Comparing the Performance, Workload, and Usability of a Gamepad and Joystick in a Complex Task

Michael A. Rupp, Paul Oppold, & Daniel S. McConnell
University of Central Florida

Complex human system interface systems such as unmanned vehicles utilize controls that range from standard joystick and keyboard interfaces to Xbox controllers. However, few research studies have been conducted to compare Xbox controllers with other types of interfaces, showing mixed results between those controllers. The current study compared the performance of a joystick and keyboard interface with that of an Xbox controller in both a low and high difficulty task. The results indicate that the Xbox controller had lower tracking errors and trended to lower workload and a slightly higher usability score as measured by the system usability scale (SUS). An interaction between difficulty and controller was not found, however allowing for a longer practice time with the interfaces may have shown significant differences which are planned in a future study. Overall these results indicate favorable results that an Xbox controller may be a viable control interface for complex human system interaction tasks.

 DOI 10.1177/1541931213571398

INTRODUCTION

In addition to automation, remote operation of complex systems has become characteristic of modern technology including telerobotics, telesurgery, and remotely operated vehicles such as the Mars rover and unmanned vehicles for military and surveillance use (DoD, 2009; Fong & Thorpe, 2001; Neumann & Durlach, 2006).

Despite the attention given to the technology, the human operator remains a critical component of remotely operated systems (Fong & Thorpe, 2001; Mouloua, Gilson, & Hancock, 2003). One important factor in this human-system interaction is the input device used for control inputs via an interface. Poor design can affect performance, create high cognitive workload, and increase errors, all of which may contribute to the loss of the vehicle (Hancock, Mouloua, Gilson, Szalma, & Oron-Gilad, 2007; Mouloua et al, 2003; Williams, 2004). Input devices can range from the keyboard and mouse, joysticks, touchscreens, a variety of unique proprietary devices, and of more recent interest, video game controllers (Axe & Olexa, 2008; Lenz, Chaparro, & Chaparro, 2008; Neumann, 2006; Rackliffe, 2005; Shively, Brasil, & Flaherty, 2007).

Compared to other devices, gamepads offer some specific advantages for portability, adaptability to the task, prior exposure for new operators, and usability to the overall population (Axe & Olexa, 2008). They are small in size, lightweight and comfortable, easily transported, they are handheld and do not require a desk or surface to sit upon (Pettitt, Redden, Pacis, & Carstens, 2008), and can be operated using haptic feedback without looking at the controller (Pettitt, Carstens, & Redden, 2012). The game controllers contain an assortment of buttons, triggers, and thumb-operated joysticks that can be programmed to perform a variety of tasks at varying sensitivities for the thumb operated joysticks, which could reduce the difficulty in learning, perceived workload, and overall task times, when compared to switching the functions of multifunction controllers (Pettitt et. al, 2008), and Individuals with prior experience using the controller can achieve a high level of performance using them. Many unmanned vehicle pilots are young and have high levels of video game play and experience with controllers (Shively, et al., 2007; Pettitt et al., 2012). Pettitt et. al. (2010) found that soldiers using a wireless Xbox 360 controller found it to be easy to learn, easy to use, and familiar based on large amounts of time previously spent using an Xbox 360. An implication of this familiarity is that the game controller not only benefits the physical aspects of its operation, but also the cognitive aspects. Familiarity may reduce the working memory demands involved in remembering the location and function of controls. The military is hoping to leverage this familiarity to reduce training time and increase performance in operating unmanned vehicles (Shively et al., 2007).

Neumann and Durlach (2006) compared a gamepad to a mouse in a simulated remote vehicle control task and showed that the gamepad resulted in better performance and lowered workload. However, other studies have indicated an increase in workload using gaming controllers, specifically for pilots (Shively et al., 2007). In driving tasks the controller was preferred for reconnaissance course maneuver tasks, manual control of the robot, and close maneuvering around obstacles, but not for driving multiple waypoints, simultaneous radio tasks, and simultaneous camera arm operation while driving (Pettitt et. al, 2012). Using shooting-based video games as a model for tracking tasks, other studies have found that gamepads result in either poor performance (Isokoski & Martin, 2007) or no difference (Lenz, Chaparro, & Chaparro, 2008) compared to other devices such as a mouse or joystick.

