

An investigation of the validity of the virtual spatial navigation assessment

Matthew Ventura, Valerie Shute, Tim Wright, and Weinan Zhao

Florida State University

Corresponding Author:

Matthew Ventura
College of Education
3205A Stone Building
1114 West Call Street
Tallahassee, FL, 32306-4453

KEYWORDS:

Environmental spatial ability
Vista spatial ability
Figural spatial ability

Abstract

This correlational study investigated a new measure of environmental spatial ability (i.e., large scale spatial ability) called the Virtual Spatial Navigation Assessment (VSNA). In the VSNA, participants must find a set of gems in a virtual 3D environment using a first person avatar on a computer. The VSNA runs in a web browser and automatically collects the time taken to find each gem. The time taken to collect gems in the VSNA was significantly correlated to three other spatial ability measures, math standardized test scores, and choice to be in a STEM (science, technology, engineering, or math) career. These findings support the validity of the VSNA as a measure of environmental spatial ability. Finally, self-report video game experience was also significantly correlated to the VSNA suggesting that video game may improve environmental spatial ability. Recommendations are made for how the VSNA can be used to help guide individuals toward STEM career paths and identify weaknesses that might be addressed with large scale spatial navigation training.

Introduction

Spatial ability has been shown to play a significant role in achievement in science, technology, engineering, and mathematics (STEM) disciplines. For instance, Wai, Lubinski, and Benbow (2009) showed that spatial ability was a significant predictor of STEM degree attainment, even after controlling for mathematical and verbal skills. Thus a thorough understanding of spatial ability and how it can be improved should be considered paramount in understanding how to engage students in STEM related fields.

One way spatial ability can be improved is through playing action video games (e.g., De Lisi, & Wolford, 2002; Dorval, & Pepin, 1986; Feng, Spence, & Pratt, 2007; Green & Bavelier, 2006; Subrahmanyam & Greenfield, 1994; Spence, Yu, Feng, & Marshman, 2009; Uttal et al., 2012). For example, Feng et al. (2007) found that playing an action video game improved performance on a mental rotation task. After only 10 hours of training with an action video game, subjects showed gains in spatial ability via mental rotation tasks, with females performing equal to males after training. Control subjects who played a non-action game showed no improvement. Recently, Uttal et al. (2012) conducted a meta-analysis of 206 studies investigating the effect of training on spatial ability. Of these 206 studies, 24 used video games to improve spatial ability. The effect size for video game training was .54 (SE = .12). Findings like these have been explained due to the visual-spatial requirements of 3D action games which may enhance spatial abilities (e.g., Feng et al., 2007). However, others have found a lack of transfer effects between action video game playing and basic cognitive functions and skills (e.g., Boot, Kramer, Simons, Fabiani, & Gratton, 2008) and have raised questions regarding the methodology of studies that observe transfer (Boot, Simons, Stothart, & Stutts, 2013; Kristjánsson, 2013).

Of particular importance in understanding the malleability of spatial ability is the distinction between figural, vista, and environmental related spatial abilities (Montello, 1993; Montello & Golledge, 1999). Figural spatial ability is small in scale relative to the body and external to the individual. It can be apprehended from a single viewpoint in both flat pictorial and 3D space (e.g., small, manipulatable objects). It is most commonly associated with tests such as mental rotation and paper folding tasks. Vista spatial ability is the ability to imagine oneself in different locations within a small space without locomotion. Vista spatial ability is useful when trying to image how the arrangement of objects will look from various perspectives (Hegarty & Waller, 2004). Environmental spatial ability is large in scale relative to the body and is useful in navigating around large spaces such as buildings, neighborhoods, and cities, and typically requires locomotion (see Montello, 1993, for a discussion of other scales of space). Environmental spatial ability may require a person to mentally construct a cognitive map, or internal representation of the environment (Montello & Golledge, 1999).

