Comparing Motor-Cognitive Strategies for Spatial Problem Solving with Tangible and Multi-touch Interfaces

Alissa N. Antle
School of Interactive Arts and Technology
Simon Fraser University
250 -13450 102 Avenue
Surrey, BC, Canada V3T 0A3
aantle@sfu.ca

Sijie Wang
Google Shanghai
60F, Shanghai World Financial Center
100 Century Avenue, Pudong New Area,
Shanghai 200120, China
wangsj.xm@gmail.com

ABSTRACT
We present the results from a mixed methods comparison of a tangible and a multi-touch interface for a spatial problem solving task. We applied a modified version of a previous framework to code video of hand-based events. This enabled us to investigate motor-cognitive strategies as well as traditional performance and preference constructs. Sixteen adult participants completed jigsaw puzzles using both interfaces. Our results suggest that the 3D manipulation space, eyes-free tactile feedback, and the offline workspace afforded by the tangible interface enabled more efficient and effective motor-cognitive strategies. We discuss the implications of these findings for interface design; including suggestions for spatial and visual structures that may support epistemic strategies, and hybrid interfaces where tangible handles may be used as structural anchors as well as controls and representational objects.

Author Keywords
Epistemic actions, hands-on interaction, tangible interaction, multi-touch interaction, interactive tabletops.

ACM Classification Keywords
H5.2. User Interfaces: Theory & Methods.

General Terms
Design, Human Factors, Theory.

INTRODUCTION
Tangible user interfaces (TUIs) enable direct, hands-on interaction with physical objects. Multi-touch interfaces enable direct, hands-on interaction with digital objects. An embodied perspective on cognition suggests that when it comes to thinking, all hand actions are not equal [1, 6]. The function of some hand actions is to move objects in order to achieve a goal. While other hand actions serve to improve or change the process of thinking related to achieving a goal. How might TUIs support the kind of hand actions that are part of the thinking process? How might multi-touch interfaces support such hand actions? What are the similarities and differences in how TUI and multi-touch interfaces enable thinking with and through the hands? In this paper, we summarize findings from an observational video study about how tangible and multi-touch tabletop interfaces support different kinds of hand actions important in spatial problem solving. To investigate hands-on motor-cognitive strategies we adapt and apply an existing video coding framework about how hand actions help us think. We discuss results in terms of the relations between performance, preference, temporal sequences of events, and motor-cognitive strategies. We compare our results to previous work with children solving similar spatial problems. Based on our findings, we end with a discussion of design considerations for hands-on interaction with tangible, multi-touch, and hybrid tabletop systems.

TEI 2013, Feb 10-13, 2013, Barcelona, Spain.

RELATED WORK
TUIs and multi-touch implemented on tabletops both involve direct hands-on, bimanual interaction. There are several differences between these two interface styles related to the manipulation space and feedback modality. The manipulation space of TUIs is three-dimensional; the third dimension enables not only more interaction space but also richer spatial information. The manipulation space for multi-touch is two-dimensional; input and output are restricted to the display surface. TUIs enable tactile feedback while multi-touch does not. The absence of tactile feedback may require a higher reliance on visual feedback. However, both interfaces have benefits as well as limitations. Multi-touch interfaces are increasingly available commercially and implementation costs are can be less than for TUIs. Multi-touch devices are typically more portable (e.g. iPad) and robust (e.g. Microsoft Surface) than custom-made TUIs.

TUIs and multi-touch interfaces are not mutually exclusive. A hybrid form is becoming more common, especially on interactive tabletops like Surface and reacTable [9, 10]. These devices have the ability to recognize both physical objects and touches. For example, Surface allows users to put a cell phone on the tabletop, which serves as a tangible agent in the subsequent interaction. The phone triggers the display of a graphical menu that surrounds the phone on the multi-touch screen.

For spatial problem solving activities involving many elements, knowing which interface to choose, or how to combine the two interfaces is challenging. Examples of such spatial activities include: facilities planning, interior design, bridge design, construction toys, models for identifying drug binding sites on cell surfaces and learning manipulatives. To address this problem, designers need to understand more than just how the differences between these two interface styles may affect motor performance (e.g. task time, accuracy) and user preference. They need to understand how each interface style affects the kinds of actions used as well as the sequence of actions, and how these in turn affect motor-cognitive problem solving strategies.

