Supporting Spatial Skill Learning with Gesture-Based Embodied Design

Po-Tsung Chiu, Helen Wauck, Ziang Xiao, Yuqi Yao, Wai-Tat Fu
University of Illinois at Urbana-Champaign
Urbana, IL, USA
{ptchiu2, wauck2, zxiao5, yuqiyao2, wfu}@illinois.edu

ABSTRACT
Prior research has shown that spatial abilities are crucial for STEM achievement and attainment. The connection between the digital and physical worlds provided by embodied interaction has been shown to enhance performance and engagement in educational contexts. Spatial reasoning is a domain that lends itself naturally to embodied, physical interaction; however, there is little understanding of how embodied interaction could be incorporated into educational technology designed to train spatial reasoning skills. We propose several guidelines for gestural interaction design in spatial reasoning education games based on an empirical study with students at a local afterschool program using a custom-built computer game for training spatial skills. We present a series of gesture sets derived from an iterative design approach that are easy for children to acquire, show sufficient congruency to specific spatial operations, and enable robust recognition from the system. We also compared children's behaviors when playing the game with our gestural interface and a traditional mouse-based interface and found that children take more time but fewer steps to complete game levels when using gestures.

Author Keywords
Embodied Learning; Embodied Cognition; Gestural Interaction; Spatial Reasoning; STEM Education; Design Guidelines.

INTRODUCTION
There has been a rising recognition that training one's spatial reasoning skills at an early age is a key to success in STEM disciplines [16]. They are believed to be malleable and can be transferred to different tasks with spatially enriched education [15]. Embodied learning, on the other hand, has become a popular topic as researchers believe that it could enhance the effectiveness and engagement of learning [10] by leveraging the embodied nature of human cognition to help enact abstract knowledge and concepts through body activity. Together, [5] presented evidence showing that the relationship between body, action, and space help shape our spatial cognition and identified the elements of interaction in embodied systems that engage spatial cognition. This opens up the opportunity to investigate widely adopted embodiment techniques for spatial ability training in educational interventions.

We explored this direction in the present work by incorporating human gestures with Leap Motion sensors[1] into a video game designed specifically for training children's spatial reasoning skills. We used Leap Motion sensors since they are portable, affordable, and accurate on small-scale movements. The video game was a construction game where players build items by attaching the virtual constituent parts together, tapping into several categories of spatial skills like *mental rotation* [12]. On top of that, we followed the theoretical principle of gestural congruency [14] to map the gesture control scheme for manipulating virtual objects to humans' cognitive framework underlying these manipulations in the real world. We included a user study where nineteen participants were recruited from an afterschool program in the Midwestern US to try out our construction game with both gesture and mouse control schemes, during which we refined the gesture control scheme iteratively from feedback and observations in gameplay.

When designing gesture-based interfaces, it is important to distinguish between whether gestures are used as controllers, by which interface objects are manipulated, and whether gestures can help students learn. This distinction is important because as controllers, gestures are likely not as good as other input devices, such as the mouse, since current gesture sensing technologies cannot provide the level of precision of control as these traditional devices. Our goal is to understand how gestures can help students learn spatial skills, assuming that sensorimotor experiences during the learning process can provide a strong foundation for spatial skills as they are integrated into the cognitive operations of spatial reasoning, providing a stronger perceptual-motor grounding of these operations [2].

Our evaluation is therefore more focused on the extent to which we find evidence that the gesture-based interfaces encourage learning, and how they help engage students during the game. Our in-game behavior data suggested that gestures may encourage more deliberation when performing spatial

IUI'18, March 7–11, 2018, Tokyo, Japan

DOI: https://doi.org/10.1145/3172944.3172994

[1] Leap Motion, Inc. 2017 https://www.leapmotion.com

Game Feature	*Spatial Operation*
Object Alignment	Spatial Visualization
Object Rotation	Physical Rotation, Mental Rotation
Location Matching	2D to 3D Projection, Shape Matching

Table 1. The mapping of construction game features designed and specific spatial operations.

operations, since participants performed fewer rotations when using gestures while completing a similar number of part attachment tasks. The post-game survey results also suggest that gestures were as engaging as mouse control, which aligns with the notion that physical engagement in embodied interaction can have conceptual development benefits in education [10].

