Augmenting Spatial Skills with Semi-Immersive Interactive Desktop Displays: Do Immersion Cues Matter?

Erin Treacy Solovey[1], Johanna Okerlund[2], Cassie Hoef[2], Jasmine Davis[2], Orit Shaer[2]

[1]Drexel University
Philadelphia, PA, USA
erin.solovey@drexel.edu

[2]Wellesley College
Wellesley, MA, USA
{jokerlun, choef, jdavis4, oshaer}@wellesley.edu

ABSTRACT

3D stereoscopic displays for desktop use show promise for augmenting users' spatial problem solving tasks. These displays have the capacity for different types of immersion cues including binocular parallax, motion parallax, proprioception, and haptics. Such cues can be powerful tools in increasing the realism of the virtual environment by making interactions in the virtual world more similar to interactions in the real non-digital world [21, 32]. However, little work has been done to understand the effects of such immersive cues on users' understanding of the virtual environment. We present a study in which users solve spatial puzzles with a 3D stereoscopic display under different immersive conditions while we measure their brain workload using fNIRS and ask them subjective workload questions. We conclude that 1) stereoscopic display leads to lower task completion time, lower physical effort, and lower frustration; 2) vibrotactile feedback results in increased perceived immersion and in higher cognitive workload; 3) increased immersion (which combines stereo vision with vibrotactile feedback) does not result in reduced cognitive workload.

Author Keywords

3-D displays; stereoscopic displays; haptic feedback; vibrotactile feedback; fNIRS; zSpace.

INTRODUCTION

Recent advances in digital display technology are making 3-dimensional stereoscopic (3DS) displays increasingly popular in gaming and entertainment, bringing 3DS capabilities into home televisions, laptops, and handhelds. Beyond entertainment, 3DS desktop displays also promise advantages for users in industry, science, and education. In particular, semi-immersive interactive 3DS displays which support stereo viewing combined with head tracking and a 6 degree-of-freedom (DOF) vibrotactile stylus offer a combination of functionality, footprint, and cost that make them attractive for use in classrooms and in the workplace. Figure 1 shows the zSpace semi-immersive interactive 3DS display.

Figure 1: A user completes a spatial reasoning task using the zSpace 3D stereoscopic display.

By tracking the user's head and hand position, interactive 3DS displays provide enhanced immersive cues to the user that leverage three perceptual abilities: *stereo vision* (perception of depth and 3-dimensional structure obtained from combining slightly different images from each eye), *motion parallax* (depth cue that results from updating the virtual environment in response to head motions), and *proprioception* (sense of relative position of limbs and the effort involved in movement which is supported by direct interaction with 3D objects using a 6DOF stylus). Researchers have suggested that the congruence among these three perceptual abilities can support natural viewing and interaction with 3D scenes that in turn could augment spatial understanding [29, 12]. Research indicates that for certain spatial tasks *fully immersive* CAVE systems significantly improve performance [3, 14, 30]. However, there is little empirical evidence proving whether *semi-immersive* 3DS desktop displays provide an advantage for spatial problem solving, as most research has shown mixed results [22, 24]. We propose that this may be due to the mismatch of traditional evaluation methods in fully capturing the user experience with such interactive tools.

Interaction styles such as virtual reality, which go beyond traditional graphical user interfaces, often pose a challenge for evaluation [18]. Traditional measures such as completion time and error rate do not provide the full spectrum of the user experience during tasks designed to take advantage

AH '15, March 09 - 11, 2015, Singapore, Singapore
Copyright 2015 ACM 978-1-4503-3349-8/15/03…$15.00
http://dx.doi.org/10.1145/2735711.2735797

of these emerging interfaces. Our goal is to explore deeper aspects of the user experience by augmenting traditional user evaluation methods and data collection with emerging wearable sensors for assessing the user's cognitive state. Recently, non-invasive brain and body sensors have been explored for use in the evaluation of emerging interaction paradigms [17, 26, 18].

We present a study that examines whether and how semi-immersive displays with different immersion cues (3D stereoscopic display and vibrotactile feedback) augment spatial problem solving. We use an experimental task, which requires various spatial abilities: determining spatial relationships among objects, manipulating objects in 3D, and visualizing a path. To evaluate users' experience, performance, and task load, we integrate functional near-infrared spectroscopy (fNIRS) (Figure 2) to augment the more traditional metrics including task completion time and standard post-task questionnaires. Using fNIRS for evaluating immersive cues in 3DS displays can shed light on immersion components that augment spatial problem solving. At the same time, such evaluation requires the development of fNIRS analysis techniques that advances the use of fNIRS technology for evaluating novel interaction techniques.

This paper makes two main contributions: 1) We describe findings from a comprehensive user study with 48 users, which integrates quantitative measures with brain activity data to explore whether and how semi-immersive displays with different immersion cues (3D stereo and vibrotactile feedback) augment spatial problem solving. Our findings indicate that both 3D stereo and vibrotactile feedback have significant effect on various aspects of the user experience; 2) We validate and demonstrate the use of fNIRS brain imaging technology in the evaluation of emerging interaction techniques and show how the real-time, continuous brain measures can provide supplemental information about the user experience. Going beyond previous fNIRS work, we provide methods for interpreting the signal over variable length task times, and without specific calibration tasks.