Studies that focus on the relationship between users' performance using TUI or multi-touch interfaces are increasingly being published (e.g. [2, 5, 13]). These studies reveal important features of interfaces that designers should consider to enable efficient performance. For example, the presence of tactile feedback on TUIs enables faster acquisition of tangible control widgets than graphical widgets [8]. A tangible keyboard supports faster typing than a virtual one because of its tactile feedback [15]. These studies focus on motor behavior in interaction and often involve users' performance during simple manipulation or acquisition tasks. Other studies focus on the elicitation and description of gestures that particular interfaces afford. For example, in [14] the authors study a sorting task on a physical, touch and TUI table. They describe gestures enacted by novice users and observe that interaction with the physical table enabled transfer of some effective gestures to the TUI. While their work provides insight into usability, they do not look at the motor-cognitive function of gesture with respect to the task.

There is little work that investigates how specific interface properties support users' motor-cognitive processes (i.e. thinking with hands) during tasks such as problem solving. Preliminary work has provided some insights into the relations between interface styles, hands actions and problem solving. In a comparative study between multi-touch and physical objects, Terrenghi et al. examined how users organize objects in space. They found that that few users took advantage of the bimanual interaction provided by multi-touch tabletop [17]. They also found that in the physical mode, users' two hands worked together in diverse ways. For example, the non-dominant hand often provided a frame of reference for the dominant hand's actions when selecting or placing objects, as predicted by Guiard's Kinematic Chain Theory [7]. In addition, in the physical condition, users often organized objects spatially into groups in order to solve the problem more easily (i.e. they used epistemic actions to simplify the task). In another comparative study of TUI, GUI and physical interfaces, Antle reported on the ways in which different interface styles affected children's spatial puzzle solving strategies [1]. She found that TUI features enabled more exploratory actions, many of which were epistemic, early in the task. This helped children develop a more efficient strategy for solving the puzzles as the task progressed.

In order to further exploration of this design space, we conducted a study looking at how tangible and multi-touch interfaces affect individual adult users' performance *and* motor-cognitive strategies for a spatial problem solving task. By doing this, we identify key interface features that are efficient to use but also support effective problem solving.

THEORY AND FRAMEWORK

An embodied perspective on cognition provides theoretical grounding for understanding motor-cognitive strategies for problem solving [3]. The theory of complementary actions provides key concepts and explanatory details about how and why using our hands to manipulate objects augments spatial problem solving [11]. While we use the terms motor behavior and cognitive processes, it is important to understand that embodied cognition posits that using our hands during thinking results in an integrated process. We can however delineate different kinds of hand actions based on their contribution to problem solving. In order to understand the role that the hands play in thinking during interaction, we need ways to identify, classify and code hand-based behaviours. We used the theory of complementary actions to develop a framework that distinguishes different kinds of hands actions based on their role in problem solving.

An individual or group of individuals can improve their cognitive strategies for solving a problem by adapting the environment. A *complementary action* is an interleaved sequence of mental operations and physical actions that results in a problem being solved more efficiently than if only mental or physical operations had been used [11].

Complementary actions can be epistemic or pragmatic. *Epistemic actions* are those used to change the world to simplify the task [11]. They don't necessarily bring an individual directly closer to their goal. Epistemic actions can do this in three ways:

1. Reduce space complexity by reducing the memory load in mental computation.
2. Reduce time complexity by reducing the number of steps needed in mental computation.
3. Reduce unreliability by reducing the probability of error in mental computation.

An example of epistemic actions is when players in Tetris rotate pieces not to place them, but to better understand how they look in different orientations. Such an action does not directly lead to the solving of the puzzle but makes the subsequent play easier by offloading mental rotation steps to physical rotation. *Pragmatic actions* are those used to bring an individual closer to their goal (e.g. solving the puzzle, winning the Tetris game) [11]. An example of pragmatic actions is when players in Tetris rotate pieces as they fall to place them correctly.

We have developed a categorization scheme for analyzing hand actions for spatial problem solving [1]. There are three main categories of actions: *direct placement actions, indirect placement actions* and *exploratory actions*. Direct placement is when a user has already determined where a piece fits, picks up or selects the piece, and directly places it in the correct position. The action directly progresses the task. It is *pragmatic* and *non-complementary*.

