

# The Impact of Time Efficacy on Air Traffic Controller Situation Awareness and Mental Workload

Jiazhong Yang,<sup>1</sup> Esa M. Rantanen,<sup>2</sup> and Kan Zhang<sup>3</sup>

<sup>1</sup>*Flight Technology College, Civil Aviation Flight University of China  
Guanghan, China*

<sup>2</sup>*Department of Psychology, Rochester Institute of Technology,  
Rochester, New York*

<sup>3</sup>*State Key Laboratory of Brain and Cognitive Science, Institute of  
Psychology, Chinese Academy of Sciences  
Beijing, China*

Time efficacy reflects people's perception of control time and evaluation of time management behaviors, and potentially influences situation awareness (SA) and mental workload experienced by air traffic controllers. In this study, a total of 31 air traffic controller trainees, assigned to a high or low time efficacy group, completed 2 scenarios with few or many aircraft simultaneously in the sector. The results showed that the high time efficacy group maintained better SA and experienced lower mental workload. The findings have implications for the selection, training, and work shift arrangements of air traffic controllers.

Time is an essential characteristic of air traffic control (ATC) as well as the principal constraint in the air traffic controller's job, which routinely involves dividing attention among multiple sources of information and performing multiple tasks nearly simultaneously under, at times, severe time pressure. Attention and switching attention require time (Hopkin, 1995), as does performing control actions. The latter are predominantly mediated by voice communications between controllers and pilots and between controllers in different sectors. The highly redundant com-

munication and other procedures followed by controllers—necessary to minimize opportunities for error—further aggravate the temporal demands of ATC. A busy air traffic controller must therefore always be aware of what tasks need to be performed when, how these tasks should be prioritized, and how to deal with possible distractions that could interfere with task completion. This kind of performance requires good situation awareness (SA; Endsley, 1995; Wickens, 2002), and in particular Level 3 SA, or prediction of future events and planning of actions accordingly. Good temporal awareness, which can here be understood as a special case of SA, or Level 3 SA more narrowly defined, might also offer effective means for controllers to safeguard against excessive workload by prioritizing their tasks, shedding tasks that are not critical, and using all of their available time to perform tasks as they unfold in time. It has been shown that controllers indeed mind their time and regulate their workload by employing a set of strategies that are more economical in a demanding situation than under lighter workload (Sperandio, 1978). Furthermore, Rantanen and Nunes (2005) showed that controllers tended to opt for the quickest and least effortful solutions in determining possible conflicts between pairs of aircraft, exhibiting time management with a definitely conservative bias.

Although time management implicitly is necessary in many of the battery tasks (e.g., job sample; Ramos, Heil, & Manning, 2001a, 2001b), the Federal Aviation Administration (FAA) Air Traffic Selection and Training (AT-SAT) test battery does not include any tests specifically measuring time management aptitudes. Some studies have been designed to assess the effect of students' time management on self-reported academic performance (Britton & Tesser, 1991; Macan, Shahani, Dipboye, & Phillips, 1990). These studies, however, were typically not directly relevant to ATC, and might deal with time scales (hours and days) markedly different from those experienced by controllers (seconds and minutes). So far, few empirical studies have attempted to test the relationships between time management and air traffic controllers' mental workload, SA, and performance.

Available means to assess time management predominantly involve subjective questionnaires (Britton & Tesser, 1991; Huang & Zhang, 2001; Macan et al., 1990). Macan et al. (1990) developed the 46-item Time Management Behavior Scale (TMBS) and used it to explore the relationship between college students' academic performance and their time management practice. The results showed that the TMBS consisted of four relatively independent factors accounting for 72% of the common variance and had moderate internal consistency ( $\alpha = .68$ ). The most predictive factor was perceived control of time, which represented a person's perception of control of time and is less "behavioral" than the other three factors; students who perceived having control of their time reported significantly greater evaluations of their performance.