RELATED WORK

Researchers have previously applied embodied interaction to systems in the educational context based on the theory of grounded cognition and attempted to enhance the effectiveness of learning conceptual knowledge. Embodied interaction powered with diverse technologies including interfaces designed for gestural input involves more of our senses than traditional interfaces like mouse and keyboard [14]. Lindgren and Moshell used MR (mixed reality) to experiment with adopting body-based metaphors into the process of learning conceptual domain knowledge, where the body movements of students were represented by the behavior of virtual asteroid in a planetary astronomy simulation [11]. They found that students created more dynamic mental models to understand the core concept of the content and proposed that MR could physically engage their learning of a domain. Zhang et al. investigated the effect of five interaction modalities, including gesture interaction, on solving spatio-visual problems, and suggested that physical expressions would affect the quality of the solutions to such problems [19]. Also, [4] incorporated a virtual reality headset and tangible interface in a spatial puzzle-solving game to study perspective taking ability, a subskill of spatial cognition. To further untangle the relationship between embodiment and learning, a hierarchical design framework for embodied learning games was proposed in [13]. However, to the best of our knowledge, no one has yet studied how different factors in gesture design could tap into the cognitive operations for spatial reasoning skills.

CONSTRUCTION GAME

The goal of our construction game focused on understanding players' spatial skills with manipulating virtual objects, in particular their abilities of *mental rotation* and *projection*. The primary game mechanic was to build a set of target objects starting with one "base" part and attaching additional parts one by one. We derived our game mechanics from a prior work of Wauck et al. [17] to shape the relations between different game features and specific spatial operations as summarized in Table 1. For example, determining a part's correct attachment location mapped to *2D-to-3D projection* and *shape matching*, while determining its correct orientation mapped to *physical* and *mental rotation*.

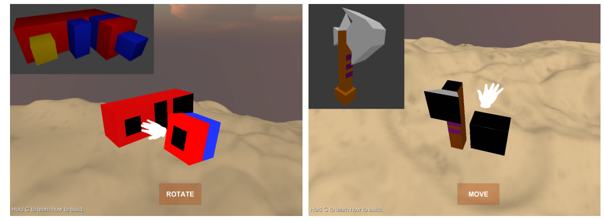

Figure 1. Left: *Easy* level. Right: *Sledgehammer* level. The 2D images on the upper left corners show what the 3D objects would look like when all parts are attached correctly. Rotate and Move modes indicate the spatial operation the player can perform at the moment.

To complete each level, players needed to attach all parts at correct locations and with correct orientations. However, they would not need to decide which part to control, as the parts would appear in a predefined order. The game's interface, as shown in Figure 1, displayed a 2D picture of the complete target object in the upper left corner to guide the player's construction. We designed 4 game levels, each with slightly different training focuses:

- The *Tutorial* level (3 parts) contained basic part shapes, all of the same color, focusing players' attention on distinguishing different shapes to keep things simple. It served as a tutorial level for learning about the game's control.
- The *Easy* level (4 parts) used only cuboids and cubes for part shapes, but varied their colors, forcing players to pay attention to the picture of the completed target object in the upper left corner of the screen when building.
- The *Rocket Boots* level (6 parts) used more complex shapes. The player needed to observe the 2D picture carefully to build the object, as some parts looked like they could be attached at multiple positions on the base part.
- The *Sledgehammer* level (11 parts) was also considerably more challenging than the *Tutorial* and *Easy* levels. All parts except one had the same color. Some parts had similar shapes, as some were occluded by other parts in the 2D picture. To complete this level, players would need to notice subtle details and reason carefully about how the 3D parts matched the picture.

GESTURE DESIGN

Our gesture design focused on two specific operations – *Movement* and *Rotation* – for moving and rotating parts in the game. Initially, we investigated the example in [18] prototyped with swipe gestures of Leap Motion API for the *Rotation* gesture. However, a crucial pitfall of this example we found was the difficulty in distinguishing between the actual swipe gesture and the following movement of the hand returning back to its initial position. This natural inclination to return the hand to its initial position confused the gesture recognition system and caused unintended rotations in game to the frustration of players. Also, the gesture used in the prototype discretized the rotation process, where players would swipe in either of the 6 directions (left/right, up/down, or forward/backward) to rotate the subject in 90 degree increments in the direction of the swipe. This mitigated the fact that rotation in the real world is continuous and the ability to control with continuous visual flow could benefit spatial learning [8]. We present three versions of our gesture control design, V1, V2, and V3, in the following paragraphs. The first three versions were tested with 2 participants each, and we iterated the development of the new version based on feedback received from the post-game survey and from observation during gameplay.

Version 1

The major difference between this version and the aforementioned prototype was *Rotation*. We addressed the two drawbacks of the prototype by changing the scheme based on the following idea. Using hands to build things in the real world, such as a LEGO set, we can manipulate the pieces freely and rotate them in any direction. We redesigned *Rotation* to support such freedom by binding player's hand angle to that of the piece being manipulated. Both *Movement* and *Rotation* would be activated as the player formed a closed fist, and deactivated when the hand opened near a piece in the game. A piece would return to one of the 6 possible orientations parallel to the x, y, and z axes when not rotating. This embodied the player's hand more holistically as the virtual object, in contrast with swiping gestures in the previous prototype.