Indirect placement is when the user picks up or selects a piece, manipulates it (e.g. rotates, translates) in order to determine where it fits, and then places it in the correct position. In this case, the user offloads a portion of the mental activity of determining where the piece fits to physical activity of moving the piece through space. Thus it is a *complementary action*. Because this action results in a correct placement, it is also *pragmatic*.

An exploratory action involves moving a piece but not placing it where it fits. Thus, the action does not directly bring the user closer to their goal (finishing the puzzle), but it may make subsequent solving easier. It is *epistemic* and *complementary*. A prototypical example is sorting pieces into piles of edge pieces. We present the specific rules for video coding using this framework under Measures in the Methodology section below.

METHODOLOGY

We conducted a within-subjects design study with sixteen adult participants who solved equivalent tangible and multi-touch jigsaw puzzles. We abbreviate the multi-touch puzzle as TOUCH in the remainder of this paper. A jigsaw puzzle is a prototypical problem solving task that involves complementary actions [3]. Successful completion requires a coupling of mental operations and physical actions on objects. Instead of imagining where pieces might fit, a user can execute these operations physically by rotating or moving pieces and visually checking the result. Action and vision are used instead of mental imagery, projection and memory [12]. A jigsaw puzzle is familiar to many people, and the nature of the task is simple enough for people to engage with immediately.

The Puzzle Interfaces

We used three puzzles to reduce order effects and to enable interface training. We implemented the three jigsaw puzzles in both TUI and TOUCH interfaces. All the puzzles were similar in style, image content, size, and shape and number of pieces. Each puzzle had 54 pieces and a size of 35 x 35 cm. We chose a puzzle with a medium number of pieces to ensure that two puzzles could be completed in an hour long session. Larger puzzles (e.g. 1000 pieces) can take many days to solve. In pilot studies we found players used similar strategies during medium and large size puzzles.

The TUI and TOUCH interfaces were designed and implemented to ensure they were comparable, and that the interface style could be isolated from other factors. The two interface styles of the puzzles were implemented on the same tabletop surface. The tangible and digital pieces were the same size. Pieces could overlap in both interfaces. The completed puzzle (35x35 cm) did not take up the whole display space (89x69 cm), leaving a similar sized working space for both puzzles. The auditory feedback was identical. Since multi-touch does not support tactile feedback, we were limited to visual feedback for the multi-touch puzzle. In line with haptic-visual modality comparative studies, our goal was for the connection feedback to be equivalent in terms of information content and salience of stimulus.

Figure 1 shows one of the TUI puzzles. It was similar to a regular jigsaw puzzle except that each piece had a portion of a reacTIVsion fiducial marker on each edge. When two pieces were assembled correctly, a complete fiducial marker was made and sensed by the digital table. The system detected the position and angle of the marker and used this data to display a white circle beneath the correct connection. Since it is possible to force two incorrect TUI pieces together, but not possible to do this in the TOUCH puzzle, we display a white circle for correct matches in the TUI puzzle (Figure 1). A "bing" sound was played as the auditory feedback. If the connection was incorrect then the complete marker was not made and neither visual nor auditory feedback was given. The TUI pieces had a print resolution of about 150 DPI.

A reference image showing the complete puzzle was provided on the screen since it is common to have a reference image when solving a jigsaw puzzle (usually on the front of the box). A physical block marked with a fiducial marker was provided to control the reference image. By manipulating the block, participants could move or rotate the reference image on the tabletop. To prevent participants from building the puzzle on top of the reference image, which might change the task type from a spatial task to a visual search task, we disabled resizing the reference image. The reference image was locked to 80% of the actual size of the finished puzzle.

The TOUCH interface (Figure 2) controls were designed to allow users to use an unlimited number of fingers to manipulate the pieces. The system recognized common finger gestures used on current multi-touch interfaces (e.g. [5, 17, 18]). For example, one finger was used to drag an object, and two-fingers were used to pan or rotate. The side of the hand was used to spread or move a group of pieces. When two pieces were moved close to each other (<20 pixels), and there was a correct fit between the two facing edges, the two pieces snapped together and triggered a "bing" sound. If the pieces were not a correct match, the system did not connect them. The projector had an approximate density of 30 DPI. The reference image in the TOUCH condition was controlled using the same touch gestures as the pieces. It was also locked at 80% of the actual puzzle size. The similarities and differences between the TUI and TOUCH puzzles are shown in Table 1.

Figure 1. TUI puzzle.

