The cognitive implications of virtual locomotion with a restricted field of view

William E. Marsh[a], Jonathan W. Kelly[b], Veronica J. Dark[b], James H. Oliver[a]

[a]Virtual Reality Applications Center, Iowa State University, Ames, Iowa, United States;
[b]Department of Psychology, Iowa State University, Ames, Iowa, United States

ABSTRACT

A study was conducted to examine the impact, in terms of cognitive demands, of a restricted field of view (FOV) on semi-natural locomotion in virtual reality (VR). Participants were divided into two groups: high-FOV and low-FOV. They were asked to perform basic movements using a locomotion interface while simultaneously performing one of two memory tasks (spatial or verbal) or no memory task. The memory tasks were intended to simulate the competing demands when a user has primary tasks to perform while using an unnatural interface to move through the virtual world. Results show that participants remembered fewer spatial or verbal items when performing locomotion movements with a low FOV than with a high FOV. This equivalent verbal and spatial detriment may indicate that locomotion movements with a restricted FOV require additional general cognitive resources as opposed to spatial or verbal resource pools. This also emphasizes the importance of this research, as users of a system may allow primary task performance to suffer when performing locomotion. Movement start and completion times were also measured to examine resource requirements of specific aspects of movements. Understanding specific performance problems resulting from concurrent tasks can inform the design of systems.

Keywords: Virtual reality, locomotion, cognitive psychology, working memory, field of view

1. INTRODUCTION

Virtual reality (VR) is used in a wide range of domains. Many of its applications, such as command and control of unmanned aerial vehicles or firefighting simulation, involve scenarios that place high demands on a user's finite cognitive resources. Primary tasks in these scenarios often require a user to navigate to distant areas of the virtual environment (VE), completing tasks along the way. Due to the constrained movement area afforded by VR systems, users must typically manipulate a locomotion interface of some sort in order to affect movement through the VE. In the "real world," humans use sensory feedback, particularly visual, to guide locomotion activities. In a VR system, sensory feedback is often limited, possibly causing users to resort to more cognitively demanding strategies.

1.0.1 Navigation and locomotion

In this paper, navigation refers to large-scale movements through the VE (i.e., walking to an unseen location across town). The term locomotion is used to describe the small, atomic movements that navigation is composed of (i.e., sidestepping to the left). Due to system constraints, navigation through an infinite VE requires an unnatural locomotion interface. This places additional cognitive demands on the user.[1]

A locomotion interface (P2V), depicted from the top down in Figure 1, similar to the Virtual Motion Controller[2] or the Magic Barrier Tape[3] has been created for use in the C6 CAVE at the Virtual Reality Applications Center at Iowa State University. In the center of the C6, there is a virtual "dead zone," within which all movement is completely natural. Once the user steps outside the dead zone, a velocity is set based on the vector from the center of the C6 to the user's physical position. The farther from the center that the user moves, the higher

---

Further author information: (Send correspondence to William E. Marsh.)
William E. Marsh: E-mail: william@wemarsh.com, Telephone: 1 937 505 0836
Jonathan W. Kelly: E-mail: jonkelly@iastate.edu, Telephone: 1 515 294 2322
Veronica J. Dark: E-mail: vjdark@iastate.edu, Telephone: 1 515 294 1688
James H. Oliver: E-mail: oliver@iastate.edu, Telephone: 1 515 294 2649

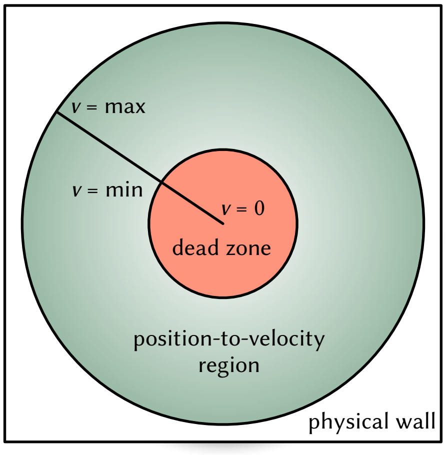

Figure 1. Top-down depiction of the P2V locomotion interface.

the virtual velocity will be. To stop, the user must return to the dead zone. This type of interface is particularly appropriate for use in a six-sided CAVE, such as the C6, because all rotation can be completely natural, yet there are tight bounds on the physical movement area.

