# The Cognitive Implications of Semi-Natural Virtual Locomotion


William E. Marsh[†*] Marisa Putnam[‡] Jonathan W. Kelly[§] Veronica J. Dark[§] James H. Oliver[†]

[†]Virtual Reality Applications Center, Iowa State University
[‡]Mathematica Policy Research
[§]Department of Psychology, Iowa State University



## Abstract

This study incorporated a dual-task paradigm, in which participants were asked to perform basic locomotion tasks with one of three interfaces while remembering a sequence of either spatial or verbal items. Interfaces varied in similarity to natural body movements. Stopping performance was compromised when concurrently remembering a spatial, but not verbal, sequence. Also users exhibited lower performance on spatial memory tasks while using more unnatural locomotion interfaces. These results confirm that semi-natural locomotion interfaces require spatial working memory resources and thus locomotion interfaces compete with ongoing spatial tasks, as opposed to those requiring verbal resources or general attention resources.



**Index Terms:** H5.2 [Information interfaces and presentation]: User Interfaces—Input devices and strategies;


## 1 Introduction

Designing interaction techniques for use in virtual reality (VR) requires compromises as developers attempt to provide a natural interface. This interaction is confined by physical constraints inherent in the hardware and software implementations of the interface. Natural interfaces are generally defined as using techniques similar to those found in real-world movement, whereby the same body segments that actuate the interface produce similar cause–effect relations as in the real world [13]. A natural interface is more transparent to the user, enhancing presence and potentially increasing the effectiveness of the virtual environment (VE) [16].

### 1.1 Navigation and Locomotion

In this paper, "virtual locomotion" refers to the atomic movements that a user makes when navigating through a VE (stepping, turning). There are two basic objectives when designing interfaces for virtual locomotion: allow navigation between any two virtual points and maximize naturalness. Virtual navigation requires the use of atomic locomotion actions while building, maintaining, and using a spatial mental model of the virtual world. It is not possible to navigate an infinite VE within a finite VR system using completely natural locomotion techniques. The required unnatural locomotion mechanisms have the potential to negatively impact a user's experience and success. Navigation, and thus locomotion, is seldom the purpose of a VE. These activities serve a support role as users simultaneously perform other, primary, tasks [3][5]. Before a user can perform the desired primary tasks, basic navigation knowledge must be learned. Before navigating, the user must learn to locomotion using the provided interface.

The cognitive components of working memory must be considered when choosing a locomotion interface as these resources are thought to be required when humans learn a novel skill. Unnatural locomotion mechanisms may require additional working memory to construct and maintain a mental model that maps possible actions to expected outcomes. These cognitive resources used for locomotion cannot be directed to completing a simultaneous primary task. Conversely, working memory that is in use by primary tasks is unavailable for locomotion [6]. An understanding of the bi-directional impact of different types of cognitive tasks during a simultaneous locomotion task can inform the design of locomotion interfaces.

A natural locomotion technique maximizes the match between proprioceptive information corresponding to actions and sensory feedback generated by the VR system. A good match will allow the user to develop a predictive model of interaction within the environment [10]. An untrained user already possesses natural perceptual-motor abilities and knowledge of interaction in the real world, and it is beneficial for virtual interfaces to replicate these interactions. There are compelling reasons that natural locomotion techniques are usually preferred. Real walking provides the translational and rotational information needed to accurately update position automatically with perceptual processes [7] and the choice of movement technique impacts cognition [17] and presence [14].

There are many handheld locomotion interfaces currently in use, involving hardware such as a wand or gamepad. These devices are not usually considered natural because they use completely different muscle groups than real walking, yet they can sometimes be ideal, particularly when coupled with a flying metaphor.

Incorporating wide-area trackers to allow for real walking is best in terms of naturalness [14]. However, the confines of the physical environment usually limit the use of natural locomotion techniques while navigating through a large virtual world. Overcoming these constraints often involves scaling and automation [8], which could disrupt locomotion cues such as visual-vestibular coupling.

