The Impact of Single-Operator Versus Team Tele-Operation of a Search Vehicle

Adrian Matheson, *Member, IEEE*; Birsen Donmez, *Member, IEEE*; Faisal Ansari; Saad Ahmed
Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Canada
donmez@mie.utoronto.ca

Abstract—This paper evaluates the tele-operation of a mobile sensor platform. In current operations, this land vehicle is controlled by a team consisting of a driver and a sensor operator. Our experiment is the first attempt, for this system, to assess the impact, if any, of using a single operator versus a team of two operators, in order to inform staffing decisions as well as the design of future automated functions. The experiment was conducted using a simulator and 24 participants: 8 single operators and 8 teams of two operators. Lower mission completion times arose for the two-operator condition in spite of any extra time and workload generated by communication required. Operators in teams also assessed the system as more usable. These findings give support to the team strategy for tele-operating mobile sensor platforms, and have implications for the staffing and design of similarly complex uninhabited vehicle systems.

Tele-operation; CBRNE; human-robot interaction; first responder; team performance

I. INTRODUCTION

Fielding of uninhabited vehicles allows for safe performance of tasks that can otherwise be life threatening, and tele-operation is one way of controlling such vehicles. This technique is used in space, underwater, and near more localized and transient hazards [1], [2]. Our study focuses on tele-operation of a sensor platform designed to identify chemical, biological, radiological, nuclear, and explosive (CBRNE) threats. In particular, we evaluate single- versus two-operator tele-operation of a CBRN Crime Scene Modeller (C2SM), a tele-operated sensor platform land vehicle developed by MacDonald, Dettwiler, and Associates (MDA) to help crime scene investigators locate hazard sources before first (human) responders are sent into the environment, thereby minimizing harm or risk of harm [3].

The workload of operators is crucial in determining the success of tele-operation missions [4]. Operators have several concurrently running tasks, such as understanding the vehicle position and orientation (through the monitoring of different camera angles, as well as through mapping), manoeuvring the vehicle, and monitoring sensor data. The number of operators as well as the division of tasks in multi-operator cases can influence workload and hence mission performance. Further, employing teams of operators in differentiated roles can enable operators to rely on each other's judgments in their respective areas of responsibility and make their own decisions promptly, thus potentially reducing mission completion times [5].

In current operations, C2SM is controlled by two-person teams consisting of a driver and a sensor operator. Our experiment is the first attempt, for this particular system, to assess the advantages, if any, of the current arrangement over a single operator. The overarching goal is to inform staffing decisions as well as those on the design of future automated functions. Previous research focussing on optimal numbers of operators for tele-operation is limited. In a simulated tele-operation task which consisted solely of manoeuvring, teams of two operators were found to outperform single operators in terms of path completion times and damage caused by collisions [6]. These two measures are of particular importance in tele-operation tasks involving hazard identification and localizing, as these tasks are likely to be time-critical, and manoeuvring around obstacles during tele-operation can be challenging.

II. METHODS

A human subject experiment was conducted at the University of Toronto. The C2SM simulator software, designed by MDA, was used by subjects to locate hazardous objects via moving the vehicle and carrying out sensor operations. Two conditions were considered in the evaluation, the first involving a single operator responsible for manoeuvring the vehicle (driving) as well as operating sensors and navigating (choosing destinations and routes). The second condition involved two operators with defined responsibilities, with a "driver" manoeuvring the vehicle and a "sensor operator" operating sensors and relaying navigation instructions to the driver. In this experiment, there were no formal strictures on communication in pairs of operators, leaving operator specialization itself to shape it.

A. Participants

24 University of Toronto students (16 male and 8 female; age mean 22.9 years; age standard deviation 1.02 years; effects of age and gender not investigated) were recruited for this experiment. Of the 24, 8 participated in the single-operator condition, and the remaining 16 participated in the two-operator condition. 17 of the participants reported playing video games on a weekly basis and therefore may have had

some experience manoeuvring or navigating a vehicle like that used in the experiment. Five of these participants were in the single-operator and 12 were in the two-operator condition (effects not investigated). None of the participants had previous experience operating the C2SM simulator or vehicle. Participants were, however, trained as detailed below.

