# LEVELING THE FIELD: REDUCING DIFFERENCES IN LEARNING FROM VIRTUAL ENVIRONMENTS

**Barbara Barnett and Bruce Perrin**
**The Boeing Company**
**St. Louis, MO**

## ABSTRACT

In recent years, Boeing has investigated the effectiveness of virtual technologies for aircraft maintainer training. In these studies, we collected data comparing transfer performance following various types of virtual maintenance training to that achieved with more traditional hardware-based training. These data revealed that performance following virtual environment (VE)-based training was marked by significant inter-subject variability. In addition, our results indicated that this transfer performance was mediated by the spatial visualization aptitude. Individuals with higher spatial visualization aptitudes learned effectively from VE-based training, while those with lower visualization aptitude did not. Results for mockup training showed no similar pattern -- mockup trainees performed equally well on the training transfer task regardless of their spatial visualization aptitude. The current study further investigates the role of spatial visualization with regard to the effectiveness of an immersive VE-based trainer. Does the same spatial visualization relationship exist within an immersive virtual training environment? And, if so, are there additional factors that could help to minimize the impact of spatial visualization, therefore making VE-based training more effective for individuals with a broader range of aptitudes? Results indicate that training transfer for the immersive VE-based training was similar to that of other VE-based training conditions. Performance in the training transfer task was highly variable and was mediated by spatial visualization aptitude. Additional experience factors of general maintenance/tool familiarity and 3-D video game exposure did not significantly reduce the effect of spatial visualization aptitude on training transfer performance. Although additional practice within the immersive VE prior to maintenance training suggests a trend towards minimizing the association between visualization and post-training performance, the relationship between visualization and post-training performance was still statistically significant. Implications for future research and training are discussed.

## ABOUT THE AUTHORS

**Dr. Barbara Barnett** has had fourteen years of human factors experience at Boeing, both in applied research and advanced design applications. Her primary focus has been in the areas of workload prediction and assessment, modeling pilot performance and decision-making, cognitive ability assessment, performance measurement, training transfer evaluation, and development of principles for decision aiding and display design. Her efforts have involved contractual and internal research and development, as well as direct support to aircraft programs. Currently, Dr. Barnett is an Instructional Designer with Boeing's Informational and Technical Data Systems group. She is currently investigating the effectiveness of virtual environments for training as part of an internally funded research and development activity.

**Dr. Bruce Perrin** is an Associate Technical Fellow, assigned to the Sustainment Data Systems area of Boeing Aircraft and Missile Systems. Bruce has been employed at Boeing for over 18 years, during which time he has been responsible for the analysis, design, development, and evaluation of training and decision support systems. He currently leads the company's research and development of Advanced Distributed Learning (ADL) technology and methods, and is principal investigator on the Advanced Instructional Systems Technology Independent Research and Development project. His research interests and publications include work in the areas of cognitive analyses of training environments, expert systems for training and job-aiding, training system evaluation, adaptive training methods, and decision-making and -aiding under uncertainty.

# LEVELING THE FIELD: REDUCING DIFFERENCES IN LEARNING FROM VIRTUAL ENVIRONMENTS

**Barbara Barnett and Bruce Perrin**
**The Boeing Company**
**St. Louis, MO**

## BACKGROUND

In studies dating from 1997 and spanning applications from underwater Remotely Operated Vehicle (ROV) pilot training to aircraft maintenance training, researchers at Boeing have documented significant individual differences in the effectiveness of Virtual Environment (VE)-based training (Perrin & Barnett, 2001; 2002). This effect appears in the significant difference in transfer performance variance between groups trained using VE techniques compared to those trained using more traditional methods. In effect, some individuals learned effectively under VE-based training, others did not, while the impact of more traditional approaches (e.g., hands-on training, training based on simulated controls and displays) was more consistently positive.

Over roughly the same time period, spatial navigation became perhaps the most widely studied application of VE technology to training. Spatial navigation training is a natural extension of VE technology, as the researchers' physical settings (e.g., office buildings or outdoor areas like near-by parks or university campuses) are often represented in the VE. Training transfer is then assessed in the corresponding, real-world setting. In a 2000 article, Durlach and his coauthors reviewed the findings of 39 studies of spatial navigation involving VE technology, 12 of which directly addressed the question of their training effectiveness (Durlach, Allen, Darken, Garnett, Loomis, Templeman, & von Wiegand, 2000). One of the primary conclusions of this review was that the effectiveness of the VEs for training was obscured by the significant variability among trainees in post-training performance, e.g., "…as a consequence of the large inter-subject variation, any attempt to show that the performance of one system is significantly better than that of another is hopeless" (p. 596).