1.1 Field of view

Locomotion through the physical world is guided by vision,[4] augmented by proprioceptive and vestibular feedback. Because of the scale of movements and the muscle groups used, proprioceptive and vestibular feedback are often unreliable in VR. Vision provides the highest acuity and, thus, is trusted more than the others in case of conflict.[5] However, visual feedback is also limited by VR systems. For example, the NVIS nVisor SX HMD has an FOV of $47° \times 38°$.[6] Modern CAVE-like systems offer nearly vision-limited resolution with no FOV restriction. However, the use of CrystalEyes stereo shutter glasses does limit the FOV to about $140° \times 90°$.[7]

Users of VR systems have problems interpreting distances and other spatial information, and this can hinder navigation and locomotion activities.[8] Some studies show limited FOV and HMD properties can distort distances,[6,9] but it is generally unclear which aspects of VR influence these problems.[10]

Studies indicate that peripheral vision is important for locomotion,[11] and FOV-related navigation deficits have been shown.[12,13] It is likely that a reduced FOV leads to alternate cognitive strategies, which may be costly in terms of resource demands. The locations of visible objects do not need to be represented in memory because they are perceptually available in the world. For this reason, limiting FOV may place an additional burden on memory.

Optic flow is the pattern of optical stimulation generated by self-motion and it provides information about displacement within the environment. For example, forward and backward movement cause expansion and contraction flows, respectively, and the rate of flow corresponds to velocity. The rate has been shown to influence the transition from running to walking and preferred walking speed.[5] Steering involves moving so the center of the flow pattern is in the desired direction of travel.[4] Humans can perform locomotion activities with no flow information,[14,15] but alternate cognitive strategies may be employed, changing resource demands.

1.2 Cognitive resources

A human's activities are bound by a finite set of cognitive resources which must be shared by simultaneous tasks. Many models have been created in attempts to better understand how these resources are structured. Models vary in specifics but they generally draw a distinction between verbal and non-verbal systems. Many of the existing models are based on Baddeley and Hitch's multi-component model of working memory.[16] In the original model, there are two subsystems: the phonological store and the visuo-spatial sketchpad. The phonological store handles storage and manipulation of verbal items, while the visuo-spatial sketchpad is responsible for storing and manipulating items that are visual or spatial in nature (non-verbal). According to the model, access to both of these systems is dependent on attention, a limited resource mediated by a third component called the central executive.

Some tasks have been devised that are known to rely on these resources individually. Spatial tasks involve remembering a sequence (span) of locations or movements of cues, such as a ball, through the VE.[17] According to the Baddeley and Hitch model, such a task would require resources from the visuo-spatial sketchpad (non-verbal pool) and/or central executive (general attention resources), which may be utilized to encode the input and reconstruct it when recalled.[16] Spatial resources are likely to be used for all but the most natural aspects of virtual locomotion.

A common task known to tax verbal resources is remembering a random sequence of digits or what is known as an n-back task, where a participant must state if the current digit (or other stimulus) is the same as that presented some distance (n) back in the sequence. In a similar vein as above, these types of tasks should rely primarily on the verbal and general attention resources.[16]

A dual-task selective-interference paradigm has been used to identify the resources required by a given task. In the paradigm, a participant is asked to perform a cognitive task with known resource demands while concurrently performing a task of interest. If performance on either task is compromised, then it can be concluded that the task of interest utilizes the same pool of resources as the cognitive task.

2. EXPERIMENT OVERVIEW AND HYPOTHESIS

An experiment was conducted in order to investigate the impact of a restricted FOV on cognitive resource demands while using a virtual locomotion interface. It was expected that the restriction would cause participants to resort to strategies that required working memory resources. Based on the literature describe above, it was possible that spatial resources would be required if strategies required storing and manipulating spatial information in the head that may normally be available in the world. Alternatively, strategies requiring verbal resources (e.g., counting to keep track of distance traveled) or more general attention resources may be involved when FOV is reduced. If spatial resources are required, performance at movements and/or memory tasks should decrease when given a concurrent spatial task and a reduced FOV. The same effect is true for verbal resources, and if general attention resources are required then an equal decrease should be seen in locomotion or memory task performance with either a verbal or spatial concurrent task.

