

Influence of Perspective-Taking and Mental Rotation Abilities in Space Teleoperation

M. Alejandra Menchaca-Brandan, Andrew M. Liu, Charles M. Oman, and Alan Natapoff
Man-Vehicle Laboratory, Massachusetts Institute of Technology
77 Massachusetts Avenue, Room 37-219. Cambridge, MA 02139
+1 (617) 253-7805
menchaca@mit.edu, amliu@mit.edu, coman@mit.edu, natapoff@mit.edu

ABSTRACT

Operator performance during Space Shuttle and International Space Station robotic arm training can differ dramatically among astronauts. The difficulty making appropriate camera selections and accurate use of hand controllers, two of the more important aspects for performance, may be rooted in a problem mentally relating the various reference frames used by the displays, hand controllers and robot arm. In this paper, we examine whether the origin of such individual differences can be found in certain components of spatial ability. We have developed a virtual reality simulation of the Space Station Robotic Workstation to investigate whether performance differences can be correlated with subjects' perspective-taking and mental rotation abilities. Spatial test scores were measured and correlated to their performance in a docking robotic task. The preliminary results show that both mental rotation strategies and perspective-taking strategies are used by the operator to move the robot arm around the workspace. Further studies must be performed to confirm such findings. If important correlations between performance and spatial abilities are found, astronaut training could be designed in order to fulfill each operator's needs, reducing both training time and cost.

Categories and Subject Descriptors

H.1.2 [User/Machine Systems]: Human factors.

General Terms

Performance, Experimentation, Human Factors.

Keywords

Space teleoperation, robotic arm, perspective-taking, mental rotations, spatial ability.

1. INTRODUCTION

1.1 Space Teleoperation

Teleoperation is used in a wide variety of areas such as medicine, underwater exploration and space activities to perform tasks in

environments that are hazardous or inaccessible to human beings. In the case of space activities, the Shuttle and International Space Station (ISS) Remote Manipulator Systems (RMS) are the main teleoperation systems currently being used. Both systems provide a set of camera viewpoints that are used by the operators to complete the task while avoiding any collisions with the surrounding structure of the Shuttle or Space Station. On the Shuttle, operators also have a direct view of the arm through a flight deck window facing onto the payload bay.



Figure 1. ISS Robotic Workstation. The operator has only three camera viewpoints that provide visual feedback when controlling the robot arm.

Many operational difficulties arising from the current space manipulator systems have been reported. Typically, operators can have problems determining clearance from structure because the fixed camera locations generally do not provide "optimal" views of the work space [12]. Only three camera viewpoints are seen at any moment, although more viewpoints are available for display, and the location of these cameras must be memorized by the operator. These additional views also likely increase the mental workload during operations. Thus, current procedures require a second operator to provide additional monitoring of the scene to avoid collisions during the task. Because of the danger of a collision with structure or a payload, arm movements are made at slow velocities and after the operators have established spatial awareness of the situation. These spatial difficulties are also one reason that multi-arm operations have not been performed during a spaceflight.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HRI '07, March 9-11, 2007, Washington, DC, USA.

Copyright 2007 ACM 978-1-59593-617-2/07/0003...\$5.00.

1.2 RMS Training and Performance Assessment

NASA astronauts begin their initial teleoperation training using a generic robot arm simulation called BORIS, or the Basic Operational Robotics Instructional System. This simulation consists of a 6 degrees-of-freedom robot arm located in a cubic room and with different camera views available. As in the Shuttle and Space Station systems, the manipulation of the arm is performed using three camera views and two hand controllers. Candidates are taught to choose the ideal camera views based on clearance visualization, as well as to correctly use the hand controllers while avoiding collisions and singularities.

Teleoperation performance is evaluated by a group of instructor astronauts based on standard criteria covering all aspects of operations. Each aspect of operation is given a different weight depending on its importance to achieving mission success. The criteria given the most importance during evaluation include: spatial/visual perception (i.e., proper camera selection and real time tracking, end position and attitude correctly visualized), situational awareness (i.e., collision and singularity avoidance), and appropriate input of the controls (i.e., ability to control multi-axis movements, motion smoothness). Astronauts that do not reach the proscribed level of skill must go through additional training and practice. Often, different trainers are assigned to help, each providing their own “personal” strategies for visualizing the workspace and accomplishing the training exercises, with the hope that one of these suggestions will enable the astronaut to succeed. (J. Young, personal communication) These strategies are likely based on the spatial skills of the trainer, and perhaps when a trainer and astronaut of similar spatial skills are matched the strategies are more readily learned.

Astronauts in robotics training exhibit significant differences in their final level of performance after initial training as well as the rate at which they acquire the necessary skills. The initial level of skill during training is not a reliable predictor of a trainee’s final level of performance [S. Robinson, personal communication].

1.3 Spatial Abilities

Spatial ability can be defined as our ability to generate, visualize, memorize, remember and transform any kind of visual information such as pictures, maps, 3D images, etc. This ability is divided into several subcomponents, which relate specifically to each of the different mental functions for image processing.