Body-based locomotion techniques can often allow for semi-natural movement because they incorporate the same muscle-groups used in analogous real-world movements. Many hardware based solutions have been developed, such as treadmills, unicycles [5], and large hamster balls [9]. These solutions still suffer from vestibular and proprioceptive mismatches.

It is often desirable to mix real-world locomotion movements with less natural virtual techniques [12]. In this vein, hybrid rate/position-control systems have been created, such as a barrier-tape metaphor [4] in which the region at the center of the physical environment allows precise, completely natural movements. When the user steps past a certain threshold distance, depicted by a virtual "barrier tape," the interface becomes rate-controlled and the user's virtual velocity increases as a function of the user's distance from the center of the physical environment. This scheme allows for rapid movement over large distances as well as natural fine-position control. A similar (without barrier tape) position-to-velocity interface (P2V), depicted from a top-down view in Figure 1, is in use in the Virtual Reality Applications Center at Iowa State University for locomotion in the C6 CAVE. Such a technique, which maps position to velocity, is particularly well-suited for a six-sided CAVE, like the C6, as the walls are very restrictive but the 360° display allows for natural virtual rotation.

### 1.2 Cognitive Resources

Human cognitive resources are finite and must be shared between ongoing tasks, each of which has different processing and resource demands [6]. Models of working memory generally distinguish verbal and non-verbal systems. Baddeley and Hitch devised the


*e-mail: william@wemarsh.com



IEEE Virtual Reality 2012
4-8 March, Orange County, CA, USA
978-1-4673-1246-2/12/$31.00 ©2012 IEEE

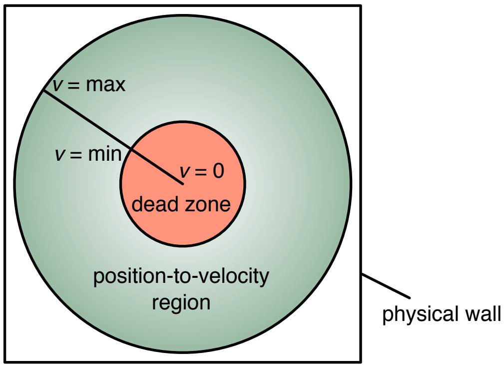


Figure 1: Top-down diagram of the P2V locomotion interface.

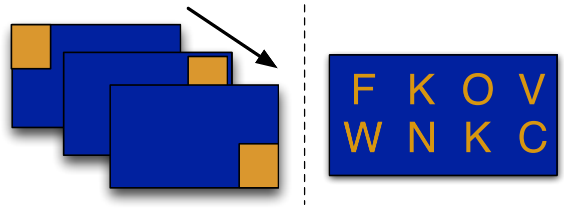


Figure 2: Spatial task of span 3 (left) and spatial recall card (right).

most widely accepted multi-component working memory model [2]. The original model included two systems: the visuo-spatial sketchpad, used for maintaining visual and spatial information, and the phonological store, used for verbal information. In the model, access to both of these is dependent on attention, a limited resource mediated by a third component, the central executive.

Tasks requiring spatial resources that have been used in the past often involve remembering locations or movements of cues, such as a ball [15], through space. A task known to tax verbal resources is remembering a random sequence of verbal items, such as digits. A dual-task selective-interference paradigm is used to isolate which resource is utilized for a particular task. The idea is to load a participant's working memory with a task that is known to tax a specific sub-system while the user simultaneously completes another task. If performance at the other task decreases while the user performs a spatial span task, but not a verbal task, for example, then one concludes that the second task requires visuo-spatial resources.

After extensive training with a consistent stimulus–response mapping, the skill becomes proceduralized into long-term memory and no longer requires working memory [1]. Body-based interfaces are usually based on actions that have already been proceduralized (e.g., real walking), and thus should require minimal resources. This is one benefit to choosing a body-based technique over a gamepad. However, no system that allows for infinite virtual locomotion in a constrained physical space is completely natural. Some aspects will be more natural than others.

## 2 Experiment Overview and Hypothesis

An experiment was conducted to investigate the impact of concurrent cognitive tasks (spatial, verbal) on performance using two interfaces: a variant of the P2V interface described above and a gamepad interface (GP). As a baseline, a "real-world" group (RW) moved about as in the physical world. The working memory tasks required participants to remember a sequence of either spatial or verbal items and then recall those items at a later time.

