# EMERGENCY AUTOMATED RESPONSE SYSTEM (EARS)

**Joshua Phillips, Sachin Kogekar, and Julie A. Adams**
Department of Electrical Engineering and Computer Science
Vanderbilt University, Nashville, TN 37235
[joshua.l.phillips, sachin.kogekar, and julie.a.adams]@vanderbilt.edu

This paper presents a comparison of two versions of an Emergency Automated Response System (EARS), a fully manual version and a partially automated version. User evaluations involving both versions of the system were conducted using a low workload task and a high workload task. The results indicate that the automation employed by the partially automated system decreased overall response time and perceived workload for both tasks, but accuracy decreased and response times increased from low workload to high workload with both versions.

## INTRODUCTION

In the context of emergency response systems, response resource balancing and response time issues are significantly influenced by the ability of the operator and the system to appropriately handle varying workload levels. It is expected that as workload increases, the addition of automation should assist the operator with monitoring activities. This paper describes an automated emergency response system (EARS) designed to aid response teams.

Several systems similar to EARS are in operation in a wide variety of domains. Automated dispatch technology offers communication service providers an effective framework to efficiently manage and respond to multiple types of work – service, emergency, construction, and maintenance (Neeley, 2001). Such technology is used in dispatching emergency response units (Aleman, 1999). The Los Angeles Emergency Operations Center automatically differentiates between event types and dispatches them to the appropriate departments (L.A., 1996). It is important to understand how workload and automation level can affect emergency response times and dispatcher accuracy.

The purpose of EARS was to analyze workload variation across two system versions. These two versions were Manual Control (MC) and Automated Control (AC). MC demanded more physical attention and user effort. In this version, the user monitored a screen showing a map of the Vanderbilt University campus. Randomly generated emergencies appeared as small red dots. The users dispatched emergency response units to the indicated emergencies. The dispatched emergency response units were visible and moved about the screen. The user had to monitor the screen for any emergencies where a response unit failed to respond within a fixed time period, and repeat the same procedure for these, and other, emergencies. Figure 1 provides a system interface example.

The AC mode was similar to the MC mode. However, a certain automation level was provided to aid user monitoring and unit dispatching. This system automatically dispatched the closest response unit when the user initially selected an emergency. Once the unit was selected and the system dispatched the unit, the user was responsible for monitoring for any new emergencies and for response unit failure.

Level of system automation versus perceived workload is an area that has been studied across many domains including robotics (Ruff et al., 2002), and air traffic control (Kaber et al., 2002). Endsley and Karber (1999) defined a level of autonomy taxonomy based upon results in various domains. In general, it is believed that higher levels of automation reduce perceived workload but may also induce boredom and complacency. The purpose of this work was to better understand the trade-offs between automation level and workload based upon two automation levels and two workload levels.

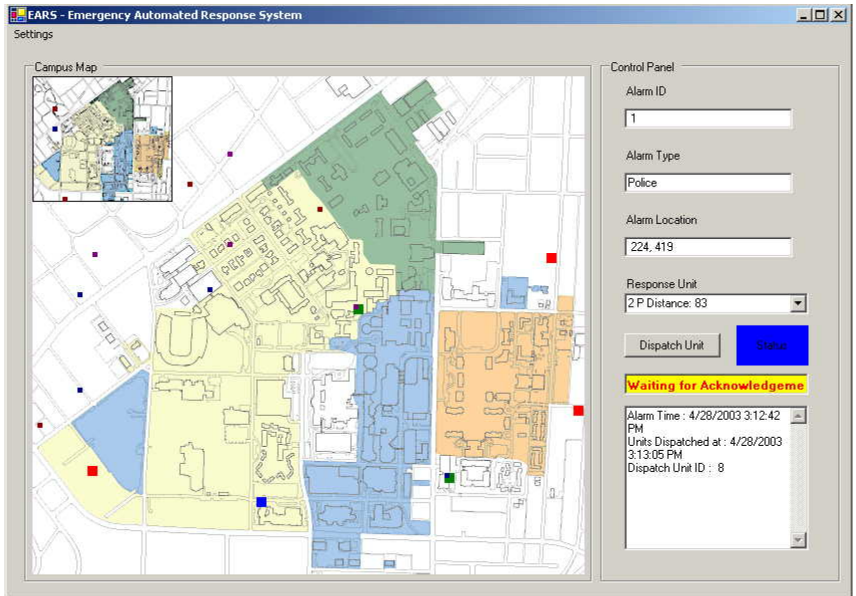


***Figure 1. EARS example screen.***

## METHOD

Twelve volunteers (3 females, 9 males) from the Vanderbilt University community completed the user evaluation. Eleven participants were graduate students and one was an undergraduate. Each evaluation began with a functionality explanation followed by a three-minute training run that allowed participants to become familiar with the MC and AC versions of the EARS interface.