Britton and Tesser (1991) used their Time Management Questionnaire (TMQ) to investigate the effects of time management practices on college grades as well. The principal components analysis revealed three factors, which accounted for

36% of the variance. One of these components, time attitudes, reflected the extent to which one's time was used constructively and who felt in charge of the way his or her time was spent. From their perspective, the time attitudes factor seemed very much like self-efficacy. According to Bandura (1999), such feelings of efficacy are generally associated with better performance. This postulate was supported by the regression analyses showing that time attitudes were the most significant predictor of cumulative grade point average.

Based on the aforementioned two time management questionnaires, Huang and Zhang (2001) developed the Time Management Disposition Inventory (TMDI) for Chinese adolescents in academic settings. The exploratory factor analysis revealed three components: time value, time control, and time efficacy, which accounted for 57.2%, 61.8%, and 54.6% of variance, respectively. Confirmatory factor analysis also revealed that the goodness of fit of the three-dimensional model was better than other unconstrained models. In TMDI, time control involved a series of explicit behaviors, such as planning, setting of goals, task prioritization, time allocation, and results checking, which were commonly considered to be specific time management behaviors taught in the training seminars. Time efficacy reflected a person's perception of control time and evaluation of one's own time management behaviors (Huang & Zhang, 2001). If a person has a strong sense of control over the valued commodity of time and is confidently able to carry out one's time management behavior, then that person is likely to develop a sense of efficacy with respect to time control. Hence, time efficacy is very similar to perceived control of time in TMBS and time attitudes in TMQ.

The assessment of time efficacy in TMDI includes two aspects: efficacy of time management and efficacy of time management behavior. The former includes five items:

1. I think time can be effectively managed.
2. I am regretful for the time that has been not used constructively.
3. I make constructive use of my own time.
4. I efficiently use the class time to study.
5. I feel in control of my time.

The latter also includes five items:

1. My time allocation between study and extracurricular activities is proper.
2. I am full of confidence to attain the goals by time management after I set them.
3. I often fail to achieve the goals at the deadline I set.
4. My plans to perform tasks are proper.
5. The priorities I set to determine the order in which I will perform tasks are proper.

From these items, it appears that perceived control of time in TMBS, time attitudes in TMQ, and time efficacy in TMDI might be counterparts of the sense of self-efficacy, which Bandura (1999) argued generally supports better performance. People who score higher on the time management scales also report a greater amount of self-efficacy for time control.

Calabresi and Cohen (1968) argued that attitudes toward time reflected basic features of the individual personality, and Wessman (1973) maintained that individuals' experience of time and the way people order and structure their lives within the temporal context is "a basic feature of human personality" (p. 103). The subsequent research by Macan et al. (1990), Britton and Tesser (1991), and Huang and Zhang (2001) further suggested that self-reported time management was multidimensional and that it was important to distinguish among the different facets of it. As a distinct dimension of time management, time efficacy probably most reflects the basic features of individual personality in relation to time. Similarly, one should expect large individual differences in time efficacy. These features of time efficacy might make it a useful tool for selection and screening purposes, especially for jobs with clear requirements for effective time management, such as ATC.

## PURPOSE OF THE RESEARCH

Time management skills in academic settings, where tasks can span hours, days, and even weeks or months (Britton & Tesser, 1991; Huang & Zhang, 2001; Macan et al., 1990), however, might be very different from those required in much faster paced task environments, such as ATC. To our knowledge, no research has been reported on the effects of individual differences in time management aptitudes on performance in tasks as fast-paced as ATC. The purpose of our research was to correct this deficiency in the research literature. Due to the demonstrated predictive power of time efficacy on performance in previous research, the objective of this study was to explore whether time efficacy scores from an established inventory (TMDI by Huang & Zhang, 2001) would predict air traffic controllers' time management aptitudes—manifested in experienced SA and mental workload—in realistic settings.

## SUMMARY AND HYPOTHESES

The literature on time efficacy already reviewed reveals a strikingly disparate picture. On one hand, a compelling case can be made for the necessity of good time management skills for operators of complex and dynamic systems (e.g., air traffic controllers). On the other hand, the research on time efficacy has focused on non-

critical everyday tasks, such as management of college studies, over long time periods of days, weeks, and months. Furthermore, the measures of time efficacy consist predominantly of subjective questionnaires. Yet, at least a theoretical link can be forged between the construct of time efficacy as defined in the literature and the performance of air traffic controllers in their operational tasks.