Figure 2. TOUCH puzzle.

	TUI	TOUCH
Tactile feedback	Yes	No
Manipulation space	3D	2D
Connectivity	Physical	Digital
Resolution	Medium	Lower
Visual feedback	White circle	Snapping effect
Audio feedback	"Bing" sound	
Bimanual interaction	Yes	

Table 1. Comparison of TUI and TOUCH puzzles.

Assumptions

An assumption in our work is that feedback about a correct connection in both conditions was equivalent in both information content and salience. In both cases, similar feedback information was given about correct and incorrect connections. In terms of salience, the combination of visual and tactile feedback in the TUI puzzle was assumed to be equivalent to the visual motion feedback in the TOUCH puzzle since motion cues are stronger than static visual cues. We validated this in our pilot study in which users reported that they easily understood both versions of the feedback. Although the digital resolution was lower than the resolution of TUI pieces, users could adequately see details. We acknowledge that a jigsaw puzzle is a simple spatial task. However, it enables us to look for key interface features that may be important to implement in interfaces for more complex spatial tasks. Future work needs to address this.

Participants

Sixteen participants were recruited from graduate and undergraduate students at Simon Fraser University. The within-subjects study design was balanced with interface styles (TUI, TOUCH), puzzle themes (pirates, witches), order (TUI first, TOUCH first) and gender.

Procedure

After a brief introduction to the TUI or TOUCH control functions, participants were given a practice session with a training puzzle on the first interface. Then they were given a different puzzle for the first interface to start solving it. After the first puzzle was completed, there was a short break, followed by was an introduction and a practice session for the second interface, again using the training puzzle. When the second puzzle was completed they were given a post-questionnaire.

Measures

The study was designed to collect both quantitative and qualitative data, including video, questionnaires and observation notes. The majority of the quantitative data used in the statistical tests was obtained by coding the video records of the sessions using the hand action framework. The answers to the questionnaires and observation notes were used to assist the interpretation of the statistical test results.

Performance

Users' overall performance was measured by the completion time and the total count of actions used in each task. Task completion time was computed as the total duration of a participant's on-task activities solving each puzzle. During video coding, the on-task activities were split into a sequence of mutually exclusive events, so the total count of user action events could be calculated.

Hand Actions & Motor-Cognitive Strategy

Hand action data was derived from video records for each participant. An "action" was the base unit for coding hand-based behaviors. It was delimited based on the objects acted upon and the type of hand action. For actions that had a single target (i.e. an individual piece, a group of connected pieces, or the reference image), an action started when a participant touched the target and lasted until they let go of the target. If a participant held two or more pieces simultaneously, the action referred to the object that was moved and/or visually attended to. We noticed that in some cases there was ambiguity about the exact beginning and ending of some events. We estimate a tolerance of +/- 150 milliseconds. To account for possible cumulative coding errors in any class, we interpret small differences with caution.

To achieve a satisfactory inter-rater reliability, four segments of video were randomly chosen and coded individually by three raters until the inter-rater consistency had reached 75%. Next 20% of the videos (i.e. 6 out of 32 sessions) were coded individually by two raters. The inter-rater reliability for all was above 75%. Finally the remaining sessions were coded by a single rater. A summary of the rules for coding each type follows. A flowchart for classifying action types can be found at: www.antle.iat.sfu.ca/Physicality/ThinkingWithHands/Papers/HAF.pdf.

Direct Placement: Non-complementary & Pragmatic

A *direct placement* (or connection) is when a participant mentally determines where a piece fits and physically manipulates the piece to connect it, resulting in a correct placement or connection. A direct placement can be visually identified in several ways. A prototypical situation is when a participant picks up a piece and connects it without any hesitation. Participants often visually fixate between the piece and its correct location, rather than scanning the puzzle. There is minimal repetitive rotation or translation of the piece.

Indirect Placement: Complementary & Pragmatic

An *indirect placement* (or connection) is when a participant physically manipulates a piece in order to determine where it will fit, and then places it correctly. The participant offloads some of the mental activity of determining where a piece fits by manipulating the piece. An indirect placement can be visually identified in several ways. Prototypical examples include moving a piece in order to visually compare it to the reference image or to the partially completed puzzle, and then putting it in the correct position. Repetitive rotations or translations of the piece accompanied by visual scanning are hallmarks of this category.