The participants completed four simulations, each lasting five minutes. Two of the simulations were completed under MC while the remaining two required the AC mode. The MC simulations involved the user moving the mouse cursor to a particular emergency on the screen and then selecting it by clicking the mouse button. The user then selected an appropriate emergency response unit from a list displayed in a combo box and dispatched it by clicking on the dispatch button on the screen. The details of selected emergencies were provided to the right of the screen. The most appropriate (closest) response unit was to be selected and dispatched. The distance to the selected emergency was provided in order to guide the user's selection. Upon dispatch, the emergency icon turned blue to indicate that the system was attempting to obtain acknowledgement that the unit would handle the emergency. If the selected response unit did not acknowledge within a certain time frame, the emergency color changed to black. The user then had to select another emergency response unit and dispatch it. The user repeated this process until there was a successful acknowledgment. When an acknowledgement succeeded, the emergency color changed to green and the chosen unit traveled to the emergency to handle the situation. The unit stayed at the response location until the situation was neutralized. After neutralization, the emergency's map icon disappeared. A dispatched unit could only handle a single emergency at a time. Therefore, the user could encounter situations where there were no available units for dispatch. In this case, the user waited and monitored the remaining emergencies until an appropriate unit became available. The user repeated the assignment process for all appearing emergencies. During unit dispatch or emergency handling, the user could simultaneously dispatch units to other emergencies.

In the AC mode, all emergencies were first handled by the automation by dispatching the closest response unit. However, if an acknowledgement failure occurred, the user had to specifically select the most appropriate response unit and dispatch it. This form of automation was chosen primarily because it aids the user in selecting the closest appropriate vehicle and because it performs this task quickly. Therefore, it should aid the users in two ways: by decreasing time spent deciding which vehicle to route and by decreasing total emergency processing time due to accurate vehicle selection. Thus, it directly addresses the two primary concerns of operators in this domain without affecting how operators interact with the system.

When using the AC mode, the user had to monitor the screen for only those emergencies with failed acknowledgements and then dispatch (via manual selection) the appropriate (preferably closest) unit as described above. Thus, MC simulations extracted more physical activity and situational awareness from the user than AC simulations.

Both MC and AC simulations had two task difficulty levels; one simulation was of low task difficulty and the other of high task difficulty. Low task difficulty simulations generated the emergencies at a lower rate compared to high task difficulty simulations. Thus low task difficulty simulations had lower workload than high task difficulty simulations.

After each simulation, a NASA TLX Workload (Hart and Staveland 1988) evaluation containing six scales representing the TLX factors was completed. A final TLX questionnaire containing the fifteen paired-comparisons was completed at the end of the experimental session.

Participants were divided into two groups of six. One group completed the MC simulations first while the other group used AC first. The two groups were further subdivided into four subgroups of three where two groups completed low task difficulty simulations first and the remaining two groups first completed high task difficulty simulations. Balancing the presentation of the task difficulty and automation levels across participants controlled order effects.

The hypothesis applied to this evaluation was that the addition of automation should decrease operator response and processing times while reducing perceived workload, independent of task difficulty.

## RESULTS

The response times for various measures were collected during all simulations. The time related measurements were gathered in milliseconds and were then averaged.

***Table 1 - Average Emergency Processing Time in seconds.***

| | **Automatic** | | **Manual** | |
|---|---|---|---|---|
| **Task Difficulty** | Mean | Std. Dev. | Mean | Std. Dev. |
| Low | 32.74 | 16.20 | 43.11 | 24.34 |
| High | 62.89 | 34.35 | 74.87 | 44.49 |

The average emergency processing time was measured from emergency onset until the routed response unit completed handling the emergency. These results allow the analysis of the operators' overall efficiency. Proper emergency response unit routing should minimize this overall time. Thus lower emergency processing time shows that the operators were more efficient as shown in Table 1. The fastest processing time occurred for the automatic, low difficulty simulations while the slowest processing time occurred in the manual, high task difficulty simulations.