In a simplified telerobotics posture control task, Guo and Sharlin (2008) reported a benefit of the Wiimote game controller compared to a keyboard. Despite these mixed results, unmanned vehicles have been controlled in the field with Xbox controllers (Axe & Olexa, 2008). This reality mandates further study of workload and performance issues of game controllers in complex human-system interface tasks. Due to the cognitively demanding nature of remote operation and other human system interaction, it is important to conduct further usability and performance testing of these devices.

METHODS

Participants

Thirty-six right-handed participants (20 female and 15 male) between the ages of 18-22 with normal color vision and visual acuity (both near and far) were recruited from a large university in the southern United States to participate in this study. One participant was removed due to having a Root Mean Square Deviation from Center (RMSD-C) of over 4 standard deviations above the mean, leaving a total of 35 participants. They received course credit in exchange for participation. All participants also reported no motor impairments.

Materials and Design

The Multi-Attribute Task Battery (MATB; Comstock & Arnegard, 1992) is a task designed to simulate several aspects of a flight task. The system uses several tasks to create a dynamic workload environment including system monitoring, resource management, and tracking. The system monitoring task simulates the demands of monitoring gauges (perceptual task). In this task, individuals have to monitor for both the absence of a green light and the presence of a red light, both, which indicate deviations from normal system control. When a deviation is observed, participants have to press a button to reset the system (Comstock & Arnegard, 1992). The compensatory tracking task simulates control of manual flight. Participants have to keep a pointer as close to the center marker as possible. The resource management task simulates a fuel tank and pump system. The user can turn multiple pumps on and off to direct fuel flow through the system. The goal of this task is to keep two main fuel tanks at an optimal level while balancing the demands of variable fuel flow of the pumps and errors that may arise.

In a fully automated (low difficulty) condition, all tasks except the tracking task were controlled by the computer. Participants only had to attend to the tracking task. This condition allowed an assessment of the psychomotor control of the pointer. In the non-automated (high difficulty) condition, participants had to perform the tracking task while also monitoring and responding to the system monitoring and resource management tasks. This condition allowed for the assessment of psychomotor control in the tracking task while under high cognitive load from secondary tasks.

The study was a 2 × 2 between subjects factorial design. The two levels of task difficulty were combined with two input devices. Participants used either a Microsoft Sidewinder II joystick and a Dell standard 110-key keyboard combination or a wired Microsoft Xbox controller.

In the joystick and keyboard condition, the default button mapping for MATB was used. In this configuration, participants completed the tracking task using the joystick, the F5 and F6 button controlled the system monitoring task, and the number buttons 1-8 controlled the fuel pumps for the resource management task.

In the Xbox controller, buttons on the interface were mapped on to functions as follows. The left thumb stick controlled the tracking task, buttons A, X, Y, B controlled pumps 1-4, the left shoulder bumper buttons controlled pumps 5 and 6, the shoulder trigger buttons controlled pumps 7 and 8 for the fuel management task, and the left and right arrow buttons controlled F5 and F6 of the system monitoring task. The Xbox controls were selected to maintained spatial compatibility between the controls and the actions they generate in the simulated system (e.g., pump 7 moves fuel left to right therefore, the right trigger was used, etc…). The program was conducted on a standard PC running Windows 7.

Additionally, Participants rated their experience with keyboards, joysticks, and Xbox controllers on a Likert type scale 1-5.

Dependent Measures

The System Usability Scale (SUS) developed by Brooke (1996) was used to measure of the usability of the input device. The NASA-TLX (Hart & Staveland, 1988) was used to obtain subjective workload ratings regarding the task. Root mean square deviation from target center (RMSD-C) was sampled every 10 seconds from MATB and used as a measure of accuracy in the tracking task.

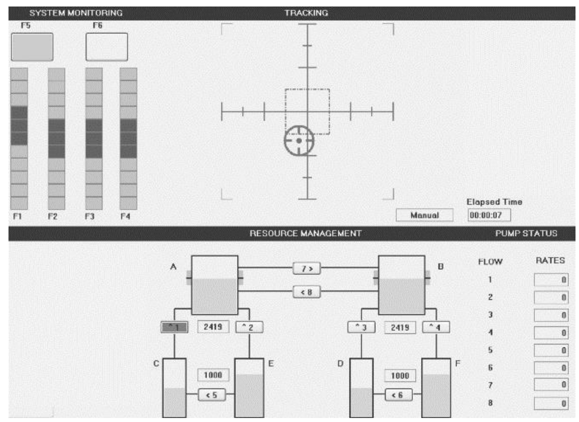

Figure 1. Picture of the MAT-B II test battery showing the experimental tasks used in the experiment. (Image retrieved and adapted from http://matb.larc.nasa.gov/tasks.php)

Procedure

Participants completed a demographics survey, and then were given a screening for visual acuity and color blindness. Following these measures, participants completed a 5-minute practice of the task followed by a 10-minute test session. Following the task participants filled out the NASA-TLX and usability survey.

