Evaluating the Embodiment Benefits of a Paper-Based TUI for Educational Simulations

Tia Shelley
Computer Science
University of Illinois at Chicago
851 S. Morgan (M/C 152)
Chicago, IL 60607 USA
tshell2@uic.edu

Leilah Lyons
Computer Science, Learning Sciences
University of Illinois at Chicago
851 S. Morgan (M/C 152)
Chicago, IL 60607 USA
llyons@uic.edu

Emily Minor
Biological Sciences, IESP
University of Illinois at Chicago
845 W. Taylor Street (M/C 066)
Chicago, IL 60607 USA
eminor@uic.edu

Moira Zellner
Urban Planning & Policy, IESP
University of Illinois at Chicago
412 S. Peoria St. (M/C 348)
Chicago, IL 60607 USA
mzellner@uic.edu

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract

Many claims have been made regarding the potential benefits of Tangible User Interfaces (TUIs). Presented here is an experiment assessing the usability, problem solving, and collaboration benefits of a TUI for direct placement tasks in spatially-explicit simulations for environmental science education. To create a low-cost deployment for single-computer classrooms, the TUI uses a webcam and computer vision to recognize the placement of paper symbols on a map. An authentic green infrastructure urban planning problem was used as the task for a within-subjects with rotation experiment with 20 pairs of participants. Because no prior experimental study has isolated the influence of the embodied nature of the TUI on usability, problem solving, and collaboration, a control condition was designed to highlight the impact of embodiment. While this study did not establish the usability benefits suggested by prior research, certain problem solving and collaboration advantages were measured.

ACM Classification Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces, Input Devices and Strategies, Interaction Styles, K.3.0 [Computers and Education] General.

General Terms

Design, Human Factors

Introduction

The fields of ecology and urban planning make use of Agent-Based Modeling (ABM) software to model and test hypotheses about complex human-natural systems [13]. In keeping with this trend, the College Board's redesign of science education standards for high school Advanced Placement tests adopted a more systems-based perspective on environmental science. In both ecology and urban planning, the relative spatial positions of elements (buildings, permeable surfaces, habitats) are critical to system functions [6], which is why the simulations created to study such systems are considered "spatially explicit." However, the most common method for specifying spatial arrangements in educational ABMs is via a programming interface. Theories of embodied reasoning claim that our abstract visual and spatial concepts are acquired from embodied sensorimotor experiences [3]. This implies that making use of a sensorimotor Tangible User Interface (TUI) to manipulate the spatial positions of simulation elements may better align with users' schema for perceiving and reasoning about spatial relationships, lessening the overhead and streamlining their problem solving.

Cost is a barrier to adoption of technology in schools, so we also aimed to create a *low-cost* TUI using paper symbols as manipulables. A webcam reads the symbols and feeds their position to the simulation, an approach we call *Paper-to-Parameters* (*PtP*), allowing students in even single-computer classrooms to collaboratively interact with a simulation.

This work builds on a small trial that found that using *PtP* to specify a (preset) configuration of 16 elements was over 7 times faster (1m 11s) than an expert user manually programming the same configuration (8m 18s) [10]. While this demonstrates an efficiency benefit *vis a vis* programming, it doesn't allow us to claim that the benefit was afforded by the *embodied* nature of *PtP* alone, as programming is not a direct input method. Surprisingly, we could find little experimental work that tested the embodied nature of TUIs.

Related Work

We are far from the first to suggest that the linkage of perception and cognition via TUIs may benefit usability for spatial problem spaces [e.g., 9], or for collaborative learning, owing to the increased ability of partners to monitor one anothers' actions [e.g., 5]. There are many TUI implementations targeting collaborative learning to be found in the literature [e.g., 2], but most take the form of *qualitative* studies that explore TUI affordances, which illuminates but does not experimentally establish the benefits of TUIs for learning. One laudable exception is a multi-method experiment that compares the performance of pairs of children solving a jigsaw puzzle with a TUI and with a mouse [1]. The between-subject experiment found that the children employed different strategies and were faster using the TUI. However, the control condition was a one-mouse desktop setup, which introduces asymmetries in participation that can affect results. Studies comparing single-mouse vs. multi-mouse desktop collaboration suggest that more symmetric participation results from adding a mouse [11], which can improve collaboration and problem solving. Our work is the first to contrast a TUI against a multi-mouse condition as a control to