If unnatural locomotion requires additional cognitive resources, performance should decrease at either locomotion actions or working memory tasks, when both are presented concurrently. If spatial resources are used more so than verbal or general attention resources, then results should show an additional performance decrease associated with a concurrent spatial task.

An interface may be useful for different types of tasks under different concurrent task conditions. Performance on the less natural aspects of an interface should be more affected by the addition of a concurrent cognitive task. For example, the P2V interface allows for completely natural rotation within the dead zone so there should be no detriment when a concurrent cognitive task is added. However, stopping with that same interface is unnatural, requiring a user to locate and return to the center of the CAVE. This should lead to increased stop time and possibly lower memory performance. Previous work has shown that greater resources are required when using unnatural locomotion interfaces [11], but that work was not designed to distinguish the relative contributions of spatial and verbal working memory. The contribution of this study is the focus on competition between specific interface aspects and specific working memory resources.

## 3 Methods

Fifty-one undergraduate students were recruited from the Department of Psychology participant pool and word of mouth, and randomly assigned to one of three interface groups. Participants were required to have 20/20 (corrected) binocular vision and all played less than or equal to 3.5 hours of video games per week. Users in GP group were further restricted to no more than than 1.5 hours of first-person video games per week. The gamepad is intended to be an unnatural locomotion interface so it was important to ensure that it was unnatural.

Participants were randomly assigned to one of the following groups, listed in order from least to greatest expected naturalness:

**GP.** Participants used a Logitech Wingman gamepad. The left stick translated while the right stick rotated in the VE. Full-body rotation was not allowed, so participants stood in the center of the C6, facing the front wall for the entire experiment.

**P2V.** Participants used the P2V interface. Since movement tasks were axis-aligned, velocity vector components were computed separately based on the distance from each axis. Calculating velocity in this way could simplify movements by reducing the chance of drifting off course on axis-aligned tasks. The dead-zone radius was 1.4 feet.

**RW.** Participants moved around the C6 using real walking.

The flow of the spatial memory task presentation phase and recall are depicted Figure 2. For recall, a random matrix of letters was displayed and the participant stated the letters that corresponded to the order of the boxes. Recall letters were randomized to prevent users from encoding the locations verbally. The verbal memory task presentation was a random sequence of numbers. For verbal tasks, the word "Recite" was displayed, alerting the participant to recite the sequence. For the control task the word "Wait" was displayed for both presentation and recall, indicating to stand still and wait for the next task.

Participants first completed a pre-questionnaire with questions on demographic information, video game experience, and athletic activities. Then they entered the C6, a six-sided $10' \times 10'$ CAVE, and were given instructions and a demonstration of how to complete memory tasks. They were then given a series of six verbal memory tasks of increasing difficulty to assess individual verbal spans. The difficulty was increased from three to five items, with two trials at each span. Next, participants were trained on spatial tasks and given six practice tasks, again increasing from three to five items. The lowest span (three) was for practice, to ensure that participants knew what they should do. Performance on the five-item tasks was used to customize difficulty in the experimental blocks. If a participant was incorrect on either five-item task, the span for that particular task type (verbal or spatial) was dropped to four in the experimental blocks. This was done to ensure that the span during the locomotion tasks was sufficient to tax the resource in question but not impossible. These pre-tests also provided practice so that participants would perform at a high level on the memory tasks in the experimental blocks.

The experimental blocks took place in a virtual room with a yellow-on-black grid texture. One virtual wall was purple and participants were instructed to always face that wall with their bodies and to stand still in the center of the C6 between tasks. Translation tasks involved moving to a virtual "golden nugget" with a radius of 1 foot, centered 4.25 feet above the floor and 5 feet from the participant. Locomotion practice was not allowed, to prevent learning from occurring, but before beginning the experimental blocks, the experimenter gave a demonstration of all movement tasks. The experiment comprised six blocks of locomotion tasks with concurrent working memory tasks. In each block, a participant was presented with a memory sequence, followed by a sequence of movement tasks, and finally asked to recite the memory sequence. Each block had a verbal, a spatial, or no memory task, assigned randomly such that each participant received two of each condition. The movement phase lasted at least 70 seconds to reduce incentive for participants to rush through the movements to reach the recall quicker. Each sequence of locomotion tasks was also randomly ordered. Between locomotion tasks, there was a six-second pause, in which participants stood in the center of the C6 awaiting the next task. The following locomotion tasks were each performed once during each sequence: step left, step right, step forward, rotate left, rotate right, and duck. This paper describes only the translation tasks and results.