The subcomponents of spatial ability that are most relevant to teleoperation are *perspective-taking* and *mental rotations*. Perspective-taking (also known as *spatial orientation*) is the ability to imagine how an object or scene looks from perspectives different to the observer’s. Mental rotations (also known as *spatial relations*), refers to the ability to mentally manipulate an array of objects. While these two abilities are logically equivalent, the critical difference lies in the coordinate frame which is manipulated to obtain the final view. Perspective-taking requires a change in the egocentric reference frame within a fixed world coordinate frame, whereas mental rotations and spatial visualization of objects occur within a fixed egocentric reference frame. Recent work by Kozhenikov and Hegarty [10] and Hegarty and Waller [9] have shown a measurable distinction between mental rotation and perspective-taking, although performance is also highly correlated.

Spatial visualization is among other subcomponents of spatial ability, and is defined by Ekstrom [5] as “the ability to manipulate or transform the image of spatial patterns into other visual arrangements.” Carroll [2] includes other factors such as *closure speed* (ability to rapidly access representations from long-term memory), *flexibility of closure* (ability to maintain the representation of an object in working memory while trying to distinguish it in a complex pattern), *perceptual speed* (ability to rapidly compare or find symbols or figures), and *visual memory* (ability to remember the spatial distribution of objects or figures). Recently, a separate set of factors, dynamic spatial performance, has been suggested [3, 15], where the abilities to perceive and extrapolate real motion, predict trajectories and estimate the arrival time of moving objects are assessed. These factors are not considered in this paper.

Many tests have been developed in order to measure these various factors of spatial ability. Many widely used tests such as the Cube Comparisons test (mental rotation), Paper Folding test (spatial visualization), and Spatial Orientation test can be found in the Kit of Factor-Referenced Cognitive Tests (ETS, Princeton, NJ) [6]. Among the perspective-taking tests are the Perspective-Taking Ability (PTA) Test, a computer-based test developed from the work described in [10], and the Purdue Spatial Visualizations test: Visualization of Views (PSVV), a paper-and pencil test found in [8]. A three-dimensional version of the PTA test, called the Pictures test, was used in [9] but has not been validated with a large population of subjects.

1.4 Spatial Abilities in Teleoperation

Indirect evidence that spatial abilities contribute to teleoperation performance comes from experiments that have manipulated the display and control reference frames during teleoperation. Lamb and Owen [11] evaluated the differences in space teleoperation performance when using egocentric (end effector) and exocentric (world) frames of reference, concluding that higher performances were obtained when an egocentric frame of reference was used. DeJong, Colgate, and Peshkin [4] also showed that reducing the number of rotations between the different reference frames can lead to improved performance. Some of the manipulations to eliminate frame rotations, such as physically moving displays, are not practical in RMS operations, because camera views can be changed during the course of the task.

The relationship between spatial abilities and teleoperation has been studied in a few previous studies. Eyal and Tendick [7] studied the effect of spatial ability when novice subjects learned to use an angled laparoscope, a medical form of teleoperation. They measured spatial ability with the Card Rotation, Paper Folding, and Perspective-Taking tests and found significant correlations of laparoscopic performance with all three measures of spatial ability. Lathan and Tracey [13] found a correlation between spatial abilities and mobile robot teleoperation performance when navigating a maze using a single camera for visual feedback. Spatial ability was measured with four tests: the Complex Figures and Stumpf Spatial Memory Tests, which gauge spatial recognition, and the Block Rotation and Stumpf Cube Perspective Tests, which measure spatial manipulation ability. Tracey and Lathan [17] examined the effect of spatial ability on the transfer of training from simulation to a real teleoperation task and found subjects with lower spatial scores showed increased transfer of training. Spatial ability was measured by the Paper Folding tests

and Stumpf's Cube Perspectives Test. All three of these tasks differ from RMS operations, in that they only provide a single view of the workspace.

We believe that both perspective-taking and mental rotation abilities play a major role in the astronaut's performance of RMS tasks. Mental rotation ability is used to understand the individual movements of the arm/payload that is seen in a single view. Perspective-taking ability is probably used when mentally imagining the different camera perspectives displayed when selecting an appropriate view of the space; it is also likely engaged when integrating the multiple viewpoints into a single representation.

This paper presents a first effort to investigate the correlation perspective-taking and mental rotation spatial abilities with teleoperation performance. By manipulating the spatial distance between camera views shown to the operators, we hope to show that operators with higher perspective-taking test scores produce better performance on our simulated task. If our hypothesis can be supported and space teleoperation performance can be predicted to some extent, astronaut training could perhaps be tailored to each astronaut's spatial needs, making the learning process more efficient.

2. METHODS

2.1 Environment

We created a simulated RMS workspace similar in nature to the BORIS training software using the Vizard VR development package (WorldViz, Santa Barbara, CA). It consisted of a 6 degrees-of-freedom (DOF) PUMA-like robotic arm of similar dimension to the Shuttle RMS mounted on one end of a fixed truss. The kinematics were computed using the RRG Kinematix v.4 software library (Robotics Research Group, Univ. of Texas) as a plug-in module for Vizard. A simulated cargo module and ISS node were created as the objects for grasping and docking. (Figure 2) The dynamics of the arm and other objects were not modeled in this simulation.