Exploratory Action: Complementary & Epistemic

An *exploratory action* is when a participant physically manipulates a piece to explore where it might fit on the puzzle or organizes the puzzle space, but does not end up with a correct connection. Prototypical examples include making piles of like pieces (e.g. edges, corners), and making repetitive rotations or translations that do not result in correct placement. Coding exploratory actions can require searching forward in the video record to determine if the action resulted in a benefit that plays out later (e.g. the first placement of a group of like pieces can be identified only after several pieces are piled together). Placement does not result in a correct connection.

Other Event Types

We also coded three classes that were observed during puzzle solving. *An adjustment action* is when a participant quickly manipulates one or more pieces or the reference image in order to immediately further task goals. Adjustment actions are typically less than 3 seconds. An *on-task-non-touch event* is when a participant stops touching any object on the screen for at least 2 seconds but is still attending to the task. An *off-task event* is when a participant temporarily switches to non-task related affairs (e.g. answering a phone call).

Analysis

To explore the effect of interface style on participants' overall performance and strategy both descriptive summaries and inferential statistics were used. For normally distributed data sets we used a t-test for correlated measures, otherwise we used a Wilcoxan comparison of means for correlated measures. Problem solving strategy was analyzed quantitatively by examining the average time spent for each hand action class, the average duration for each action event, the average number of each type of action and the temporal sequence of events. Since the average total task time and number of actions varied significantly between TUI and TOUCH, we also report relative (normalized) time and count results. We ran inferential statistics on normalized data to ensure a fair comparison between TUI and TOUCH. We also report qualitative data based on participants' self-reports and our observation notes.

RESULTS

Performance

All the sixteen participants completed both puzzles. On average, participants took 20:14 minutes (SD=6:58 min) to finish the TUI puzzle and 29:04 minutes (SD=8:05 min), or 44% more time, to finish the TOUCH puzzle. They used 194 hand actions (SD=74.8) to finish the TUI puzzle and 268 actions (SD=97.3), or 40% more actions, for the TOUCH. A t-test showed that the difference in total completion time between TUI and TOUCH was statistically significant ($t(15) = 3.31$, $p < 0.01$). The difference in the total count of actions between TUI and TOUCH was also statistically significant ($t(15) = 2.41$, $p < 0.05$). These results are in line with previous studies that show a performance advantage for TUI compared to TOUCH (e.g. [2, 5, 13]). We suggest that efficiencies due to the use of physical structures for organization (e.g. table edges) and 3D, tactile interaction are features of TUIs that improve overall performance.

Hand Actions and Motor-Cognitive Strategies

On average, participants spent 2:49 minutes (TUI) and 6:11 (TOUCH) making *direct placements*. Once we normalized the data to account for difference in total task time, the relative proportion of time spent making direct placements was 15% (TUI) and 23% (TOUCH). The relative difference is statistically significant ($t(15) = 3.01$, $p < 0.01$). The average duration of a direct placements for TUI (M=5.2 sec, SD=0.8 sec) was significantly shorter ($t(15) = 8.2$, $p < 0.01$) than for TOUCH (M=9.4 sec, SD=1.9 sec) (Figure 3). The relative proportional number of direct actions was 17% (TUI) and 15% (TOUCH), which was not significantly different (Figure 4). In summary, both groups made the same number of direct placements, however, they were made relatively more quickly with the TUI, and temporal analysis showed different distribution of these events (see below).

On average, participants spent 3:32 minutes (TUI) and 3:07 minutes (TOUCH) making *indirect* placements. The proportion of time spent making indirect placements was 19% (TUI) and 11% (TOUCH). The relative difference is statistically significant (t(15) = 3.58, p < 0.01). The average duration of an indirect placement was 10.6 seconds (SD=1.9 sec) for TUI, significantly shorter (t(15) = 7.1, p < 0.01) than for TOUCH, which was 16.6 seconds (SD=2.8 sec). The proportional number of indirect placements was 10% (TUI) and 5% (TOUCH), which was significantly different (W(15) = 367.5, Z = 3.9, p < 0.01). In summary, participants made many more, shorter indirect placements in the TUI condition. The TUI better enabled users to offload the pragmatic part of determining where pieces fit. We suggest that this is in part due to the 3D, tactile interactional space that enabled users to easily connect pieces while also visually attending to the next step in the task. Temporal analysis (see below) indicated that many of these actions took place early in the task, when the puzzle was most difficult to solve.