At the beginning of each translation task, the nugget appeared in front or to the appropriate side of the participant. The participant was tasked with moving to that location. If the nugget was to either side, an arrow appeared on the floor, pointing to it. Since participants were required to always face the purple wall, sidestepping was necessary to reach nuggets to the sides.

Participant head positions were tracked and recorded. A moving average of positions was used to determine if the participant was stopped. Several performance metrics were calculated for each task: time elapsed from presentation until completion; time elapsed from presentation until movement started; time elapsed from completion until movement stopped; distance traveled; and memory items missed.

After exiting the C6, participants completed a post-questionnaire and answered questions in an unstructured interview. Questions were intended to uncover strategies or any particular problems encountered, specifically involving competition for cognitive resources.

## 4 Results

Analysis focused on the effects of the locomotion interface and memory task combinations on movement performance and on working memory performance. Because in many cases users had problems that led to the movements not being atomic, left, right, and forward data were combined for analysis. For example, if a participant sidestepped left but passed the nugget, a right sidestep (or some other mix of movements) was required in order to complete the task. Thus it is not appropriate to treat that movement as a left-sidestep action as initially intended, rather it is more reasonable to treat it as a translation movement. Stop time and memory items missed were most influenced by the interface type and/or memory task, so the following analysis focuses on those measures. A drop in performance on either memory tasks or a movement task is considered evidence of competition for resources, meaning that both tasks had similar (competing) resource demands.

Some data did not exist and some recorded data were removed for experimental consistency reasons. In some cases, according to the moving average of head positions, participants were already moving before a task was presented, and start time was not logged. Similarly, some users didn't completely come to a stop before the next task was presented. Application crashes or other errors led to incomplete data for some participants. Head position data were manually inspected and some data were discarded when it was clear that the user had passed the nugget and stopped, waiting for the next task, before realizing the mistake. Finally, in either the post-questionnaire or exit interview, some users reported using a verbal strategy for the spatial tasks (i.e. coding the locations as numbers). Since the spatial task was intended to tax spatial resources, the data from the affected experimental block was discarded if a participant reported using such a strategy.

### 4.1 Stop Time

Mean stop time is plotted in Figure 3. A mixed model analysis was performed with fixed effects for interface group and working memory task combinations (9 means) and random effects for participants. Significant main effects of interface [$F(2,45)=298.74$, $p<.001$] and memory task [$F(2,641)=4.22$, $p=.02$] were identified as well as an interaction between interface and memory task [$F(4,641)=3.36$, $p=.01$]. A significant difference between interface groups was expected, since stopping with the gamepad (let go of stick) or real-world locomotion is trivial, while stopping with the P2V interface requires locating and returning to CAVE-center. This expectation is supported by the analysis. Also, since stopping times were so low in the GP and RW groups, one should not expect to see a difference between memory tasks in those groups. This is also supported by the analysis. A Markov chain Monte Carlo (MCMC) simulation from the posterior distribution for the model is used to obtain estimates and p-values for comparisons of interest. The most interesting stopping results are in the P2V group. Users of this interface stopped significantly faster when performing a spatial memory task than when performing no task ($p=.04$), and significantly faster when given a verbal memory task as opposed to a spatial memory task ($p=.02$).

