Analyzing Children's Hand Actions using Tangible User Interfaces

Alissa N. Antle
School of Interactive Arts and Technology
Simon Fraser University, Surrey, B.C., Canada V3T 0A3
aantle@sfu.ca

ABSTRACT

We present the theory and mixed methods approach for analyzing how children's hands can help them think during interaction. The methodology was developed for a study comparing indirect with direct input methods for object manipulation activities in digitally supported problem solving. We propose a classification scheme based on the notions of complementary and epistemic actions in spatial problem solving. In order to overcome inequities when comparing mouse input with the multi-access, bimanual input, we develop a series of relative measures based on our classification scheme. This methodology is applicable to a range of computationally augmented activities involving object manipulation.

Author Keywords
Input methods, tangible computing, embodied interaction, bimanual manipulation, video analysis, methodology.

ACM Classification Keywords
H5.2. User Interfaces: *Evaluation/methodology*.

INTRODUCTION

The embodied nature of tangible user interfaces has become of increasing interest to designers of children's educational technologies [1-3, 5, 6, 11, 12]). This interest is predicated on the view, common in education, that learning through hands-on manipulation of physical manipulatives may be beneficial (e.g., Montessori Method, Frobel's Gifts) [16]. However, there is little empirical evidence to date to support such claims in the realm of children's tangible computing [1, 11]. Understanding the role that the hands play in supporting certain mental processes during tangible interaction can help guide design decisions about how to design such interfaces. Studying children (aged 7-10) provides a window on such interaction and may highlight results that can be generalized to adult populations.

There are many open questions which concern the interrelation between input style and resulting interaction for a task that requires manipulation of objects or pieces (e.g., jigsaw puzzle, block construction, tesselation). For example: What are the differences between how physical objects are manipulated with the hands compared to how digital representations of those objects are manipulated with a mouse? Does supporting users to manually handle augmented physical objects change how they problem solve? How can we design interfaces to support children to offload difficult mental tasks to physical interactions with environment through using their hands? Does physical or digital manipulation take longer? If it takes longer does this mean it is harder? Does direct physical interaction allow more opportunities for actions which support task learning?

In this paper we provide a description of a mixed quantitative and qualitative methodology for comparing the type, number, and duration of children's hand-based physical actions. We focus on an age appropriate spatial problem solving task which involves objects that can be represented both physically and digitally, and can be manipulated with a mouse and by the hands. A large size jigsaw puzzle is such an activity. The puzzle can be implemented in its traditional cardboard form, in a PC-based graphical user interface style with a single mouse and on a tangible tabletop [15]. We present our methodology using a jigsaw puzzle task for illustrative purposes.

THEORETICAL FRAMEWORK

Object Manipulation

Computational objects can be manipulated using indirect (e.g., mouse) and direct (e.g., touch, tangible) input methods. Proponents of tangible and physical interaction claim that the role of *direct* physical action on physical computational objects can make abstract concepts more accessible [13]. Less widely appreciated is the value of actions that can simplify mental tasks which involve abstract concepts or symbolic representations [9]. There is a benefit to supporting physical actions on computational objects which can make difficult mental tasks easier to perform. For example, the physical manipulation of jigsaw puzzle pieces makes the requisite mental tasks of visual search, image visualization and spatial rotation easier to perform. Task completion requires the tight coupling of mental and physical operations. As the

CHI 2009, April 3–9, 2009, Boston, MA, USA.

proportion of physical to mental operations is increased the task becomes easier to perform (up to a threshold). As users' skill development proceeds through practice they may reduce the proportion of physical to mental operations to an optimal level as they develop the requisite mental skills.

The value of using the hands to manipulate objects in problem solving is not necessarily confined to direct input methods. Objects and digital representations of objects can be manipulated indirectly with a mouse. In order to compare the benefits of indirect and direct approaches, we require a methodology that can be equality applied to both. The methodology must take into account the cognitive benefits of object manipulation in problem solving in general.