The arm and working environment could be viewed from up to three camera viewpoints at a time displayed on three separate monitors. The simulation was run on two Windows computers using Vizard's networking capabilities. The main simulation server (1.5GHz Pentium4 PC with dual head nVidia GeForce 6600 graphics) performed the kinematic calculations, graphics processing and hand controller I/O. The second computer (550MHz Pentium3 PC with nVidia GeForce3 graphics) rendered the third camera viewpoint.

The subjects controlled the robotic arm using two 3DOF joysticks in a manner similar to that used in spaceflight. A translational hand controller was custom built in the lab from a 2DOF joystick, a linear potentiometer and USB controller card to provide three DOF control of translation. A Logitech Sidewinder 3DOF USB joystick was used as the rotational hand controller. Throughout the experiment, operators moved the end effector position using velocity control in a fixed world coordinate frame, similar to the "external frame mode" of actual RMS operations. End effector rotations were made with respect to a reference frame fixed to the end effector.

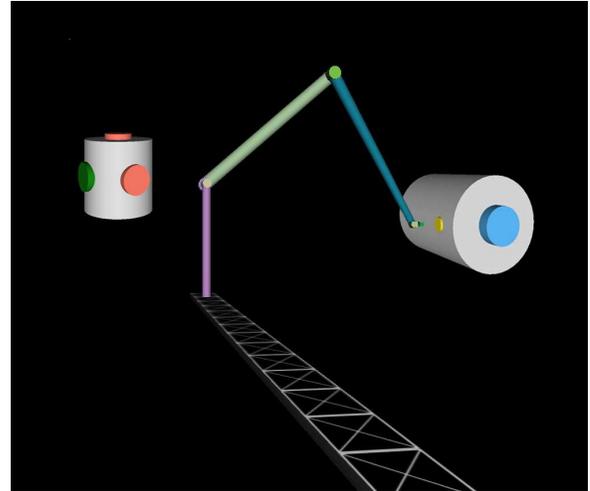


Figure 2. A view of the simulated teleoperation environment showing the robotic arm, truss, a module near the free end of the arm and the node.

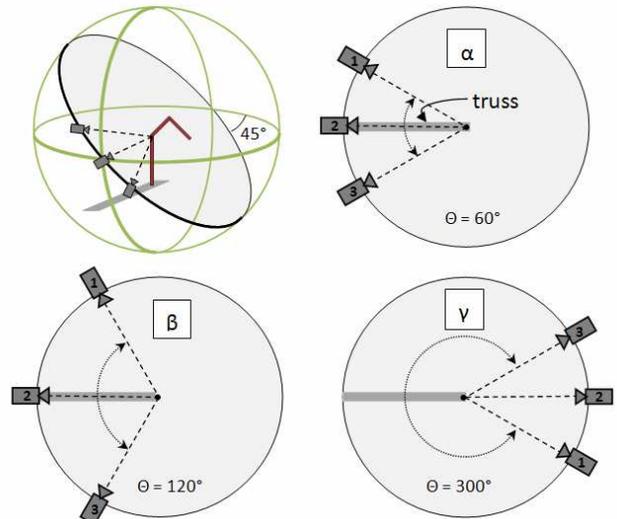


Figure 3. The three cameras are positioned along a circumference (black contour) tilted 45° from the horizontal plane. Camera configurations α , β and γ are defined by the angular distance between cameras 1 and 3.

For this experiment, three configurations (α , β , γ) were created, each consisting of three camera views. In all three configurations, the three cameras were located at the same distance from and pointed toward the shoulder joint of the robot arm. The central camera (#2) was placed just above the truss and was pointed in a direction along the truss. The other two cameras (#1 and #3) were placed along a line tilted 45° from the XZ plane (Figure 3). These two cameras were separated by 60° (measured from the shoulder joint) in camera configuration α and 120° in camera configuration β . The third camera configuration, γ , was the same as configuration α but rotated 180° about the base of the robot. This rotation produced similar camera views as configuration α but with a left-right reversal of the scene.

2.2 Task

On each trial, subjects had to manipulate the arm to capture the module then dock it onto the node. The module was automatically captured when the end effector touched any point of the capture port (Figure 4, left). Docking the module to the node required the subject to align and overlap both docking ports and press the space bar when the final position of the module was considered to be properly aligned. The ideal docking position was defined as when both ports were coaxial and in contact. The axial separation between the docking ports was not considered to be an accuracy factor.

The initial pose of the arm as well as the position and orientation of the node was the same in each trial. The cargo module was initially located in one of four possible positions (location and orientation). Each starting position was used four times, resulting in 16 trials for one camera configuration. The four initial positions were distributed in the workspace to balance their relative locations with respect to the base of the arm (right/left, front/back, up/down).

