Emerging Technologies and Future of Work, Vol. 117, 2023, 206–212
https://doi.org/10.54941/ahfe1004420

Exploring the Effects of Speech Speed and Environmental Noise on Human and Machine Performance in Civil Air Traffic Control Communication Tasks

Manrong She[1,2,3], Yi Hu[1,4], and Zhizhong Li[1,2]

[1]Department of Industrial Engineering, Tsinghua University, Beijing 100084, China
[2]National Key Laboratory of Human Factors Engineering, Beijing 100084, China
[3]Teaching Centre for Writing and Communication, School of Humanities, Tsinghua University, Beijing 100084, China
[4]Huaxin Jushu (Hangzhou) Microelectronics Co., Ltd., Hangzhou 310051, China

ABSTRACT

Considering the relative strengths of humans and machines may not be static, this study investigates the effects of speech speed and environmental noise on human and machine performance in the context of civil air traffic control communication. 32 participants were recruited to perform route selection, parameter setting and radio adjustment according to the voice commands from the control tower. Their performance was evaluated with respect to varying levels of speech speed, environmental noise and time pressure. Additionally, human performance was compared to that of a machine (i.e., a voice recognition software). The experimental results showed that both speech speed and environmental noise had significant effects on human performance in terms of recognition accuracy and operation accuracy. Humans excel in situations with high noise and low speech speed, while machines outperform humans when dealing with high speech speed and low noise. The findings demonstrate that a static human-machine function allocation method may not always yield optimal results. Suggestions are provided on how to develop a dynamic allocation method.

Keywords: Human-machine function allocation, Speech speed, Environmental noise, Civil air traffic control

INTRODUCTION

In recent years, there has been a rapid advancement in automation and autonomy. Humans and machines are increasingly collaborating as team members to accomplish tasks. Within this context, how to allocate the tasks between the humans and machines becomes a vital concern. Function allocation was first introduced by Fitts in 1951, and refers to the process of assigning a task or function to either human(s) or machine(s) (Fitts, 1951). If the tasks/functions are excessively assigned to humans, they may be overloaded. Conversely, if too many tasks/functions are distributed to machines, humans may get over-reliance on automation, leading to a decline in skills or a loss of situation awareness, and human may encounter difficulties in handling unexpected situations (Endsley, 2017; Sarter and Woods, 1995). Therefore, a proper function allocation is necessary to ensure system safety and team performance.

Considering the distinct capabilities of humans and machines, researchers have proposed to assign the tasks/functions to humans or to machines in which they excel, a concept known as MABA-MABA (men are better at, machines are better at, Fitts, 1951). The MABA-MABA list was widely applied in early industrial automated systems. Subsequent researchers updated the list as automation advanced. Some researchers also used the list as a basis for human-machine function allocation, and consider various objectives such as effectiveness, efficiency, safety, reliability, feasibility and cost (e.g., Price, 1985; Older et al., 1996). Other researchers based their methods on levels of automation (e.g., Sheridan and Verplank, 1978; Riley, 1989; Parasuraman et al., 2000) or on the results of cognitive task/work analysis (Roth et al., 2019). The methods share a common drawback – the allocations derived from them are static (Feigh et al., 2012). Due to the dynamic nature of the situation and the task, and the unpredictable variability of human and machine states, static methods cannot ensure that the allocation always remains optimal. Consequently, the comparison of human and machine performance in various conditions is requisite to establish the guidelines pertaining to the allocation of functions in these conditions.

This study selected the task context of communication between the control tower and pilots in civil air traffic control (ATC), considering many accidents have been attributed to communication issues. Communication plays an important role in preserving situation awareness, preventing conflicts, and ensuring the safe landing, take-off and movement of the aircraft. The pilot sometimes receives ATC instructions in a noisy environment, making it difficult to clearly hear the specific instructions. Moreover, when the ATC is busy, the instructions are likely to be delivered at a rapid pace, further increasing the difficulty to accurately receive the instructions. In light of this, the current study compared human performance and machine performance under different levels of speech speed and environmental noise. The results were expected to provide insights into the appropriate allocation between humans and machines within the said conditions.