The simulation, depicted above, was identical in both conditions, and ran on a separate computer to ensure that in both conditions the manipulation phase was separated from the simulation feedback phase. The swale configuration read from the text file (which is exported by the two UIs) is shown on the right. The lower left shows the graphs that track the progress of stormwater drainage while the simulation executes, and the upper left displays the cost and drainage scores obtained when the execution is complete.

better isolate the potential benefits the embodied nature of TUIs may confer.

Software Development

Thus, for this work we built a dual-mouse, drag-and-drop interface as a control condition[1]. Our experiment contrasts the performance of pairs of users employing the *embodied* direct-manipulation *PtP* against their use of *non*-embodied direct-manipulation dual-mouse input with respect to usability, spatial problem solving, and collaboration. The vision system was trained with 1000 positive images of manipulable paper symbols using the Haar Trainer packaged with OpenCV [8]. The *PtP* interface is a large paper map where users can place re-stickable symbols backed in Post-It® Note Glue. The control dual-mouse interface allows users to drag and drop symbol sprites onto a digital version of the map. In both conditions users press a key to save symbol coordinates to a file that the simulation, implemented in *NetLogo* [7], uses on set-up.

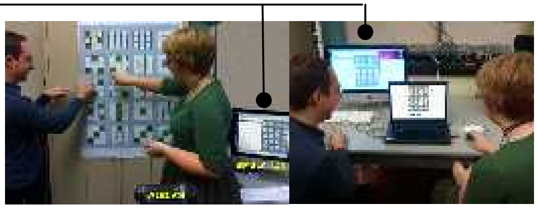


Figure 1. *Paper-to-Parameters* condition depicted on left, dual mouse condition depicted on right.

Experiment Design

Participants

Forty college students were recruited through fliers, classroom visits, and email lists. The participants were paired by preference if one was stated; otherwise they were paired randomly. Three pairs were not included in the analysis owing to data collection errors.

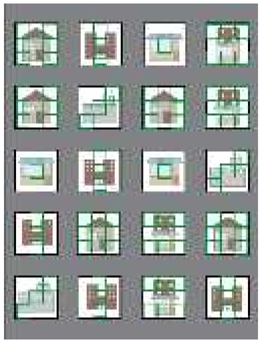
Task

The simulation was adapted from a green infrastructure planning simulation, *L-GrID*, developed for the Illinois EPA-funded *Green Infrastructure Plan for Illinois* project. Swales are areas of permeable land that aid in reducing standing water after rainfall, a common problem in urban areas. Users were asked to place symbols representing swales within a gridded map of an urban landscape. As in real life, only certain lots are available for conversion into swales. Placing the swales to optimize drainage requires strategic spatial placement, as their effectiveness is impacted by the locations of storm drains and by other swales.

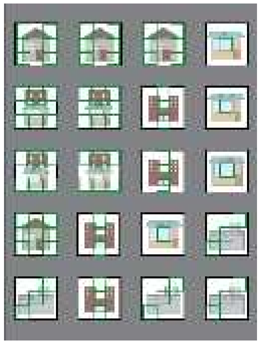
Their challenge was to balance two competing objectives: to reduce the amount of time it takes rain water to drain from the landscape, and to keep the costs to the county low (cost per swale: $10,000.00). To motivate participant performance, we offered a financial incentive that was additively computed from two scores: drainage time and cost. The time score ranged from 0 – 1.25, with 0 corresponding to the maximum possible amount of time for water to drain and 1.25 mapping to the minimum possible time (e.g., all possible swales on the map). The cost score also ranged from 0 – 1.25, with 0 representing the maximum cost (e.g., all possible lots converted to swales) and 1.25 representing the minimum cost (no swales). Participants were paid with respect to the trial with the maximum summed score, rounded up to a whole dollar amount.