B. Apparatus

The Vehicle Control Station (Figure 1), shown on the screen of the laptop running the simulator (left side on Figures 2 and 3), displayed the view from the C2SM vehicle forward camera and was used for manoeuvring the vehicle. The C2SM Control Client (Figure 4), displayed on an external monitor (right side on Figures 2 and 3), showed the instantaneous intensity of cues used to detect hazards (e.g., radiation level, for a radioactive hazard) at the current vehicle location. This client was used to roughly determine the locations of the hazards and to perform sensor recording operations in order to more precisely map hazard positions.

The 2-D View (Figure 5), also displayed on the external monitor, to the right of the C2SM Control Client, was used for navigating the vehicle, annotating the locations of hazard cue sensor readings, and mapping hazard locations.

Figure 3. Condition 2: A driver (left) manoeuvres the vehicle while a sensor operator (right) operates sensors and provides navigation instruction.

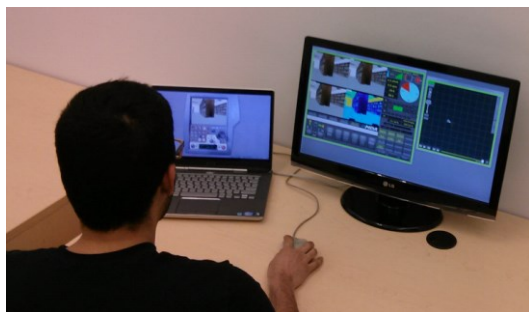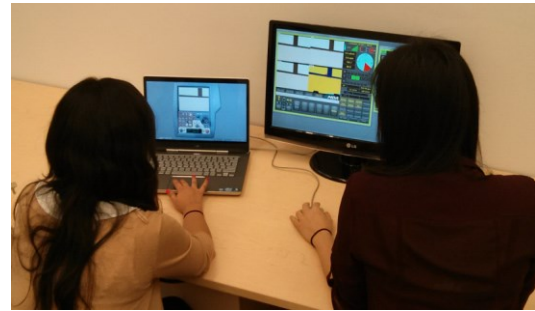Figure 4. C2SM Control Client.

Figure 1. Vehicle Control Station.

Figure 2. Condition 1: A single operator controls both manoeuvring of the vehicle (via keyboard) and operation of the sensors (via mouse).
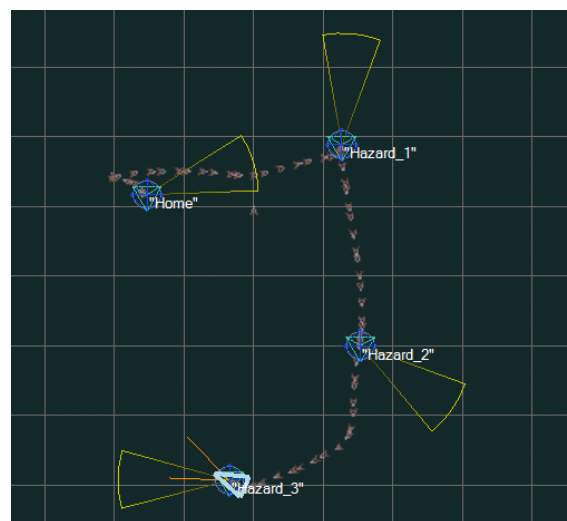

Figure 5. 2-D view for one of the experimental scenarios (Scenario 2), with the start point and three hazards marked by a participant.

Operators controlled the vehicle with the laptop arrow keys and conducted sensor operations with a mouse. Participants in the single-operator condition (Figure 2) were of course responsible for all tasks and controls, and those in the two-operator condition (Figure 3), as they were designated as either the driver or the sensor operator for the duration of the experiment, operated either the arrow keys or the mouse.

C. Procedure

Each participant was given a brief description of C2SM, and was then asked to complete a questionnaire on previous experience with similar systems, and on demographic data. Next, each participant received a 15 minute training session in which he or she was familiarized with functions of the C2SM simulator and with the appropriate user interface(s). All participants received the same training material and guidance from the facilitator apart from this different in UI used. Participants in the two-operator condition were randomly assigned to either the driver or the sensor operator position and were paired with an operator of the opposite type for the duration of the experiment. Pairs were trained and completed missions together. After training, single-operator participants or teams completed a practice mission. Participants were instructed to complete each mission as quickly and accurately as possible while minimizing collisions with obstacles.