Although the relevant research is limited and somewhat inconsistent, several studies suggest that the effectiveness of VE-based spatial navigation training may be mediated by spatial aptitude or ability. Waller (1999), for example, traced learning to navigate within a VE to individual differences in spatial ability. In his research, spatial ability was measured by a composite of the visualization and spatial orientation aptitudes and a way-finding performance measure. Based on findings from these studies, one might wish to argue that learning to navigate large-scale areas is inherently spatial, and therefore, should be mediated by spatial aptitude. In fact, until recently, the prevailing interpretation for the well-documented difficulties people have in navigating within a VE were based on theories and models of how people learn to navigate in the real world (Colle & Reid, 1998). The explanation, however, is not that simple, as spatial aptitude does not predict navigation learning in the real world. As Waller (1999) notes, "spatial ability has no such relationship with spatial knowledge acquired in the real world – even when that knowledge is assessed with the same types of measures" (p. 170-171). Put simply, spatial aptitude does not mediate learning to navigate; it mediates learning to navigate within a VE.

Our own research findings parallel these results. Across the studies we have performed, which examined a host of background characteristics, abilities, and aptitudes, only the visualization aptitude (Ekstrom, French, Harman, & Dermen, 1995) was found to consistently and significantly correlate with transfer performance for VE-trained individuals (Perrin & Barnett, 2001; 2002). Additionally, consistent with Waller's findings, the visualization aptitude was not correlated with performance following training on physical systems or simulations of them, i.e., the visualization aptitude mediated learning in the VE, not learning the task using other training approaches.

Our research that extends the mediating effect of spatial abilities to maintenance training differs from most of the spatial navigation research in one important respect, i.e., the demonstrated effect in our studies occurred using desktop VE training systems. That is, unlike the spatial navigation studies, which primarily used head-mounted displays or similar technologies to support stereoscopic viewing of 3-D environments, i.e., an immersive VE, our systems showed 3-D aircraft design models on a computer screen. Thus, our systems were limited to those depth cues that are available from 2-D displays, such as shading, texture gradients, and monocular motion parallax.

Consequently, the first objective of the current study is to assess whether the visualization aptitude mediates transfer performance following maintenance training in an immersive VE, as it has shown to do with desktop training. While a positive finding seems likely, given the extensive research on spatial navigation training, our previous research has shown that increasing the graphical detail used in our desktop systems reduced the mediating effect of the visualization aptitude. Presumably, the increased number of depth cues available in an immersive VE might further reduce the constraining effect of visual ability. On the other hand, a positive demonstration involving a task that occurs within a very limited area further supports the growing body of research that indicates that spatial ability constrains learning within a VE to a degree not seen in learning in the real world.

Assuming that visual/spatial abilities do constrain what is learned in an immersive VE, the next question is whether these abilities can be changed via experience. Although work on this issue is limited, research generally suggests that environmental experiences can alter at least some aspects of spatial performance (Baenninger & Newcombe, 1995), and general exposure to computers is often suggested as a factor that may influence the effectiveness of VE-based training (e.g., Waller, Hunt & Knapp, 1998). To address this possibility, we surveyed our participants about the time they have spent playing 3-D computer and video games, as well as the time that they have spent performing common maintenance/repair tasks involving the use of hand tools. Should these experiences be found to correlate with both post-training performance and the visualization aptitude, it would be difficult to know if prior experience increased spatial abilities, or if individuals with higher spatial abilities were more likely to seek the challenge of these activities.

Apart from long-term experience with computers and 3-D game environments, exposure to a VE has been proposed as a way to train the specific spatial skills that it requires (Durlach et al., 2000). Indeed, some initial positive findings have been reported from giving students additional training time in the VE (Waller, Hunt & Knapp, 1998). It is difficult to know, however, if these positive results are due to developing specialized spatial skills or just increased practice on the training problems. To account for this possibility, we have included in the current study a training group that receives increased exposure to the immersive VE system, but by practicing a maintenance task unrelated to the task under study. Thus, our participants are afforded an opportunity to gain skill in understanding and manipulating the VE, without gaining information specific to the training and transfer maintenance task.