---

[1] It should be noted that current common educational ABMs do not have multi-mouse drag-and-drop features.



The layout of Map A



The layout of Map B.
While the exact arrangement of lots available for swale conversion is different, to prevent users from memorizing a favorite pattern and applying it across conditions, the maps are equivalent in terms of scoring possibilities.

Conditions

This study used a 2 x 2 (interface x map design) within-subject with-rotation design. The independent variables were the interface (*Paper-to-Parameters* or dual mouse) and, to prevent memorization of the spatial solution strategy, map design (map a or map b). These conditions were both rotated to account for any learning or inherent differences in the maps. The dependent variables we observed were the number and scores of solutions produced, the time to produce them, the amount and nature of conversation, and Likert ratings of usability and collaboration.

Hypotheses

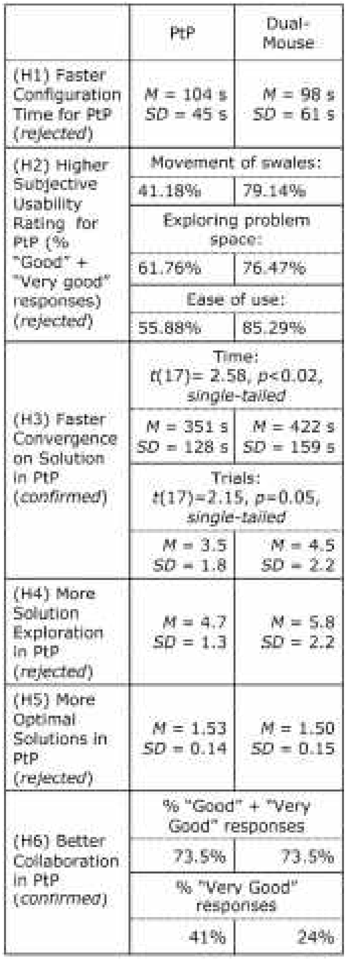
Usability. The sensorimotor alignment of TUIs should promote speedier manipulation [11, 1] and allow users to produce *(H1) Faster Configuration Times in the PtP condition versus Dual-Mouse*. This was measured by how much time is spent executing each spatial rearrangement. We also predicted *(H2) Higher Subjective Usability Ratings in the PtP condition versus Dual-Mouse*, measured on a Likert scale.

Problem Solving. Another supposed advantage of TUIs is that the cognitive work of spatial problem solving can be offloaded to the interface representation itself [5], although some work has failed to show this effect [1]. In theory, this should allow *PtP* users to achieve *(H3) Faster Convergence on Best Solutions in the PtP condition versus Dual-Mouse*, measured by the trials and time taken to reach their highest-scoring solution. This may occur partially because TUIs afford exploration of the solution space [1], so we would also expect *(H4) More Exploration in the PtP condition versus Dual-Mouse*, which would be measured by the number of solutions users attempted. By trying more configurations, we expected *PtP* users to be able to detect appropriate patterns about the system, and therefore attain *(H5) Higher Optimality of Solutions in the PtP condition versus Dual-Mouse*, measured by the highest score attained in each condition.

Collaboration. TUIs should support collaboration by making actions very visible and placing them within a shared physical space [5]. We expect *PtP* users to show *(H6) Better Collaboration in the PtP condition versus Dual-Mouse*, measured by the number and type of comments made during the study, and by the subjective ratings of the participants.

Results

Usability. We found no significant difference in the average time users spent working on each individual configuration *(H1)*, with *PtP* spending slightly more (M=104s, SD = 45s, Min = 40s, Max = 200s) time than dual mouse users (M=98s, SD = 61s, Min = 48s, Max = 279s). This may be less of a judgment levied on TUIs than on the lack of habituation with the input method, which added to configuration times. In ongoing work, we are using video data to measure placement times in each condition's final trial to account for this analysis complication. Occasional computer vision glitches (factored out of the timing data above) may explain the poor responses *PtP* received on a 5-point Likert questionnaire (where 1= very bad, 5= very good, and 3= neutral). For swale movement, only 41.18% rated *PtP* positively (a 4 or 5), while 79.14% rated the dual mouse positively *(H2)*. 61.76% of users rated *PtP* positively for exploring the problem space, compared to 76.47% positive ratings for the dual mouse. For overall ease of use, 55.88% rated *PtP* positively, compared to 85.29% for the dual mouse.