***Table 2 - Average Emergency Recognition Time in Seconds.***

| | **Automatic** | | **Manual** | |
|---|---|---|---|---|
| **Task Difficulty** | Mean | Std. Dev. | Mean | Std. Dev. |
| Low | 4.24 | 8.32 | 10.11 | 15.50 |
| High | 10.08 | 12.85 | 33.16 | 34.83 |

The average emergency recognition time was measured from emergency onset until the operator first selected the emergency. These results provide an understanding of the operators' vigilance-workload relationship. The operators should respond more quickly to each emergency in the lower task difficulty simulations. However, if the operators were less vigilant (i.e. bored or distracted) due to the lower workload levels, then this time may increase. Table 2 summarizes these measurements. The emergency recognition times indicate that the fastest recognition occurred with the automatic, low task difficulty simulations. As task difficulty increased, the recognition time increased and continued to increase in manual mode.

***Table 3 - Average Response to Acknowledgement Failure Time in seconds.***

| | **Automatic** | | **Manual** | |
|---|---|---|---|---|
| **Task Difficulty** | Mean | Std. Dev. | Mean | Std. Dev. |
| Low | 12.49 | 11.84 | 24.21 | 23.39 |
| High | 38.82 | 38.11 | 42.13 | 41.45 |

The response to acknowledgement failure time was measured from first instance of acknowledgement failure until an emergency response unit acknowledged it would handle the emergency. This provides a measure of the operators' vigilance level, but is also related to the emergency processing time. If operators failed to quickly address acknowledgements, then their average performance degrades due to the extra time emergencies remained active. These times vary greatly from operator to operator, but the results indicate how the average operator reacts to changes in the current emergency state. The best response time occurred with the automatic, low task difficulty simulations as shown in Table 3. An interesting result is that all acknowledgement failure times are longer than emergency recognition times.

***Table 4 - Average of percentage response units correctly chosen.***

| | Automatic | | Manual | |
|---|---|---|---|---|
| **Task Difficulty** | Mean | Std. Dev. | Mean | Std. Dev. |
| Low | 97.85 | 13.97 | 88.43 | 33.52 |
| High | 92.675 | 24.89 | 85.26 | 34.62 |

The percentage of correct response units chosen is a measure of whether the closest relevant response unit was chosen. This measure provides an understanding of the accuracy of operators' response choices. An increase in accuracy was anticipated for the automatic simulations, but operators tended to rely heavily on automation in high task difficulty simulations. When the automation fails, as in the acknowledgement failures, then the operators' accuracy may not be as high as anticipated. The results are summarized in Table 4. While there is a larger difference between the automatic and manual simulations, there is a much smaller difference between the task difficulty variations within an automation level. The automatic simulations provided the most accurate responses for both task difficulty conditions.

***Table 5 - Average workload values.***

| | Automatic | | Manual | |
|---|---|---|---|---|
| **Workload** | Mean | Std. Dev. | Mean | Std. Dev. |
| Low | 25.341 | 14.41 | 51.258 | 16.56 |
| High | 46.75 | 21.39 | 69.7 | 23.52 |

Participants provided feedback regarding perceived workload via the NASA TLX tool. It was anticipated that the highest perceived workload would occur with manual modes and more difficult tasks. As anticipated, the highest workload levels were found with the manual simulation modes, as shown in Table 5. The highest level was found for the high task difficulty, manual condition while the automatic, low task difficulty condition provided the lowest perceived workload.

For both the AC and MC tasks, the standard deviations of all processing times were consistently lower for the low difficulty task than for the high difficulty task. During the low difficulty task, many fewer overall emergencies were under consideration. As a result, all emergencies received a relatively similar amount of attention (based on response time) throughout these tasks. However, for the high difficulty task, there were many more emergencies to be processed in the same amount of time. Participants appeared to switch between two different strategies during the high difficulty task. For a while, they would attempt to process emergencies as quickly as they appeared. This would lead to many fast response times. However, this strategy would not allow the participant to process all of the emergencies in this manner. After some time, the number of unprocessed emergencies would start to increase. At some point, the participants would finally become aware of this build-up of emergencies that had not yet been attended to. Now the participants would change to a new strategy where they attempted to process all of the older emergencies that they perceived had been left waiting for some time. The response times for these emergencies were large in comparison to the mean. Eventually, the users would go back to first strategy when they felt that they had handled most of the older emergencies. Thus, the strategy switching that participants used to process emergencies in the high difficulty tasks caused the response times for emergencies in these tasks to be much more varied than in the low difficulty tasks where either strategy tended to consistently work well throughout the tasks.