The study incorporated a 2×3 design, with FOV (high, low) as a between-subjects variable and memory-task type (spatial, verbal, none) as a within subjects variable.

2.1 Participants

Thirty-one undergraduate students (20 males) were recruited from the Iowa State University Department of Psychology participant pool (SONA), word of mouth, and an announcement in an undergraduate course. All participants were required to have (corrected) 20/20 visual acuity. They were assigned randomly to two FOV groups. The high-FOV group wore CrystalEyes stereo shutter glasses, providing an FOV of $140° \times 90°$. The low-FOV group wore the same type of glasses, but with cardboard attached to the lenses, reducing FOV to approximately $60° \times 45°$.

2.2 Stimuli and Design

All locomotion tasks took place in a virtual room with a yellow-on-black grid texture on the floor and all walls except one, rendered in Figure 2. The texture was intended to be simple but provide good visual feedback for locomotion. The remaining wall was textured with a yellow-on-purple grid pattern.

Participants performed three basic types of locomotion tasks: translation, rotation, and ducking. Two of each (left, right, forward) translation task was assigned during each block. One of each (left, right) rotation task was assigned. One ducking task was assigned. All of the above tasks were in random order within each block. A memory task (spatial, verbal, none) was presented at the beginning of each block and recalled at the end of each block, such that there were two blocks of each type of memory task, randomly ordered.

Users were required to face the purple wall with their bodies at all times, so when the environment rotated by 90°, a physical rotation was necessary. The translation tasks involved left, right, or forward movement, and

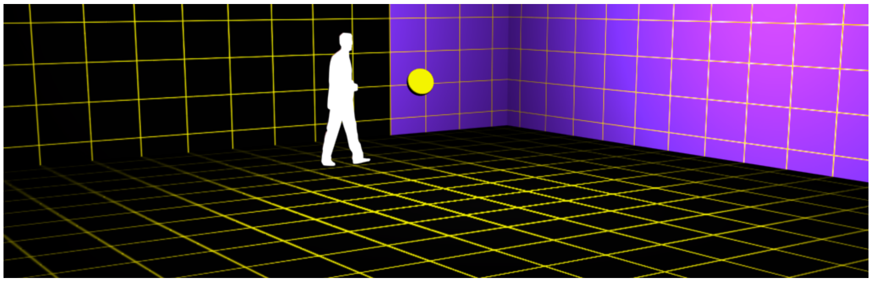

Figure 2. Rendered grid room.

required the participant to retrieve to a virtual "golden nugget" in the VE, shown in Figure 2. The nugget was centered 7.0 feet from the participant, 4.25 feet above the floor, and had a radius of 1 foot. For nuggets presented to the left or right, users were required to sidestep in order to keep their bodies facing the purple wall. Finally, once per block a virtual I-beam came flying overhead, requiring ducking to avoid being hit.

The spatial and verbal memory tasks were presented as a sequence of cards in the VE, as shown in Figure 3. Verbal tasks were a span of numbers and spatial tasks were a span of box positions. During the recall phase, another card was displayed for each type of task, as shown in Figure 4 When it was time to recall a verbal span, a card appeared with the word "Recite" on it. Participants were then supposed to recite the most recent verbal span, in order. When it was time to recall a spatial span, a card appeared with a random array of letters. The participant was then supposed to recite the letters that corresponded to the positions of the boxes in the original span, in order.

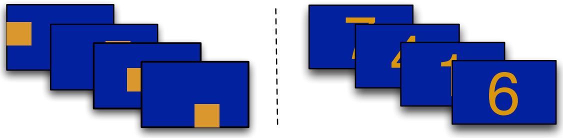

Figure 3. Sample spatial (left) and verbal (right) memory task presentation sequences.

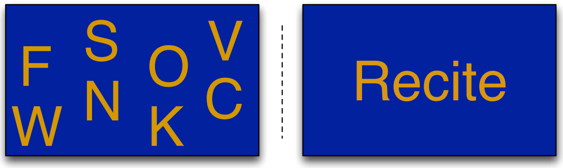

Figure 4. Sample spatial (left) and verbal (right) memory task recall cards.