Specific processes in environmental spatial ability may result from the accumulation of three main types of knowledge of the environment: landmark, route, and configurational knowledge (Siegel & White, 1975; Tolman, 1948). First, landmark knowledge is acquired of perceptual objects through visual encoding. These perceived landmarks are then assimilated and are connected sequentially along a traversed path into route knowledge. Configurational knowledge is formed through the amassing of route knowledge, as a map-like representation of the environment is formed that allows for navigational inferences (Siegel & White, 1975). For example, new routes and distance and direction judgments can be formed as a result of a navigator's configurational knowledge. The details of these environmental representations depend on a

number of factors including the perception of environmental information, the speed in which the information is encoded, and how the information is maintained (Ittelson, 1973).

Existing measures of environmental spatial ability include map retracing, distance estimation, direction estimation (Hegarty et al., 2006), and self-report measures (Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002). Map retracing and distance and direction estimation (Hegarty et al., 2006) require a participant first to navigate through an environment (real world, 3D virtual environment, or first-person video). Afterwards, a person can be asked to (a) judge the distance among various features in the environment, (b) provide direction estimates among features in the environment, or (c) draw a 2D map of the environment. While these measures may seem distinct, factor analysis has revealed that these three measures were highly correlated and loaded on one factor, suggesting they measure a common ability (Hegarty et al., 2006). Additionally, factor analysis has shown measures based on virtual and video environments load on one factor (video factor) while measures based on real environments load on a second factor (real environment factor). The correlation between the two factors was high ($r = .61$) suggesting that the cognitive processes being used in virtual simulations are similar to the ones being used in real environments (Hegarty et al., 2006).

The present study

Hegarty et al. (2006) proposed three main sources of variance in environmental spatial ability: (1) ability to encode spatial information from sensory experience, (2) ability to maintain a high quality internal representation of that information in memory, and (3) ability to perform spatial transformations in order to make inferences from this spatial information. In line with this theory, we developed the Virtual Spatial Navigation Assessment. Advances in game devel-

opment software now enable researchers with little programming experience to create virtual environments. These virtual environments are increasing being used to assess large scale spatial ability. For example, Herting and Nagel (2012) created a virtual water maze task to assess visuospatial memory. The Virtual Spatial Navigation Assessment (VSNA) requires a participant to explore a virtual 3D environment using a first person avatar on a computer. One significant difference between the VSNA and traditional measures of environmental spatial ability (e.g., map retracing) is that the VSNA collects data while a person is *in* the environment itself as opposed to collection of data post hoc *outside* of the environment. Measures (e.g., direction and distance estimation) based on one's memory of an environment may be a source of construct irrelevant variance in the assessment of environmental spatial ability (Hegerty et al., 2006). Additionally, measurement outside of navigation in the environment requires individuals to make inferences that were not made within the environment (Montello, Waller, Hegarty, & Richardson, 2004). For example, the ability to point accurately to locations or estimate distances requires one to remember spatial configurations encoded in an environment. Assessing navigational performance in the task itself removes the additional burden of memory requirements that may contaminate the assessment of environmental spatial ability.

In each VSNA environment, a person must collect a set of brightly colored gems which are scattered throughout the environment. Participants need to complete the task twice for each environment. The first collection of gems is the training phase, which is intended to familiarize the person with the environment. The second collection of gems is the testing phase, which requires the person to obtain all the gems again as fast as possible. While no questions are explic-

itly asked about distance or direction between objects, the time to complete the VSNA require distance and direction estimation.

In this correlational study, we address two research questions centered around the validity of the VSNA as a measure of environmental spatial ability. The first question refers to how the VSNA relates to other measures of spatial ability. We compare the VSNA to a measure of figural spatial ability (Mental Rotation Test; adapted from Vandenberg & Kuse, 1978), vista spatial ability (Spatial Orientation Test; Hegarty & Waller, 2004) and a self-report measure of environmental spatial ability (Santa Barbara Sense of Direction; Hegarty, et al., 2002). In the Mental Rotation Test (MRT), participants view a 3D target figure and four test figures. The task is to determine which two test figures are correct rotations of the target figure as quickly and accurately as possible. The Spatial Orientation Test (SOT) requires the participant to make direction estimations from different perspectives relative to a 2D picture. For example, the person may be asked to give the direction of a car from the perspective as if the person is standing at a tree facing a traffic light. The degree to which a person can give the correct direction of objects from various perspectives is proposed to assess vista spatial ability. The Santa Barbara Sense of Direction (SBSOD) scale measures a person's self-report belief about various navigation abilities in the real world (e.g., I don't have a very good "mental map" of my environment, I enjoy reading maps). The SBSOD has been found to correlate (e.g., $r = .44$) with tests of spatial knowledge that involve orienting oneself within real-life environments (Hegarty, et al., 2002). We also compare the VSNA to verbal and math Scholastic Aptitude Test (SAT) scores since spatial ability has been shown to correlate to math achievement but not verbal achievement (e.g., Hegarty et

al., 2006). Thus we expect to show divergent validity of the VSNA by showing it does not relate to verbal SAT scores.

