DIVIDED ATTENTION IN USING ROBOTIC SURGICAL SYSTEMS

Webster, J.L. & Cao, C.G.L.
Department of Mechanical Engineering, Tufts University
Medford, Massachusetts

Current surgical robots require surgeons to divide their attention between performing the surgery and driving the robot. Our goal was to examine how this division of attention, as imposed by the new technology, affects the surgeon's performance and memory. We expected that combining the various sources of attentional demands on the robotic console would lead to better performance. Twelve subjects were tested using three interface designs: 1) LCD menu separated from operative site, 2) menu overlay on the operative site, and 3) voice control. Results showed that voice control was the fastest in delivering a command to the robot. Error rates and confidence levels were not significantly affected by the interfaces. This research has implications for designing future surgical robots to enhance the efficiency and safety of surgical procedures.

INTRODUCTION

Minimally invasive robotic surgery has been proposed as a solution to some of the difficulties in laparoscopic surgery. While robotic surgery offers many benefits, such as elimination of the fulcrum effect (Ben-Porat, Shoham, & Meyer, 2000) and reduced surgeon fatigue (Berguer, Forkey, & Smith, 1999), introducing a robot into the operating room changes the surgeon's tasks and responsibilities (Webster, 2004; Nio, Bemelman, Busch, Vrouenraets, & Gouma, 2004). Our comparison of a robotic procedure to a conventional laparoscopic procedure showed that instead of standing over the patient as in laparoscopic surgery, the surgeon sat at a robotic console, removed from the patient. Not only was the surgeon responsible for performing the surgery (i.e., cutting and suturing), but she was also responsible for driving the robot, which was a master-slave teleoperator. For the latter task, an LCD menu display was required, necessitating a high degree of coordination and integration of information from various sources and locations. When using the robotic system, the surgeon was visually attending to several different locations: the monitor with the operative site, the input handles, the LCD display for controlling the robot, and the operating room as a whole. Thus, the surgeon's spatial working memory was divided many ways.

With these increased attentional demands placed on the surgeon, it is important to look at how the surgeon's performance might be affected. According to the predictions of the multiple resource theory (Wickens, Vidulich, & Sandry-Garza, 1984), it would seem that performing the many spatially demanding tasks in robotic surgery would lead to the detriment of the primary task of performing surgery. Visual momentum (Hochberg & Brooks, 1978) is also an important factor in robotic surgery. Since the console's display format splits the surgeon's attention amongst the LCD menu screen, the monitor of the operative site, and the handles, visual momentum could decrease and have a negative effect on surgical performance (Woods, 1984). Much research is being done into how information can be better represented. Yeh and colleagues (Yeh, Merlo, Wickens, & Brandenburg, 2003) found that when targets are more salient or cued, head mounted displays benefit performance. Meier and colleagues (Meier, Fong, Thorpe, & Baur, 1999) also combined information from sensors to create interfaces for the teleoperation of a vehicle. They suggest that sensor fusion interfaces can help convey information more efficiently, aid in understanding remote environments, and improve situational awareness.

The goal of this study was to examine the effects of increased attentional demands as a result of using the robot for laparoscopic surgery. Since the LCD interface requires the surgeon's attention to be diverted from the surgical site, it is thought that memory of the operative site may be adversely affected. We hypothesized that by combining multiple attentional demands into fewer locations, and diversifying the types of attentional demands, memory of the operative site would improve. A controlled experiment was conducted to test the hypothesis using alternate interface designs where the LCD menu was overlaid on the operative site or where the LCD menu was replaced by voice control technology.

METHODS

Subjects

Twelve subjects (5 female, 7 male) participated in this experiment. They ranged in age from 21-33. One was left-handed, 11 were right-handed. They were not experienced users of the robotic system and were not paid for their participation.