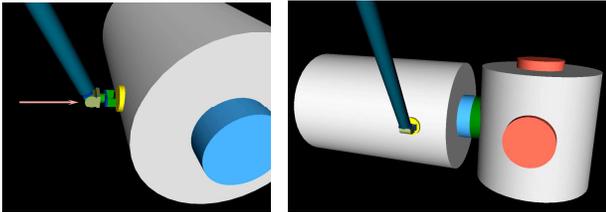


Figure 4. Capture of the module (left) and docking with the node (right)

Before starting the trials, subjects were reminded to avoid collisions between the arm and any of the structural elements in the space (e.g. node, truss) and singularities. No feedback was given to the subjects in the case of a collision and the number and type of collision was not collected at the time of the experiment. Subjects were also reminded to avoid moving the arm beyond its full extension of the arm, which would cause the simulation to crash. This condition required a restart of the simulation and the trial was repeated.

2.3 Spatial Ability Metrics

Spatial ability of the subjects was measured using the Cube Comparisons (CC) test, the Perspective-Taking Ability (PTA) test, and the Purdue Spatial Visualization of Views test (PSVV). Subjects were shown pairs of labeled cubes and asked if the cubes could be identical. The test required mentally rotating one of the cubes to make the comparison. It was completed in 2 three-minute sessions, each with 21 pairs of cubes. Test scores were calculated as the number of correct answers minus incorrect answers. The PTA test was administered on a Windows PC. Subjects were shown a top-down plan view of seven objects distributed in a circular space; they were instructed to imagine they were visualizing one of those objects, and to imagine the object array from that perspective. They then had to indicate the direction to another target object in the array in their local reference frame. The test consisted of 58 trials and scores were based on direction accuracy and response time. In the PSVV test, subjects were

shown a three dimensional object at the center of a “glass” cube. The task was to determine which one of five alternative views corresponded to the designated viewpoint, shown by a black dot at a specific corner of the cube. The original test is self-paced, however, a five-minute constraint was set for this experiment, in order to assign some weight to the response time, and avoid the development of strategies different to perspective-taking. The test had 30 trials and the scores were calculated as the number of correct answers minus one fifth of the number of incorrect answers.

2.4 Task Performance Metrics

The data acquired from each trial included instantaneous location and orientation of the end effector, time to capture (t_0 coincided with the beginning of the trial), and time to dock. From these data, we calculated the following performance metrics listed in Table 1. These metrics were selected from a larger set described in [1] that characterized operator performance in the BORIS training system.

Table 1. A description of the task performance metrics used in the experiment.

Metric	Description
Observation time (t_{obs})	Time between t_0 and the first hand controller input
Total time (t_{total})	Time required to complete the task
%Motion	Percent of t_{task} during which the end effector was moving
Axial DOF input (DOF_{Ax})	Average of simultaneous use of axial degrees of freedom (DOF) during t_{task} ($DOF_{Ax}=1$ if the subject never moved on more than one axis at a time; $DOF_{Ax}=3$ if the subject moved on the three axes every time throughout the task)
Angular DOF input (DOF_{Ang})	Average of simultaneous use of angular DOF during t_{task}
Docking-position offset (YZ offset)	Distance between the axes of the two docking ports (Figure 5, top)
Docking-attitude offset (angular offset)	Angle between the axes of the two docking ports (Figure 5, bottom)

2.5 Subjects

Seven subjects, three female and four male, participated in the experiment. Their ages ranged from 23 to 33, and all had an Aerospace Engineering background. On average, they spent the same daily amount of time in front of the computer (7.6 hrs) and on videogames (0.5 hrs), however their past gaming experience varied from 2 to 10 hours a week. They did not receive any monetary compensation.

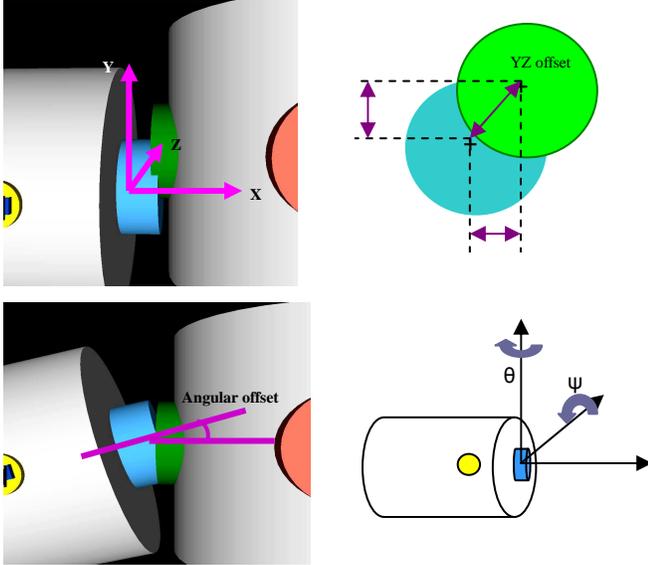


Figure 5. Performance metrics: (Top) docking position offset (YZ offset) and (Bottom) docking angular offset ($\Theta+\Psi$)