Finally, regarding criterion related validity, we will investigate the relationship between the VSNA scores and choosing a STEM career path. Addressing this question expands on the work by Wai, Lubinski, and Benbow (2009) showing that spatial ability predicts STEM degree attainment.

We make the following hypotheses regarding question one:

1. *The VSNA should relate more to the SBSOD scale than the SOT and the MRT.* While the SOT and the MRT have been shown to relate to the SBSOD scale (e.g., Kozhevnikov & Hegarty, 2001), the VSNA should more accurately assess environmental spatial ability than the MRT or the SOT.
2. *The VSNA will relate higher to math SAT scores than verbal SAT scores.* Spatial skills correlate to math achievement (e.g., Hegerty et al., 2006). Therefore the VSNA should relate more to math SAT scores than to verbal SAT scores.
3. *The VSNA, SOT and MRT will relate to STEM career path and achievement after controlling for gender, verbal and math ability (via SAT scores).* Spatial ability has been found to predict STEM degree attainment (Wai, Lubinski, & Benbow, 2009). Thus we expect to see a similar result for the VSNA as well as for the MRT and SOT.

The second question refers to the extent to which video game use influences environmental spatial ability (as measured by the VSNA). The question further addresses the malleability argument that video game use can impact spatial ability and specifically, environmental spatial

ability. While this study focuses on correlational relationships, it is informative since it may show that even casual video game use can have a potential effect on environmental as well as vista and figural spatial ability.

Despite the large body of work investigating the role of video games on spatial ability, we are aware of only three studies that have specifically investigated the relationship between video game use and environmental spatial ability (Rehfeld, Jentsch, Curtis, & Fincannon, 2005; Richardson, Powers, & Bousquet, 2012; Schuster et al., 2008). Schuster et al. (2008) found that video-game experience correlated with college students' ability to plan routes for unmanned vehicles in a 3D virtual simulation.

We make the following hypothesis regarding question two:

4. *Video game use will relate to performance on all measures of spatial ability (figural, vista, environmental).* Higher video game use will be associated with in better spatial abilities compared to less video game use.

Method

Participants

323 undergraduate students (129 males, 194 females) enrolled in introductory psychology and education courses at a large southeastern state university volunteered to participate in the study for course credit.

Measures

Virtual Spatial Navigation Assessment. The VSNA was created in Unity, a free video game development software tool. In the VSNA, a participant explores a virtual 3D environment using a first person avatar on a computer. The avatar is controlled by a single key on the key-

RUNNING HEAD: virtual spatial navigation assessment

board and the mouse. Pressing the key moves the avatar forward and the mouse controls the direction of the avatar. Participants are instructed that the goal is to collect three gems in an environment and return to the starting position. Participants first complete a short familiarization task that requires them to collect three gems in a small room. The VSNA consists of: (a) a small indoor environment consisting of halls (easy indoor), (b) a larger indoor environment (hard indoor), (c) a small outdoor environment (easy outdoor), and (d) a larger outdoor environment (hard outdoor). In each environment three gems are strategically located in the four environments so that an optimal path can be used to collect all the gems. In each environment the participants so that an optimal path can be used to collect all the gems. In each environment the participant must collect the gems twice. The first collection is the training phase and the second collection is the testing phase. Figure 1 displays a screenshot of the easy indoor environments.