The paper continues with background on spatial reasoning, immersive cues in virtual environments, and fNIRS brain imaging, followed by experimental methods and findings.

BACKGROUND AND RELATED WORK

Spatial Reasoning

Spatial reasoning is defined as the ability to make judgments and reason about objects and their spatial relations [13]. There are two key aspects of spatial reasoning [2]: (1) *visual*, which entails perceiving the visual details of a scene accurately (e.g. color, size, shape), performing transformations upon one's initial perceptions, and recreating aspects of one's visual experience; and (2) *spatial*, which involves making judgments about the position or location of an object, determining the spatial relations, and manipulating objects in three dimensions [7].

While traditional graphical user interfaces lack sufficient support for spatial reasoning, Patterson [30] suggests several ways in which interactive 3DS displays support spatial reasoning, by allowing users to:

- Manipulate an object to scan its features
- Manipulate displayed objects to compare their features
- Manipulate an object to examine its parts
- Manipulate objects to determine their spatial layout
- Align an egocentric representation of a scene with an allocentric representation

We study whether and how these capabilities are supported by interactive 3DS displays with different configurations, which vary in their immersive cues. The configurations are: i) stereo vision with vibrotactile feedback; ii) stereo vision with no vibrotactile feedback; iii) no stereo feedback with vibrotactile feedback; iv) no stereo feedback with no vibrotactile feedback.

Immersion in Virtual Environments

A common measure of the effectiveness of virtual environments is the level of immersion in the experience [32]. In a *fully immersive* virtual environment, the user can move freely, look in all directions, and manipulate objects, giving the perception of reality. Successful immersive environments make the user feel as though they are not interacting with digital information, but with real objects in real space [32]. There are many factors that can contribute to immersion, such as display size, display resolution, scene refresh rate, the look of the 3D objects, stereoscopy, and motion parallax [4]. Feedback when the user interacts with digital objects also increases the realism. To provide these immersive cues, head worn devices are usually required.

Desktop virtual reality devices are limiting compared to head worn devices, and typically offer a *semi-immersive* experience. Despite this, these displays are increasingly popular due to the smaller footprint and lower cost. To determine their value in educational and business settings where fully immersive systems may be impractical, there is a need to understand whether these semi-immersive 3DS displays still provide advantages over standard displays. However, it has been challenging to evaluate such systems.

Evaluation of Immersive Cues in Virtual Environments

Previous work has compared the effects of particular immersion components on user performance. To frame our study, we review this work, focusing on stereo vision and haptic feedback, which are directly relevant to our study.

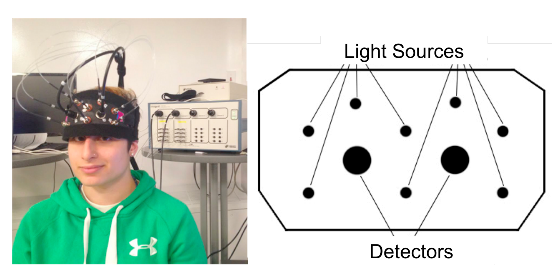

Figure 2. Left: A user wearing the fNIRS device. Right: Sensor geometry with 8 light sources and 2 detectors, which provide 10 source-detector pairs with 3cm distances.

Stereo Vision

Stereo vision has been evaluated in several studies with a *path-tracing task* [3, 33, 34, 41, 42]. These studies indicate that combined head coupling and stereo provide a significant enhancement compared to 2D computer graphics, and that head coupling is more important than stereo in 3D visualization. McKenna [27] found that head coupled perspective improves performance in 3D positioning tasks. Ware and Franck [42] evaluated nine different viewing modes for a path-tracing task and found that combining motion parallax with binocular parallax is important. La Viola [24] explored how user performance of a rotation task is affected by different display modes and rotation techniques. Their findings indicate that stereo viewing with and without head coupling provides an advantage compared with no head tracking and no stereo regardless of rotation technique. Barfield [5] investigated the effects of stereo viewing and head tracking on presence and performance of a task, which required the understanding of a simple 3D object. Results indicated that neither stereo viewing nor head tracking improved accuracy, but head tracking significantly improved the reported level of presence. While these studies focused on the performance of simple isolated spatial tasks, they do not investigate whether semi-immersive interactive displays provide an advantage for high-level, complex spatial problem solving.

A large body of work focuses on the comparison of real-world non-immersive and immersive games and applications. Real-world games are games that have the look and feel of an environment set in the real non-digital world. Stereo has been found helpful in playing simple games where a user is manipulating a single object at a time [12]. Studies showed no significant advantage for 3D stereo in user performance over a 2D display in modern PC-based games [25], but increased user engagement. Kulshreshth [23] studied user performance of games using 3D motion controllers in 3D stereoscopic vision compared to monoscopic viewing and found a positive effect of stereo on performance for particular tasks, depending on game expertise.