To manage mental workload effectively and to acquire and maintain good SA, controllers need to prioritize tasks according to their relative importance, identify the time required as well as time available for individual tasks, and allocate appropriate time and resources accordingly. If a controller is confident with his or her time management behavior and attains good performance through proper task prioritization, he or she likely develops a strong sense of time efficacy. Conversely, and because SA and mental workload are key factors influencing controller performance, controllers with high time efficacy can be differentiated from those with low time efficacy by the former having fewer operational errors (losses of separation or procedural violations), better temporal awareness manifested in timely performance of tasks, and lower subjective workload ratings. To the best of our knowledge, however, no research has been done to explore the relationship among time efficacy, SA, and mental workload in ATC settings.

This research sought to examine the predictive validity of existing time efficacy questionnaires in dynamic multitask settings, such as ATC; individuals who scored high on such questionnaires should have an advantage over those with lower scores. Specifically, we hypothesized that (a) individuals with high time efficacy scores would have higher SA and lower subjective workload ratings than those scoring low in time efficacy; (b) individuals who score high in time efficacy would be able to better control their taskload in an ATC scenario than those with low time efficacy scores, showing how good time management skills allowed the high scorers to manage their workload more effectively than the low scorers; and (c) individuals with high time efficacy scores would have better performance on the task, mediated by better time management skills than those with low scores and manifested in fewer operational errors (e.g., loss of separation between aircraft) and timely resolution of potential conflicts. If these hypotheses were supported, a case could be made for further research on time efficacy, its assessment, and the use of time efficacy measures in selection of controller trainees.

## METHOD

### Participants

Altogether 51 air traffic controller trainees from the Civil Aviation Flight University of China were recruited for this study. All gave their informed consent to participate in the study voluntarily and were paid ¥10.00 per hour for their time. No

screening tests related to time management and any sort of dynamic task had been administered to them at the time they were recruited as air traffic controller trainees. Although after the trainees graduate and seek formal positions in ATC units some sort of dynamic testing might be used by some employers, the trainees in our study represented a reasonably random sample from the general population in terms of time management skills (apart from a hypothetical effect of aptitude for time efficacy on self-selection; i.e., willingness to seek employment as a controller). The participants had received about 32 hr of ATC simulation training from the same instructor by the time they were recruited for the study.

The participants filled in the time efficacy questionnaire to assess their level of time efficacy. A total of 31 participants were further selected for the study based on their time efficacy scores and assigned to a high time efficacy group ( $n = 16$ ) or to a low time efficacy group ( $n = 15$ ), according to their scores. The criteria for assigning participants to the high and low time efficacy trainees groups were that they scored in the upper 27th percentile and lower 27th percentile on the scale, respectively. All participants were men, with ages ranging from 20 to 24 years ( $M = 22.05$  years).

### Time Efficacy Measure

The time efficacy subquestionnaire used in the study was from the TMDI for Chinese adolescents (Huang & Zhang, 2001). The test included 10 items, each answered on a 5-point Likert scale, and consisting of responses ranging from 1 (*do not agree at all*) to 5 (*completely agree*). The test assessed the two aspects of time efficacy: efficacy of time management and efficacy of time management behaviors. Internal consistency (as measured with Cronbach's  $\alpha$ , a measure of the reliability of a psychometric instrument) were  $\alpha = 0.74$  and  $\alpha = 0.61$ , respectively. Scores in the time efficacy questionnaire can range from 10 to 50, with higher values on the scale corresponding to better time efficacy practices or disposition. Our participants' scores ranged from 30 to 48 ( $M = 39.29$ ,  $SD = 4.03$ ).