Figure 1. Indoor environment in the VSNA

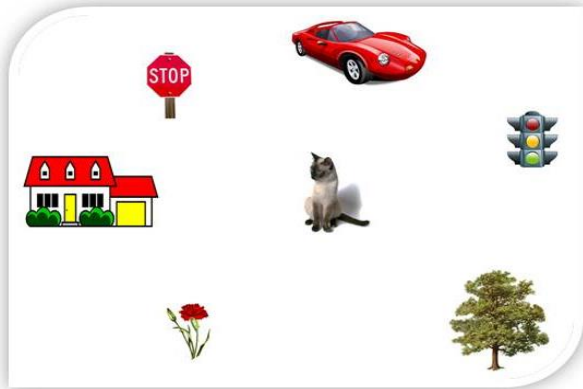
The VSNA records the time taken to complete the training and testing phases per environment (i.e., time taken to collect the gems). The training phase is intended to measure one's ability to *search* and *encode* information in the environment, while the testing phase is intended to measure one's ability to *retrieve* and *apply* the encoded information. There is a 5 minute time limit (per phase) in the easy indoor environment. The hard indoor and both outdoor environments each have a 10 minute time limit per phase. If a person times out in a training phase, the participant skips the testing phase and goes to the next environment. The automated skip was done to keep the testing phase a recall task not a searching task. Lower score indicate higher environmental spatial ability. For ease of reading we reverse coded the VSNA scores so higher scores mean better performance.

The Santa Barbara Sense of Direction Scale (Hegarty, et al., 2002). This test consists of 15 self-report items pertaining to environmental spatial ability (e.g., I am good at reading maps) that are rated on a 5 point likert scale. Higher score indicate higher environmental spatial ability.

Spatial Orientation Test (Hegarty & Waller, 2004). This test consists of 12 questions requiring the participant to estimate directions of objects from different perspectives in one picture. The participant must draw the direction in a circle relative to an arrow in the circle that is always pointing to the 12 o'clock position. Each response is scored as a distance between the participant's angle and the correct angle (ranges from 0-180 degrees; see Figure 2). For example, the correct angle that is formed between the two landmarks for the example problem is approximately 60 degrees. In order to score each item, the degree of the angle from the participant's response is subtracted from the correct angle. Thus lower scores indicate lower vista spa-

RUNNING HEAD: virtual spatial navigation assessment

tial ability. For ease of reading we reverse coded the SOT scores so higher scores mean better performance.



Example:
Imagine you are
standing at the **flower**
and facing the **tree**.
Point to the **cat**.

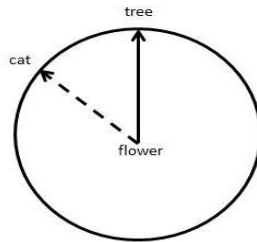


Figure 2. Spatial Orientation Test (Hegarty & Waller, 2004)

Mental Rotation Test (adapted from Vandenberg & Kuse, 1978). In this test, participants view a three-dimensional target figure and four test figures. Their task is to determine which of the test figures are rotations of the target figure. The MRT has two correct answers for each of the 10 items. The total score is based on the total number of items where both correct objects are found. Higher score indicate higher figural spatial ability.

Video game use and VSNA-video game similarity. Participants answered one question about general video game use: *How often do you play video games?* 1 = not at all, 2 = about once a month, 3 = a few times a month, 4 = a few times a week, 5 = everyday, but for less than 1 hr, 6 = every day for 1–3 hr, 7 = every day for more than 3 hr (Jackson, von Eye, Fitzgerald,

Witt, & Zhao, 2011). Additionally, participants were asked a question about the similarity between the VSNA and the video games they play: *How similar was the navigation task to a video game you play (not at all, somewhat similar, similar, very similar, completely identical)?*

Procedure

The study was conducted online in a web browser without supervision. Participants first reported their GPA, SAT scores, academic major, and completed the SBSOD scale. Then they completed the VSNA, the SOT, and the MRT. Finally, they completed some questions about video game use and usability of the VSNA. No tests were counterbalanced since we wanted to see how participants performed on the VSNA without the influence of fatigue from other spatial ability tests.

Results

Table 1 displays the means and standard deviations of the SBSOD, times across the 8 phases of the VSNA, the SOT, and MRT (listed in the order they were presented). Due to the difficult nature of the VSNA hard environments, and the study being unproctored, not all participants completed all tests. In some cases a participant timed-out of a training environment which results in the participant skipping its corresponding testing phase (as described in the VSNA measures section). In order to maximize power we still included participants in analysis who completed the training phase for each environment.