Finally, several fully immersive CAVE-based applications for visualization [3] and path planning [15, 33] reported significant performance gains compared with their desktop counterparts. While these studies indicate that for certain spatial tasks immersion improves performance, findings are limited to specialized settings. More focus is needed on understanding whether semi-immersive 3DS desktop displays provide an advantage for spatial problem solving.

Haptic Feedback

Mine [28] et al. discuss the importance of one's physical presence in a virtual environment, and explore ways of incorporating this into the interaction design. They conclude that without *haptic feedback*, it is difficult to give users a sense of any physical objects they manipulate.

Implementing systems that provide haptic feedback in 3D virtual environments introduces technical challenges, but has been proven in several studies to make a difference in performance and task time [39, 34]. Our study focuses on vibrotactile feedback provided by a stylus. To date, little work has been done to explore vibrotactile feedback from a 6-DOF stylus in 3DS semi-immersive environments. We explore such feedback when combined with stereo vision.

Evaluation of Immersive Cues with Brain Data

While our study aims to better understand the effect of immersive cues on user experience, the key expected differences are often not measureable using traditional metrics designed for work-related contexts and standard graphical user interfaces. Novel interaction techniques pose a challenge, as subtle qualities or internal processes are not easily measured with traditional research methods. Because of this, research methods have been explored that analyze brain activity to provide additional insight on the user's internal state. For example, Frey et al. [11] used electroencephalography (EEG), the measure of brain electric activity, to evaluate user comfort when viewing 3D objects at different depths.

In addition to EEG, functional near-infrared spectroscopy (fNIRS) brain imaging has been explored in computer interfaces, both as real-time user input and in user evaluations. It is a non-invasive head-worn device used to measure mental workload by detecting changes in blood oxygen levels in the prefrontal cortex. The device sends near infrared light through the forehead, which is partially absorbed by oxy- and deoxy- hemoglobin in the brain. The rest of the light is diffusely reflected back to the detector and is inversely related to the amount of oxygen in the blood. The changes in oxygen concentration in the brain reflect the hemodynamic response due to brain activity, and has been shown to be related to changes in mental workload.

The sensors are held in place with a foam headband, making the device easy to put on and allowing the user to move. Previous work has designed and tested protocols and analysis methods for the use of fNIRS in the evaluation of interfaces [18, 37, 38]. These usually employ standardized *benchmark cognition* tasks that the user performs to provide training data in well-understood tasks for later classifying brain data during actual user interface tasks. However, it can be difficult to identify appropriate benchmark tasks, as well as impractical to require users to perform unrelated tasks to build the model. However, these methods

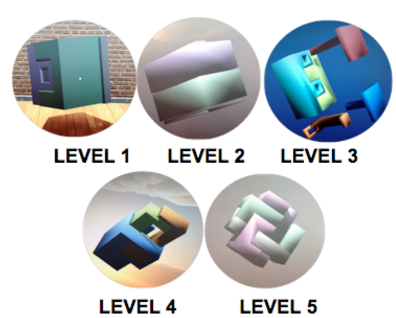

Figure 3. The 5 levels of zPuzzle

	C1	C2	C3	C4
Display size	23.6 in			
Display angle	30 degrees			
Resolution	1920x1080			
Mouse	X	X	X	X
6-DOF stylus	X	X	X	X
Keyboard	X	X	X	X
Stereoscopy		X		X
Head tracking	X	X	X	X
Vibrotactile Feedback			X	X

Table 1. Comparison of experimental conditions

show promise as a complementary measure in user studies.

EXPERIMENT

The goal of our experiment was to investigate how immersion cues, specifically stereo vision and vibrotactile feedback from a stylus, affect user interaction during a spatial task with semi-immersive interactive 3DS displays. We chose to focus on these immersion cues since they are both prominent in spatial reasoning tasks and could be easily controlled in semi-immersive interactive 3DS displays. Our experimental task combines the following spatial tasks: determining spatial relationships among objects, manipulating objects in 3D, and visualizing a path. Our task also increases complexity across 5 levels (Figure 3). We used traditional evaluation methods combined with emerging brain-imaging tools that may be able to identify more subtle differences in the user experience. We consider zSpace as a means of augmenting our understanding and ability to reason about 3D space and the fNIRS as a device that is capable of providing extra information while complex thought processes take place.

Experimental Task: zPuzzle

Participants were asked to solve 3D spatial reasoning puzzles. Each puzzle presented a 3D structure consisting of interlocked blocks. To solve the puzzle participants needed to sort out the various blocks and free them into individual pieces until the structure was dismantled. This task was inspired by the popular game Interlocked [20]. We used five puzzles in increasing levels of complexity (Figure 3). The number of pieces and the spatial relations among them determined the complexity of the puzzles. Complexity was validated with a pilot study of 26 users completing 5 levels of the puzzle under different conditions [31].