METHOD

Participants

Thirty-two students from Tsinghua university (14 males and 18 females) were recruited as participants. Their average age was 22.2. The experimenter first informed them of the experiment details, and the participants signed the consent form voluntarily. All the participants had normal hearing, and normal or correct-to-normal vision.

Tasks and Scenarios

The experimental platform simulated the dialogues between a pilot and the ATC, and the follow-up operations that the pilot ought to perform. The platform was developed with Python. It offered three types of tasks: route selection, parameter setting, and radio adjustment.

Route selection: The ATC informed the pilot of the route to be chosen, the direction of travel, and the altitude to be maintained. The pilot was required the select the correct route and enter the correct values according to the voice instructions.

Parameter setting: During the scenario, the pilot might be asked to adjust the parameters such as the pitch angle and the sea level pressure to keep the aircraft in a safe flight.

Radio adjustment: During take-off and landing, the pilot might need to shift among different control towers. The voice instructions informed the pilot of the channel to be selected and the frequency to be tuned. The pilot was required to select the expected channel and enter the expected value of frequency.

The system interface is shown in Figure 1. The system log and the participants' performance data were recorded in the platform. The pilot was played either by a human participant or by a voice recognition software (i.e., WeChat Voice Recognition). Many voice recognition products were tested before the experiment, and WeChat Voice Recognition showed a superior performance than the others. The experimental tasks for the participant and the software involved listening to the voice instructions from the ATC, repeating the instructions, and then executing the required operations accordingly.

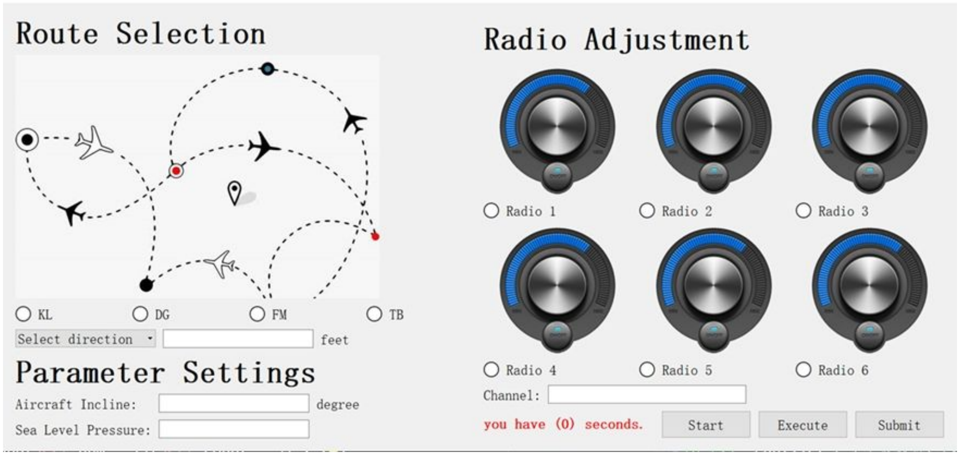

Figure 1: The interface of the platform.

Independent Variables

There were three independent variables – speech speed, environmental noise and time pressure in the experiment. They were all within-subject variables. Speech speed consisted of two levels: 145 words per minute as the lower speed and 1.3 times faster as the higher speed. Environmental noise also had two levels. In situations with the high noise level, the platform played the noise at the same decibel level as the voice command. The noise was extracted from the cockpit of Flying Tiger Line Flight 66, a flight that crashed in 1989 partially due to the miscommunication between the pilot and the ATC. In situations with the low noise level, the platform did not play any noise. Time pressure was applied during the recognition stage (i.e., repeating the voice instructions) in every trial. The baseline of time pressure was set as the average time (12.3 seconds) that participants spent on recognition in the pilot study. The length for the high time pressure level was 10 seconds (0.8 times the baseline). There was no time limit for the low time pressure level.

Dependent Variables

The speech recognition accuracy, operation accuracy and completion time were recorded to evaluate human (participant) and machine (voice recognition software) performance. Recognition accuracy was calculated as the percentage of the number of correctly repeated key words (by the participant or the software) to the total number of key words in the given voice instructions. Operation accuracy was calculated as the percentage of the number of correctly performed operations to the total number of required operations. Completion time was the time length from the start of the voice instruction to the time when the operations were completed.