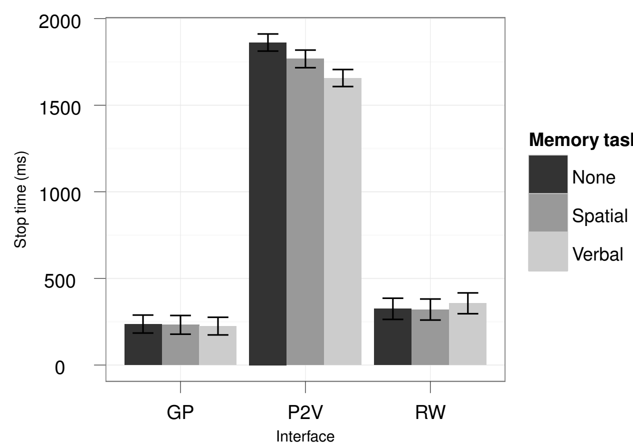


Figure 3: Stop time. Error bars show $\pm 1$ std. error of the mean.

One explanation for stopping performance being slowest when there was no concurrent memory task is that users were motivated to move faster in order to end the resource competition between the locomotion task and the cognitive task. This conclusion is supported by participant feedback, which indicated a subjective sense that the tasks competed for resources. It is also supported by a visual inspection of completion times for sidestepping tasks, which also show slower performance in the P2V group when there was no task than when there was a verbal or spatial task.

The intriguing result is the difference in stop times when performing spatial and verbal tasks. There are at least two possible explanations. First, users could have been more motivated when given a verbal task than a spatial task. Second, the users could have been equally motivated during both types of memory task but they may have been incapable of stopping as fast during the spatial task, due to competition for spatial resources. The second possibility is supported by a visual inspection of the start times, which show the same general trend for the P2V group but did not reach statistical significance. The start time data provide additional information in the GP group where no-task performance is much slower and there is no meaningful difference between the performances during a concurrent spatial or verbal task. Because stopping is trivial when using a gamepad, the stop time data do not show any pattern. These patterns support the notion that users are equally motivated when given spatial and verbal tasks. Feedback from post-questionnaires and exit interviews also supports the idea that spatial memory tasks interfered with locomotion to a greater degree than verbal memory tasks.

### 4.2 Memory items missed

When a user simultaneously performs two tasks requiring common resources, a performance detriment can be expected at either task, or both. The number of memory items missed is plotted in Figure 4. For this analysis, memory task performance is combined into two means (spatial and verbal), so performance on the two spatial tasks was combined into one number, for example. A two-way ANOVA revealed a significant effect of memory task [$F(1,40)=20.61$, $p<.001$]. A possible interpretation of the significant effect of memory task is that the spatial tasks are harder than the verbal tasks so participants missed more items. However, during the pre-test users were remembering 100% of items on both types of memory tasks when administered with no concurrent task. Based on expected results, overall patterns in the data, and participant feedback, the difference is likely due to the concurrent locomotion movements. Because of these expectations and patterns, analysis proceeds on the spatial results in isolation. Recall that the gamepad was expected to be the least natural interface, the P2V interface to be slightly more natural, and real world locomotion to be a completely natural baseline. Analysis proceeds by testing the memory performance patterns across interfaces using contrast weights (1, 0, -1) determined by the hypothesis. The predicted contrast significantly describes the data [$F(1,41)=4.34$, $p=.04$].

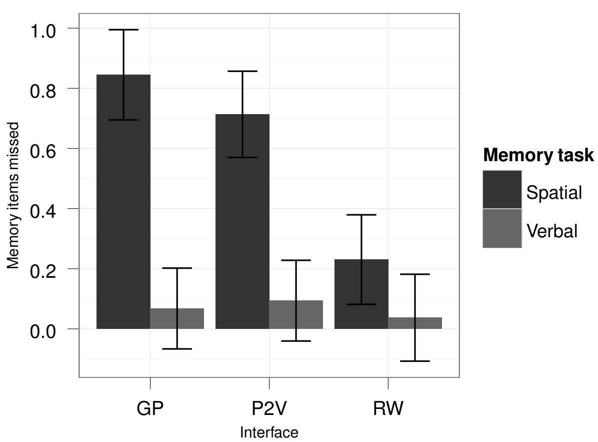


Figure 4: Memory items missed. Error bars show $\pm 1$ std. error of the mean.