## EXPERIMENTAL DESIGN

For the study described in this paper, two different immersive VE conditions were used to train a specific aircraft maintenance task. Training transfer was evaluated on a hardware mockup of the task. A description of these training conditions, the maintenance task, and the mockup follows.

### Maintenance Task

The maintenance-training task used in this study was the removal and replacement of an aircraft fuel valve. This 24-step procedure provided the appropriate amount of complexity, and imposed both cognitive and physical demands on the trainees. Physical and visual obstructions made this task ideal to evaluate some of the unique capabilities of visual simulation, e.g., the ability to see or move through solid structures. The pre-existing computer models of the necessary parts aided in rapid development of the training and mockup materials. This maintenance task has been used in a number of previous VE-based training studies, so the data from the current study may be compared to that from previous research. Participants completed a single pass through the training materials, and then completed the transfer test, identical to participants in the previous studies.

### Training Conditions

Two immersive VE training conditions were established for this study: immersive VE with basic practice, and immersive VE with extended practice. For the basic practice condition, participants were given a brief familiarization with the immersive environment, which included instructions on maneuvering within the environment, grasping objects, and moving from step to step in the maintenance procedure. In the extended practice condition, participants were provided with extra practice on a separate and distinct maintenance task within the immersive environment, in addition to the basic familiarization. This extended practice task involved using the same VE navigation and manipulation conventions, and so, should have allowed the participants to develop a more complete understanding of the visual representation of spatial information this VE provided. The maintenance task used in extended practice, however, had no similarities to the eventual training task used in the study. Participants in the basic practice condition received an average of 10 minutes familiarization prior to the start of their training, while those in the extended practice condition received an average of 30 minutes of familiarization and extended practice. For the training portion of the study, conditions in both training groups were identical. The training conditions differed only in the amount of practice that the trainees had in the immersive environment prior to the maintenance-training task.

***Figure 1. (a) Participant using the immersive visor and 3-D mouse; (b) View of the maintenance task from within the visor.***

**The Immersive Environment**

In both of these immersive VE training conditions, participants donned a head mounted display (HMD) and used a 3-D mouse to interface with the virtual environment (see Figure 1). The study was run using dVISE software on a Silicon Graphics Reality Monster. A stereoscopic HMD manufactured by Virtual Research was used to present the images to the participants. The HMD consisted of a 640 x 480 display with a refresh rate of 30hz. Ascension magnetic tracking devices were mounted to the HMD and in the 3-D mouse to monitor the position of the participant's head and hand in space. When the 3-D mouse was in the participant's field of view, a "virtual hand" appeared within the VE that could be used for grasping tools or manipulating parts.

Prior to the study, an extensive series of pilot tests was conducted to optimize the 3-D mouse interface to the VE for training purposes, thus minimizing the complexity of the trainee's interaction with objects. Some of the main features were as follows: changes in a part's color were used to indicate when a part was selected with the 3-D mouse; tools were automatically placed in the virtual hand at relevant steps in the task, parts could be activated only at the appropriate step in task procedure; and parts were "snapped" to predefined locations to eliminate the need for minute position corrections using the mouse. These VE interface implementations allowed the trainee to become proficient rapidly within the immersive environment -- eliminating the need for lengthy familiarization training.

**Hardware Mockup**

The hardware mockup consisted of a full-scale mockup of the appropriate aircraft section, shown in Figure 2. The mockup was built from actual production models and drawings, using off-the-shelf hardware and simulated aircraft parts built to specifications from epoxy resin. For this study, the mockup was used to assess training transfer. Participants were asked to use the mockup to remove and replace the fuel valve, exactly as they had been trained.

**Procedure**

Participants were Boeing employees who volunteered to participate in a training study via solicitation in the Boeing daily news. These volunteers were not pre-screened for skill or ability, although people who had previous experience with either aircraft maintenance or fuel systems were not allowed to participate in the study. Each participant was randomly assigned to one of the two training conditions. There were 40 participants total--20 in each condition.