2.6 Procedure

The experiment involved three sessions completed on separate days. During the first session (one hour maximum), subjects were given a questionnaire to obtain data such as gender, age, background, previous gaming experience, and current use of computer. Then subjects completed the Cube Comparisons test followed by the Perspective-Taking Ability test. The second session (from two and a half hours on average, four hours in one case) started with a Powerpoint presentation providing theoretical training about the main elements of the simulation (e.g. the hand controllers) and the instructions for the experiment task. After the theoretical training, the subjects completed a practical training session where they performed similar tasks to the ones in the experiment, but with a different set of objects than the module and the node. The eight training trials were designed to help the participants learn how to manipulate the arm, to capture objects, and to dock them. No feedback was provided during training. Subjects were instructed to do the task as fast and accurately as possible, avoiding any kind of collisions or singularities. Astronauts in actual robotics training are similarly evaluated on their ability to avoid arm singularities [J. Young, personal communication].

After the training, subjects began the main experiment. A total of 48 trials was divided into three blocks of 16 trials. Within each block, only a single camera configuration was used and four repetitions of each starting location of the module were completed. The order of starting locations of the module was balanced to reduce any effects of order. The subjects were randomly divided into two groups to examine any effects of the order in which the camera configurations were seen. Group A (n=4, 2m, 2f) performed the experiment using the configurations in the order α - β - γ , whereas Group B (n=3, 2m, 1f) followed the sequence α - γ - β . Subjects were allowed a short break between blocks. Finally, a post-experiment questionnaire was given to assess the possible discomforts caused by the test, and to get feedback on the subjects' strategies to perform the tasks.

During a third session, subjects completed the Purdue Spatial Visualization test. This test was given in a separate session because it was added after the main experiments had been completed.

3. RESULTS

3.1 Spatial Ability Test Scores

The Cube Comparisons test (CC) scores ranged from 23 to 42 (mean: 34.43, SD: 7.18). However, these scores were clearly separable into two groups. Three subjects' scores were above the mean (40 to 42) and four subjects' scores were below (23 to 33). Interestingly, all the female subjects were in the low-scoring group while three of the four male subjects were in the high-scoring category. This result is consistent with previous findings that women's spatial visualization abilities are weaker than men's [14].

The scores for the Purdue Spatial Visualizations test ranged from 7.2 to 23.4 (mean: 17.29, SD: 6.82). These scores were bimodally distributed with two subjects scoring very low (7.2 and 7.8) and five subjects scoring above the mean.

The scores for the Perspective-Taking Ability test ranged from 20.6 to 27.4 (mean: 23.11, SD: 2.59). The scores were unimodally distributed about the mean. No obvious difference between genders was apparent for PSVV or PTA.

Two male subjects had high scores for all three tests, whereas one female scored below the mean on all three tests. No significant correlation was found between CC and PTA scores ($R^2 = 0.022$), or between CC and PSVV scores ($R^2=0.0831$). PTA and PSVV scores were found to be correlated with $R^2=0.5940$.

3.2 Task Performance Results

We used a mixed regression model (Systat v.11) to statistically analyze the relationship between spatial scores and task performance. The fixed effects considered were: test score, camera configuration (α , β , γ), group (A, B), the cross effect of group and camera configuration. The only random effect was subject. No effects were found to be significant due to age, daily hours of computer use nor daily hours of gaming. For the regression analysis, in order to fit the model assumptions of comparable variance and normal distribution of the residuals, the time and accuracy data were transformed to their logarithms, the %Motion and DOF data were transformed by $\arcsin(\sqrt{\cdot})$, transformation often used for percentages [18].

The data show an effect of the camera configurations on the total time for a trial (t_{total}), observation time (t_{obs}), %Motion and axial degrees of freedom (DOF_{ax}). On average, t_{total} was lower while using configuration β (cameras 1 and 3 separated by 120°) than with either configuration α (60° separation) or γ (60° "left-right reversed"). This pattern is consistent for both the time to capture the module and the time to dock it to the node. The average t_{obs} needed by the subjects in both groups remained approximately constant between configurations α and β , and increased on configuration γ . The high t_{total} and t_{obs} obtained for configuration γ was an expected result, given the left-right reversal in the camera views. %Motion and DOF_{ax} were higher for configuration β . The order of the camera configurations in which the subjects completed the task did not significantly affect their performance.

PTA scores did not show any effect on performance different from PSVV; this is not surprising given the narrow distribution of test scores with our limited set of subjects. The data were consequently grouped by the subjects' CC and PSVV test scores into high scoring and low scoring groups, relative to the means. When grouped by CC scores (Figure 6, left), the average t_{total} and t_{obs} (not shown) for the high scoring group are statistically shorter than for the low scoring group across all three camera configurations (t_{total} : $p=0.007$; t_{obs} : $p=0.003$). Even if not significant, differences between the CC score groups are also evident for the %Motion and angular docking offset measures, although not consistently across the camera configurations. The high scoring group generally kept the arm in motion during a greater portion of the trial (configurations β and γ) and showed better angular docking performance except in configuration γ . No differences between the CC score groups were evident for the remaining performance metrics.