	PtP	Dual-Mouse
(H1) Faster Configuration Time for PtP (*rejected*)	M = 104 s SD = 45 s	M = 98 s SD = 61 s
(H2) Higher Subjective Usability Rating for PtP (% "Good" + "Very good" responses) (*rejected*)	Movement of swales:	
	41.18%	79.14%
	Exploring problem space:	
	61.76%	76.47%
	Ease of use:	
	55.88%	85.29%
(H3) Faster Convergence on Solution in PtP (*confirmed*)	Time: $t(17)$= 2.58, $p<0.02$, *single-tailed*	
	M = 351 s SD = 128 s	M = 422 s SD = 159 s
	Trials: $t(17)$=2.15, p=0.05, *single-tailed*	
	M = 3.5 SD = 1.8	M = 4.5 SD = 2.2
(H4) More Solution Exploration in PtP (*rejected*)	M = 4.7 SD = 1.3	M = 5.8 SD = 2.2
(H5) More Optimal Solutions in PtP (*rejected*)	M = 1.53 SD = 0.14	M = 1.50 SD = 0.15
(H6) Better Collaboration in PtP (*confirmed*)	% "Good" + "Very Good" responses	
	73.5%	73.5%
	% "Very Good" responses	
	41%	24%

Summary of Results

Problem Solving. Users did reach their maximum score significantly faster *(H3)* using *PtP* when measured by both absolute time (M = 351s, SD=128s) for *PtP* versus (M = 422s, SD = 159s) for dual mouse, $t(17)$= 2.58, $p<0.02$, *single-tailed*, and number of trials, with (M=3.5, SD = 1.77) for *PtP* versus (M=4.5, SD= 2.21) for dual mouse, $t(17)$=2.15, p=0.05, *single-tailed*. (Any computer vision read errors were factored out of timing data). This supports our claim that TUIs aid problem solving. However, we saw slightly more attempted solutions with the dual mouse method (M = 5.8) than *PtP* (4.7) $t(17)$ = 2.48, $p < .02$, *two-tailed*. This result did not match our hypothesis *(H4)*, and it is not in keeping with the improved problem solving indicated by the *(H3)* results, so we decided to investigate further.

To better understand the nature of the intermediate solutions, we looked at the percent of available lots holding swale symbols each trial, which showed no significant difference in the *overall percent of lots used* between the two interfaces. However, when we computed a *trial-to-trial delta*, measured as the percent of swales changed (added, removed, or moved) between trials, there was a 10% average trial-to-trial delta between map configurations in the *PtP* condition, while the mouse condition had an average trial-to-trial delta of 17%, $t(17)$ = 2.83, $p < .02$, *single-tailed*. The combination of this finding with the finding of shorter time to best solutions suggests that users may have made more careful, incremental changes in the *PtP* condition. We found corroborating evidence by computing the *solution efficiency*, calculated by dividing the best score obtained by the number of trials, which was significantly higher for *PtP* (M = .34) than for the dual mouse (M = .28), $t(17)$= 2.82, $p<$ .02, *single-tailed*. Although they test fewer solutions, each solution is more useful for achieving a high score. This emulates authentic scientific research practice in its controlled alteration of variables, as opposed to wild testing that doesn't provide deeper knowledge. Other studies of learners' use of ABMs find that oscillatory testing, while not necessarily helpful, is very common [4].

That said, the users did not find closer-to-ideal solutions in either case *(H5)*. The average optimal score for the paper condition was 1.54 (SD = .14, min = 1.31, max = 1.76), while the average score for mouse was a slightly lower 1.5 (SD = .15, min = 1.25, max = 1.74), with no significant difference.