Thinking with Hands -- Complementary Actions

An individual or group of individuals can improve their cognitive strategies for solving a problem by adapting the environment. One of the ways individuals do this is through a complementary strategy. Kirsh defines a complementary strategy as any organizing activity which recruits external elements to reduce cognitive loads [7]. A complementary action can be recognized as an interleaved sequence of mental and physical actions that result in a problem being solved in a more efficient way than if only mental or physical operations had been used. The external elements may be fingers or hands, pencil and paper, stickies, counters, or other entities in the immediate environment. Typical organizing activities include arranging the position and orientation of nearby objects, manipulating counters, rulers or other artifacts that can encode information through manipulation. Complementary strategies involve actions which can be either pragmatic or epistemic as described below.

Thinking with Hands -- Epistemic Actions

Individuals can use physical action in the environment to lighten mental work through epistemic actions. Epistemic actions are those actions used to change the world in order to simplify the problem-solving task. This is often subtly misstated or misinterpreted as manipulating something in a task to better understand its context. However, the defining feature is that the action changes the world in some way which makes the task easier to solve. The classic example involves a user manipulating pieces in the computer game Tetris -- not to solve the task at hand but to better understand how rotated pieces look [9]. Physical action transforms the difficult task of mentally visualizing possible rotations and offloads it to the world, making it a perceptual-motor task of physically rotating pieces in order to make the subsequent play of the game easier. In this case actions aren't directly related to solving the current falling pieces in Tetris but instead make it easier to understand how pieces look when they are rotated in general so subsequent game play is easier. In contrast, pragmatic actions are those actions whose primary function is to bring the individual closer to his or her physical goal (e.g., winning the game, solving the puzzle, finding a solution).

From a methodological standpoint, it is often hard to prove that an individual performs a particular action for epistemic rather than for pragmatic reasons. An action can serve both epistemic and pragmatic purposes simultaneously. In the realm of jigsaw puzzles, players typically organize pieces into groups containing: corner pieces, edge pieces, same colored pieces, or pieces of similar shape. These intermediate steps support visual search, but their function is epistemic, in that they do not bring players physically closer to their pragmatic goal of placing pieces to complete the puzzle [8].

A Prototypical Example – Jigsaw Puzzle

A jigsaw puzzle is a visual search activity that is traditionally solved by two or more players using a combination of single and two handed manipulation of physical objects. Solving a jigsaw puzzle requires a combination of purely internal mental operations with physical operations on objects [4, 8]. From an embodied cognition perspective, a jigsaw puzzle is a prototypical activity that requires the combination of purely internal mental operations with physical operations on objects [4, 8]. Solving the puzzle requires that mental operations be tightly coupled with physical actions in the environment to test hypotheses and generate new states of information.

Physical manipulation may serve three intertwined roles in jigsaw puzzle solving. First, players may manipulate pieces simply to move pieces into their correct positions. We call these *direct placement actions*. Second, players may use a complementary strategy to manipulate pieces on route to their correct placement because doing so makes the mental operations of visual search, image visualization and/or spatial rotation easier to perform by offloading part of each operation to physical action in the environment [7]. These actions are often part of a trial and error approach to visual search and as such, their function is pragmatic. We call these *indirect placement actions*. Third, players may use a complementary epistemic strategy in which they explore the problem space (e.g., organize puzzle pieces into groups containing corner pieces, edge pieces, or pieces of the same colour or shape). These actions result in a simplification of the task through changing the environment. Their function is epistemic [8, 10]. We call these *exploratory actions*.

These three kinds of actions are found in a range of other kinds of activities involving object manipulation. For example, In the URP urban planning tabletop [14], when a user moves a building (which can be represented either digitally or physically) to determine wind flow, we can interpret the nature of the action on the building based on the role moving it plays in problem solving. We can interpret the action that results in the movement of a building as *direct placement* when the user knows where they want to place the building and does so. We can interpret the action as *indirect placement* when the user moves the building until a desired wind flow state is achieved. We can interpret the action as an *exploratory* move when the user moves the building in order to explore how the system responds for various buildings locations and orientations.

METHODOLOGY

The coding and quantizing of action events in object manipulation tasks requires a theoretically based methodology that defines classes of observable behavioral events based on the role that hands-on action plays in thinking. We provide our methodology for pairs of subjects working together. It can be used for a single user or extended to accommodate any number of multiple users.