### Apparatus: ATC Simulator

A custom-designed ATC simulator was used for the experiments. The simulation program was written in C++ and ran on desktop PCs with 21-in. TFT displays and  $1280 \times 1024$ -pixel resolution. A regular mouse was provided for control inputs. The simulator display mimicked the display system replacement, which is currently replacing the plan view displays in U.S. en route centers. A data link (DL) capability provided controllers a digital communications path between them and the computer-generated targets via an electronic link. The DL employed in the simulation used a variable response time but was probably faster and more reliable than voice communications in real-world ATC task environments. However, the

benefits of the DL system, that is, the controlled variability in pilot-controller communications, far outweighed the cost of somewhat reduced realism in an experimental setting.

The experimental task mimicked the job of air traffic controllers. The ATC simulator presented participants with an en route sector with four airways. Aircraft traveled across the sector along these airways, three of which crossed at three locations. These intersections were possible conflict points between aircraft, in addition to opposite or overtaking aircraft on the same airway. The participants were required to perform five basic tasks to successfully control the traffic: (a) receive handoff of aircraft prior to their entering the sector, (b) monitor for possible conflicts between aircraft (a few were programmed in the scenarios) and change the aircrafts' headings, altitudes, or speeds via DL communications to resolve the conflicts, (c) respond to downlinked messages from aircraft requesting altitude changes, (d) initiate handoffs before the aircraft exited the sector, and (e) transfer communications and the next sector accepted the handoff. The aircraft speeds, altitudes, vertical speeds, and flight dynamics (e.g., turning radius) were realistic. A simulation block lasted for 30 min.

## Design

The independent variables were traffic density (number of aircraft or aircraft count in the sector at any given time throughout the duration of the simulation; a within-subjects variable) and time efficacy group (between-subject variable). The parameters of the scenarios appear in Table 1; the values were calculated from the data yielded by the scenarios run by the participants and demonstrate that the desired manipulation (i.e., traffic density) was also maintained in the actual experiment.

There were four main classes of dependent variables to be examined: SA, subjective workload, actual taskload, and objective performance, measured by the number of true conflicts in the scenario and the time to loss of separation between two aircraft (the time from action taken to resolve a potential conflict to the time of predicted loss of separation).

SA was measured using the three-dimensional SA rating technique (3-D SART; Taylor, 1990). The three-dimensional bipolar scale, demand on attentional resources, supply of attentional resources, and understanding of situation, was subjectively rated on a continuous 100-mm line with the endpoints *low* (0 mm) to *high* (100 mm) by participants at the end of each scenario. An overall score for SA can be computed by using the algorithm  $SA(\text{calc}) = \text{Understanding} - (\text{Demand} - \text{Supply})$ . Higher scores indicated better SA.

Two kinds of subjective workload estimates were recorded. First, a concurrent workload probe (similar to the Air Traffic Workload Input Technique [ATWIT]) appeared adjacent to the main display every 5 min; the participants were required

TABLE 1  
Parameters Used to Calculate Time Required (TR) for Performing the  
Tasks in the Experimental Scenarios and the Nominal Taskload (TL<sub>N</sub>)  
Values for Each Scenario

<i>Task</i>	<i>Few Aircraft</i>	<i>Many Aircraft</i>
No. aircraft	17	30
Mean no. of simultaneous aircraft	4.59	8.39
<i>SD</i> no. of simultaneous aircraft	1.38	2.09
Max no. of simultaneous aircraft	7	11
Median TR Comm_xfer (sec)	0.10	0.10
Median TR conflict (sec)	8.48	3.85
Median TR Downlink (sec)	5.15	12.54
Median TR InitHand (sec)	0.10	0.10
Median TR RecHando (sec)	0.10	0.10
TR (for each aircraft) (sec)	107.57	155.05
TA (= block duration) (sec)	1800.00	1800.00
TR/TA (Nominal taskload, TL <sub>N</sub> )	0.06	0.09

to click on a value that best represented the workload they were experiencing at that moment on a scale from 1 (*low*) to 10 (*high*; Stein, 1985). The mean of six ATWIT ratings from throughout the scenario was computed as the overall measurement of workload in each scenario. Second, the participants completed the NASA-TLX workload survey (Hart & Staveland, 1988) at the end of each scenario. The total workload experienced by participants in the scenario was calculated with the formula

$$\text{workload} = \frac{\sum_{i=1}^6 e_i w_i}{15}$$

where  $e$  is the elements included in NASA-TLX and  $w$  represents the weight of individual element (Bunce & Sisa, 2002).