Collaboration. The evaluation of the collaboration through dialog transcription and analysis is ongoing. When asked to use a 1-5 Likert scale to rate their experiences collaborating with both interfaces, in both cases, 73.5% reported positive responses (ratings of 4 or 5). A breakdown of positive responses shows that a higher proportion rated *PtP* "Very Good" (41%) than in the Dual Mouse condition (24%), suggesting that groups found *PtP*'s increased partner-monitoring affordances especially useful. Because familiarity with one's partner might affect collaboration, we collected data on the extent of their perceived friendship. Of the 34 users, 20 considered the people they were working with to be friends or colleagues, while 14 were only acquaintances. When asked to perform a forced-ranking on the interfaces, 50% of acquaintances reported preferring *PtP* for control, while only 10% of friends preferred *PtP* for controlling the simulation. This decrease in preference for *PtP* with increasing familiarity suggests that the partner-monitoring affordances of TUIs may be most useful when partners do not yet have an established working relationship.

Discussion and Future Work

While our usability hypotheses were contradicted, this may be due more to habituation and occasional computer vision read errors. This study did isolate and illustrate some potential benefits associated with the embodied nature of TUIs for problem solving, which seemed due to the use of more deliberate and efficient solution strategies in the TUI condition. The expected collaboration benefits were also found, confirming prior TUI studies (which did not isolate embodiment). We also found that users without an existing relationship may especially benefit from the joint attention and monitoring affordances of TUIs. We will assess this claim with ongoing video and dialogue coding.

Future work will involve constructing a vision system which can handle live, real-time manipulations of the simulation’s spatial parameters, so that users can receive more immediate feedback on their actions rather than waiting to run a new scenario. This system will need to deal more intelligently with the observed jitter problems. We also plan to embed *PtP* in an actual classroom’s curriculum so we can study its impact on content-area learning gains.

Acknowledgments

We thank Brian Slattery and Jingmin Shi for their contributions. This work was supported by NSF grant 1020065.

Citations

[1] Antle, A., Droumeva, M., and Ha, D. Hands on what?: comparing children's mouse-based and tangible-based interaction. In *Proc. IDC 2009*, ACM Press (2009), 80-88.

[2] Falcão, T. P., & Price, S. What have you done! The role of ‘interference’ in tangible environments for supporting collaborative learning. In *Proc. CSCL 2009*, ISLS (2009), 324-334.

[3] Johnson, M.L. Embodied Reason. In Weiss, G., & Haber, H.F. (Eds), *Perspectives on Embodiment : The Intersections of Nature and Culture*. Routledge, London, GBR, 1999, 81-102.

[4] Levy, S., and Wilensky, U. Students' patterns in exploring NetLogo models, embedded in the Connected Chemistry curriculum. At *AERA 2005*, AERA, 2005.

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

[6] Minor, E. and Urban, D. A graph-theory framework for evaluating landscape connectivity and conservation planning. *Conservation Biology 22*, (2008), 297-307.

[7] NetLogo. "http://ccl.northwestern.edu/netlogo/".

[8] OpenCV. “http://opencv.willowgarage.com/wiki/”.

[9] Sharlin, E., Watson, B., Kitamura, Y., Kishino, F., & Itoh, Y. On tangible user interfaces, humans and spatiality. *Personal and Ubiquitous Computing*, *8*, 5, (2004), 338-346.

[10] Shelley, T., Lyons, L., Shi, J., Minor, E., & Zellner, M. (2010). *Paper to parameters*: designing tangible simulation input, Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing. Copenhagen, Denmark: ACM.

[11] Stanton, D., Neale, H., & Bayon, V. Interfaces to support children's co-present collaboration: multiple mice and tangible technologies. In *Proc, CSCL 2002, ISLS (2002),* 342–351.

[12] Tan, K., Lewis, E., Avis, N., and Withers, P. Using augmented reality to promote an understand*ing of materials s*cience to school children. In *Proc. SIGGRAPH Asia 2008*, ACM Press (2008), 1-8.

[13] Zellner, M. Embracing Complexity and Uncertainty: The Potential of Agent-Based Modeling for Environmental Planning and Policy. *Planning Theory and Practice 9*, 4 (2008), 437-4.