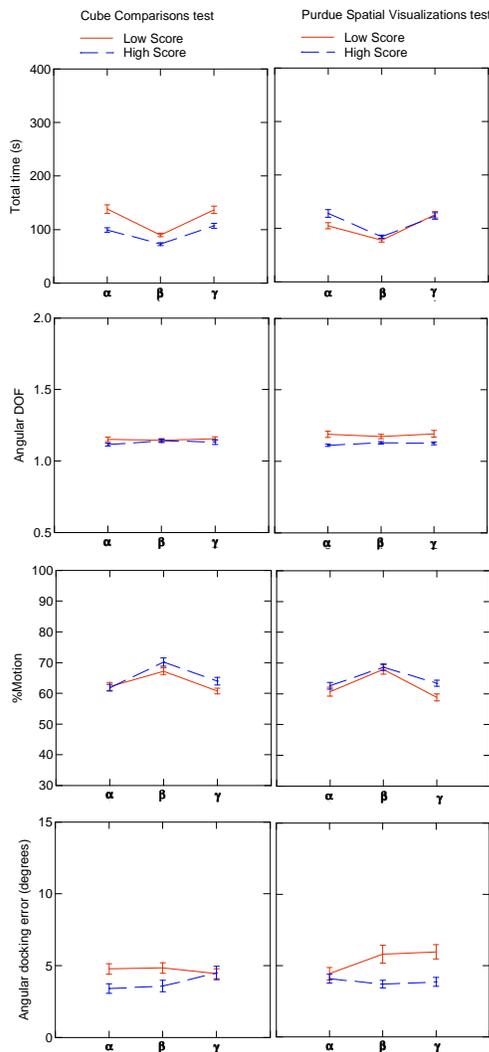


Figure 6. Average task performance per trial grouped by subjects scoring low (solid lines) or high (dashed lines) on the Cube Comparison and Purdue Spatial Visualization tests.

When task performance is grouped by PSVV scores (Figure 6, right), a clear difference in the average DOF_{ang} score is apparent across the three camera configurations. Surprisingly, the high PSVV score group had a significantly lower DOF_{ang} score ($p=0.002$), indicating that they tended to rotate the end effector around one axis at a time. The high PSVV score group also showed longer t_{total} in configuration α , higher %Motion in configuration γ , and showed better angular docking performance for configurations β and γ . Differences between PSVV score groups were not evident for the remaining performance measures.

Comparing task performance for each repetition of a trial provided some insight into the learning curve for the subjects. There was a significant learning trend across repetitions, with decreasing t_{total} and t_{obs} , and increasing %Motion and DOF_{ax} ($p=0.0005$, all), the effect being similar for both high and low CC score groups. Low scoring subjects on the PSVV test showed a big decrease in angular accuracy, a %Motion decrease and a slight YZ offset increase in the two last trial repetitions for the last two repetitions. Performance for other metrics show very small changes over repetitions.

3.3 Post-experiment Questionnaire Summary

Subjects reported low levels of mental fatigue (1.8 out of 5), disorientation (1.5), and eye strain (2.5) despite the length of the experiment. They reported that translating the end effector was generally intuitive, but they often resorted to trial and error to determine how to control the rotation. A general task strategy (five subjects) was to move in only one axis at a time and mainly use the central camera (#2) for general movement control and the other two cameras for depth information and final alignment. Subjects moved the arm slowly and tried to watch all three views to avoid collisions, although two subjects reported ignoring between the module and arm.

Finally, subjects were requested to describe and sketch the location of the cameras in the three configurations. One subject could not recall any of the camera configurations. All other subjects correctly recalled configuration γ , but only one subject correctly identified a difference between configurations α and β . They also properly described the central position of camera 2. Three subjects correctly matched the right/left location of cameras #1 and #3 in configuration γ , and two of them in α . Finally, only one subject mentioned the up/down location of cameras 1 and 3, but incorrectly recalled that both cameras were above camera 2.

4. DISCUSSION AND FUTURE WORK

Our primary goal was to determine if a correlation between human spatial abilities and space teleoperation performance could be established. Our data does suggest a link between spatial ability and performance, and also provides some important insights into the separate aspects of the tasks and spatial tests. However, with only seven subjects, we must be cautious of the conclusions drawn from the study. For example, the distribution of subjects' PTA scores was markedly different from the bimodal distributions of the CC and PSVV scores. The difference could simply be due to the small number of subjects and nature of our subject population, although it could be reasonably argued that the aerospace background of the subjects is comparable to many current astronauts. Another possible explanation is based on the nature of the spatial ability being tested. The PTA test is performed in a two dimensional space, which is typical for Earth-

bound navigation, but less representative of the 3D environments in space teleoperation. Since we all have lifelong experience with 2D navigation, differences in this type of perspective-taking are much smaller across the population. Subjects may utilize other learned strategies from their experience to perform the task – a common problem with spatial tests. The PSVV test, in contrast, tests perspective-taking in three dimensions, which is probably a less frequently utilized skill so scores may reflect an individual’s innate spatial ability. Further experiments with a larger subject population are clearly needed.