**Figure 2. Participant using the hardware mock-up**

Each participant completed the study individually. At the beginning of the test session, the participant completed a short background questionnaire containing basicdemographic questions, including questions about the time participants spent playing 3-D computer and video games and performing basic repair and maintenance tasks. The questionnaire was followed by a paper-and-pencil test to assess spatial visualization aptitude (Ekstrom, French, Harman, & Dermen, 1995). After the background information was collected, each participant was given instructions regarding the training procedures. The participant was then outfitted with the HMD and 3-D mouse and then completed the maintenance training in either the basic or extended immersive practice condition. Trainees listened to audio instructions presented through a speaker, and performed the maintenance task within the immersive environment as per the instructions. In all cases, progression through the training course was self-paced and only a single pass through the training materials was allowed. After the training was finished, the participant immediately completed a written test of their task knowledge. Some of the questions were detailed and specific, so that subtle differences in acquired knowledge could be determined.

The participant next completed the training transfer task on the full-scale hardware mockup without the aid of written or audio instructions for reference. They were asked to work as quickly and as accurately as possible, and to complete the steps in the order in which they had been trained.

**Performance Measures**

A number of performance measures were collected during the study. Knowledge test scores were recorded, as was time spent completing the transfer task. Errors in transfer task performance were categorized and recorded. The following actions were considered errors:

**Omitted step** – A step from the training was omitted.

**Step out of order** – Steps were completed in a sequence other than as presented in the training.

**Incorrect action** – An action was taken that was inconsistent with the trained procedure. For example, if a participant tried to install a component in the wrong orientation.

**Wrong tool** – The wrong tool was used to complete a step. The most common tool error was using a regular screwdriver rather than the speed handle with the screwdriver adapter.

**Forgotten procedure** – The participant simply forgot the procedure, and had no idea what came next. The experimenter had to step in and "suggest" what to do next.

**Major Procedural Errors** –These are errors that would result in damage to the task components or would compromise the integrity of the aircraft flight. For example, forgetting to reconnect the fuel lines.

These error categories were later combined to distinguish important errors from minor, inconsequential errors. Errors that were considered to be important were: omitted step, incorrect action, forgotten procedure, and major procedural error. These three primary measures of performance – important errors, task performance time, and knowledge test scores – were combined into a single performance index by converting each to z-scores and summing them.

**Data Analysis**

Data from this study were analyzed by comparing them to data collected earlier using the same maintenance-training task, but different training conditions. The additional training treatments were as follows:

**VE-High Detail**. Under this treatment, equipment components were presented on a computer screen with considerable detail, providing depth cues such as shading, texture gradients, and monocular motion parallax (Figure 3). This view was highly similar to that of the immersive VE, but the images were displayed from fixed camera points on a computer screen instead of a HMD, and stereoscopic presentation was not supported. Highlighting and animation were used to emphasize specific parts or actions.

***Figure 3. Sample display under VE-High Detail training.***

**VE-Low Detail.** This treatment used line drawings (Figure 4) that were taken directly from the existing Interactive Electronic Technical Manual (IETM) for this task. While these drawings would not generally be classified as a VE, they do provide a 3-D perspective, but only supported by the most rudimentary object perception cues such as the principles of closure and good continuation. As with the High Detail treatment, the relevant parts were highlighted with color, and animation was used to illustrate a step. The low- and high-detail graphics were both based on the same underlying Unigraphics models of the fuel system, and were created with their differing visual properties by translating the models into different formats and by using different presentation software.

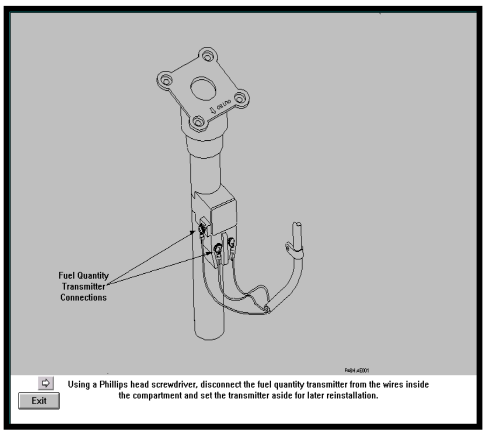


***Figure 4. Sample display under VE-Low Detail training.***

**Mockup**. This condition consisted of "hands-on" training using the same full-scale mockup that was described earlier in the training transfer task. For purposes of training, the participant received the same audio instructions as the other training conditions, but actually performed the maintenance task on the hardware mockup as they progressed through each step in the training. In addition, participants received coaching from an instructor as they performed the task. A door was provided on the side of the mockup through which trainees could see the inside of the fuel compartment. This door was later closed during transfer testing.