After training, each participant completed two missions, each of which involved a different type of hazardous material (detectable only by a particular sensor), number of hazardous objects, and scenario layout (map). None of these parameters were divulged before running a given scenario, and the order of the scenarios was counterbalanced to control for a possible differential severity of learning effects in the single-operator versus dual-operator conditions.

Scenario 1 was an underground tunnel network with a single noxious gas source as the hazard, as shown in Figure 6. Scenario 2 was an indoor school area with three radiation sources as the hazards, as shown in Figure 7. Scenario 2 contained far more obstacles than did Scenario 1, making it more difficult to manoeuvre the vehicle, and making it more likely that a benefit of two operators over a single operator would be realized.

Figure 6. Scenario 1: A tunnel network with one source of noxious gas.
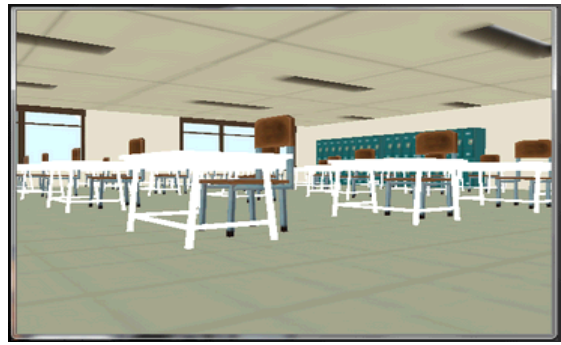

Figure 7. Scenario 2: An indoor school area with three sources of radiation.

In each mission, participants were required to perform 4 different types of tasks. Introductory tasks involved the operator activating appropriate sensors as well as activating sensor alarms to receive auditory feedback based on proximity to hazards, and setting these alarms to pre-defined sensitivity levels. Navigation tasks required operators to move the C2SM vehicle to pre-determined locations. Measurement tasks involved use of the 2-D map to take distance measurements between points of interest. These tasks also required the use of a simulator feature which marks a location on the 2-D map. Sensor operations involved operators searching for hazards by moving the vehicle and monitoring sensor readings from the Control Client, and employed a simulator feature which records instantaneous sensor readings along with vehicle location. A mission was complete after all hazards were located.

After participants completed both scenarios, they took part in an informal interview and completed NASA Task Load Index (TLX) and System Usability Scale (SUS) questionnaires [7], [8], [9].

D. Parameters Measured

Data were collected through paper-based questionnaires, though timing and real-time note-taking by the experimenters, and through the recording of computer screens, participant faces, and audio communications.

A timer was started as soon as the participants began reading about a task, and was stopped when the participants began reading about the following task or stated that the previous task had been completed. Mission completion times were generated by aggregating these task completion times.

The number of collisions of the vehicle with walls and other obstacles were identified through the video feed. It was difficult to identify minor collisions, so only very evident collisions, recognizable by the erratic movement they generated in the video feed, were counted.

'Errors' committed by each operator were noted by one of the experimenters in real time, where a deviation from the instructions on the instruction sheet—either missing a task or performing it incorrectly—was considered an error.

Gaze switches between the two monitors (laptop and external) were identified and counted for each participant during post-processing of video recordings. Each instance of an operator switching his or her gaze from one screen to

another was counted as a view change. Gazes toward the instruction sheet were not included in this assessment.

For the two-operator condition, the number of times participants communicated with each other was extracted from audio recordings. A communication was defined as a spoken sentence relevant to the task at hand. Any small talk between participants was not included in the overall count.

III. RESULTS

Here we present the preliminary analysis and the associated findings from this experiment.