Hypotheses

H1) The gamepad will result in better tracking accuracy compared to the joystick and keyboard.

H2) Users will report lower workload and better usability with gamepad compared to the joystick and keyboard.

H3) The high difficulty condition will have poorer usability ratings than the low difficulty condition.

H4) Usability and performance benefits of the gamepad will be more pronounced in the high difficulty condition.

RESULTS

Reports of experience using joysticks or Xbox controllers were not significantly correlated with participants' judgments of post task workload or usability ratings. However, there was a significant relationship between Xbox controller experience and RMSD-C in the tracking task, $r = -.46$, $p < .01$. Those who had more experience using Xbox controllers had significantly lower tracking deviations than those who had less experience.

In order to determine if the experimental groups had significantly different levels of experience with the controllers, t-tests were conducted using the controller type as the grouping variable. Participants in the Joystick and Keyboard condition and the Xbox controller conditions did not vary for experience of keyboards ($p = .076$), or Xbox controllers ($p = .58$). However, they did significantly vary for joystick experience (t (33) = -2.07, $p = .047$); indicating that participants who were assigned to the Xbox condition ($M = 2.76$, $SD = 1.03$) had slightly more joystick experience than those in the Joystick condition ($M = 2.11$, $SD = .83$) Additionally, those in the low difficulty task condition did not vary from those in the high difficulty task on levels of Xbox controller experience ($p = .12$), keyboard experience ($p = .92$) or joystick experience ($p = .26$).

The independent variables were task difficulty (high and low) and controller (joystick + keyboard vs. Xbox). A 2 x 2 between-subjects MANOVA was performed on the three dependent variables associated with the performance of the task: NASA-TLX total post task workload, RMSD-C of the tracking task, and post-task SUS score. Also, pre task workload, extraversion, and controller experience were identified as possible covariates; however, during further analysis none of the possible covariates met inclusion criteria, thus a MANCOVA was not used. Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity with no serious violations noted.

Using Wilks' Lambda criterion, the combined DVs were significantly related to the IVs, F (3, 29) = 7.59, $p = .001$, partial $\eta^2 = .44$ for task difficulty, F (3, 29) = 14.06, $p < .001$, partial $\eta^2 = .60$ for controller, however, the interaction was not significant, $p = .322$.

When the results for the dependent variables were considered separately a significant effect for RMSD-C as a function of task difficulty was found. RMSD-C was higher in the difficult task condition ($M = 36.30$, $SD = 13.70$) compared to the low difficulty task ($M = 25.78$, $SD = 7.14$), F (1, 31) = 20.48, $p < .001$, partial $\eta^2 = .40$. Post task workload measures were marginally higher in the high difficulty condition ($M = 67.30$, $SD = 15.40$) than the low difficulty condition ($M = 55.14$, $SD = 22.56$), F (1, 31) = 3.43, $p = .074$, partial $\eta^2 = .10$. Subjective reports of usability did not significantly differ between the task difficulty conditions ($p = .31$).

Manipulation of input device resulted in a higher RMSD-C for the joystick ($M = 38.87$, $SD = 11.71$) compared to the Xbox controller ($M = 23.06$, $SD = 5.32$), F (1, 31) = 44.56, $p < .001$, partial $\eta^2 = .59$. There were no other main effects of input device on the other dependent variables ($p = .43$ for workload and $p = .31$ for SUS score).

There were no significant interactions between automation and controller, except for a marginal interaction on the RMSD-C, F (1, 31) = 3.43, $p = .074$, partial $\eta^2 = .10$.