Our data indicate that CC, PTA and PSVV test scores may be predictive of performance measured by the total and observation times, but this was not true for the docking accuracy or DOF metrics. The Perspective-Taking Ability test scores were found to be significantly correlated to the use of angular degrees-of-freedom. A trend for higher PSVV scorers to require longer total and observation times, as well as to use less angular DOF and higher %Motion was also identified. The fact that the two spatial abilities do not present significant effects on the same performance metrics is somewhat surprising since Hegarty and Waller [9] found perspective-taking and mental rotation abilities to be highly correlated, while also dissociable.

The fact that high CC test scorers required shorter total and observation times to perform the task whereas high perspective-taking ability scorers required longer times suggests that subjects with better perspective-taking ability dedicate more time to analyze the workspace from the multiple viewpoints before manipulating the robotic arm. Subjects with weaker perspective-taking ability would have more difficulty integrating the viewpoints, so like the subjects in [16], they might use only one display and simply begin moving the arm.

Subjects did not change their use of angular degrees of freedom throughout the experiment, which could be explained by their minimal training with the rotational hand controller and resulting lack of mastery. This is supported by the fact that many subjects reported using trial and error as their main strategy to rotate the end-effector. Astronauts spend over 30 hours for basic robotic arm training but this is also not practical. Spending a preceding day solely on training, such as in [1] could be sufficient to train subjects performance to a sufficient level. It is interesting to note, however, that high perspective-taking scorers consistently used single axis rotation movements more often than the low scorers. It could reflect a trial-and-error strategy for the low PSVV score subjects, or their greater lack of proficiency controlling the arm. The measurement of arm-object collisions or direction reversals in future experiments could help distinguish these possibilities.

The different experiment camera configurations were chosen on the assumption that subjects with better perspective-taking ability would be better able to integrate highly disparate views and thus complete the task faster and more accurately, but this effect was not found. Quite possibly the camera views in configuration β (120° separation) were not sufficiently different to force the operator to use their perspective-taking ability, so no effect was seen. Instead, operators might have simply relied on mental rotations of the objects to understand the relationship between the three displays. The suggestion is supported by the significant effect of CC test score on total time, and by the fact that subjects could not report the particular camera locations of configuration β . Thus the improvement in total time or %Motion performance

for this camera setting was simply a learning effect. The similarity of the effect between the two CC score groups suggests that the effect of mental rotation ability was the same for all three camera configurations. For our next series of experiments, we plan to investigate camera configurations that are more widely dispersed around the working environment. An experiment examining operator performance when selecting appropriate camera views might be more likely to show a correlation of performance with perspective-taking ability tests. As mentioned in the Introduction, this is one of the criteria used to evaluate operator performance during RMS training.

Angular docking accuracy was higher for subjects with superior spatial abilities, and it decreased from configuration α to configurations β and γ for poor perspective-takers. This could be explained by a speed-accuracy trade-off that also led to a decrease in %Motion score and a slight YZ offset increase in the two last trial repetitions. Results of angular docking error between configurations suggest that with camera configurations, such as α , which do not provide much depth information and are relatively simple to understand, only mental rotation ability is used to perform the docking. With a configuration that provides more depth cues (β), either mental rotation or perspective-taking can be utilized; finally, with a hard camera setting, such as γ , subjects only rely on their perspective-taking ability. Although this result need to be confirmed with a greater pool of subjects, it may imply that operators would need strength in both spatial abilities in order to guarantee high accuracy in teleoperation involving a wide set of cameras.

Further research to understand the spatial skills that underlie teleoperation task performance could be helpful in improving current training procedures for astronauts. Training programs could be individualized according to their spatial skill set, and overall training time might even be shortened. Knowledge of the mechanisms that support spatial reasoning could also have a direct impact on improving the design of interfaces for human-robotic interaction. For example, Trafton, Cassamatis, Bugajska, Brock, Mintz and Schulz [18] applied the concept of perspective-taking to improve the interaction with autonomous robots, such as NASA’s Robonaut. Understanding individual differences in spatial ability may suggest guidelines or new methods for displays that can be customized to support the spatial abilities of users, or to lead to insights that improve the training methods for robotic systems.

5. CONCLUSIONS

This study investigated the influence of spatial abilities on space teleoperation performance. We have identified a number of task performance metrics that seem to be correlated with our chosen measures of both mental rotation and perspective-taking ability. More specifically, the total task time and observation time were clearly correlated with mental rotation ability, but were inversely correlated with perspective-taking ability. Rotational control behavior of the end-effector also seems to change according to perspective-taking ability, with high scoring subjects exhibiting more single axis control movements. Finally, subjects with higher spatial ability seemed to keep the arm in motion for a higher percentage of time during the task, perhaps reflecting better awareness of the robotic arm in the workspace. Further studies involving a larger subject population are needed to confirm these results.