Apparatus

Task space. A *Simulab* endoscopic training box measuring 39 cm x 39 cm x 22 cm was used to simulate the abdominal cavity. The task space was a 10 cm x 10 cm cardboard square with 10 pushpins (5 pink, 5 blue) arranged in a random pattern (see Figure 1). Nine different patterns were created. A matching set of nine task spaces was constructed, placing 10 pins in the same locations, but using only one pin color (all purple). The second set of task spaces was used for the recall test. Several 2.5 cm diameter rubber bands were provided for the pick-and-place task.

Figure 1*:* 10 cm x 10 cm task space with 10 pins.

Endoscope and monitor. A 10mm zero-degree endoscope was used to obtain a view of the task space in the box. The endoscopic view was projected onto a 20" monitor, placed at eye-level in front of the subject.

Robotic manipulator. The LaproTek Remote Robotic Surgical System (endoVia Medical, Inc.) was used. There was no force feedback available. The motion scaling on the master-slave system was 2:1.

Reading material distraction task. Subjects read aloud an excerpt from a calculus text as a distraction task.

Video camera. A video camera was used to capture the image of the robotic console's LCD menu screen, to be overlaid on the endoscopic monitor.

Video mixer. A video mixer was used to combine the image of the menu screen with the endoscopic view for display on the endoscopic monitor.

VCR. A VCR was used to record the endoscopic view with the LCD menu screen as an inset.

VoiceXML & speech recognition software. A short VoiceXML program was run through VoiceGenie phone line on the DeveloperPlatform (VoiceGenie VoiceXML Gateway Software, v.6.1 Beta).

ATT telephone. A telephone was used to transmit the subject's voice command to the speech recognizer.

Task

Subjects were asked to perform a simulated divided attention task using the robotic system. The task was to place rubber bands on all the pink pushpins in the task space and then use the menu system to disable the robot for a tool change. After an 80-second period of performing a distraction task (reading aloud from calculus text), subjects were asked to recall the previous blue pin positions. As well, subjects were asked to rate the confidence of their answers.

Procedure

Subjects sat at the robotic console and viewed the task space on a monitor located directly in front of them, 30 cm above the LCD screen. Subjects practiced using the robotic instrument to place rubber bands on pushpins until they felt comfortable (usually several minutes). The first practice task space had pink and blue pins, and then they were given a practice task space with all purple pins. The rubber bands were located along the right edge of the 10 cm x 10 cm task space. Subjects were instructed to place rubber bands as quickly as possible onto the pink pins using the graspers held by the robotic arm, with the right hand. Subjects were informed that a recall test would follow. Subjects were further told that after placing five rubber bands on the five pushpins, they were to disable the robotic tool. The tool could be disabled in one of three ways:

1. Menu condition: Using the input handles to navigate the menu on the LCD screen at the robotic console.
2. Overlay condition: Using the input handles to navigate the menu which was overlaid on the monitor of the operative site.
3. Voice condition: Using voice control by saying aloud: "Disable right tool". The word choice for this command was selected because it most directly represented the action. When the subject's command was recognized, the system responded with "Ok, disabled." There was no visual feedback.

The method of control was pre-determined for each trial. As soon as the tool was disabled, the subject received an 80 second long distraction task (reading out loud), after which they were presented with a memory recall task. A replica of the 10 cm x 10 cm space with an identical pushpin pattern was placed in the training box. Instead of two pin colors, there was only one: purple. The subject was instructed to place rubber bands on the pins that they had *not* previously put rubber bands on. After completing the recall test, subjects were asked to rate their confidence about the correctness of each pin: 5=positive, 3= somewhat sure, 1=a guess.

Experimental design

This was a within-subjects experiment; subjects disabled the tool using menu navigation on the LCD screen (Menu), menu navigation on the main monitor (Overlay), or voice command (Voice). Each of the participants completed all 3 conditions 3 times. The conditions were counterbalanced. The 9 task spaces (pushpin patterns) were randomized for each subject.