Classification of Observable Behavior Events

For a user manipulating pieces to solving a puzzle, we have identified several kinds of observable behavioral events. Each type of event can occur using the mouse to manipulate a digital puzzle piece or the hands to directly act on a physical puzzle piece. We acknowledge that this classification scheme may need to be "tuned" to suit other object manipulation activities. However, the three main manipulation classes as described in the next paragraph are appropriate for many activities and contexts.

Subjects' behaviors in video segments can be coded using an event based a unit of analysis called a "touch." A *touch event* begins when a puzzle piece is first "touched" (by cursor or hand) and ends when the piece is "let go." Based on the roles of object manipulation in spatial problem solving, we used three classes of touch events: direct placement, indirect placement and exploratory. A *direct placement* touch event is when manipulation only serves to orient the piece to the correct location. We can visually identity direct placement event when a user picks up a specific piece and immediately places it, often with the hands directly following eye gaze. There is no hesitation. An *indirect placement* touch event occurs when the subject manipulates the piece in order to determine where it fits and then places it. In this case, physical manipulation serves to offload some portion of mental operation to physical action. A prototypical example is when a subject picks up or selects a random piece and moves the piece across the display, visually comparing it to the puzzle image in order to see where it might fit using a trial and error approach. An *exploratory* touch event is when a user touches or moves a piece but does not place the piece in the puzzle. A prototypical example is when a subject organizes edge pieces by placing them in a pile.

We also included *on-task but non-touch* events (e.g., gazing at the puzzle; verbal or gestural communication related to the task) and *off-task* events into our coding scheme. Our scheme is mutually exclusive. The three classes of touch events (i.e., direct, indirect and exploratory) combined with the non-touch but on-task and off-task classes constituted all observable behaviors. We did not observe users simultaneously but independently placing two pieces into the puzzle, one with each hand, so we confine our analysis scheme to the dominant hand that is manipulating an object. For paired interaction all video was coded twice, once for each subject. Video examples of each action event class can be found online. (Due to ethical considerations with minors, please contact primary author for details).

Relative Measures

In order to compare single mouse input with multi-user input we developed relative measures. Manipulation time (MT) is the absolute amount of time that pairs spend "touching" a puzzle piece, using either their hands on tangible objects or the mouse on digital objects. MT includes direct, indirect and exploratory touches. CT is completion time. For an activity that can be done multiple times, CT_n is the nth completion time. The value of MT for a session exceeds completion time (CT) since the MTs for each subject in a pair is summed. From this we can derive relative manipulation time for a pair of subjects for their first puzzle completion (RMT_{CT1}). In general RMT is the summed MTs for each subject in a session divided by n times the CT_1 (where n = number of subjects). For a pair of subjects we have,

$$RMT_{CT1} = \frac{[MT_{CT1} \text{ subject a} + MT_{CT1} \text{ subject b}]}{[2*CT_1]}$$

RMT_{CT1} gives a relative proportion of the puzzle first completion time that participants spent manipulating puzzle pieces. For example, RMT_{CT1}= .75 means that 75% of the time taken to complete the puzzle the first time was spent with one or both subjects manipulating puzzle pieces. We can also calculate relative measures for other event classes. For example, $ROTNT_{CT1}$ is the relative time during first completion spent in on-task but in non-touch activity (OTNT). Similarly, $ROffT_{CT1}$ is the relative time spent during first completion time in off task activity (OffT).

In order to further examine the proportion of touch activity spent in direct, indirect and exploratory actions we develop a second relative mean time metric. We can calculate RMT for each kind of touch event as a percentage of active manipulation time only. We then have relative measures of direct placement (RMT_1.DP), indirect placement (RMT_1.IP), exploratory (RMT_1.Ex). These variables give us an indication of the breakdown of manipulation time (MT) into direct placement, indirect placement and exploratory actions only for active manipulation time. For a pair of subjects we have,

$$RMT_1.XX = \frac{[MT_1.XX \text{ subj a} + MT_1.XX \text{ subj b}]}{[2*MT_1]}$$

For example, $RMT_1.DP = 15\%$ means that 15% of the time actively manipulating objects was spent with one or both subjects taking direct placement actions on puzzle pieces. Using these variables we can compare the single-controller mouse group with the multi-access tabletop groups.