2.3 Procedures

First, participants were asked to complete a questionnaire including demographic information and topics such as video game usage. Then they completed the Perspective-Taking and Spatial-Orientation Test (PTSOT)[18] to investigate the role of individual differences in spatial ability.

Next participants entered the C6 and were given training on how to complete verbal memory tasks. Then the participant was given a series of verbal practice tasks. The practice tasks increased in difficulty from three to six items, with two at each difficulty level. A participant's performance on the practice tasks was used to customize the task difficulty during the experimental blocks as follows. If a participant was 100% correct on all tasks with span of four, five, or six, then the difficulty was set to six. If the participant was 100% correct on the tasks with a span of four or five, but missed one or more sequences with a span of six, then the difficulty was set to five. If the participant missed any items with a span of five or six, then the difficulty was set to four.

Participants were then trained on how to complete the spatial memory tasks. This was followed by a series of spatial practice tasks, administered exactly as described above for the verbal practice tasks.

Next the experimenter gave a demo of how to complete the locomotion tasks in the VE. Extensive practice with an interface would result in the actions becoming proceduralized, therefore requiring fewer cognitive resources. Since this study is interested in the nature of those working memory resources, participants were not allowed to practice the locomotion tasks.

Each of the six experimental blocks started with a memory task presentation, such that each participant experienced two of each memory task condition (in random order): spatial, verbal, and none. Then all locomotion tasks were performed in random order. Finally, the user was asked to recall the memory sequence.

2.4 Response variables

Several measures were recorded for each translation task. The application used a moving average of participant head positions to determine when the user started and stopped movement. The start time was defined to be the time from task presentation until movement started. Other variables, such as total movement time, stopping time, rotation time, and ducking performance were also recorded but are not reported here.

3. RESULTS

Results showed that, when given competition for cognitive resources, participants primarily sacrificed performance on the memory tasks as opposed to the locomotion tasks.

As described above, numerous measurements were collected for each participant. If a participant was already moving when a translation task was presented, no start time was recorded. For consistency, experimenters did not attempt to record times manually. In some cases, due to hardware problems, software problems, or participant confusion, data points were discarded if the problem was likely to impact a user's ability to correctly complete the task. Reasons were:

- head tracker interference;
- graphical anomalies;
- some participants got close enough to the virtual walls that a nugget was displayed on the other side; and
- one participant reported disobeying directions and playing around.

Additionally, some participants reported using a verbal strategy to remember spatial sequences (i.e., assigning numbers to spatial positions). Since the task was intended to tax spatial resources, affected data were removed whenever a participant reported using such a strategy.

3.1 Analysis of working memory performance

Recall that, in the case of a conflict for resources, performance may decrease at either the locomotion tasks or the concurrent memory task. Responses on the memory tasks were scored according to the minimum number of replacements or swaps needed to make the participant's answer correct. The number of items missed was treated as a Poisson variable and a two-factor mixed-model analysis showed significant effects of FOV group [$F(1, 27) = 4.27$, $p = .049$] and memory task [$F(1, 69) = 25.26$, $p < .001$]. The means are plotted in Figure 5. It appears that restricting the FOV led to an equivalent decrease in both verbal and spatial memory task performance. In terms of the multi-component models of working memory, this pattern of results indicates that additional general attention resources are required when the FOV is reduced. Alternatively, an equal amount of both verbal and spatial resources may be required, which would also explain the observed results.

3.2 Analysis of movement start times

Times for each phase of movement were also analyzed. Start time for sidestepping tasks did not return significant results but a plot of the means, shown in Figure 6, seems to provide additional evidence in support of the conclusion above. The pattern indicates that verbal and spatial times increased approximately equivalently when the FOV was reduced. The start time reflects the time required for identifying the task to be performed, motor planning, and initiation of movement so future research should investigate differences in planning strategies when FOV is restricted.

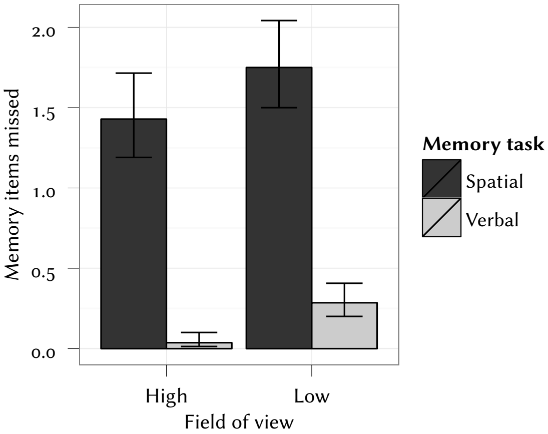

Figure 5. Memory items missed. Error bars show ±1 standard error of the mean.