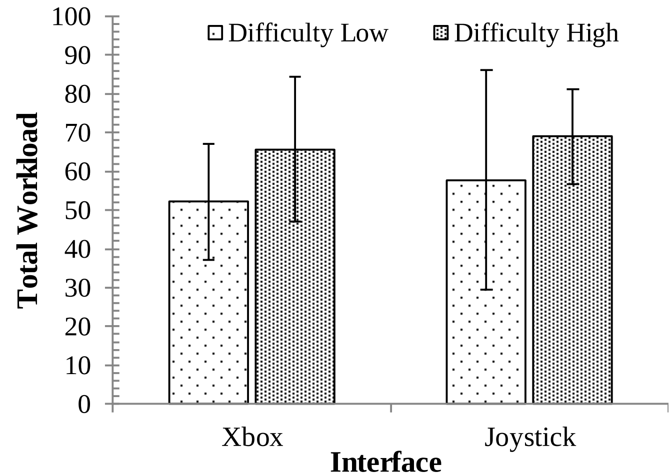

Figure 2. Total post task workload is displayed for all conditions. Error bars are standard deviations.

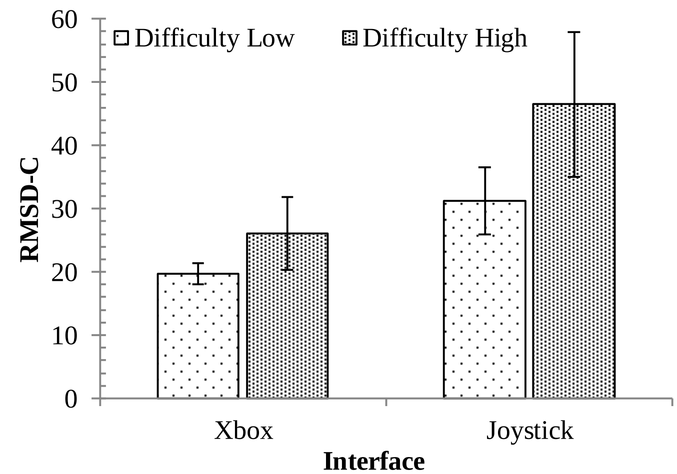

Figure 3. The root mean square deviation from center for the tracking task is displayed. Error bars are standard deviations.

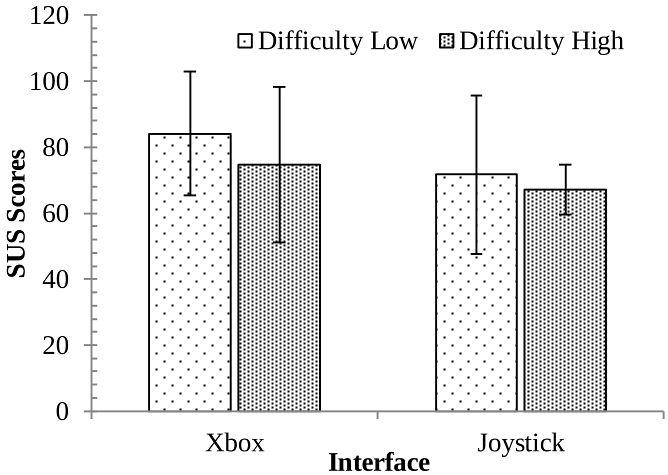

Figure 4. The mean scores for the System Usability Scale are shown. Error bars are standard deviations.

DISCUSSION

Hypothesis 1 was supported. The Xbox controller overall outperformed the Joystick for tracking accuracy, as well as participants in the low difficulty task committed fewer errors that the in the harder task condition; however, the other hypotheses were not fully supported. Hypotheses 2, 3, and 4 were not supported significantly by the data. Figure 1 shows slightly lower workload and Figure 3 shows that the SUS scores were slightly higher in the Xbox conditions; however these differences were not significant. No interactions were found between the task difficulty and interface, which was hypothesized by hypothesis 4.

Why were only significant results found for tracking accuracy and marginally for workload? One possible explanation is that although the Xbox controller matched the task better than the joystick, individuals' experienced similar amounts of workload and preferences for the devices. If this explanation is correct the finding of improved performance for the Xbox controller still supports the use of these controllers in the field.

It is also possible that the design of the study was not sensitive to the between group differences of the different interface devices. Participants only interacted with one device and one task; thus, they were not able to compare the devices against each other. Allowing particpants to interact with both devices may have yielded more sensitive results.

Additionally, participants were only given limited amount of time to practice and interact with the devices prior to measuring performance. A design that allowed additional interaction time to minimize practice effects would have contained more power to identify differences between the other measures.

Future Studies

Future studies are planned that include adding stressors such as noise and vibration to simulate the complex environments in which these controls are sometimes used. Additionally, we are pursuing studies using other psychomotor tasks such as the Fitts' pointing task (Fitts 1954; Fitts & Peterson, 1964) to compare the bandwidth of the devices (e.g., MacKenzie, 1992) as well as examine the separate physical and visual characteristics of the pointing task (e.g., Bohan, McConnell, Chaparro, & Thompson, 2010).