Reliability was good for the SBSOD ($\alpha = .89$), MRT ($\alpha = .76$), and SOT ($\alpha = .87$). Based on the high correlation between VSNA testing and training times ($r = .61$), we took the average score across training and testing. Easy and hard times were also highly related ($r = .56$) so we took the average score across easy and hard times to yield an indoor and outdoor VSNA score.

While the correlation between indoor and outdoor environments was also high ($r = .53$), we report results for them separately since the sample size differs between the indoor and outdoor environments. Additionally, combining the time data across the indoor and outdoor environments could give an added advantage to participants who did not complete the outdoor task (i.e., give lower means to a person who did not complete the outdoor versus a person who did complete the outdoor).

Table 1: Means and SDs of the SBSOD, VSNA phases, SOT and MRT

	N	Mean	SD
SBSOD	323	3.16	0.79
Easy indoor train*	323	132.69	55.59
Easy indoor test*	310	102.64	41.89
Hard indoor train*	322	206.64	115.66
Hard indoor test*	308	161.36	82.87
Easy outdoor train*	300	279.49	178.35
Easy outdoor test*	252	93.70	32.47
Hard outdoor train*	282	350.94	177.97
Hard outdoor test*	212	118.51	93.54
SOT	273	38.11	27.30
MRT	271	4.77	2.73

* = measured in seconds

We recoded students' self-reported major into two categories: STEM related and non-STEM related. Examples of STEM related majors include: biology, engineering, computer science, and chemistry. Examples of non-STEM related majors include: English, education, business, communication, and history. Non-majors ($n = 36$) were excluded from the STEM major variable. Table 2 displays the correlations between STEM career path (0 = non-STEM, 1 = STEM), gender (males = 0, females = 1), SAT math scores, MRT, SOT, and the indoor and out-

RUNNING HEAD: virtual spatial navigation assessment

door VSNA scores (time data, where less time is better). GPA was omitted from Table 2 since it did not relate to any spatial ability measures.

Regarding hypothesis 1 (i.e., VSNA should relate more to the SBSOD scale than the SOT and the MRT) both the indoor and outdoor VSNA scores significantly relate to the SBSOD, MRT, and the SOT. However only the indoor VSNA scores appear to support hypothesis 1: indoor VSNA scores are more highly correlated to the SBSOD ($r = .37$) relative to the MRT ($r = .24$) and the SOT ($r = .18$). The VSNA indoor and outdoor times both account for unique variance in the SBSOD after controlling for MRT and SOT scores ($pr = .36, p < .001$; $pr = .18, p < .001$).

Regarding hypothesis 2 (i.e., VSNA will relate higher to math SAT scores than verbal SAT scores), all spatial ability measures related to SAT math scores. No spatial ability measures related to GPA or verbal SAT.

Table 2: Correlations (r) among gender, STEM major, SAT, spatial measures, and video game experience

	gender	STEM	SATm	SATv	SBSOD	MRT	SOT	Indoor	Outdoor
STEM	-.12*								
SATm	-.17**	.16*							
SATv	.05	.10	.62**						
SBSOD	-.33**	.14*	.17**	-.01					
MRT	-.23**	.10	.24**	.14	.17**				
SOT	-.24**	.08	.24**	.01	.17**	.45**			
Indoor	-.44**	.22**	.16**	.02	.37**	.24**	.18**		
Outdoor	-.37**	.14*	.15*	-.04	.19**	.26**	.18**	.53**	
VG use	-.62**	.03	.19**	-.00	.18**	.17**	.29**	.37**	.33**

* = $p < .05$; ** = $p < .01$; SATm = SAT math; VG use = video game experience

Regarding hypothesis 3 (i.e., VSNA, SOT and MRT will relate to STEM major and achievement after controlling for gender, verbal and math ability), gender, SAT math scores, SBSOD, and both indoor and outdoor VSNA scores significantly relate to STEM majors. A hierarchical regression was run to predict STEM major. Predictors were entered in the following order: gender, math SAT scores, SBSOD scores, and finally VSNA indoor scores. Only VSNA indoor was a significant predictor of STEM interest after controlling for all other predictors (std $\beta = .24, p < .01$; R^2 change = .04, $F(1, 219) = 10.13, p < .05$). The same analysis was conducted entering VSNA outdoor scores last (gender, math SAT scores, SBSOD scores, VSNA outdoor scores) but the R^2 change was not significant. GPA did not relate to any spatial ability measures for the STEM majors ($n = 119$).