As in previous studies, post-training performance variance was assessed with Hartley's F-test. We then conducted an analysis of covariance, which included a test for the covariate by treatment interaction. This analysis specifically tests whether the relationship between the covariate and the performance measures differ across training treatments. Finally, for those measures with a significant treatment by covariate interaction, we then performed a t-test of the within-class correlation between the covariate and performance to determine the treatments for which the relationship was significantly greater than chance.

## RESULTS

Results from the F-tests for homogeneity of variance among the treatments for the combined measure of performance are summarized in Figure 5. As seen in previous studies, there were significant differences in performance variance under the different training conditions. Variance following training for all VE-based training conditions was high, with variance resulting from mockup-based training the lowest. A similar pattern of results was observed for scores on the knowledge test and the number of important errors, but not for test time.

| Type of Training | Variance | F-test |
|---|---|---|
| Low-Detail VE | 5.768 | $F$ (5,19) = |
| High-Detail VE | 5.998 | 10.74, |
| Immersive: Basic Practice | 7.055 | p < .001 |
| Immersive: Extended Practice | 6.211 | |
| Mockup | 0.657 | |

***Figure 5. F-test on combined performance variance following different training conditions.***

In order to investigate this lack of homogeneity of variance further, we conducted an analysis of covariance for each of our dependent measures, using training condition as the independent variable and experience, computer/video game exposure, and spatial visualization as covariates. Looking first at total experience, we observed significant main effects on each of the performance measures assessed: knowledge test ($F$ (1, 84) = 19.25, $p$ <.001), test time ( $F$ (1, 84) = 10.22, $p$ =.0019), important error

($F$ (1, 84) = 24.55, $p$ <.001), and combined performance ($F$ (1, 84) = 31.39, $p$ <.001).

Our analysis of the experience by training condition interaction, however, was not significant for any of the performance measures. This pattern of significant main effects but no interaction indicates that previous experience with tools and maintenance improved performance uniformly across all training conditions.

Looking at computer/video game exposure, we observed no significant relationship between hours spent playing games and any of our performance measures. This finding was the same for main effects for each of the performance measures, as well as the interaction among different treatment conditions.

Our analysis of spatial visualization provided a third pattern of results. For most performance measures, there was a significant visualization by condition interaction: knowledge test ($F$ (4, 85) = 2.93, $p$ = .0254), important errors ( $F$ (4, 85 ) = 4. 33, $p$ = .0031), and combined performance ($F$ (4, 85) = 3.37, $p$ = .0131). These significant interactions indicated that spatial visualization impacted training transfer performance differently, depending upon which type of training the participant received.

In order to further define the nature of this relationship, we calculated within class correlations between each of the different performance metrics and spatial visualization for each training condition. Figure 6 shows the results from the t-test of the within class correlations between combined training transfer performance and the visualization aptitude. The Low-Detail VE condition as well as the two Immersive conditions produced post-training performance that was significantly associated with the visualization aptitude. With the High-Detail VE condition, the correlation between post-training performance and this aptitude was no longer significant ($r$ = .257), and it drops even more ($r$ = .1379) when training was provided on the physical mockup.

| **Type of Training** | **Within Class r's** | **t-test** |
|---|---|---|
| Low-Detail VE | 0.485 | t (18) = 2.35, $p < .05$ |
| High-Detail VE | 0.257 | t (18) = 1.13, n.s. |
| Immersive VE: Basic Practice | 0.811 | t (18) = 5.89, $p < .01$ |
| Immersive VE: Extended Practice | 0.444 | t (18) = 2.10, $p < .05$ |
| Mockup | 0.138 | t (18) = 0.59, n.s. |

***Figure 6. Within-class correlations between visualization and combined performance following different types of training.***

Additional performance metrics revealed a similar pattern of results where spatial visualization aptitude was significantly correlated with performance for some, if not all, of the VE-based training conditions, while not correlated with performance for the mockup-based training condition. These results are indicated in Figure 7, where within class correlations are listed by training condition for each performance measure.