Procedure

Only one participant took part in the experiment at a time. The participants were first informed of the experiment purpose and signed the consent form. Then they were introduced about the experiment background, the tasks to be performed, and the usage of the platform. After that, they performed the exercise with tasks similar to those they would perform during the formal experiment. The exercise consisted of eight trials. Both the speech speed and environmental noise during exercise were set at a "medium" level, which were between the high and low levels adopted in the formal part. There was no time limit during exercise. Afterwards, the participants were asked to perform eight (2*2*2) formal trials. Each trial consisted of six voice instructions, and each trial corresponded to a combination of levels of speech speed, environmental noise and time pressure. The sequence of the trials was randomized. There was a 40-second rest between every two trials.

RESULTS

Table 1 summarises the descriptive statistics of speech speed, environmental noise and time pressure on the participants' recognition accuracy, operation accuracy and completion time.

Table 1. Descriptive statistics for all variables: Mean (SD).

	Experiment conditions							
Dependent variables	High time pressure				Low time pressure			
	High speed		Low speed		High speed		Low speed	
	High noise	Low noise	High noise	Low noise	High noise	Low noise	High noise	Low noise
Recognition accuracy	0.515 (0.356)	0.646 (0.340)	0.788 (0.286)	0.830 (0.261)	0.609 (0.366)	0.736 (0.338)	0.679 (0.371)	0.842 (0.269)
Operation accuracy	0.515 (0.362)	0.614 (0.355)	0.726 (0.325)	0.782 (0.304)	0.556 (0.381)	0.716 (0.340)	0.634 (0.380)	0.799 (0.309)
Completion time	14.513 (5.170)	15.304 (0.614)	14.470 (5.446)	13.225 (4.661)	17.244 (7.973)	14.591 (6.119)	15.966 (6.336)	15.041 (6.204)

The recognition accuracy and the operation accuracy data satisfied the normality tests, thus repeated measures ANOVA was used (see Table 2 for the results). The completion time data violated the normality assumption, so the Wilcox rank sum test was used to examine its main effects on three independent variables.

Table 2. Repeated measure ANOVA results.

Factors and interactions	Recognition accuracy		Operation accuracy	
	F	*p*	*F*	*p*
Speech speed	18.873	<0.001**	10.953	0.002**
Environmental noise	10.099	0.003**	8.689	0.005**
Time pressure	0.354	0.555	0.181	0.673
Speech speed * Environmental noise	0.136	0.714	0.050	0.824
Speech speed * Time pressure	3.718	0.061*	1.798	0.188
Environmental noise * Time pressure	0.655	0.423	1.088	0.303
Speech speed * Environmental noise * Time pressure	0.734	0.397	0.089	0.767

** $p < 0.05$, * $p < 0.1$

Effects on recognition accuracy: Both speech speed ($p < 0.001$) and environmental noise ($p = 0.003$) were found to have significant effects on the participants' recognition accuracy. To be specific, with higher speech speed, the participants' recognition accuracy was significantly decreased. In situations with higher environmental noise, the participants' recognition accuracy significantly dropped down. The effect of time pressure on recognition accuracy was not significant ($p = 0.555$). The interaction effect between speech speed and time pressure was marginally significant. No other interaction effects were found.

Effects on operation accuracy: Both speech speed ($p = 0.002$) and environmental noise ($p = 0.005$) were also found to have significant effects on participants' operation accuracy, i.e., with higher speech speed or higher environmental noise, the participants' operation accuracy was significantly lower. The effect of time pressure on operation accuracy was not significant ($p = 0.673$). No interaction effects were found on the participants' operation accuracy.

Effects on completion time: According to the results of non-parametric tests, there were no significant effects of speech speed or environmental noise on the participants' completion time ($ps > 0.1$). The effect of time pressure on completion time was marginally significant ($p = 0.080$).