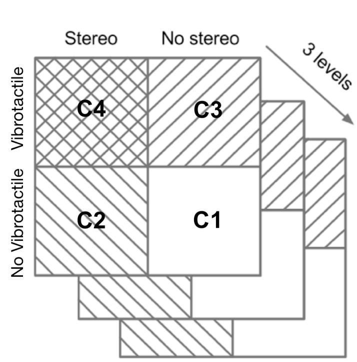

Figure 4. 2x2 matrix of conditions across 3 levels for 2x2x3 mixed design

We selected this task because it tests various spatial abilities including: understanding spatial relationships among objects, manipulating objects in 3D, and visualizing a path. Its game-like nature is simple enough for participants to engage with it immediately. However, at increased levels of complexity, these puzzles provide models of real-world problems such as: the identification of small molecule protein binding sites and the modeling of mechanical systems.

We developed an application called zPuzzle for viewing and manipulating the puzzles. A puzzle is presented at the center of the screen. Upon solving the puzzle, the system moves to the next puzzle. zPuzzle is implemented in C# using Unity Game Engine and the zSpace SDK.

Experimental Design

We use a 2x2x3 mixed design (Figure 4) and study differences between and within subjects across tasks.

Independent Variables

There were two between-subject independent variables: *stereoscopy* and *vibrotactile feedback,* each with two levels, and one within-subject independent variable: *difficulty,* with three levels. Thus, each of the participants was randomly assigned to one of the following configurations: 1) no stereoscopy with no vibrotactile feedback; 2) stereoscopy with no vibrotactile feedback; 3) no stereoscopy with vibrotactile feedback; and 4) stereoscopy with vibrotactile feedback; and completed several puzzles of varying difficulty. All conditions support bimanual interaction, which combines a mouse for scene rotation and a 6-DOF stylus for direct manipulation of 3D objects. Table 1 summarizes the experimental conditions.

Considering findings from related work, which indicate that combined head coupling and stereo provide a significant enhancement in 3D tasks and that head coupling is more important than stereo in 3D viewing [3, 33, 35, 41, 42], we did not study binocular and motion parallax separately.

Difficulty is the within-subjects independent variable, and

Dimension	Quantitative measures
Performance	Completion time per level
Spatial Presence	MEC-SPQ post task questionnaire [40]
Perceived Workload	NASA Task Load post task questionnaire [15]
Measured Workload	fNIRS oxy-hemoglobin and deoxy-hemoglobin readings

Table 2. Measures for evaluating spatial problem solving.

there were three levels (L3, L4, L5). Our focus in this study was on investigating the performance and workload of trained users, and not in measuring learning. Thus, each participant did eight practice trials to familiarize themselves with the problem. These practice trials consisted of two simple levels (L1 and L2) to get them acquainted with the puzzles. Then, each participant was presented with two puzzles of L3, L4, and L5 in the following order: L3, L4, L3-r1, L5, L4-r1, L5-r1, where r denotes repetition. Every time a particular puzzle level was repeated, we rotated the puzzle and used a different set of colors for the blocks. This allowed users to learn how to use the interface as well as to understand the demands of task and to develop problem-solving strategies in the levels that would be tested. Finally, they were presented with the three experimental puzzles (L3, L4, L5), counterbalanced across participants.

Dependent Variables

Traditional Measures. The traditional dependent variables we use are completion time, subjective workload rating (NASA Task Load Index [16]), and subjective presence rating (MEC-Spatial Presence Questionnaire [40]). We collected quantitative data from the user from post-task questionnaires and from logging noted points during the study (e.g. level completion). Table 2 summarizes the measures we used in this study.

Brain Data. In addition, we collected fNIRS brain data throughout the experiment (Figure 5). For pre-processing of the data, we used HOMER2, an interface built on top of Matlab scripts made specifically for processing fNIRS data [19]. We first pruned any channels with a signal to noise ratio less than 2. The raw light intensity signal was converted to optical density units. We used a low-pass filter of 0.10 Hz to remove any high frequency noise in the signal. Using the modified Beer-Lambert law, and partial pathlength factors of 6.5 and 5.9, the optical density change units were converted to concentration values for oxyhemoglobin (blood carrying oxygen to the brain) and deoxyhemoglobin (blood in which oxygen has been consumed). These measures reflect the hemodynamic response in the brain and are the basis of both fNIRS and functional magnetic resonance imaging (fMRI) techniques. We expect increases in oxyhemoglobin (HbO) and decreases in deoxyhemoglobin (HbR) during periods of activation [19].

After preprocessing the brain data, we calculated features of interest in the data. While previous fNIRS experiments measure a user performing a task for a set amount of time (e.g. 30-second task period), our experiment allowed participants to complete tasks at their own pace. Thus, completion times for each participant and each task varies. To account for this variation in time, we computed the following key features in the brain data over the entire task period, regardless of duration: i) average HbO and HbR values over the task period for each channel; and ii) maximum HbO and minimum HbR value during the task period for each channel (since we are interested in positive peaks in HbO and negative peaks in HbR). We wanted to look at overall activation, and thus combined the multichannel data into one value by averaging across the 10 channels. Our analysis focused on the average across channels of the following features: maximum HbO, minimum HbR, average HbO and average HbR during the varying-length task.