Dependent Measures

Five measures of performance were recorded: 1) time to task completion, 2) length of time to disable the tool, 3) the length of time to complete the recall test, 4) the number of errors on the recall task, and 5) confidence rating.

RESULTS

In terms of time to disable a tool, voice was fastest (9.69 sec. +/-5.49), then Menu (13.03 sec. +/- 5.54), and then Overlay (13.33 sec. +/- 5.69) (see Figure 2). These differences were statistically significant (F(2,105)=4.720, =0.011) and a post-hoc analysis showed significant differences specifically between Voice vs. Menu and Voice vs. Overlay. Average time to task completion was not significantly different among interface conditions: Overlay=78.00(+/-17.75) seconds, Voice=80.00(+/-25.23) seconds, and Menu 81.56(+/-22.78) seconds, neither was the time to complete the recall test: Menu=81.39(+/- 25.89) seconds, Voice=85.33(+/-25.68) seconds, and Overlay=85.61(+/- 27.45) seconds. Total error in the recall test varied between 0.94 (+/- 1.88) (Menu), 1.06(+/- 2.37) (Overlay), and 1.61 (+/- 2.78) (Voice), but the error as a function of interface was not statistically significant. Confidence, though not significantly different, was highest with the Overlay (4.62, +/- 0.58), closely followed by Voice (4.44, +/-0.82), and lowest with the Menu (4.34, +/- 0.86).

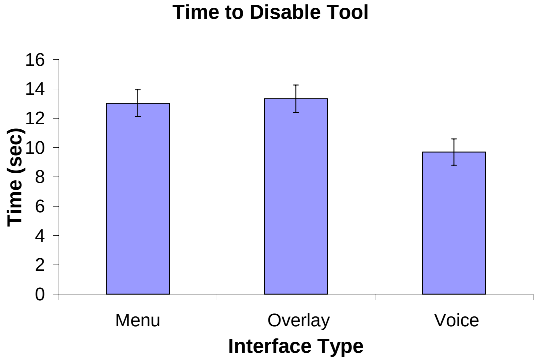

Figure 2. Time to disable the tool.

DISCUSSION

As expected, the results showed that the speech interface was faster for disabling the tool than the LCD menu or the overlay menu interfaces. The time to disable the tool included time for error recovery in all cases. The speech interface did not negatively impact recall time, the amount of errors in recall, or confidence. While the Menu and Overlay conditions did not affect those variables either, their task completion times were significantly longer. This suggests that designing a speech interface could make surgeries faster without compromising errors in memory. A faster surgery is beneficial to the patient because it means the patient receives less anesthesia.

It was expected that subjects would be more confident about their memory in the Overlay condition than in the LCD menu conditions since there was no need to divide their visual attention between two different locations. Furthermore, they could also see the pegboard pattern for longer than they could during the LCD menu condition (13 seconds more on average). However, our results did not show any difference between conditions for confidence, suggesting that an added 13 seconds of viewing time did not affect the subject's confidence.

It was somewhat surprising that the voice condition did not result in a higher confidence than the LCD menu condition. This could be due to the fact that people would often look away from the operative site to the telephone (used for speech recognition) which was situated on the console to their left. The phone had to be close to the subject for better recognition, but it could not be placed directly in front of them due to lack of space. If a speech recognition system were embedded in the robotic console, we would recommend placing it directly under the monitor in front of the surgeon. Then, the surgeon would not be tempted to turn to some other location, taking his eyes away from the monitor and losing the benefit of a non-divided attention interface.

There was no significant interaction or correlation between the interface type and recall, suggesting that the time to change a tool (accounted for in this experiment by the 80 second distraction task) was too short to show memory decay effects. This length of time was used because it was the average time for tool changes with the robotic system in the operating room (Webster, 2004). While this experiment was representative of a real-life situation and time duration, an interesting follow-up study would be to extend the distraction time. This could be important for instances when unexpected interruptions in the OR distract the surgeon for an extended period of time during surgery.