Comparison of human and machine performance: In this experiment, only if the machine could recognize the speech accurately, could it be adopted to perform the operations, regardless its completion time, since it is always quicker than human. The performance of the machine was stable under each experiment condition. Thus only recognition accuracy of humans and machines was evaluated using the descriptive data, as shown in Table 3. This is very different from that of human participant which had remarkable variations. The results of one-sample t-test showed that with low environmental noise, the average speech recognition accuracy of machines was significantly better than that of humans. In situations with high environmental noise, the average speech recognition accuracy of humans was significantly better than that of machines.

Table 3. The comparison of human and machine recognition accuracy (descriptive data).

Human/Machine	Experiment conditions							
	High time pressure				Low time pressure			
	High speed		Low speed		High speed		Low speed	
	High noise	Low noise	High noise	Low noise	High noise	Low noise	High noise	Low noise
Human	0.515 (0.356)	0.646 (0.340)	0.788 (0.286)	0.830 (0.261)	0.609 (0.366)	0.736 (0.338)	0.679 (0.371)	0.842 (0.269)
Machine	0.000	0.944	0.083	1.000	0.000	0.944	0.000	1.000

CONCLUSION

This study investigates the effects of speech speed and environmental noise on the performance of both humans and machines in the context of civil air traffic control communication. The findings demonstrate that both high speech speed and high environmental noise significantly impaired the participants' recognition accuracy and operation accuracy. Nevertheless, the relative performance of humans and machines exhibits variability. Humans outperform machines in situations with high noise and low speech speed, while machines perform better when dealing with high speech speed and low noise. This insight is noteworthy because it highlights that the relative strengths between humans and machines are not static. A rigid function allocation based on these relative strengths may not always be optimal. Consequently, the development of a dynamic allocation method becomes necessary. However, the current findings are qualitative rather than quantitative. While we prove that the relative strengths of humans and machines fluctuate under different conditions (in this study, speech speed and environmental noise), we cannot precisely determine at what specific values this shift occurs. We propose two potential directions of future research. First, a comprehensive comparison of human and machine performance across various conditions should be conducted. This would help establish the guidelines pertaining to the allocation of functions in these conditions. Second, continuous monitoring of the situational and task changes, as well as the human and machine states, should be implemented. This is beneficial to determine when task/function re-allocation is needed.

ACKNOWLEDGMENT

This study was supported by the National Natural Science Foundation of China (Grant No. T2192932).

REFERENCES

Endsley, M. R. (2017) "From Here to Autonomy: Lessons Learned from Human-Automation Research", Human Factors, 59(1), pp. 5–27.

Feigh, K. M., Dorneich, M. C., Hayes, C. C. (2012) "Toward a characterization of adaptive systems: A framework for researchers and system designers", Human Factors, 54(6), pp. 1008–1024.

Fitts, P. M. (1951) "Human Engineering for an Effective Air-Navigation and Traffic-Control System", Oxford, England: National Research Council, Div. of (Human engineering for an effective air-navigation and traffic-control system), p. xxii, 84.

Older, M., Clegg, C., Waterson, P. (1996) "Report on the revised method of function allocation and its preliminary evaluation", Institute of Work Psychology, University of Sheffield.

Parasuraman, R., Sheridan, T., Wickens, C. (2000) "A Model for Types and Levels of Human Interaction with Automation", IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans: A Publication of the IEEE Systems, Man, and Cybernetics Society, 30, pp. 286–297.

Price, H. E. (1985) "The Allocation of Functions in Systems", Human Factors, 27(1), pp. 33–45.

Riley, V. (1989) "A General Model of Mixed-Initiative Human-Machine Systems", proceedings of the Human Factors Society Annual Meeting, 33(2), pp. 124–128.

Roth, E. M. et al. (2019) "Function Allocation Considerations in the Era of Human Autonomy Teaming'\", Journal of Cognitive Engineering and Decision Making, 13(4), pp. 199–220.

Sarter, N. B., Woods, D. D. (1995) "How in the world did we ever get into that mode? Mode error and awareness in supervisory control", Human Factors, 37(1), pp. 5–19.

Sheridan, T. B., Verplank, W. L. (1978) "Human and Computer Control of Undersea Teleoperators", Massachussetts Institute of Technology, Cambridge, Massachussetts: Man-Machine Systems Laboratory, Department of Mechanical Engineering.