Thinking with Hands: An Embodied Approach to the Analysis of Children's Interaction with Computational Objects

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

Milena Droumeva
School of Interactive Arts and Technology
Simon Fraser University, Surrey, B.C., Canada V3T 0A3
mdroume@sfu.ca

Daniel Ha
School of Interactive Arts and Technology
Simon Fraser University, Surrey, B.C., Canada V3T 0A3
dha@sfu.ca

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA
ACM 978-1-60558-247-4/09/04.

Abstract
We present the theory and mixed methods approach for analyzing how children's hands can help them think during interaction with computational objects. The approach was developed for a study investigating the benefits of different 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 in number of access points when comparing different input methods, we develop a series of relative measures based on our classification scheme.

Keywords
Input methods, video analysis, physical interaction, complementary actions, epistemic actions, children.

ACM Classification Keywords
H5.2. [Information interfaces and presentation]: User Interfaces –*Evaluation/methodology; interaction styles.*

Introduction
The embodied nature of tangible user interfaces has become of increasing interest to designers of children's educational technologies [1, 7,8]. 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). However, there is little empirical evidence to date to support such claims in the realm of children's tangible computing [1, 7]. Understanding the role that the hands play in supporting certain mental processes can help guide design decisions about how to choose an input method and design representations for a particular activity. Studying how children use their hands to augment developing cognitive abilities provides a window on physical 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 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 the theory and overview of a mixed methods approach for comparing the type, duration and temporal sequencing of children's hand actions on objects. We focus on an age appropriate spatial problem solving task for children aged 7 to 10. To ensure that our approach can be widely applied we chose a task which can be supported computationally using different input methods and using either digital or physical objects. A jigsaw puzzle is such an activity as described in [10]. We present our method using a jigsaw puzzle task for illustrative purposes.

Theoretical Framework

Computational Manipulation

Computational objects can be manipulated using indirect (e.g., mouse, touchpad) and direct (e.g., touch screen, 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 [8]. Less widely appreciated is the value of actions that can simplify mental tasks which involve abstract concepts or symbolic representations [5]. There is a benefit to supporting physical actions on computational objects which can make difficult mental tasks easier to perform.

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 or using a touchpad. In order to compare the benefits of indirect and direct approaches, we require a method that can be equality applied to both. The method 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 [4]. 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. 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 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. A classic example involves a user manipulating pieces in the computer game Tetris -- not to solve the task at hand but to better understand how difficult task of mentally visualizing possible rotations and offload it to the world, making it a perceptual-motor task. Epistemic physical rotations aren't directly related to placing the current falling pieces but instead make it easier to understand how pieces look when they are rotated, making subsequent game play 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 [5].

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 or two handed manipulation of physical objects. 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 [3, 4]. 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 [4]. 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 or structure the problem space (e.g., compare pieces; organize pieces into groups containing corner pieces, edge pieces, or pieces of the same colour or shape). These actions often result in a simplification of the task through changing the environment. Their function is often epistemic [5, 6]. 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 [9], when a user moves a building 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. If the user knows where they want to place the building and does so, we can interpret the action as *direct placement*. If the user moves the building to various locations until a desired wind flow state is achieved, we can interpret the action as *indirect placement.* If the user moves the building simply to explore how the system responds for various buildings locations and orientations, we can interpret the action as *exploratory*.

Thinking with Hands Method

The coding of hand-action events from a video record requires a theoretically based method that defines role that hands-on action plays in thinking. Once action events are coded by class, they can be quantized (e.g., frequency, duration). The temporal sequence of events can also be analyzed. We used Noldus Observer for our analysis. Other video analysis packages that support event based analysis can be used. For illustrative purposes we describe our method for pairs of subjects. Each subject is first coded separately and later combined with their partner's data. This method can be used for a single user or extended to accommodate any number of multiple users.

Classification of Observable Behavioral 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, tangible or digital 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 involving object manipulation.

Subjects' behaviors 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 action. A direct placement touch event is when manipulation only serves to orient the piece to the correct location. We can visually identify 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 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

Once video data is coded it can be quantized. For example, video analysis software (e.g., Noldus Observer) can calculate the duration of each event. In order to compare single mouse input with multi-user input we developed relative measures for event duration. Manipulation time (MT) is the absolute amount of time that pairs spend "touching" a puzzle piece, using either their hands or the mouse. 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 are 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. Similarly, we can calculate relative measures for other event classes.

In order to further examine the proportion of touch activity spent in direct, indirect and exploratory action we develop a second relative mean duration metric. We can calculate RMT for each kind of touch event as a percentage of active manipulation time only. For a pair of subjects we have,

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

where XX is the event class. 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.

Temporal Analysis

After classification, Noldus Observer can be used to create temporal visualizations of subject events for each session. This allows for the visual identification of patterns of event classes within and between sessions. We also suggest calculating average frequency and durations for each event class, and running lag sequential analysis in order validate observed sequential patterns. 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.

Contribution and Future Work

Understanding how to design children's computational activities requires new approaches that investigate the role of the hands in human computer interaction. We contribute such an approach based on an embodied perspective on cognition. Through relative measures and temporal analysis of "thinking with hands" event classes, we can objectively compare the benefits of different design approaches. This short paper contributes a description of our method in the context of a jigsaw puzzle task. We hope this work will support other designers and researchers to consider the role the hands play in thinking during physical, graphical and tangible interaction in order to choose and design appropriate input and representational supports for thinking.

Publications under review present the details and results of the application of this method to the jigsaw puzzle study [2]. Future work is required to validate the method and extend it in order to analyze other object manipulation tasks and applications.

Acknowledgements

This research was funded by NSERC and SSHRC.

References

[1] Antle, A.N. The CTI framework: Informing the design of tangible systems for children. In *Proc.TEI 2007*, ACM press (2007), 195-202.

[2] Antle, A.N., Droumeva, M. and Ha, D. Hands on what? Comparing children's mouse-based and tangible-based interaction. In *Proc. IDC 2009*, ACM Press (2009), submitted (under review).

[3] Clark, A. *Being There: Putting Brain, Body and World Together Again*. Bradford Books, MIT Press, Cambridge, MA, USA, 1997.

[4] Kirsh, D., Complementary strategies: Why we use our hands when we think. In *Proc. of Annual Conference of the Cognitive Science Society*, (1995), 212-217.

[5] Kirsh, D. and Maglio, P.P. On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18* (1994), 513-549.

[6] Klemmer, S., Hartmann, B. and Takayama, L., How bodies matter: Five themes for interaction design. In *Proc. DIS 2006,* ACM Press (2006), 140-149.

[7] Marshall, P., Do tangible interfaces enhance learning? In *Proc.TEI 2007,* ACM Press (2007), 163-170.

[8] Resnick, M. Computer as paintbrush: Technology, play, and the creative society. In (eds.) Singer, D., Golinkoff, R.M. & Hirsh-Pasek, K., *Play = Learning*, Oxford University Press, 2006.

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

[10] 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 *Proc. TEI 2008,* ACM Press (2008), 191-198.