| **Type of Training** | **Knowledge Test** | **Important Error** |
|---|---|---|
| Low-Detail VE | $r$ = .429<br>t (18) = 2.01<br>$p < .05$ | $r$ = -.471<br>t (18) = -2.26<br>$p < .05$ |
| High-Detail VE | $r$ = .413<br>t (18) = 1.92<br>$p < .05$ | $r$ = -.025<br>t (18) = -.106<br>n.s. |
| Immersive VE: Basic Practice | $r$ = .795<br>t (18) = 5.56<br>$p < .01$ | $r$ = -.662<br>t (18) = -3.75<br>$p < .01$ |
| Immersive VE: Extended Practice | $r$ = .446<br>t (18) = 2.12<br>$p < .05$ | $r$ = -.520<br>t (18) = -2.58<br>$p < .05$ |
| Mockup | $r$ = -.051<br>t (18) = -0.22<br>n.s. | $r$ = .077<br>t (18) =0.33<br>n.s. |

***Figure 7. Within-class correlations between visualization and individual performance measures following different types of training.***

Similar to the pattern observed for the combined performance measure, we observed significant correlations between the performance metrics and spatial visualization aptitude for both of the Low-Detail VE immersive extended practice, and immersive basic practice metrics. Alternatively, only one metric, the knowledge test score, correlated significantly with spatial visualization for the High-Detail VE condition, and no performance metrics correlated with spatial visualization for the mockup-based training condition.

The only performance measure for which there was no significant training condition by spatial visualization interaction was test time ($F(4, 85) = 0.49$, $p$ = n.s.). For this measure, there was a significant main effect of spatial visualization ($F(1, 85) = 6.71$, $p$ = <.01), but this effect did not vary as a function of training condition.

Figure 8 plots post-training maintenance performance as a function of visualization score for individuals who received training using the Basic Practice Immersive VE or the mockup. All of the individuals trained on the mockup, regardless of their visualization score, performed at or above average. Following training based on the Basic Practice Immersive VE, however, performance was much more variable. Additionally, it was correlated with the visualization aptitude, such that individuals with high visualization aptitude scores tended to perform the transfer task more skillfully than those with relatively low scores. This pattern would be replicated if we were to plot additional performance metrics or training conditions.

We did observe a marginally significant correlation between computer/video game experience and our measure of spatial visualization ($r$ $(98)$ = .222, $p$ <.05). Because this data is correlational, we have no ability to determine if spatial visualization ability is developed as a result of playing the 3-D games, or if people with high visualization ability are more likely to enjoy and therefore play the games. Either way, the correlation between these

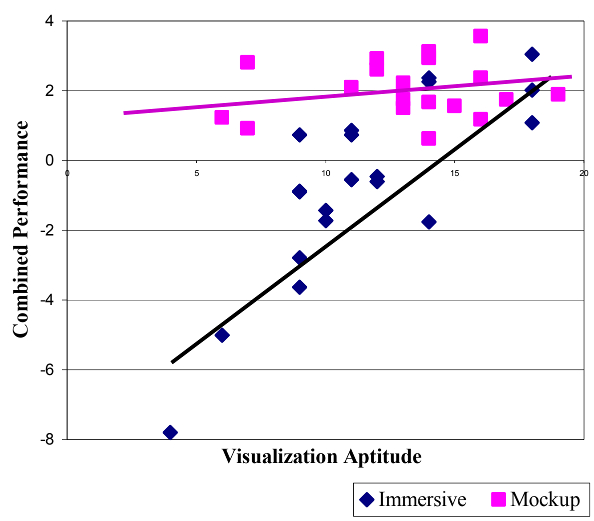


***Figure 8. Transfer performance as a function of visualization after Immersive Basic Practice Training or Mockup training.***

two potential performance predictors does not carry over to observed performance, as only the spatial visual metric is predictive of performance for the VE-based training treatment conditions.

## CONCLUSIONS AND FUTURE DIRECTIONS

The goal of the current research effort, along with our previous efforts, was to systematically evaluate the impact of virtual environments on training effectiveness. In doing so, we have evaluated a number of different VE-based training implementations, comparing training effectiveness against a baseline mockup training condition using a training transfer task. In addition, we have assessed the impact of potential mediators of performance for each of the training methods that we have evaluated. Several significant findings have emerged from this research effort.