Upon closer examination of the data for individual differences, a correlation was found between individual's confidence and amount of error (Pearson r= -0.769). This shows that individuals who made fewer errors were also more confident. The high correlation of confidence and errors also suggests the importance of designing interfaces which will enhance the user's confidence.

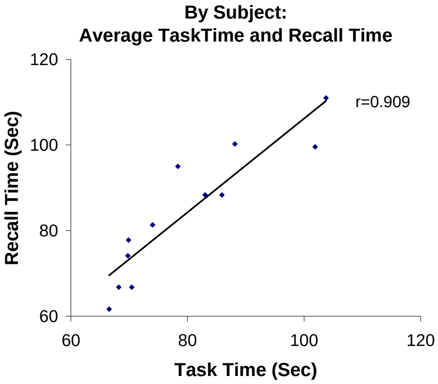

Figure 3: Individual's task time vs. recall task time.

Total task completion time and recall time were also highly correlated: r= 0.909 (see Figure 3). This shows that a subject's motor movement capability was a good predictor of the recall task time. One might have thought that if the task was done quickly and that the task space was viewed for less time, the recall of that task space might be more difficult and therefore take longer in the recall task. This was not the case. Most likely, if people were fast at maneuvering in the task, they were fast at maneuvering in the recall task. If they were slow with their motor coordination in the task, they were then slow in the recall task. Average task time and recall time were related to the subject's physical abilities rather than memory. Given the huge variance in individuals' averages, it would be interesting to look into training for improving motor skill ability.

Conclusions

The use of voice commands was significantly faster than other interface types, with no loss of performance or recall accuracy. If surgical procedures can be done more quickly without sacrificing or risking performance, this would be a benefit to patients, the operating room team, and the hospital. More research is needed to tease apart other interactions between interface type and memory/performance. Though our research did not find as many significant results as expected, it does provide a basis for considerable amounts of future research.

Acknowledgements

The authors acknowledge the support of W. Lee who allowed us to use the robot. This work was supported by the National Science Foundation under Grant #0238284.

References

Ben-Porat, O., Shoham, M., & Meyer, J. (2000). The effect of hand-controller configuration on task performance in endoscopic tele-operation. *Presence,* 9: 256-267.

Nio, D., Bemelman, W.A., Busch, O.R.C., Vrouenraets, B.C., Gouma. D.J. (2004). Robot-assited laparoscopic cholecystectomy versus conventional laparoscopic cholecystectomy: a comparative study. *Surgical Endoscopy,* 18(3): 379-82.

Berguer, R., Forkey, D.L., & Smith, W.D. (1999). Ergonomic problems associated with laparoscopic surgery. *Surgical Endoscopy* 13(5): 466-468.

Hochberg, J. & Brooks, V. (1978). Film cutting and visual momentum. In J. W. Senders, D. F. Fisher & R. A. Monty (Eds.), *Eye Movements and the Higher Psychological Functions* (pp. 293-313): Lawrence Erlbaum.

Meier, R., Fong, T., Thorpe, C., & Baur, C. (1999). Sensor Fusion Based User Interface for Vehicle Teleoperation, *Field and Service Robotics*, Pittsburgh, PA, August 1999.

Webster, J. (2004). *Introducing a Robotic Surgical System to the Operating Room: Effects on Information Flow and Attentional Resources.* Unpublished Master's Thesis, Tufts University, Medford, MA, USA.

Wickens, C.D., Vidulich, M., & Sandry-Garza, D. (1984). Principles of S-C-R compatibility with spatial and verbal tasks: The role of display-control location and voice-interactive display-control interfacing. *Human Factors,* 26: 533-543.

Woods, D.D. (1984). Visual Momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies,* 21: 229-244.

Yeh, M., Merlo, J.L., Wickens, C.D., & Brandenburg, D.L. (2003). Head up versus head down: the costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling. *Human Factors,* 45(3): 390-407.