6. ACKNOWLEDGEMENTS

The authors thank Hiro Aoki, Jennifer Young (NASA Johnson Space Center), and NASA astronauts Jeff Hoffman, Koichi Wakata, Steve Robinson and the HRI reviewers for helpful comments. This work was supported by the National Space Biomedical Research Institute through NASA Cooperative Agreement NCC 9-58, and the National Council for Science and Technology of Mexico, CONACyT.

7. REFERENCES

- [1] Akagi, T.M., Schlegel, R.E., Shehab, R.L., Gilliland, K., Fry, T.L. and Hughes, Q., Towards the Construction of an Efficient Set of Robot Arm Operator Performance Metrics. In *Human Factors and Ergonomics Society 48th Annual Meeting*, (New Orleans, LA2004), 1194-1198.
- [2] Carroll, J.P. Human cognitive abilities: a survey of factor-analytical studies. Cambridge University Press, New York, 1993.
- [3] Contreras, M.J., Colom, R., Hernandez, J.M. and Santacreu, J. Is Static Spatial Performance Distinguishable From Dynamic Spatial Performance? A Latent-Variable Analysis. *The Journal of General Psychology*, 130, 3 (2003), 277-288.
- [4] DeJong, B.P., Colgate, J.E. and Preskin, M.A., Improving Teleoperation: Reducing Mental Rotations and Translations. In *IEEE Intl. Conf. on Robotics and Automation*, (New Orleans, LA, April 26 - May 1, 2004), 3708-3714.
- [5] Ekstrom, R.B., French, J.W. and Hartman, H.H. Cognitive Factors: Their Identification and Replication. In *MBR Monograph*, Society of Multivariate Experimental Psychology, 1979.
- [6] Ekstrom, R.B., French, J.W. and Hartman, H.H. *Manual for kit of factor referenced cognitive tests*. Educational Testing Service, Princeton, NJ, 1976.
- [7] Eyal, R. and Tendick, F. Spatial Ability and Learning the Use of an Angled Laparoscope in a Virtual Environment. In Westwood, J.D. and al., e. eds. *Medicine Meets Virtual Reality 2001*, IOS Press, Amsterdam, 2001, 146-152.
- [8] Guay, R. Purdue Spatial Visualization Test - Visualization of Views, Purdue Research Foundation, West Lafayette, IN, 1977.
- [9] Hegarty, M. and Waller, D. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32, 2004 (2004), 175-191.
- [10] Kozhevnikov, M. and Hegarty, M. A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29, 5 (2001), 745-756.
- [11] Lamb, P. and Owen, D., Human Performance in Space Telerobotic Manipulation. In *ACM Symposium on Virtual Reality Software and Technology*, (Monterey, CA, November 7-9, 2005), ACM, 31-37.
- [12] Lapointe, J.F., Dupuis, E., Hartman, L. and Gillett, R., An Analysis of Low-Earth Orbit Space Operations. In *Proceedings of the Joint Association of Canadian Ergonomists/Applied Ergonomics (ACE-AE) Conference*, (Banff, Alberta, Canada, 21-23 Oct, 2002).
- [13] Lathan, C.E. and Tracey, M. The Effects of Operator Spatial Perception and Sensory Feedback on Human-Robot Teleoperation Performance. *Presence: Teleoperators and Virtual Environments*, 11, 4 (2002), 368-377.
- [14] Linn, M.C. and Petersen, A.C. Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Development*, 56, 6 (1985), 1479-1498.
- [15] Pellegrino, J.W., Hunt, E.B., Abate, R. and Farr, S. A computer-based test battery for the assessment of static and dynamic spatial reasoning abilities. *Behavior Research Methods, Instrumentation, & Computers*, 19, 2 (1987), 231-236.
- [16] Spain, E.H. and Holzhausen, K.-P., Stereo Versus Orthogonal View Displays for Performance of a Remote Manipulator Task. In *Stereoscopic Displays and Applications II*, (1991), SPIE, 103-110.
- [17] Tracey, M.R. and Lathan, C.E. The Interaction of Spatial Ability and Motor Learning in the Transfer of Training From a Simulator to a Real Task. In Westwood, J.D. ed. *Medicine Meets Virtual Reality 2001*, IOS Press, Amsterdam, 2001, 521-527.
- [18] Trafton, J.G., Cassamatis, N.L., Bugajska, M.D., Brock, D.P., Mintz, F.E. and Schulz, A.C. Enabling Effective Human-Robot Interaction Using Perspective-Taking in Robots. *IEEE Transactions on Systems, Man, and Cybernetics, Pt. A*, 35, 4 (2005), 460-470.
- [19] Wilkinson, L., Blank, G. and Gruber, C. *Desktop Data Analysis with SYSTAT*. Prentice-Hall, Inc., Upper Saddle River, NJ, 1996.