Actual taskload ( $TL_A$ ) was calculated from the data recorded during participants' performance on the experimental scenarios. Of the total of 12 tasks or events recorded (receive handoff, initiate handoff, transfer communications, altitude-, heading-, and speed menu manipulations, datalink response, data block move, halo toggle, conflict resolution, and workload probe response) all but conflict resolution and workload probe response were included in the sum of times required ( $TR$ ) to perform the tasks. Conflict resolution was excluded because of these records' overlap with the flyout menu (altitude, heading, and speed) manipulation and workload probe response was excluded because it was not part of the controllers' tasks and common to all participants and scenarios. Some actions had



zero duration as they essentially were made up of a single mouse click; these actions were assigned a constant  $TR$  value of 0.1 sec. The  $TL_A$  values were computed for each participant in each condition (scenario) and 5-min (300-sec) epoch by the following formula:

$$TL_A = \frac{\sum_{\text{task}=1}^n TR_n}{TE}$$

where  $TR_n$  is the  $n$ th task in the epoch,  $n$  is the number of tasks in the epoch, and  $TE$  is the duration of the epoch, in this case 300 sec. As we assume that  $TE = TA$ , that is, all the time in the epoch was available to the controller to perform some task,  $TL_A$  is simply a ratio of  $TR/TA$ .

All participants controlled traffic in two 30-min scenarios with different aircraft count using the same airspace. The computer recorded and time-stamped all participants' actions. The simulator also included a conflict detection algorithm that recorded the time to loss of required separation between two aircraft.

## Procedure

Prior to the start of the study, the purpose of the experiment and task to be completed were explained to the participants. They then received a 30-min demonstration and instructions regarding SART, ATWIT, and NASA-TLX. The participants also received a 30-min practice trial. For the formal experiments, the participants completed the SART and NASA-TLX surveys at the end of each scenario. Four or five participants in high or low time efficacy groups participated in the experiment at the same time. As the participants had filled out a number of questionnaires and were administered other aptitude tests prior to the experiment, it is likely that they did not have specific expectations about the goals of the study. The two scenarios (high and low aircraft count) were counterbalanced across participants.

## RESULTS

Repeated-measures analysis of variance (ANOVA), with aircraft count as within-subjects factor and time efficacy participant group as between-subject factor, was used in the analysis of the data. Details of the dependent measures across aircraft count and time efficacy group are presented in Table 2.

### Situation Awareness

The effect of aircraft count on SA ratings was statistically significant,  $F(1, 29) = 63.03, p < .001$ . Participants rated significantly higher in SA under the few aircraft

TABLE 2  
Means and Standard Deviations of Dependent Measures by Aircraft Count  
and Time Efficacy Group

Measures	Few Aircraft				Many Aircraft			
	High TE Group		Low TE Group		High TE Group		Low TE Group	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SART rating	144.87	27.83	99.94	33.97	100.08	21.81	70.56	27.81
Attention demand	23.06	15.65	38.07	20.34	44.75	21.50	61.13	20.67
Attention supply	81.56	8.51	66.27	14.12	70.33	15.90	64.62	12.82
Understanding of situation	86.37	8.06	71.73	10.31	74.50	11.44	67.07	12.54
TLX	2.35	0.83	3.02	0.88	3.30	1.12	4.25	0.89
ATWIT	1.62	0.72	1.88	0.65	2.47	1.37	3.49	1.40
TL_A	0.088	0.050	0.095	0.052	0.108	0.032	0.136	0.085
No. conflicts	1.13	0.74	0.81	0.91	2.33	1.11	3.05	1.18
TLOS (min)	4.14	0.58	4.06	0.56	2.99	0.18	3.03	0.18

Note. TE = time efficacy; SART = situation awareness rating technique; TLX = NASA-TLX index; ATWIT = Air Traffic Workload Input Technique; TL\_A = actual taskload; TLOS = time to loss of separation.  $N = 31$ .