A. Mission Completion Time

Figure 8 presents mission completion times. For both of the scenarios (tunnel and school), the two-operator condition resulted in lower mission completion times compared to the one-operator condition (Scenario 1: $t(14) = -2.75$, $p = .02$; Scenario 2: $t(14) = -3.43$, $p = .004$).
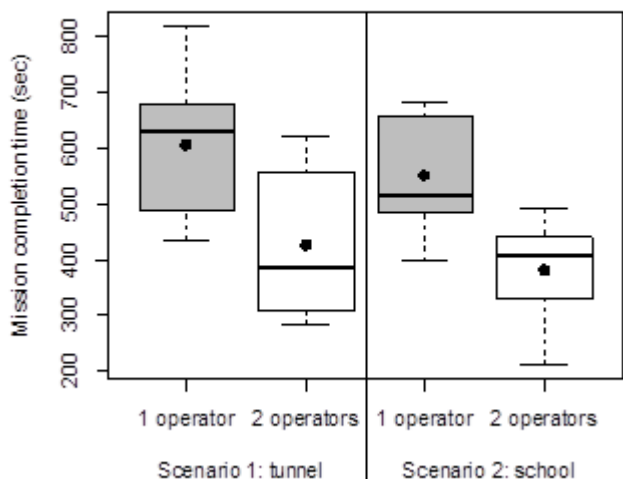

Figure 8. Boxplots of mission completion time (solid circles represent sample means).

B. Numbers of Collisions and Errors

Number of collisions within each scenario was compared between single- and two-operator conditions through the Wilcoxon rank sum test. There were no statistically significant differences between the two conditions ($p>.05$). A similar analysis was conducted for number of errors and again there were no significant findings ($p>.05$).

The number of collisions ranged between 0 and 7, with a median of 1, a mean of 1.72, and a skew of 1.01. The number of errors ranged between 0 and 4, with a median of 1, and mean of 0.91, and a skew of 0.97.

C. Number of View Changes

Figure 9 presents total number of view changes, where two-operator team member views were summed. There was a significant result only for Scenario 2, with the single operators exhibiting approximately twice the number of view changes as that exhibited by the two-operator teams ($t(14) = 2.73$, $p = .02$).
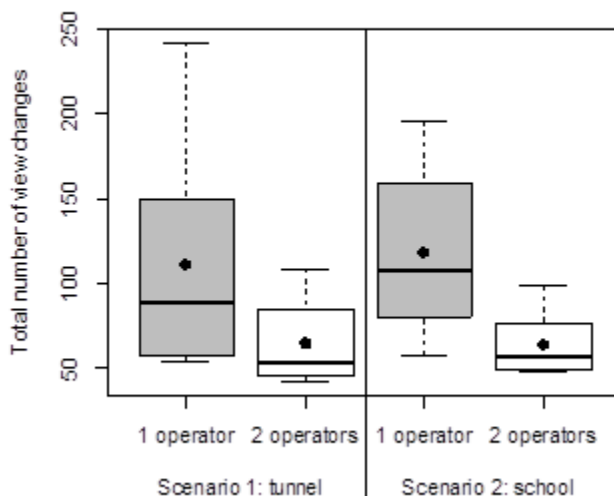

Figure 9. Boxplots of total number of view changes (solid circles represent sample means).

Correlation analyses on total number of views, total number of communications, and mission completion time for the two-operator condition revealed significant positive correlations between total number of views and communications (rho = 0.58, $t(14) = 2.68$, $p = .02$), as well as between total number of views and mission completion time (rho = 0.5, $t(14) = 2.68$, $p = .048$).

D. Subjective Measures

The NASA TLX is a subjective assessment tool developed to rate perceived workload. The tool assesses six constructs: mental demand, physical demand, temporal demand, performance, effort, and frustration level, though for this experiment physical demand was not assessed [7]. "Raw TLX" (no assignment of individual weights to the constructs) was used, and the constructs were assessed here individually. Raw TLX has been shown to be at times equal, more, or less sensitive than the original scoring system [8]. NASA TLX data showed no differences between the single-operator and two-operator conditions for any of the constructs ($p>.05$).