Regarding hypothesis 4 (i.e., video game use will relate to performance on all measures of spatial ability), video game use significantly relates to the four spatial ability measures. However, after controlling for gender and video game similarity (for VSNA only), video game use only relates to SOT ($pr = .21, p < .01$) and VSNA indoor scores and ($pr = .15, p < .05$).

Discussion

Hypothesis 1 (i.e., VSNA should relate more to the SBSOD scale than the SOT and the MRT) was partially confirmed. We found that indoor VSNA scores had moderately highly correlations to SBSOD scores than to MRT and SOT scores. Both the VSNA indoor and outdoor scores accounted for unique variance in the SBSOD after controlling for MRT and SOT scores. These findings partially supports the construct validity of the VSNA as a measure of environmental spatial ability. We did not find evidence of construct validity for the outdoor VSNA

scores (i.e., outdoor scores were more highly correlated to the MRT than to SBSOD). This finding may be due to the outdoor environments always being after the indoor environments which could cause fatigue effects on the outdoor environments. Finally, method effects (i.e., different task requirements) and well known psychometric issues related to self-report measures (e.g., social desirability) could be a reason why the correlation was not higher between the SBSOD and the VSNA.

Hypothesis 2 (i.e., VSNA will relate higher to math SAT scores than verbal SAT scores) was confirmed. We found that the VSNA, MRT, and SOT scores were all significantly related to math SAT scores and not verbal SAT scores. This is consistent with other research (e.g., Hegarty et al., 2006) that has shown a relation between mathematical and spatial abilities. This finding further supports the construct validity of the VSNA as a measure of spatial ability.

Hypothesis 3 (i.e., VSNA, SOT and MRT will relate to STEM major and achievement after controlling for verbal and math ability) was partially supported. The VSNA indoor scores significantly correlated to being a STEM career path after controlling for gender, math SAT scores, and SBSOD scores (verbal SAT scores were not related to STEM career path). Thus we established criterion related validity of the VSNA. This finding extends the work by Wai, Lubinski, and Benbow (2009) who showed that spatial ability was a significant predictor of STEM career path, even after controlling for math and verbal skills. However, we did not find the VNSA outdoor scores predicted STEM career path. This may be due to the lower number of participants who completed the outdoor VSNA. We also did not find that figural or vista spatial ability related to STEM career path. Thus environmental spatial ability may be a unique spatial ability separate from figural and vista ability that affects STEM career path. Additionally, no

spatial ability measures related to GPA. Spatial ability may not give students an added academic advantage in STEM courses. However, the GPA variable was based on a variety of courses outside of STEM subjects. Future work research should investigate how environmental spatial ability relates to grades and performance in specific STEM courses.

Hypothesis 4 (i.e., video game use will relate to performance on all measures of spatial ability) was partially supported. Video game use was significantly correlated with the SOT and the indoor VSNA after controlling for gender. Importantly, the VSNA does not just assess one's ability to play video games – video game use significantly relates to the VSNA after controlling for VSNA-video game similarity. The relation between video game use and the VSNA might be underestimated considering we only asked a broad question about video game use. Future work should consider using more detailed questions regarding video game use to further identify if specific video game use (e.g., 3D video games) relates more strongly to the VSNA.

A case can also be made that the VSNA-video game similarity question might not be sufficient to measure how similar the VSNA is to video games. For example, a particular video game player might see lots of differences between the VSNA and video games in general (e.g., lack of controls, gameplay options), while another video game player might see lots of similarities (e.g., first-person exploration in a 3D world). Future work should consider more detailed questions regarding the similarity between the VSNA and video games. However, the VSNA requires little motor control beyond skills learned by normal computer use (single button press with one hand and mouse control with the other hand). In this regard, the VSNA can be seen as a transfer task of environmental spatial ability independent from other video game play heuristics (e.g., effective use of controllers). These results are consistent with experimental evidence that

video game use can improve spatial ability (e.g., Uttal et al., 2012). Thus exposure to video games may affect one's ability to encode, store, retrieve, and apply environmental spatial information. Contrary to this theory, Richardson et al. (2012) found video game use was related to a pointing task after navigating through a virtual environment but not a pointing task after navigating through a real environment. However, Richardson et al. (2012) states the limitation of using pointing tasks to assess environmental spatial ability in that pointing tasks do not require actual navigation to targets. Thus it is possible that video game use does improve actual navigation performance *in* real environments (i.e., environmental spatial ability) but not to pointing performance *after* navigating through real environments.