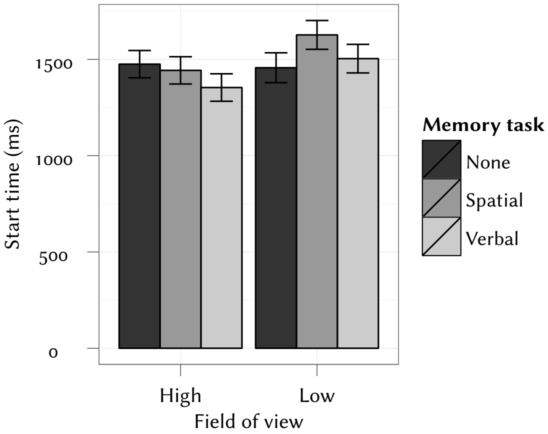

Figure 6. Mean start time for left and right translation tasks. Error bars show ±1 standard error of the mean.

3.3 Perspective-taking and spatial-orientation test

Answers on the PTSOT test were scored by averaging each participant's deviation from the correct answer. As described by Kozhevnikov et al.,[19] participants with scores in the bottom quartile (8.58° – 14.5°) were placed in the "high" ability category (7 males, 2 females) and those with scores in the upper quartile (37.82° – 109.11°) were placed in the "low" ability category (5 males, 3 females). All participants in the middle two quartiles were eliminated from the analysis that follows.

A three-factor mixed-model analysis was conducted on the left and right translation start time, adding PTSOT ability as an additional variable to the existing model from the sidestepping start time analysis above. In this model, memory task [$F(2, 263) = 5.72$, $p < .004$] and the interaction between memory task and PTSOT ability [$F(2, 263) = 4.44$, $p = .01$] are both significant. Figure 7 shows a plot of the means.

First, observe that participants with low perspective-taking ability started sidestepping tasks slower when concurrently remembering a spatial sequence than when remembering a verbal sequence or no sequence at all. This makes sense, as an individual with a lower spatial ability should be expected to perform worse on some types of spatial tasks and it is reasonable that planning and initiating bodily movements might require spatial resources.

Second, notice that participants with a high perspective-taking ability do not exhibit the same detriment from a concurrent spatial task that was present in the low-ability category. The big difference here is the slower performance when given no concurrent memory task. This may mean that users with high spatial abilities use more time-consuming strategies when planning and initiating locomotion movements, if resources are not already being used for another task.

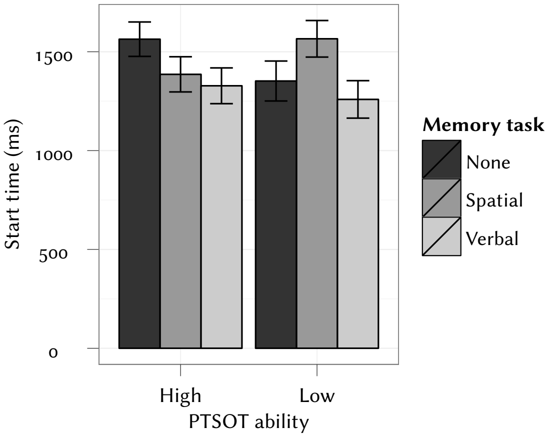

Figure 7. Mean start time for left and right translation tasks with PTSOT ability. Error bars show ±1 standard error of the mean.

4. CONCLUSION

The results presented above provide strong evidence that a reduction in FOV causes users to resort to locomotion strategies that require additional general attention resources. In the Baddeley and Hitch multi-component model of working memory, these would involve the central executive. Alternately, these results do not eliminate the possibility that an equal amount of verbal and spatial resources are required for locomotion with a reduced FOV.

The PTSOT results indicate that these findings are subject to individual differences and users with different ability levels are differentially impacted by the addition of a concurrent cognitive task. Specifically, users with a high ability level may actually take longer to plan and initiate sidestepping movements when there is no competition for resources and users with a low ability level are slower to start these movements when a concurrent spatial task is present.