Temporal Analysis

After classification it is possible to create temporal visualizations of subject events for each session. We also suggest calculating average frequency and durations for each event class, and running lag sequential analysis in order to determine common sequential patterns of actions. Our recent work suggests the importance of interpretations based on both relative measures and analysis of the temporal patterns of interaction in order to fully understand the details of interaction.

CONCLUSION

Understanding the opportunities and challenges of a tangible approach to children's computational activity design requires new methodologies that investigate the role of the hands in human computer interaction. We contribute such a methodology based on an embodied perspective on cognition.

REFERENCES

1. Antle, A.N., The CTI framework: Informing the design of tangible systems for children. In *Proceedings of Conference on Tangible and Embedded Interaction*, (Baton Rouge, Louisiana, 2007), ACM Press, New York, NY, USA, 195-202.
2. Antle, A.N., Droumeva, M. and Corness, G., Playing with The Sound Maker: Do embodied metaphors help children learn? In *Proceedings of Interaction Design for Children*, (Chicago, IL, USA, 2008), ACM Press, New York, NY, USA, 178-185.
3. Chipman, G., Druin, A., Beer, D., Fails, J., Guha, M. and Simms, S., A case study of tangible flags: a collaborative technology to enhance field trips. In *Proceedings of Conference on Interaction Design and Children*, (Tampere, Finland, 2006), ACM Press New York, NY, USA, 1-8.
4. Clark, A. *Being There: Putting Brain, Body and World Together Again*. Bradford Books, MIT Press, Cambridge, MA, USA, 1997.
5. Droumeva, M., Antle, A. and Wakkary, R., Exploring ambient sound techniques in the design of responsive environments for children. In *Proceedings of Tangible and Embedded Interaction*, (Baton Rouge, LA, USA, 2007), ACM Press, 171-178.
6. Fails, J., Druin, A., Guha, M., Chipman, G., Simms, S. and Churaman, W., Child's play: a comparison of desktop and physical interactive environments. In *Proceedings of Conference on Interaction Design and Children*, (Boulder, Colorado, 2005), ACM Press New York, NY, USA, 48-55.
7. Kirsh, D., Complementary strategies: Why we use our hands when we think. In *Proceedings of Annual Conference of the Cognitive Science Society*, (1995), 212-217.
8. Kirsh, D., Distributed Cognition, Coordination and Environment Design. In *Proceedings of the European Conference on Cognitive Science*, (1999), 1-11.
9. Kirsh, D. and Maglio, P.P. On Distinguishing Epistemic from Pragmatic Action. *Cognitive Science*, *18* (4). 513-549.
10. Klemmer, S., Hartmann, B. and Takayama, L., How bodies matter: five themes for interaction design. In *Proceedings of Designing Interactive Systems*, (University Park, PA, USA, 2006), ACM Press, 140-149.
11. Marshall, P., Do tangible interfaces enhance learning? In *Proceedings of Conference on Tangible and Embedded Interaction*, (Baton Rouge, Louisiana, 2007), ACM Press, New York, NY, USA, 163-170.
12. Price, S., Rogers, Y., Scaife, M., Stanton, D. and Neale, H. Using 'tangibles' to promote novel forms of playful learning. *Interacting with Computers*, *15* (2). 169-185.
13. Resnick, M. Computer as paintbrush: Technology, play, and the creative society. in Singer, D., Golinkoff, R.M. & Hirsh-Pasek, K ed. *Play = Learning*, Oxford University Press, 2006.
14. Underkoffler, J. and Ishii, H. Urp: a luminous-tangible workbench for urban planning and design. In *Proceedings of Conference on Human Factors in Computing Systems*, ACM Press, Pittsburgh, Pennsylvania, United States, 1999, 386-393.
15. Xie, L., Antle, A.N. and Motamedi, N., Are tangibles more fun? Comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. In *Proceedings of Conference on Tangible and Embedded Interaction*, (Bonn, Germany, 2008), ACM Press, New York, NY, USA, 191-198.
16. Zuckerman, O., Arida, S. and Resnick, M., Extending tangible interfaces for education: Digital montessori-inspired manipulatives. In *Proceedings of Conference on Human Factors in Computing Systems*, (Portland, Oregon, USA, 2005), ACM Press New York, NY, USA, 859-868.