condition than under the many aircraft condition; this result was expected and it also attests to successful manipulation of scenario difficulty through the number of aircraft in them. A significant difference was found between high and low time efficacy groups as well,  $F(1, 29) = 16.99$ ,  $p < .001$ . SA ratings were significantly higher for the high efficacy group (144.87 vs. 99.94 and 100.08 vs. 70.56 in the low and high traffic conditions, respectively). The interaction of aircraft count and time efficacy group did not have a significant effect on overall SA,  $F(1, 29) = 2.72$ ,  $p = .11$ . Figure 1 shows how overall SA and ratings on three components varied as a function of aircraft count and time efficacy group.

The 3-D SART used in this study included three components: demands on attentional resources, supply of attentional resources, and understanding of situation. Further analyses were performed on each component separately. The effect of aircraft count on all three components was significant,  $F(1, 29) = 32.98$ ,  $F(1, 29) = 6.00$ , and  $F(1, 29) = 14.62$ , respectively, all at  $p < .05$ . The few aircraft scenarios exhibited less attentional demand, more attention supply, and better situation comprehension than the many aircraft scenarios.

The separate analyses of the three dimensions of SART revealed significant differences between the low and high time efficacy groups in attention demand ( $M = 50.72$  vs. 33.80), attention supply ( $M = 68.13$  vs. 77.77), and situation comprehension ( $M = 67.84$  vs. 80.80). These differences were also statistically significant,  $F(1, 29) = 7.10$ ,  $F(1, 29) = 5.47$ , and  $F(1, 29) = 12.08$ , respectively, all at  $p < .05$ .

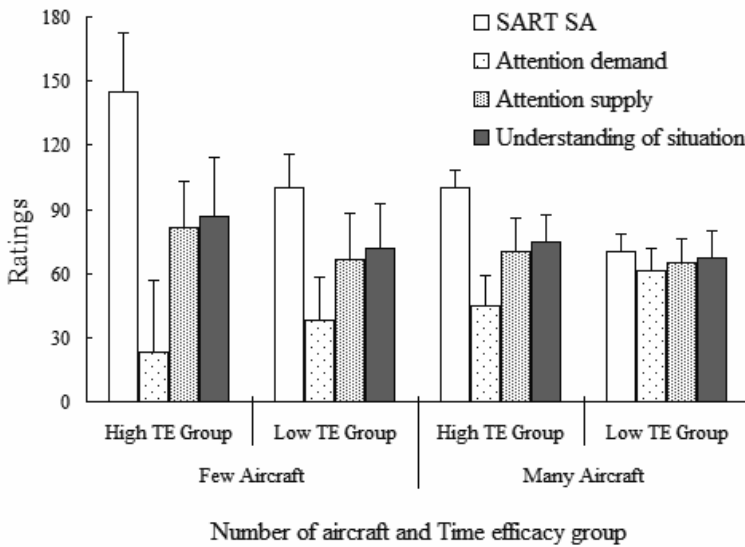


FIGURE 1 The participants in the high time efficacy group rated their situation awareness significantly higher in both traffic load conditions than those in the low time efficacy group. The decrement of situation awareness in the high traffic condition was also significant. The error bars represent standard deviations.

The high time efficacy group experienced less attention demand, more attention supply, and better understanding of situation across few and many aircraft scenarios (see Table 2 and Figure 1). The interactions of aircraft count and time efficacy group did not have significant effect on any of the three SART components (all at  $p > .10$ ).

### Subjective Workload

Subjective workload was measured both with a concurrent workload probe similar to ATWIT (Stein, 1985) and the NASA-TLX probe administered at the end of each experimental scenario. The effect of aircraft count on NASA-TLX ratings was significant,  $F(1, 29) = 55.91, p < .001$ . The TLX workload ratings were also significantly lower for the high time efficacy group,  $F(1, 29) = 7.15, p < .05$ , with mean scores of 2.35 vs. 3.02 and 3.30 vs. 4.25 in the low and high traffic conditions, respectively. Interaction between traffic condition (aircraft count) and time-efficacy group was not significant (see Figure 2).