Finally, the positive relation between video game use and environmental spatial ability also shows what lifestyle factors might indirectly affect interest in STEM (i.e., both the VSNA and SAT math scores relate to STEM major). While we did not find that video game use directly relates to STEM interest, video game use may indirectly affect interest in STEM by positively affecting environmental spatial ability and math ability.

Consistent with other work on spatial abilities (e.g., Spence, Yu, Feng, & Marshman, 2009), we found robust gender differences among the spatial ability measures. Females scored significantly worse on the SBSOD, SOT, MRT, and the VSNA (indoor and outdoor) compared to males. Follow up analysis revealed that after controlling for video game experience this gender effect was only eliminated for performance on the SOT. Future research should investigate how training can eliminate the gender gap in environmental spatial ability.

Looking forward, the VSNA could potentially be used for large scale assessment since it is scalable (i.e., run in a web browser) and quick to administer (as short as 10 minutes). This is

important due to the growing number of studies suggesting the need to assess spatial ability in our education system (e.g., Uttal et al., 2012; Wai, Lubinski, & Benbow, 2009). There are many STEM related careers (e.g., engineering, medicine) and non-STEM careers (e.g., transportation, military, tourism) that require high levels of environmental spatial ability. These fields could benefit from having an assessment to be used for selection as well as intervention work. Additionally, the VSNA is a performance-based assessment not subject to social desirability effects. This gives it an advantage over traditional environmental spatial ability measures (e.g., SBSOD). Finally, assessment studies of environmental spatial ability using the VSNA can be covert since the gem finding activity does not explicitly cue participants about the purpose of the VSNA. This can be useful in situations where test anxiety could potentially affect the validity of the test.

This study cannot rule out the selection hypothesis that people with high environmental spatial ability may enjoy playing video games more than people with low environmental spatial ability. Future work should focus on experimental research investigating how using 3D simulations or video games can improve performance in the environmental spatial ability. Another limitation in this study was students completed all tests online without proctor supervisions. Results might have been more robust if participants were directed to stay on task throughout the session.

Another limitation of this study was we were not able to investigate the latent factorial structure of the VSNA. Given the limited time we had to run reach participant we could only investigate two levels of difficulty (easy vs. hard) in two distinct environments (indoor vs. outdoor). Future work should investigate creating “forms” of the VSNA that contain multiple isomorphic environments of similar difficulty (e.g., 5 outdoor nature environments). These forms can be compared to other forms that contain other features of the VSNA (e.g., 5 outdoor nature

environments, 5 indoor environments, 5 outdoor urban environments). This design allows for confirmatory factor analysis and structural equation modeling. Additionally, counterbalancing and time spacing between forms should be implemented to control for any fatigue effects that may be occurring as a function of extended VSNA testing.

Finally, virtual environments do not provide any information to body-based senses (i.e., vestibular, proprioceptive) and thus may afford less detailed representations than real world environments (Richardson, Powers, & Bousquet, 2012; Waller, Loomis, & Haun, 2004). However, Wan, Wang, and Crowell (2009) provide evidence that participants still spatially update (e.g., remember locations of objects and landmarks) information in virtual environments much like in real environments. Future work should investigate how performance in the VSNA relates to real world navigation tasks.

Acknowledgement

We would like to gratefully acknowledge support for this research provided by the John D. and Catherine T. MacArthur Foundation. The views expressed herein are those of the authors and do not reflect the views of the funding foundation.