The findings from the current study, similar to those in our previous studies, suggest a strong relationship between VE-based training effectiveness and the spatial visualization aptitude. In these instances, participants of high visualization aptitude benefited from the VE-based training, while those with lower spatial aptitude did not. This is true for the aircraft maintenance task described in the current study, but has been demonstrated in other domains with different VE-based simulation implementations as well (Perrin & Barnett, 2002). The focus of the current study suggests that this relationship holds for immersive VE-based applications, as well as the previous desktop-based VEs. Unlike the VE-based training, performance for participants in the mockup-based training condition was not impacted by the spatial visualization aptitude. In this condition, participants of both high and low visualization aptitude benefited from the mockup training.

Within the current study, we also investigated whether additional exposure to the VE prior to the maintenance training exercise would impact this relationship between visualization aptitude and training effectiveness. Although the correlation between these two factors was smaller under the extended practice condition, both were significant. In the current study, our extended practice condition provided trainees with an average of 30 minutes exposure within the VE prior to initiating the training, compared to an average of 10 minutes for the basic practice trainees. One possibility is that with significantly more experience within the VE prior to training, we would observe further reduction in the impact of visualization on VE-based training effectiveness. Constraints of our experimental protocol did not allow us to perform such an evaluation.

Additionally, we investigated the relationship between training transfer and other potential mediators of performance, namely computer/video game experience and a general measure of tool familiarity/usage. Contrary to a commonly held belief, we observed no relationship between computer/video game experience and VE-based training effectiveness. For each of our VE-based training conditions, as well as the mockup training condition, we failed to observe any significant correlations between hours spent playing video games and training transfer performance. This finding contradicts those who have speculated that future generations raised on computer/video games will naturally be more adept at operating within VE-based simulations. Our measure of tool familiarity/usage resulted in an overall improvement in training transfer performance, but this measure was uniform across all training conditions. There were no unique impacts on the VE-based conditions.

Future research efforts will focus on identifying those unique aspects of the VE that could be contributing to the relationship between visualization aptitude and training effectiveness. Are there factors specific to the tasks selected for evaluation that have contributed to this effect? In future studies, we will be looking at different types of maintenance tasks (e.g. diagnostics and troubleshooting) to evaluate the generalizability of our findings across tasks.

## REFERENCES

Baenninger, M., & Newcombe, N. (1995). Environmental input to the development of sex-related differences in spatial and mathematical ability. *Learning and Individual Differences, 7(4)*, 363-379.

Colle, H.A., & Reid, G.B. (1998). The room effect: Metric spatial knowledge of local and separated regions. *Presence, 7(2)*, 116 – 128.

Durlach, D., Allen, G., Darken, R. Garnett, R.L., Loomis, J., Templeman, J., & von Wiegand, T.E. (2000). Virtual environments and the enhancement of spatial behavior: Towards a comprehensive research agenda. *Presence, 9(6)*, 593 – 615.

Ekstrom, R.B., French, J.W., Harman, H.H. & Dermen, D. (1995). *Manual for Kit of Factor-Referenced Cognitive Tests*, Princeton, NJ: Educational Testing Service.

Perrin, B.M., & Barnett, B.J. (July, 2001). *Virtual environments for training*. Presentation at the Image Conference 2001, Scottsdale, AZ.

Perrin, B.M., & Barnett, B.J. (April, 2002). *Is VE-based training for everyone? The mediating effect of visualization in training transfer*. Paper presented at ITEC 2002, Lille, France.

Waller, D.A. (1999). *An assessment of individual differences in spatial knowledge of real and virtual environments*. Unpublished doctoral dissertation, University of Washington.

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence, 7(2)*, 129 – 143.

## ACKNOWLEDGEMENTS

A portion this work was completed under the Maintainer-in-the-Loop Program as part of the Joint Strike Fighter Concept Demonstration Phase. The authors would like to thank Michael Hennessy and Robert Finley, Maintainer-in-the-Loop Technical Coordinators, and Mariellen Curtis, Boeing Training Integrated Product Team Point of Contact. The authors would also like to acknowledge the inputs of Glenn Hancock, Katrin Helbing, and Will Wojczyk in preparing technical materials and collecting data as part of this research effort.