The current study fails to identify the exact nature of any performance benefit from the Xbox controller compared to the joystick and keyboard devices. Familiarity and experience with the devices may account for some group differences, though our analyses suggested that they did not. While matched samples can address this problem, we plan to conduct future studies designed to measure the experience factor more directly. If participants are given a chance to practice the task with a given device for a period sufficient to reach a performance plateau, then any remaining differences in performance or usability metrics may be attributable to differences inherent to the device. Further, time taken to reach plateau can provide additional insight on device utility.

In summary, the current data shed favorable light on the hypothesis that gamepads are useful for complex human-system interaction tasks, although the full extent of the benefit has yet to be identified.

REFFERENCES

Axe, D. & Olexa (2008). *War Bots: How US Military Robots Are Transforming War in Iraq, Afghanistan, and the Future*: Nimble Books LLC.

Bohan, M., McConnell, D.S., Chaparro, A., & Thompson, S.G. (2010). The effects of visual magnification and physical movement scale on the manipulation of a tool with indirect vision. *Journal of Experimental Psychology: Applied, 16*(1), 33.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry, 189*, 194.

Comstock, J. R., Jr., & Arnegard, R. J. (1992) *The Multi-Attribute Task Battery for human operator workload andstrategic behavior research* (NASA TM-104174). Hampton, Virginia: NASA Langley Research Center.

Department of Defense. (2009). *FY2009-2034. Unmanned Systems Integrated Roadmap* (ADA522247). Washington, DC: Department of Defense.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381.

Fitts, P. M., Peterson, J. R., & Wolpe, G. (1963). Cognitive aspects of information processing: II. Adjustments to stimulus redundancy. *Journal of Experimental Psychology, 65*(5), 423.

Fong, T., & Thorpe, C. (2001). Vehicle teleoperation interfaces. *AutonomousRrobots, 11*(1), 9-18.

Guo, C., & Sharlin, E. (2008). *Exploring the use of tangible user interfaces for human-robot interaction: a comparative study.* Paper presented at the Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems.

Hancock, PA, Mouloua, M., Gilson, R., Szalma, J., & Oron-Gilad, T. (2007). Provocation: Is the UAV control ratio the right question? *Ergonomics in Design, 15*(1), 7.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human Mental Workload, 1*, 139-183.

Isokoski, P., & Martin, B. (2007). Performance of input devices in FPS target acquisition. *Proceedings of the International Conference on Advances in Computer Entertainment Technology ACE 07,* 240. ACM Press..

Lenz, K.M., Chaparro, A., & Chaparro, B.S. (2008). The Effect of Input Device on First-Person Shooter Target Acquisition. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 52*(19), 1565-1569.

MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction, 7*(1), 91-139.

Neumann, J., & Durlach, P.J. (2006). Effects of interface design and input control method on unmanned aerial system operator performance. *The*

Interservice/Industry Training, Simulation & Education Conference, I/ITSEC, 2882, Orlando, FL.

Mouloua, M., Gilson, R., & Hancock, P. (2003). Human-centered design of unmanned aerial vehicles. *Ergonomics in Design: The Quarterly of Human Factors Applications, 11*(1), 6-11.

Petitt, R. A., Redden, E. S., & Carstens, C. B. (2008). *Scalability of robotic controllers: an evaluation of controller options* (No. ARL-TR-4457). Army Research Lab Aberdeen Proving Ground Md.

Pettitt, R. A., Redden, E. S., Pacis, E., & Carstens, C. B. (2010). *Scalability of Robotic Controllers: Effects of Progressive Levels of Autonomy on Robotic Reconnaissance Tasks* (No. ARL-TR-5258). Army Research Lab Aberdeen Proving Ground Md.

Pettitt, R. A., Carstens, C. B., & Redden, E. S. (2012). *Scalability of Robotic Controllers: An Evaluation of Controller Options-Experiment III* (No. ARL-TR-5989). Army Research Lab Aberdeen Proving Ground Md.

Rackliffe, N. (2005). *An Augmented Virtuality Display for Improving UAV Usability*. DTIC Document.

Shively, R., Brasil, C., & Flaherty, S. (2007). *Alternative UAV sensor control: leveraging gaming skill.* Paper presented at the Proceedings of the ISAP'07 Conference.