The SUS is based on 10 questions with 5-point Likert scale ratings, and is used to capture a subjective assessment of usability [9]. The SUS involves a scoring formula, with equal question weights, leading to a total score out of 100. Participants in the two-operator condition rated the system as more usable than did those in the single-operator condition ($t(12.1) = -2.58$, $p = .02$).
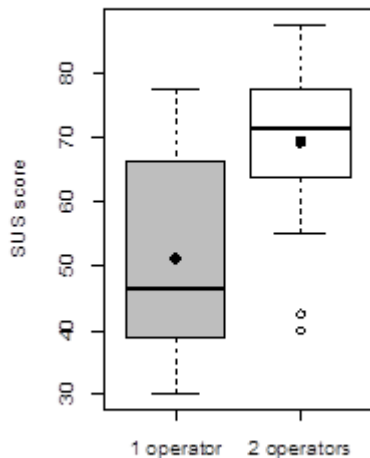
Figure 10. Boxplots of SUS score (solid circles represent sample means).

IV. DISCUSSION

The performance of two operators with defined roles, versus that of single operators, was found to be significantly better in some respects and not different in the remainder, lending clear support for the two-operator strategy in this task.

Employing two operators reduced mission completion times, but had no effect on the numbers of collisions and errors. That is, of the performance measures directly reflecting mission completion, only one was affected. This could be because the effect of employment of a second operator is truly limited to speed, but it could also be the result of a particular type of speed-accuracy trade-off on the part of operators. It may the case that operators prioritize holding collisions and errors (accuracy) constant (at a low level), allowing any kind of increase or decrease in performance to be reflected only in speed.

The objective measure of performance indirectly related to mission completion, number of view changes, was seen to be different in only one of the two scenarios. This could be an artifact of the effect being marginally significant due to a lack of statistical power, but it is reasonable to conclude that some difference between the scenarios may have modulated the effect. Scenario 2, in which the effect was detected, involved more complex navigation than did Scenario 1. More complex navigation can fairly be expected to require more operator attention paid to the navigation (sensor) view, and in the single-operator case more frequent changes to and from that view. We suggest future studies on a possible interaction effect between operator number and navigation complexity on view change number.

The absence of differences in NASA TLX measurements indicates that the enhanced performance of two-operator teams likely did not result from decreased perceived workload. The presence of differences in SUS scores does provide some insight into why the two-operator teams succeeded, however, in that the questions on the SUS questionnaire deal with system complexity and learnability. This implies that the disadvantage of employing single operator could be overcome through extra training, and we therefore suggest this as another topic for future study.

Though more staff resulting in better performance is a common effect for a variety of tasks, it is worthwhile to know that this holds true for the specific task of tele-operating this mobile sensor platform. In addition to supporting the continued use of two-operator teams for C2SM system missions, we expect that this experiment can inform decisions involving the costs and benefits of different staffing levels for similar systems and tasks.

REFERENCES

[1] T. Fong and C. Thorpe, "Vehicle teleoperation interfaces", *Autonomous Robots*, vol. 11, 2001, pp. 10–13.

[2] D. Zhu, T. Gedeon, and K. Taylor, "Exploring Camera Viewpoint Control Models for a Multi-Tasking Setting in Teleoperation", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 2011, pp. 53–62.

[3] P. Jasiobedzki, H. Ng, M. Bondy, and C. McDiarmid, "Increasing situation awareness of the CBRNE robot operators", *Proceedings of SPIE 7666, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense*, 2010, pp. 1–2.

[4] M. Prewett, R. Johnson, K. Saboe, L. Elliott, and M. Coovert, "Managing workload in human–robot interaction: A review of empirical studies", *Computers in Human Behavior*, vol. 26, no. 5, 2010, pp. 840–856.

[5] B. Sperling, "Information distribution and team situational awareness: An experimental study", *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* , vol. 50, no. 3, 2006,, pp. 477–481.

[6] S. Glynn and R. Henning, "Can teams outperform individuals in a simulated dynamic control task", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 44, no. 33, 2000, pp. 141–144.

[7] S. Hart, "NASA task load index (NASA-TLX); 20 years later." Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, vol. 50, no. 9, 2006, pp. 904–908.

[8] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research" in *Human Mental Workload*, P. Hancock and N. Meshkati (Eds.), Amsterdam: North Holland Press, 1988.

[9] J. Brooke, "SUS: A 'quick and dirty' usability scale" in *Usability Evaluation in Industry*, P. Jordan, B. Thomas, B. Weerdmeester, and A. McLelland (Eds.), London: Taylor and Francis, 1996.