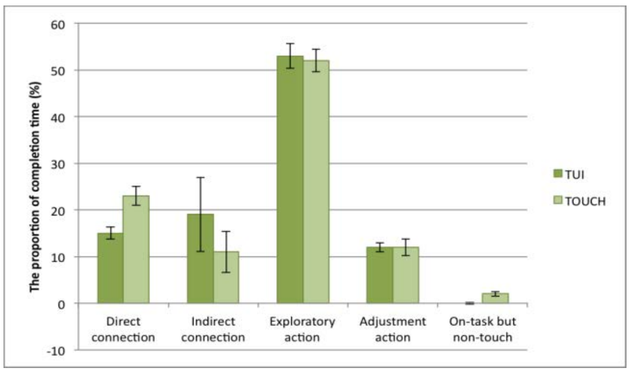

Figure 3. Proportion of time (%) for each type of action.

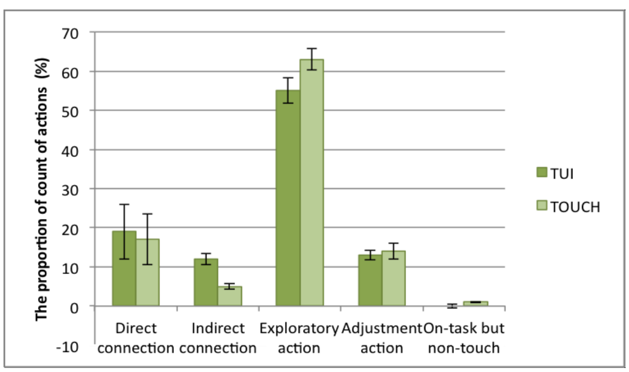

Figure 4. Proportional number of each type of action.

On average, participants spent 11:13 minutes (TUI) and 15:34 minutes (TOUCH) taking exploratory actions. The proportion of time spent taking *exploratory actions* was 53% (TUI) and 52% (TOUCH). The average duration of an exploratory action was 6.2 seconds (SD=1.5 sec) for TUI and 5.6 seconds (SD=1.4 sec) for TOUCH. The proportional number of exploratory actions was 60% (TUI) and 66% (TOUCH). None of these differences are significant.

Temporal analysis of TUI data indicated a common pattern of early exploratory and indirect placements followed by clusters of fast direct placements near completion. A common TOUCH pattern was interleaving exploratory and longer direct placements with little change over the session. This same pattern was found in [1]. The authors suggest the TUI enables a progression from complementary to mental problem solving, and TOUCH enables a steady trial and error approach that does not progress. This conjecture may help explain why the TUI puzzle solving is faster – complementary actions are supported early on, and as the task gets easier (in part due to epistemic actions, and in part due to the puzzle being completed) the strategy changes to one where more computation is done mentally, which is faster. This may have implications for learning. Further work is necessary.

Self-Reports

In response to open questions, participants made several common comments. With the TUI puzzle, participants reported that they were able to move the pieces above the reference image for visual comparison. This strategy was not as effective in TOUCH since the pieces blocked the reference image beneath it. In future designs, this feature could be enabled in a TOUCH puzzle by using transparency.

Participants also reported that they had more difficulty finding pieces in the TOUCH puzzle. As a result, they tended to choose pieces from the top of a pile to work with or they spent time sorting and keeping the pieces organized. This is consistent with the longer total time spent taking exploratory actions with TOUCH since such actions were *needed* to simplify the task. Fifteen out of 16 participants rated TUI easier to use. They liked moving pieces up to their eyes to better see the details when comparing them (even though the resolution was high). With TOUCH users had to move their eyes closer to the tabletop to achieve the same result. In [16] the authors report that participants moved themselves closer to the table and also moved objects closer to themselves. We also observed both approaches but found a preference for the latter. It is more cost effective to move pieces rather than self [12]. Participants liked tactile feedback which enabled them to complete one task (connecting) while beginning the next (looking for the next piece). In these ways, the physicality of the TUI more effectively enabled users to offload pragmatic and epistemic aspects of the task to actions as needed.