These findings can be used in the design and of VR systems. When selecting a display technology, it is important to consider the types of locomotion activities and concurrent primary tasks that a user will be likely to engage in as well as the spatial abilities of a typical user. The fact that users sacrificed on the spatial memory tasks as opposed to the movement tasks reinforces the importance of this research direction, as these tasks were intended to simulate the existence of real-world primary tasks. However, users may make sacrifices on the locomotion tasks instead if the cognitive task was of real-world importance. Future work should investigate this possibility.

Future work should specifically investigate a possible effect of concurrent task load on translation start times. Such research should also attempt to determine what role individual differences play when such competition is present. Future studies can incorporate additional measures of individual abilities as well as translation tasks designed to reveal more information on planning strategies in use.

ACKNOWLEDGMENTS

This research was funded by Air Force Research Laboratory grant FA955007-1-0475. We would also like to thank Jon Hobbs for his help with R and John Tuohy for his help when conducting the study and parsing results.

REFERENCES

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

[2] Wells, M., Peterson, B., and Aten, J., "The virtual motion controller: A sufficient-motion walking simulator," in [*Proceedings of VRAIS '97*], 1–8 (1996).

[3] Cirio, G., Marchal, M., Regia-Corte, T., and Lécuyer, A., "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*], 155–162, ACM, Kyoto, Japan (November 18-20 2009).

[4] Gibson, J. J., "Visually controlled locomotion and visual orientation in animals," *British Journal of Psychology* **49**, 182–194 (August 1958).

[5] Mohler, B. J., Thompson, W. B., Creem-Regehr, S. H., Pick, Jr., H., and Warren, Jr., W. H., "Visual flow influences gait transition speed and preferred walking speed," *Experimental Brain Research* **181**, 221–228 (2007).

[6] Willemsen, P., Colton, M. B., Creem-Regehr, S. H., and Thompson, W. B., "The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments," *ACM Transactions on Applied Perception* **6**, 8:1–8:14 (February 2009).

[7] http://www.inition.co.uk/inition/products.php.

[8] Ruddle, R. A. and Jones, D. M., "Movement in cluttered virtual environments," *Presence* **10**, 511–524 (October 2001).

[9] Kline, P. B. and Witmer, B. G., "Distance perception in virtual environments: Effects of field of view and surface texture at near distances," in [*Proceedings of Human Factors and Ergonomics Society 40th Annual Meeting*], **2**, 1112–1116, Human Factors and Ergonomics Society, Philadelphia, PA, USA (1996).

[10] Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., and Beall, A. C., "Does the quality of the computer graphics matter when judging distances in visually immersive environments?," *Presence* **13**(5), 560–571 (2004).

[11] Crowell, J. A. and Banks, M. S., "Perceiving heading with different retinal regions and types of optic flow," *Perception & Psychophysics* **53**(3), 325–337 (1993).

[12] McCreary, F. A. and Williges, R. C., "Effects of age and field-of-view on spatial learning in an immersive virtual environment," in [*Proceedings of Human Factors and Ergonomics Society 42nd Annual Meeting*], **2**, 1491–1495, Human Factors and Ergonomics Society, Chicago, IL, USA (October 5-9 1998).

[13] Alfano, P. L. and Michel, G. F., "Restricting the field of view: Perceptual and performance effects," *Perceptual and Motor Skills* **70**, 35–45 (1990).

[14] Macuga, K. L., Loomis, J. M., Beall, A. C., and Kelly, J. W., "Perception of heading without retinal optic flow," *Perception & Psychophysics* **68**(5), 872–878 (2006).

[15] Warren, Jr., W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., and Sahuc, S., "Optic flow is used to control human walking," *Nature Neuroscience* **4**, 213–216 (February 2001).

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

[17] Vergauwe, E., Barrouillet, P., and Camos, V., "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).

[18] Hegarty, M. and Waller, D., "A dissociation between mental rotation and perspective-taking spatial abilities," *Intelligence* **32**, 175–191 (2004).

[19] Kozhevnikov, M., Motes, M. A., and Hegarty, M., "Spatial visualization in physics problem solving," *Cognitive Science* **31**, 549–579 (2007).