## DISCUSSION

As task difficulty increased, emergency-processing time increased. However, the AC version provided faster processing time over the MC mode for matching levels of task difficulty. This follows the hypothesis that automation should provide better overall performance.

The AC mode provided better emergency recognition time for all task difficulty levels. As task difficulty increased, AC mode provided super-linear speed-ups in the recognition time. It is felt that the reduction in the recognition time for AC was due to the reduced workload provided by the automation's assistance to the participant. The TLX data correlated the participants' perceived workload to the reduced emergency recognition time. Therefore, these results support the hypothesis.

When task difficulty was low, AC out-performed MC. As task difficulty increased, AC continued to perform better than MC, however the marginal performance difference was reduced. It appears that for lower task difficulty, the use of automation enabled the participants to pay more attention to the emergencies. As task difficulty increased, the advantage gained from the automation was reduced, though overall the results remained better than the MC simulations.

AC provided better results than MC regardless of task difficulty level. The automation restricted the user's response unit choices for most emergencies to the best one. MC permitted users to choose a response unit, including an incorrect unit. They often accidentally selected an incorrect unit and were unwilling to make another selection. However, this phenomenon was not anticipated because participants were instructed to choose the most appropriate response unit at all times. Future studies could be performed to quantitatively assess this phenomenon.

When task difficulty was low, the users claimed that AC reduced their perceived workload by half. During the higher task difficulty simulations, AC still rated lower but the difference margin was reduced by a small amount. This supports the hypothesis that the automation aided the reduction of perceived workload.

During high difficulty tasks, the users employed two different strategies and switched between these strategies, resulting in more variation among emergency processing times. Trained operators in this domain would probably learn better strategies that would allow them to eliminate these large variations. This seems to be the case since the only difference between high and low difficulty cases was the emergency occurrence rate. Thus, the high standard deviation across high difficulty tasks appears to be dependent on a lack of experience rather than a difference between tasks. AC helps accommodate for a lack experience and the AC showed smaller differences between task standard deviations than MC, suggesting that AC helped reduce this difference some as well.

A comparison between values in Table 4 and Table 1 reveals that the high workload task not only taxed the participants' abilities to appropriately route emergency vehicles, but also taxed their abilities to handle emergencies quickly. This is a little surprising since more accurate vehicle routing choices should decrease the overall emergency processing times and the EARS AC mode was developed to leverage the benefits of this intuitive observation. However, this is the typical relationship between response time and accuracy in most, but not all, domains. Therefore, the results of this study might only generalize to other domains where it seems that increased accuracy should improve response times, but empirical studies prove otherwise.

## CONCLUSIONS

This paper has presented the two versions of an Emergency Automated Response System (EARS). This system provides two modes of operation, one fully manual and one that provides a basic level of automation but still requires user input to handle certain situations. A user evaluation that included twelve participants was conducted to assess the trade-offs between automation level and workload. The evaluation included the two automation levels and two workload levels.

All evaluation data reflected that automation proved to reduce perceived workload while improving performance. This result is consistent with our hypothesis that the addition of automation should decrease operator response and processing times while reducing perceived workload, independent of task difficulty. While this evaluation sample is small, it is anticipated that a larger study would provide similar results

## References


Aleman Jr., S. (1999). Computer-aided dispatch comes to the Ozarks. Retrieved February 2003, from The Wireless Market Place - Mobile Radio Technology, http://iwce-mrt.com/ar/radio_computeraided_dispatch_comes/.

Endsley, M. R., & Karber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, 42 (3), 462-492.

Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload***,** 139-183. Amsterdam, The Netherlands: Elsevier.

Kaber, D. B., Prinzel III, L. J., Wright, M.C., & Clamann, M. P. (2002). Workload-Matched Adaptive Automation Support of Air Traffic Controller Information Processing Stages. NASA/TP-2002-211932.

L.A. (1996). Automated massaging improves response in L.A. Retrieved February 2003, from American City & County, http://advertisers.americancityandcounty.com/ar/government_automated_massaging_improves/.

Neeley, B.J. (2001). Automated Dispatch for Communications Companies. *Proceedings of the 2001 GITA Annual Conference XXIV*, USA.

Ruff, H., Narayanan, S. & Draper, M. (2002). Human Interaction with Levels of Automation and Decision-Aid Fidelity in the Supervisory Control of Multiple Simulated Unmanned Air Vehicles. *Presence: Teleoperators & Virtual Environments*, 11 (4), 335-351.