Hardware

The physical hardware setup is the same across the four conditions (Table 1): the zSpace display is at a 30 degree angle, a keyboard is attached to the display; two mice are attached to the computer at either side of the keyboard; a stylus is positioned on the right; the fNIRS headband is attached to the user's head (Figures 1 and 2). The zSpace supports both stereo and non-stereo displays. Stereo vision uses binocular parallax and motion parallax to track the user's head and display a different angle of the object as they move. For stereo vision, the user wears glasses. The system updates the puzzle perspective based on the user's gaze given by the position of the glasses relative to the screen. Users move puzzle pieces by pointing the stylus at a piece and clicking the stylus button to "grab" the object.

For the *haptic* feedback condition, we enabled vibrotactile feedback from the stylus, which occurs when two 3D objects collide in the virtual environment. The stylus provides 6-DOF for directly interacting with 3D objects. The user can rotate the scene using the mouse. In the *non-stereo* conditions, the Stereo setting on the zSpace machine's Nvidia graphics card is switched off, but users still wear the glasses for consistency. All users manipulate puzzle pieces using the stylus and rotate the scene using a mouse.

The fNIRS device was an ISS Imagent with 8 light sources and 2 light detectors (Figure 2) arranged to provide 10 source-detector pairs with 3cm distances between them.

Experiment Procedure

The experiment began with the participant seated in front of the display and the moderator seated to their side. Participants were briefed about the task and given a standard consent form. After putting on the fNIRS headband, we recorded a baseline level of brain activity and then users completed zPuzzle's tutorial level. The users worked through the eight practice puzzles and the three experiment puzzles with a 20 second rest between each level to allow the their brain activity to return to baseline. After performing all of the tasks, the participants completed the subjective workload and presence questionnaires.

Participants

48 undergraduate students between 18-23 years old (M=20, SD=1.6) were recruited from our institution. We only used female participants because previous studies indicate that males and females approach spatial reasoning problems differently and we did not want to account for gender dif-

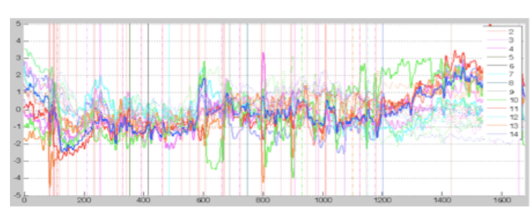

Figure 5. Unprocessed oxy- and deoxy-hemoglobin from one user over 10 channels for the length of the study.

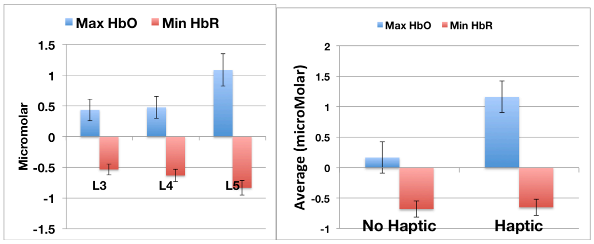

Figure 8. Means and std. error of max HbO and min HbR by difficulty (top) and haptic condition (bottom).

ferences. Participants were all right-handed. Each participant was randomly assigned to a condition (12 participants per condition). Across conditions we found no significant difference in terms of experience with 3D displays and games. All participants completed the experiment and were compensated with a 10-dollar gift card.

RESULTS

Completion Time

We explored the effect of difficulty level and immersive cues on completion time, using total time per level as a repeated measure (Figure 6). Because of skewed distribution of the residuals, we use the natural log of completion time in our analysis. We ran a mixed ANOVA with *difficulty* as a 3-level within subjects variable (levels 3,4,5 of the puzzle) and two 2-level between subject variables (*vibrotactile* and *stereo*). We found a significant effect of *difficulty* (p<0.001). Pairwise comparisons showed a significant difference between *level 3* and *level* 5 as well as between *level 4* and *level 5.* In addition, we found a significant difference effect of *stereo* (p=.033), with users in the non-stereo condition taking longer than users in the stereo condition. No other interactions between difficulty level and any of the immersive conditions were found.

Perceived Task Workload

We measured users' perceived task workload with the NASA TLX questionnaire [16]. We interpret the results of the unweighted, raw NASA TLX data, grouped by category (i.e. frustration, effort, mental demand, physical demand). The score of each category is the sum of all questions related to that theme. We used Shapiro Wilks to test the normality of the residuals. All the categories satisfied the normality assumption except physical demand, which we transformed using the natural log to normalize the data. Factorial ANOVA found that the only category that is significant is the natural log of physical demand (F(1,44)= 8.910, p=.005). Users who had no stereo display (M=1.445, SD=.517) experienced higher physical demand than users with stereo display (M=.927, SD=.661). We also found a marginally significant difference in frustration between stereo and non stereo (F(1,44)= 2.197, p=.095). Users with the non-stereo display (M=4.957, SD=2.364) experienced higher frustration than those with stereo display (M=3.840, SD=2.192). No other significant differences were found.