## 5 Conclusions and Future Work

The findings indicate that unnatural locomotion interfaces and spatial memory tasks compete for the same cognitive resources. A simultaneous spatial memory task increases stop time with the P2V interface, a task that requires finding and returning to the center of the physical environment. Conversely, users exhibited decreased ability to remember a sequence of spatial items while using unnatural interfaces.

Many of the tasks performed in VR place extreme cognitive demands on the user and they can be critical to success in the underlying scenario. The findings from the study described above, together with application-specific knowledge, can be used to inform the design of future VR systems, particularly the choice of locomotion interfaces.

## Acknowledgements

This research was funded by AFRL grant FA955007-1-0475. We thank John Tuohy, Jon Hobbs, Rosza Brown, and Jamal Brown.

## References

[1] J. R. Anderson. Acquisition of cognitive skill. *Psychological Review*, 89(4):369–406, 1982.

[2] A. D. Baddeley. Is working memory working? the fifteenth bartlett lecture. *The Quarterly Journal of Experimental Psychology*, 44A(1):1–31, 1992.

[3] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, 1st edition, 2005.

[4] G. Cirio, M. Marchal, T. Regia-Corte, and A. Lécuyer. The magic barrier tape: a novel metaphor for infinite navigation in virtual worlds with a restricted walking workspace. In *Proceedings of ACM Virtual Reality Software and Technology (VRST) 2009*, pages 155–162, Kyoto, Japan, November 18-20 2009. ACM.

[5] R. P. Darken, W. R. Cockayne, and D. Carmein. The omni-directional treadmill: A locomotion device for virtual worlds. In *Proceedings of UIST '97*, pages 213–221, Banff, Canada, October 14-17 1997.

[6] D. Gopher and E. Donchin. Workload - an examination of the concept. In K. R. Boff, L. Kaufman, and J. P. Thomas, editors, *Handbook of Perception and Human Performance*, volume 2, Cognitive Processes and Performance, chapter 41, pages 1–49. John Wiley & Sons, Inc., 1986.

[7] R. L. Klatzky, J. M. Loomis, A. C. Beall, S. S. Chance, and R. G. Golledge. Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, 9(4):293–298, July 1998.

[8] A. Kulik. Building on realism and magic for designing 3d interaction techniques. *IEEE Computer Graphics and Applications*, pages 22–33, November/December 2009.

[9] E. Medina, R. Fruland, and S. Weghorst. Virtusphere: Walking in a human size vr "hamster ball". In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*, pages 2102–2106, 2008.

[10] M. Slater, M. Usoh, and A. Steed. Steps and ladders in vitrual reality. In *Proceedings of the Conference on Virtual Reality Software and Technology*, pages 45–54, Singapore, 1994. World Scientific Publishing Co., Inc.

[11] E. A. Suma, S. L. Finkelstein, S. Clark, P. Goolkasian, and L. F. Hodges. Effects of travel technique and gender on a divided attention task in a virtual environment. In *IEEE Symposium on 3D User Interfaces*, pages 27–34, 2010.

[12] J. N. Templeman, P. S. Denbrook, and L. E. Sibert. Virtual locomotion: Walking in place through virtual environments. *Presence*, 8(6):598–617, December 1999.

[13] J. N. Templeman and L. E. Sibert. Immersive simulation of coordinated motion in virtual environments: An application to training small unit military tactics, techniques, and procedures. In G. L. Allen, editor, *Applied Spatial Cognition - From Research to Cognitive Technology*, chapter 13, pages 339–372. Psychology Press, September 2006.

[14] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks, Jr. Walking > walking-in-place > flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 359–364. ACM Press/Addison-Wesley Publishing Co., 1999.

[15] E. Vergauwe, P. Barrouillet, and V. Camos. Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4):1012–1028, 2009.

[16] M. Wells, B. Peterson, and J. Aten. The virtual motion controller: A sufficient-motion walking simulator. In *Proceedings of VRAIS '97*, pages 1–8, 1996.

[17] C. Zanbaka, B. Lok, S. Babu, A. Ulinski, and L. F. Hodges. Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment. *IEEE Transactions on Visualization and Computer Graphics*, 11(6):694–705, 2005.