Qualitative Observations

We noticed several epistemic sorting strategies in the TUI condition. It was common to use the wooden edges of the table as an offline working zone. For example, players reduced the number of steps in the task by positioning pieces on top of the border close to the area where they would eventually be placed (Figure 5). Participants also used the border for sorting pieces into groups, thereby using spatial position to reducing memory load. In TOUCH, there was plenty of offline screen space but it was used less frequently. We suggest that the TUI offline border space was perceptually and spatially distinct from the screen and that this contributed to enabling this strategy. Some players organized TUI pieces on the screen by butting pieces up against the wooden edge. TOUCH pieces could be butted up against the virtual edge but this strategy was rarely seen. Perhaps it didn't occur to TOUCH users to do this, or the tactile feedback of the TUI piece hitting the edge made this strategy more appealing.

Figure 5. Using the table edges as a sorting area.

Bimanual interaction was observed both in groups but was different in nature. Bimanual actions seen in the TOUCH group were mostly symmetric, such as using one finger of each hand to rotate or move a piece, or symmetrically sweeping some pieces to the left or right with two hands in order to see pieces beneath them. We also noticed more unimanual actions in the TOUCH condition, which is consistent with findings in [17]. Bimanual actions seen in the TUI group were more diverse. For example, some participants frequently held multiple pieces in one hand (typically the non-dominant hand), and used the other hand to choose from these pieces. One participant placed four pieces between his fingers, one piece held vertically between each pair of fingers like a filing cabinet. This kind of handling of objects in 3D space enables effective cognitive-motor strategies (sort, organize, select).

DISCUSSION & DESIGN IMPLICATIONS

Complementary actions are important because they enable integrated motor-cognitive processes that often simplify tasks. Our work provides an exemplar of one way to approach the study of complementary actions in order to inform interface design. While this work is difficult, and fraught with confounds, it enables us to focus on important details about how manipulating objects in space helps, changes or limits our ability to think with and through those objects, and how we might chose and design interfaces to support this. Choosing one interactional approach over another based solely on performance of ease of implementation may lead to designs that are fast and accurate at an interactional level but do not adequately support cognitive processes.

In a spatial task complementary actions enable a user to offload aspects of a difficult mental task to combined mental-physical action. The TUI interface resulted in more effective and efficient interleaved sequences of pragmatic and epistemic actions and enabled a strategy that reduced the reliance of offloading as the task progressed. These benefits were enabled by the 3D manipulation space, which enabled more effective visual search for pieces (pragmatic) and use of the reference image; using hands and 3D edges as sorting structures (epistemic); and by tactile feedback, which facilitated physically fitting pieces while visually attending to something else.

In general, external representations or structures can be designed to help users carry out tasks or prepare for future steps [12]. For example, a visual or spatial structure, can suggest how to order, pair or group objects effectively for an upcoming task. A line suggests a linear order, whereas a circle, suggests grouping. Structures that coordinate mental and external operations are called anchors [12]. They can be visual, spatial, gestural or textual. Our study suggests that spatial structures may have benefits for grouping and sorting tasks. However, structures that need to be readily modified or copied may need to be visual and digital. Another reason to externalize structures is so that they can be shared with others. A structure may act as a referent for shared thought processes. This is particularly important in tabletop applications made for group work.

The offline tabletop borders support epistemic actions by providing visually and spatially separate workspaces for storing (reduce memory load) and organizing pieces (reduce number of steps). Enabling use of offline requires adequate space adjacent to the sensed area so that objects can easily be moved back and forth. Task may be aided by physical or visual structures (e.g. a sorting grid, delineated storage areas) which indicate how the space could be used. For a TOUCH interface it may be possible to implement a surrogate for the physical edge using a physics function that causes pieces to bounce off an edge and then align to that edge. Or pieces could visually change to indicate that they have been moved into an offline area. If digital objects are placed on top of a non-interactive visual aid (e.g. reference image) they should become transparent to allow users to see through them to the aid beneath.

Hybrid interfaces may provide ways to enable epistemic actions and the use of offline space. For example, tangible or mobile devices may be used as "handles" for digital objects. Users could then acquire digital objects, enact complementary motor-cognitive strategies on the hybrid objects (e.g. manipulate physical objects in 3D space, receive tactile feedback), and then release them. Handles could also be used to move objects back and forth between online and offline spaces. The tangible or mobile objects themselves may provide spatial structures for the digital pieces. For example, a user could line up tangible objects in order to line up attached digital objects. Early work explored these ideas (e.g. [4]) and researchers have been coming back to them (e.g. [16]).