This is consistent with the findings from completion time where users in the non-stereo condition taking longer to complete the puzzles than users in the stereo condition.

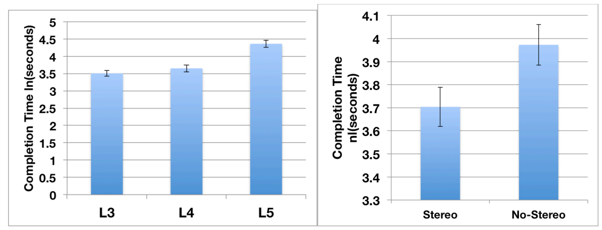

Figure 6. Average time (natural log) and standard error. We found significant differences between L3 and L5 and L4 and L5, as well as between stereo and non-stereo.

Spatial Presence

We measured users' perceived spatial presence with the MEC-SPQ standardized questionnaire [40], which consists of eight scales. Figure 7 shows the MEC-SPQ results.

We used Shapiro Wilks to test the normality of the residuals. All the categories satisfied the normality assumption. Factorial ANOVA found significant differences in two scales: Spatial Situational Model (SSM) and Spatial Presence Possible Actions (SPPA) (SSM: F(1,44)=4.981, p=.031; SPPA: F(1,44)=7.496, p=.009). The Spatial Situational Model scale is a measure of users perceived understanding of the virtual environment and the relative size and positions of all the objects in it. The Spatial Presence Possible Actions scale gauges the participant's impression of being able to act in the virtual environment.

In both scales, users who experienced vibrotactile feedback (SSM: M=4.031, SD=.652; SPPA: M=4.014, SD=.764) had higher spatial presence scores than users who did not (SSM: M=3.615, SD=.599; SPPA: M=3.417, SD=.724), meaning they felt more immersed in the virtual environment. As with the NASA TLX scores, we also looked at the interaction between stereo and vibrotactile, but found no other significant results.

These results show that vibrotactile feedback helps users to think they understand the virtual objects and to believe that they have control to manipulate the objects.

Brain Data

With the measures for average and maximum oxy-hemoglobin and average and minimum deoxy-hemoglobin (Figure 8), we conducted a mixed ANOVA analysis with *difficulty* as a 3-level within subjects variable (levels 3,4,5 of the puzzle) and two 2-level between subject variables (*vibrotactile feedback* and *stereo*). *Difficulty* had a marginal effect on the *Average HbO* (p=0.059) and *average HbR* (p=0.065) and a significant effect on *Max HbO* (p=.002) and on *Min HbR* (p=.001). The maximum oxy-hemoglobin increased as difficulty increased and the minimum deoxy-hemoglobin decreased as difficulty increased. This result verifies our use of the fNIRS data in our setup as a measure of workload for users solving zPuzzle.

Vibrotactile feedback had a significant effect on *average HbO* (p=.009) and *Max HbO* (p=.009). Users who experienced vibrotactile feedback had higher values than those without. This significance shows that in the case of vibrotactile feedback, the users with the more immersive

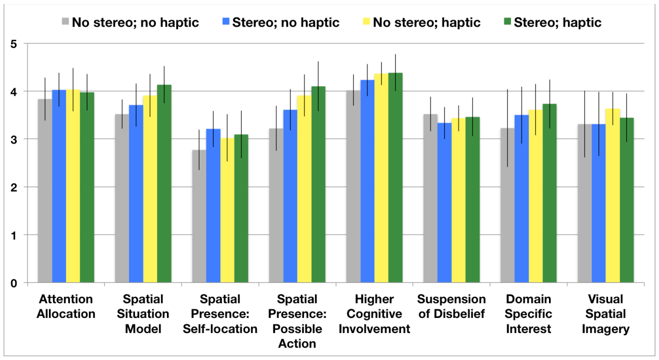

Figure 7. Average MEC-SPQ scores. Error bars show 95% confidence interval.

condition were working harder than those without it. There were no significant effects of deoxy-hemoglobin values. In addition, stereo did not show any significant effects, and there were no interaction effects for vibrotactile x stereo. There were no other significant interactions.

Physical Discomfort

18 of the 48 participants experienced some sort of headache or discomfort: 3 of these users felt dizzy from the puzzles and display. From the people with headaches, 7 felt the fNIRs device being tight or heavy and 5 felt their eyes strained from focusing on the zSpace display. No one was unable to complete the study due to discomfort. We have redesigned the fNIRs headband to be more comfortable in future studies by using more lightweight materials.