Version 2

During the gameplay session of V1, the participants had a lot of trouble getting Leap Motion to accurately distinguish between moving, rotation, and releasing from rotation or movement. Particularly frustrating for players were situations in which they had the part rotated properly, but then as they tried to move the part to the correct position, the game detected a rotation operation rather than movement, and as a result, the part became incorrectly rotated. This led us to split moving and rotation into two modes (Figure 1) for the player to switch between by hitting the space key. Both *Movement* and *Rotation* were the same as in V1, but our goal was that the player would be able to perform both operations with confidence.

Version 3

While V2 avoided the confusion encountered between moving and rotation, Leap Motion could not track the hand's orientation in a robust way. We attributed this problem to the hurdle for Leap Motion's depth camera to distinguish different orientations of a closed fist. Hence, rather than turning to other gesture sensing devices, we used a different hand position in V3, namely the opposite hand posture – palm flat and open – to enable moving and rotation. The movement of the open palm would map to the position of the piece in control when in moving mode. When the player tilted their hand in one direction (x, y, or z) as shown in Figure 2, the object would begin rotating in that direction until the player's hand closed back into a fist.

Guidelines

While embodied activities have the potential to enhance learning, we believe that they need to be designed adaptively to fit with the conceptual knowledge we expect the participant to learn. Drawing upon our experience in the user study, we propose the following design guidelines for gestural control

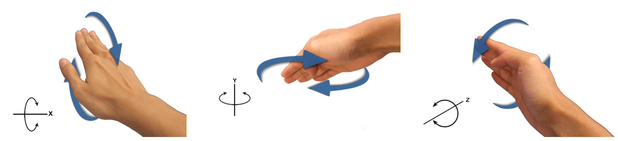

Figure 2. Final *Rotation* gestures. Left to right: *Rotation x, y, and z*.

in spatial skill training games for children: (a) using separate modes for different spatial operations, (b) maintaining gestural congruency to spatial concepts, and (c) using static hand posture for each spatial operation.

First, defining a clear picture of the correspondence between game modes and the spatial operations in Table 1 could help players concentrate on reasoning about one operation at a time, which could reduce cognitive load and frustration. Second, making sure that the gesture performed is congruent to a spatial concept aids the construction of that particular concept and strengthens the action-concept link [10] that can improve learning. We formulated the third guideline based on observations of how the children performed gestures during the study. By using a fixed, static hand posture for each spatial operation, players can focus more on the mapping between large, simple hand movements and the virtual objects being manipulated, while also reducing system errors with camera-based gesture recognition systems like Leap Motion and enabling a consistent (or at least less frustrating) user experience. These guidelines can be used by game designers to develop more effective adaptive gesture sets for spatial games, and by educators seeking to engage participants in higher-order thought processes to acquire new knowledge.

METHOD

The goals of our study were to understand how the gestures designed help and engage children to reason and learn in spatial tasks. We also compared the players' in-game behavior between 1) different versions of the gesture design, 2) the gestural interface and the mouse-based interface. We conducted the study at an afterschool program center and recruited 19 participants through the program director to participate in a 50-minute study session. Their ages ranged from 8 to 12 (median=10, mean=9.84) with 8 of them being female.

Procedure

The participants spent 20 minutes playing each version of the game (gesture versus mouse control) – 40 minutes in total. After playing, they filled out a post-game survey that asked them to indicate their age and gender and rate how easy, fun, frustrating, and tiring they thought each version of the game was and why. We counterbalanced the gameplay order of the gesture and mouse-controlled version. For each control scheme, participants played the *Tutorial* level followed by the *Easy* level, and then were assigned to play one of the more challenging levels: *Rocket Boots* or *Sledgehammer*. If the participant was assigned to play the *Rocket Boots* level for the first control scheme they tried, then they were assigned to play the *Sledgehammer* level for the second control scheme, and vice versa. We playtested gesture versions V1 through

V3 with 2 participants each in the first three sessions and all participants after played with V3, the final version.

In-Game Behavior Metrics

Our primary measures of interest were *completion time*, *number of rotations*, and *number of wrong attachments* in each level since they were related to the difficulty of the tasks and player impulsiveness. We also looked at some other metrics from the data collected, such as the *number of rotations per successful attachment* (the number of rotations divided by number of successful attachment attempts) and the *average time spent between each rotation* (the time taken to finish divided by number of rotations). These two metrics gave us a sense of how children reason when trying to determine the correct spatial relations of the parts, or the case where they perform the rotation operations directly and without bothering to mentally visualize the rotation first.