CONCLUSION

Like other researchers we have demonstrated performance benefits related to TUIs compared to multi-touch input. We have extended past work to focus on how different key features of the interfaces afford pragmatic and/or epistemic actions, and how these enable effective spatial problem solving. Our preliminary results show motor-cognitive benefits when using TUIs for a simple spatial task. These benefits appear to be supported by 3D, tactile interaction, offline space with spatial structures that simplify epistemic strategies, online space with visual aids and the ease of handling physical objects.

We do not yet know if our findings will generalize to a variety of spatial problem solving tasks that involve similar forms of searching, matching, sorting, grouping, assembly and exploration such as molecular biology research, learning manipulatives, and modeling in engineering and design fields. We encourage other researchers to use the Hand Action Framework to explore different spatial problems involving object manipulation.

ACKNOWLEDGMENTS

This research was funded, in part, by NSERC, SSHRC and the GRAND NCE. Thank you to all the participants, and to the reviewers for their helpful feedback.

REFERENCES

1. Antle, A.N. Exploring how children use their hands to think: An embodied interactional analysis. *Behaviour and Information Technology*, to appear, 1-17. DOI:10.1080/0144929X.0142011.0630415.
2. Brandl, P., Forlines, C., Wigdor, D., Haller, M. and Shen, C., Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In *Proc. AVI 2008*, ACM Press (2008), 154-161.
3. Clark, A. *Being There: Putting Brain, Body and World Together Again*. MIT Press, Cambridge, MA, 1997.
4. Fitzmaurice, G. and Buxton, W., An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input. In *Proc. CHI 1997*, ACM Press (1997), 43–50.
5. Forlines, C., Wigdor, D., Shen, C. and Balakrishnan, R., Direct-touch vs. mouse input for tabletop displays. In *Proc. CHI 2007*, ACM Press (2007), 647-656.
6. Goldin-Meadow, S. *Hearing Gesture: How Our Hands Help Us Think*. First Harvard University Press, Cambridge, MA, USA, 2005.
7. Guiard, Y. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. *J. Mot. Behav. 19*, 4 (1987), 486-517.
8. Harrison, C. and Hudson, S.E. Providing dynamically changeable physical buttons on a visual display. In *Proc. CHI 2009*, ACM Press (2009), 299–308.
9. Jorda , S., Geiger, G., Alonso, M. and Kaltenbrunner, M., The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Proc. TEI 2007*, ACM Press (2007), 139-146.
10. Kirk, D., Sellen, A., Taylor, S., Villar, N. and Izadi, S., Putting the physical into the digital: Issues in designing hybrid interactive surfaces. In *Proc. British Computer Society Conference on Human-Computer Interaction*, (2009), 35-44.
11. Kirsh, D. Complementary strategies: Why we use our hands when we think. In. *Proc. Conf. of the Cog. Sc. Soc.*, Lawrence Erlbaum Associates (1995), 212-217.
12. Kirsh, D. Thinking with external representations. *AI and Society*, 25. (2010) 441-454.
13. Lucchi, A., Jermann, P., Zufferey, G. and Dillenbourg, P., An empirical evaluation of touch and tangible interfaces for tabletop displays. In *Proc. TEI 2010*, ACM Press (2010), 177-184.
14. North, C., Dwyer, T., Lee, B., Fisher, D., Isenberg, P., Robertson, G. and Inkpen, K., Understanding multi-touch manipulation for surface computing. In *Proc. INTERACT 2009*, Springer-Verlag (2009), 236-249.
15. Roeber, H., Bacus, J., and Tomasi, C., Typing in thin air: the canesta projection keyboard-a new method of interaction with electronic devices. In. *Proc. CHI 2003*, ACM Press (2003), 712–713.
16. Terrenghi, L., Kirk, D., Richter, H., Krämer, S., Hilliges, O. and Butz, A., Physical handles at the interactive surface: Exploring tangibility and its benefits. In *Proc. AVI 2008*, ACM Press (2008), 138-145.
17. Terrenghi, L., Kirk, D., Sellen, A. and Izadi, S., Affordances for manipulation of physical versus digital media on interactive surfaces. In *Proc. CHI 2007*, ACM Press (2007), 1157-1166.
18. Tuddenham, P., Kirk, D., and Izadi, S., Graspables revisited: multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. In. *Proc. CHI 2010*, ACM Press (2010), 2223–2232.