they perform, such as using an upward hand movement to open. This
finding echoes the broader discussion of the interaction of metaphor and gesture in Cienki and Müller [1].

- **Considering sensorimotor expectations** that users will exhibit due to their lifetime of experience interacting tangibly, such as assuming that pointing gestures will align with their own perspective, rather than with a separate frame of reference defined by the device. This is consistent with the notion of embodied interaction [4,5].

- **Considering gestural affordances** for discoverable interaction, especially in cases where there is no persistent perceptual indicator that a gesture can be performed, or where no expectation of interactability exists. This relates to the notion of “symbolic correspondence” described in Price [16]. This exists separately from whether a gesture is intuitive once learnt.

**CONCLUSIONS**

This paper presents an approach to teaching the design of tangible interfaces theoretically grounded in literature on gesture and thought, presents the designs and considerations that resulted from that approach, and discusses the principles and hypotheses for future research in gestural interaction. This early exploration into the relationship between thought and gesture in tangible computing suggests that much more investigation into sensorimotor expectations, metaphorical alignment and gestural affordances and consistency is needed before gestural interaction can claim to be truly “natural”.

**ACKNOWLEDGMENTS**

omitted for blind review

**REFERENCES**