RESULTS

There were 15 participants that played at least 2 levels with both the mouse-based and gestural V3 interfaces. The other 4 quit before the end of the study due to frustration or because their parents picked them up early from the afterschool program. In addition, we excluded the behavior data from the *Tutorial* level of the game as we provided verbal assistance to the children as they learned how to play the game and demonstrated different gestures for them in this level.

In-Game Behavior Analysis

First, we analyzed differences in behavior between the gesture and mouse control. We found a significant difference in the number of rotations per successful attachment under a repeated measures t-test; players needed fewer rotations to correctly align parts when playing with gestures than when playing with mouse control ($t(27.821) = 2.560$, $p = .016$). Therefore, participants may be performing more redundant rotations with the mouse interface, suggesting that they are using the epistemic mode of immediate "doing" rather than mediated "thinking" [1] when tackling the task. This difference is also consistent with the results that while the number of successful attachments were fairly close for both interfaces in the *Easy* level (mouse-based: mean=4.00, gesture: mean=3.67) and *Rocket Boots* level (mouse-based: mean=5.11, gesture: mean=5.33), participants performed more rotations with the mouse-based interface in the *Easy* level (1.63 times more) and the *Rocket Boots* level (1.77 times more). Interestingly, the only participant who completed the *Sledgehammer* level with gesture controls used only 22 rotations, far fewer than the number of rotations used by those who completed the level using the mouse-based controls (mean=113.2, max=187, min=53). One possible explanation for these findings is that the mouse-based and gestural interfaces may guide the participants to play the game in different ways. Borrowing the analogy from Kahneman's theory of effortless intuition and deliberate reasoning [9], the mouse-based interface could be promoting finding a rapid and heuristic path to the solution, while the gestural interface may be slower yet more accurate and encourage deliberation. It is important to note that the latter over time can become more effortless and rapid [6,7]. In addition, we found that participants performed less rotations along the y axis when using the gestural interface across all game levels, possibly due to the ergonomic limitations of the required open hand posture.

Next, we compared behavior across different versions of the gesture control scheme. Of the four participants using gestural interface V1 or V2, only one completed the *Easy* level, and none of them completed the more challenging *Rocket Boots* or *Sledgehammer* level. The participants using V3 performed better; 50% completed the *Easy* level, and 30% completed the more challenging levels. Also, the number of rotations per successful attachment for V3 (mean=4.91) was less than with the first two versions (mean=9.50). This indicated that V3 facilitated better performance and was easier for the participant to learn to use.

SURVEY

Not surprisingly, most participants reported that the mouse control was easier to use, less frustrating ("*Something didn't work [using gesture]*"), and less tiring ("*Need to move hands a lot [using gestures]*" or "*Arms wouldn't get tired [using the mouse]*"). But when asked how fun they thought using the gestural interface was, participants responded with statements like "*You have lots of fun and more experiments*", "*You get to learn and have fun doing it*", or simply "*Liked to move my hand*", "*My hand is in the game*", etc., showing participants' engagement with the game and using gestures. In addition, some participants (n=4) reported that the gestural interface was either easier to use or more preferable to play with, particularly V3. Even though the participants' familiarity with mouse control may give them an advantage in performance or learnability with our mouse control scheme, the additional learnability difficulties introduced by our gesture control scheme could potentially enhance recall and transfer during the learning process, based on the theory of desirable difficulties [3].

CONCLUSION

In this study, we discussed the design process of our gesture-based spatial learning game, as well as results that supported the notion that gestures can help students learn spatial skills. Although the evaluation involved a limited number of students, the results were in general supportive of the notion that the current design approach is useful for generating embodied games for training spatial skills for children. We expect that a large scale user study with more participants and run for a longer period of time will provide more insights into the details of how these gesture sets can be designed to enhance learning.

ACKNOWLEDGMENTS

This work was partially funded by the Strategic Instructional Innovations Program (SIIP) at College of Engineering of the University of Illinois at Urbana-Champaign and by the NSF grants SES-1419297 and IIS-1441563. We thank Yuan Yao for ideation, and all staff at the Urbana Neighborhood Connections Center for their help with the user study.