References

- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*, 387-398.
- Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*(4), 445-454.
- De Lisi, R., & Wolford, J. (2002). Improving children's mental rotation accuracy with computer game playing. *Journal of Genetic Psychology, 163*, 272–283.
- Dorval, M., & Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual and Motor Skills, 62*, 159– 162.
- Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850–855.
- Golledge, R., (1999). Human Wayfinding and Cognitive Maps. In: Golledge, R. (Ed.), *Wayfinding Behavior - Cognitive Mapping and Other Spatial Processes*, Johns Hopkins University Press, Baltimore, pp. 5-45.
- Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1465–1478.

- Hegarty, M., Montello, D.R., Richardson, A.E., Ishikawa, T., Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude test performance and spatial layout learning. *Intelligence*, *34*, 151-176.
- Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*, 425–447.
- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*, 175-191.
- Herting, M., & Nagel, B. (2012). Aerobic fitness relates to learning on a virtual morris water task and hippocampal volume in adolescents. *Behavioural Brain Research* *233*(2), 517-525.
- Ittelson, W. H. (1973). Environment perception and contemporary perceptual theory. In W. H. Ittelson (Ed.), *Environment and cognition* (pp. 1– 19). New York, NY: Seminar.
- Jackson, L.A., Witt, E. A., Games, A. I, Fitzgerald, H. E., von Eye, A., & Zhao, Y. (2012). Information technology use and creativity: Findings from the children and technology project. *Computers in Human Behavior*, *28*, 370-376.
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object-manipulation spatial ability and spatial orientation ability. *Memory and Cognition*, *29*, 745–756.
- Kristjánsson, Á. (2013). The case for causal influences of action videogame play upon vision and attention. *Attention, Perception, & Psychophysics*, *75*(4), 667-72.

RUNNING HEAD: virtual spatial navigation assessment

- Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank, & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS. Proceedings of COSIT '93. Lecture notes in Computer Science, vol. 716* (pp. 312–321). Berlin: Springer-Verlag.
- Montello, D. R., & Golledge, R. G. (1999). Scale and detail in the cognition of geographic information. *Report of the Specialist Meeting of Project Varenus, Santa Barbara, CA, May 14–16, 1998*. Santa Barbara: University of California.
- Montello, D. R., Waller, D., Hegarty, M., & Richardson, A. E. (2004). Spatial memory of real environments, virtual environments, and maps. In G. L. Allen, & D. Haun (Eds.), *Remembering where: Advances in understanding spatial memory* (pp. 251–285). Mahwah, NJ: Lawrence Erlbaum Associates.
- Rehfeld, S., Jentsch, F., Curtis, M., & Fincannon, T., (2005). Collaborative teamwork with unmanned vehicles in military missions. *Proc. 11th Annual Human-Computer Interaction International Conference, Las Vegas, NV, Human Factors and Ergonomics Society, CD-ROM*.
- Richardson, A., Powers, M., Bousquet, L. (2012). Video game experience predicts virtual, but not real navigation performance. *Computers and Human Behavior, 27*, 552-560.
- Schuster, D., Fincannon, T., Jentsch, F., Keebler, J., & Evans, A. W. (2008, December). *The role of spatial ability in the relationship between video game experience and route effectiveness among unmanned vehicle operators*. Poster presented at the meeting of the Army Science Conference.

- Subrahmanyam, K., & Greenfield, P. M. (1994). Effect of video game practice on spatial skills in girls and boys. *Journal of Applied Developmental Psychology, 15*, 13–32.
- Siegel, A.W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H .W. Reese (Ed.), *Advances in Child Development and Behavior*. (pp. 9-55) New York: Academic Press.
- Spence, I., Yu, J. J. J., Feng, J., & Marshman, J. (2009). Women match men when learning a spatial skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1097–1103.
- Tolman, E. C. (1948) Cognitive maps in rats and men. *Psychological Review, 55*, 189-208.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2012, June 4). The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. *Psychological Bulletin*. Advance online publication. doi: 10.1037/a0028446
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations: Group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599– 604.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817–835. doi:10.1037/a0016127
- Waller, D., Loomis, J. M., & Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review, 11*, 157-163.
- Wan, X. I., Wang, R. F., & Crowell, J. A. (2009). Spatial updating in superimposed real and virtual environments. *Attention, Perception & Psychophysics, 71*, 42-51.