DISCUSSION

Both stereo display and vibrotactile feedback have significant effects on user experience solving spatial problems with zSpace. Users without the stereo display take longer to complete the task and experience higher physical demand and slightly higher frustration. This could be explained by the lack of a portrayed 3D image in the non-stereo condition, requiring users to spend time creating a mental image of the 3D object, while users in the stereo condition have system help with that reasoning. This could be a likely source of physical demand and extra time. Brain data did not show significant differences between stereo and non-stereo conditions. This may indicate that the stereo condition led to more task-related brain activation, improving performance. The non-stereo condition had similar activation that was split between task-related and interface-related workload, leading to increased completion time and frustration.

Ideally, higher immersion would result in lower mental workload, reducing the amount of spatial reasoning the user has to do by giving them an environment that looks and feels more like the physical world. However, the brain data indicates that higher immersion (as indicated by the MEC-SPQ) did not directly translate to a reduction in the users' mental workload. Vibrotactile feedback from the stylus, which increased perceived Spatial Presence, also increased brain activation, without a performance gain (measured by completion time). This may indicate that some of the users' workload is dedicated to understanding the vibrotactile interaction, which increases immersion, but requires them to work harder to feel this way. We also must consider that the nature of the particular vibrotactile feedback used in this experiment may have been a source of frustration, since it was provided in a centralized and minimal way.

CONCLUSIONS AND FUTURE WORK

Our main contributions are in two areas. First, we provide an increased understanding of the effects of immersive cues, specifically vibrotactile feedback from a stylus and stereo, on people solving 3D spatial reasoning tasks. We showed that stereo display with head coupling and motion parallax results in lower completion time, lower physical effort, and slightly lower frustration, but similar brain activation. We also found that vibrotactile feedback from the stylus results in increased perceived presence but also in higher brain activation. We conclude that increased immersive cues do not necessarily reduce users' mental activity when solving spatial reasoning tasks with 3D stereoscopic displays. Second, we show the promise of using fNIRS to collect complementary information when evaluating novel interaction paradigms in an HCI setting. We showed that the fNIRS data shows differences between puzzle levels of of different difficulties. We also showed a scenario where the fNIRS data was useful in helping HCI researchers augment understanding of differences not shown by traditional measures. By combining fNIRS with other measures, we can observe differences in workload that are task-related versus those that are user interface related. This is key for UI evaluation where we aim to increase task-related focus, by decreasing workload required by the user interface [18].

This work builds a foundation for continued study of various immersion cues and their ability to augment spatial reasoning abilities as well the use of fNIRS to supplement evaluations of interactive systems. Future work can explore different types of problem solving tasks with the same immersion cues as well as similar tasks with new immersive systems and variations on the immersive cues we used. The use of fNIRS allows for more extensive exploration of the interaction between immersion and mental workload. Future work can also deeply explore the changing brain signal over time to identify points of interest that may not have been detected with the aggregate measures used here. Finally, by demonstrating the feasibility of using fNIRS with 3DS displays to gain user state information, we can build interactive systems that use the brain data in real time to adapt the display and feedback based on cognitive activity.

ACKNOWLEDGMENTS

We'd like to thank Sean Bluestein, Heather Petrow, and Yi Tong for helping with this work. This research is partially funded by a Brachman Hoffman grant.