REFERENCES

1. Dor Abrahamson and Robb Lindgren. 2014. *The Cambridge handbook of the learning sciences* (2nd ed.). Cambridge University Press, Cambridge, UK, Chapter Embodiment and embodied design.
2. Lawrence W. Barsalou. 2010. Grounded Cognition: Past, Present, and Future. *Topics in Cognitive Science* 2, 4 (2010), 716–724.
3. E. L. Bjork and R. A. Bjork. 2011. *Psychology and the real world: Essays illustrating fundamental contributions to society*. Worth Publishers, New York, NY, USA, Chapter Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning, 56–64.
4. Jack Shen-Kuen Chang. 2017. The Design and Evaluation of Embodied Interfaces for Supporting Spatial Ability. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction (TEI '17)*. ACM, New York, NY, USA, 681–684. DOI: http://dx.doi.org/10.1145/3024969.3025033
5. P. G. Clifton, J. S.-K. Chang, G. Yeboah, A. Doucette, S. Chandrasekharan, M. Nitsche, T. Welsh, and A Mazalek. 2016. Design of embodied interfaces for engaging spatial cognition. *Cognitive Research* 1, 24 (2016). DOI: http://dx.doi.org/10.1186/s41235-016-0032-5
6. Hubert L. Dreyfus and Stuart E. Dreyfus. 1999. The Challenge of Merleau-Ponty's Phenomenology of Embodiment for Cognitive Science. In *Perspectives on Embodiment: The Intersections of Nature and Culture*, Gail Weiss and Honi Fern Haber (Eds.). Routledge, 103–120.
7. Efraim Fischbein. 1987. *Intuition in Science and Mathematics: An Educational Approach*. Springer, Dordrecht, Holland.
8. Corinne Holmes, Steven A. Marchette, and Nora Newcombe. 2017. Multiple Views of Space: Continuous Visual Flow Enhances Small-Scale Spatial Learning. (02 2017).
9. Daniel Kahneman. Maps of bounded rationality: A perspective on intuitive judgment and choice. (????).
10. Robb Lindgren and Mina Johnson-Glenberg. 2013. Emboldened by Embodiment. *Educational Researcher* 42, 8 (2013), 445–452. DOI: http://dx.doi.org/10.3102/0013189X13511661
11. Robb Lindgren and J. Michael Moshell. 2011. Supporting Children's Learning with Body-based Metaphors in a Mixed Reality Environment. In *Proceedings of the 10th International Conference on Interaction Design and Children (IDC '11)*. ACM, New York, NY, USA, 177–180. DOI: http://dx.doi.org/10.1145/1999030.1999055
12. Marcia C. Linn and Anne C. Petersen. 1985. Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Development* 56, 6 (1985), 1479–1498. http://www.jstor.org/stable/1130467
13. Edward F. Melcer and Katherine Isbister. 2016. Bridging the Physical Divide: A Design Framework for Embodied Learning Games and Simulations. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 2225–2233.
14. Ayelet Segal. 2011. *Do Gestural Interfaces Promote Thinking? Embodied Interaction: Congruent Gestures and Direct Touch Promote Performance in Math.* Ph.D. Dissertation. Columbia University. https://doi.org/10.7916/D8DR32GK
15. David Uttal, Nathaniel Meadow, Elizabeth Tipton, Linda Liu Hand, Alison R Alden, Christopher Warren, and Nora Newcombe. 2012b. The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. 139 (06 2012).
16. David H. Uttal, Cheryl A. Cohen, Ken Forbus, Dedre Gentner, Mary Hegarty, Madeleine Keehner, Ken Koedinger, Nora Newcombe, Kay Ramey, and Uri Wilenski. 2012a. Spatial Thinking and STEM Education: When, Why, and How?
17. Helen Wauck, Ziang Xiao, Po-Tsung Chiu, and Wai-Tat Fu. 2017. Untangling the Relationship Between Spatial Skills, Game Features, and Gender in a Video Game. In *Proceedings of the 22Nd International Conference on Intelligent User Interfaces (IUI '17)*. ACM, New York, NY, USA, 125–136. DOI: http://dx.doi.org/10.1145/3025171.3025225
18. Yuan Yao, Po-Tsung Chiu, and Wai-Tat Fu. 2017. A Gestural Interface for Practicing Children's Spatial Skills. In *Proceedings of the 22Nd International Conference on Intelligent User Interfaces Companion (IUI '17 Companion)*. ACM, New York, NY, USA, 43–47. DOI: http://dx.doi.org/10.1145/3030024.3038265
19. Ting Zhang, Yu-Ting Li, and Juan P. Wachs. 2016. The Effect of Embodied Interaction in Visual-Spatial Navigation. *ACM Trans. Interact. Intell. Syst.* 7, 1, Article 3 (Dec. 2016), 36 pages. DOI: http://dx.doi.org/10.1145/2953887