REFERENCES

1. Afergan, D., et al. Dynamic difficulty using brain metrics of workload. *Proc. CHI'14*. ACM. (2014). 3797–3806.

2. Antle, A. N., & Wang, S. Comparing motor-cognitive strategies for spatial problem solving with tangible and multi-touch interfaces. In *Proc. of TEI '13*, ACM. (2013).
3. Arns, L., Cook, D., and Cruz-Neira, C. (1999). The Benefits of Statistical Visualization in an Immersive Environment." In proc. *IEEE VR*. 1999.
4. Arthur, K. W., Booth, K. S., & Ware, C. (1993). Evaluating 3d task performance for fish tank virtual worlds. *TOIS*,11(3), 239-265. ACM.
5. Barfield, W., Hendrix, C., & Bystrom, K. (1997, March). Visualizing the structure of virtual objects using head tracked stereoscopic displays. In Proc. of VRST, IEEE
6. Blender. http://www.blender.org/
7. Brooks, L. R. (1968). Spatial and verbal components of the act of recall. Canadian Journal of Psychology/Revue canadienne de psychologie, 22(5), 349.
8. Bowman, D. and McMahan, R. Virtual Reality: How Much Immersion Is Enough? *IEEE Computer*, vol. 40, no. 7, 2007, pp. 36-43.
9. Danaher, K., & Crandall, C. S. (2008). Stereotype threat in applied settings re-examined. *J Applied Social Psychology*, 38(6), 1639-1655.
10. Durlach, N. I., Mavor, A. S. (Eds.). 1994. *Virtual reality: scientific and technological challenges*. National Academies Press.
11. Frey, J., Pommereau, L., Lotte, F., Hachet, M. Assessing the Zone of Comfort in Stereoscopic Displays using EEG. *CHI '14 Extended Abstract*s, 2014.
12. Fujimoto, M., & Ishibashi, Y. (2004). The effect of stereoscopic viewing of a virtual space on a networked game using haptic media. In *Proc. of SIGCHI ACE*. ACM.
13. Gardner, H. (1983/1993) *Frames of mind: The theory of MI*. New York: Basic Books.
14. Gibson, J. J. (1986). *The ecological approach to visual perception*. Routledge.
15. Gruchalla, K. (2004). Immersive well-path editing: investigating the added value of immersion. In Proc. VR. IEEE.
16. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX: Results of empirical and theoretical research. *Human mental workload,* 1(3), 139-183.
17. Hirshfield, L., et al., This is your brain on interfaces: enhancing usability testing with functional near-infrared spectroscopy. *Proc. CHI'11*. 373-382.
18. Hirshfield, L., Solovey, E., Girouard, A. et al. 2009. Brain measurement for usability testing and adaptive interfaces: an example of uncovering syntactic workload with functional near infrared spectroscopy. In *Proc. CHI'09*.
19. Huppert, T.J., Diamond, S.G., Franceschini, M.A. and Boas, D.A. (2009). "HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain." *Appl Opt* 48(10): D280-98.
20. Interlocked. http://armorgames.com/play/10591/interlocked
21. Jacob, R., Girouard, A., Hirshfield, L., Horn, M., Shaer, O., Treacy, E., Zigelbaum, J. Reality-Based Interaction: A Framework for Post-WIMP Interfaces, Proc. ACM CHI 2008 Human Factors in Computing Systems Conference, ACM Press, 2008.
22. Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive science*, 18(4).
23. Kulshreshth, A., Schild, J., & LaViola Jr, J. J. (2012). Evaluating user performance in 3D stereo and motion enabled video games. In *Proc. of FDG* (pp. 33-40). ACM.
24. LaViola, J. J., Forsberg, A. S., Huffman, J., & Bragdon, A. (2008). The influence of head tracking and stereo on user performance with non-isomorphic 3D rotation. In *Proc. of Eurographics Conf Virtual Environments*.
25. LaViola Jr, J. J., & Litwiller, T. (2011). Evaluating the benefits of 3d stereo in modern video games. In *Proc. of CHI'11* (pp. 2345-2354). ACM.
26. Lee, Y.-C., et al. (2014). An EEG-based approach for evaluating audio notifications under ambient sounds. *Proc CHI '14*.
27. McKenna, M. (1992). Interactive viewpoint control and three dimensional operations. Symposium on i3D Graphics
28. Mine, M., Brooks, F., Sequin, C. (1997) Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction. *SIGGRAPH 1997.* ACM.
29. Parsons, T. D. et al. (2004). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia*, 42(4), 555-562.
30. Patterson, R. E. (n.d.). Interactive Stereo Displays, White paper. Retrieved 2012, from zSpace: http://zspace.com/wpcotent/uploads/2012/08/INTERACTIVE_STEREOSCOPIC_DISPLAYS_07_30_12.pdf
31. Petrow, H. Haptic Interaction in 3D Stereoscopic User Interfaces. Wellesley College Undergraduate Thesis. 2014
32. Robertson, G., Czerwinski, M., and van Dantzich, M. 1997. Immersion in desktop virtual reality. In *Proc. UIST*.
33. Schuchardt, P., & Bowman, D. A. (2007). The benefits of immersion for spatial understanding of complex underground cave systems. In *Proc. of VRST*. ACM.
34. Sodhi, R., Poupyrev, I., Glisson, M., and Israr, A. 2013. AIREAL: interactive tactile experiences in free air. *ACM Trans. Graph*. 32, 4, Article 134.
35. Sollenberger, R. L., & Milgram, P. (1991). A comparative study of rotational and stereoscopic computer graphic depth cues. *In Proc. of HFES* (Vol. 35, No. 20,). SAGE.
36. Solidworks. http://www.solidworks.com/sw/3d-cad-design-software.html
37. Solovey, E.T., et al. Brainput: enhancing interactive systems with streaming fNIRS brain input. *Proc. CHI'12*.
38. Solovey, E.T., et al. Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines. *Proc. UIST 2009*.
39. Ullah, S., Ouramdane, N., Samir Otmane, S., et al. 2008. Augmenting 3D interactions with haptic guide in a large scale virtual environment. In *Proc. VRCAI* 2008.
40. Vorderer, Peter, et al. "MEC spatial presence questionnaire (MEC-SPQ): Short documentation and instructions for application." Report to the European Community, Project Presence: MEC (IST-2001-37661) (2004).
41. Voyer, D., Voyer, S., Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological bulletin*, 117(2), 250.
42. Ware, C., & Franck, G. (1996). Evaluating stereo and motion cues for visualizing information nets in three dimensions. *TOG*, 15(2), 121-140. ACM.
43. Weibel, D., & Wissmath, B. (2011). Immersion in computer games: The role of spatial presence and flow. *Intl J Computer Games Technology,* 6.