The ATWIT ratings showed a nearly significant interaction between traffic condition and group,  $F(1, 29) = 4.06, p = .053$ . The main effect of traffic condition was significant,  $F(1, 29) = 41.47, p < .001$ , higher workload ratings were given in

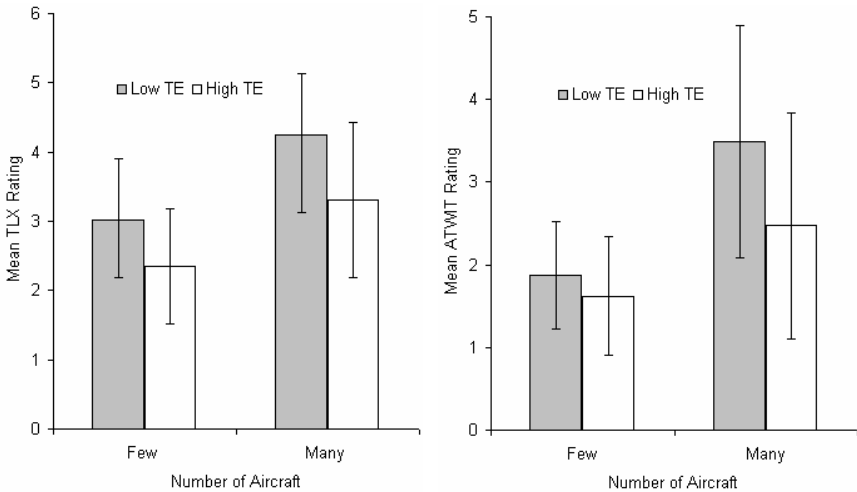


FIGURE 2 Both NASA–TLX and concurrent (ATWIT) subjective workload ratings were significantly higher in the high traffic load condition than in the low traffic condition. NASA–TLX ratings were also higher for the low time efficacy group than for the high time efficacy group. The error bars represent standard deviations.

the many aircraft scenarios, and the effect of time-efficacy group approached significance,  $F(1, 29) = 3.46, p = .073$ . To evaluate these results further, we performed one-way ANOVAs on group separately for each traffic condition. In the few aircraft condition time-efficacy group membership was not significant,  $F(1, 29) = 1.05, p = .313$ , with mean workload ratings low for both groups ( $M = 1.88$  and  $M = 1.62$  for low and high time efficacy groups, respectively). However, in the many aircraft scenarios the difference between the groups was statistically significant,  $F(1, 29) = 4.23, p < .05$ , with mean ratings of 3.49 and 2.47 for the low and high time efficacy groups, respectively (see Figure 2). Note also that these ratings were in the lower end of the 10-point workload scale.

### Actual Taskload

The results show an anticipated trend in that  $TL_A$  was significantly higher in the many aircraft scenarios than in the scenarios with fewer aircraft,  $F(1, 29) = 6.85, p < .05$ , showing that the taskload manipulation indeed was successful. The low time efficacy group also exhibited higher  $TL_A$  values than the high time efficacy group, and the difference between the groups was larger in the many aircraft scenarios than in the fewer aircraft scenario, but this effect was not significant, nor was the interaction between time efficacy group and traffic load ( $F < 1$ ). Nevertheless, these results suggested that participants who scored high on time efficacy could moderate their

taskload better than those who scored low in time efficacy. Figure 3 shows how actual taskload varied as a function of aircraft count and time efficacy group.

Yet another objective indicator of workload is the response latency to the ATWIT probes, longer response times implying higher workload than faster response times. There were no differences in response times between the traffic conditions, but the difference between time efficacy groups was quite substantial, with the low time efficacy group's average response times over 2.5 sec slower than those of the high time efficacy group ( $M$  across traffic conditions = 4.92 sec and 2.23 sec, respectively); this difference was also statistically significant,  $F(1, 29) = 32.02, p < .05$ .

### Objective Performance

There were significantly more loss of separation incidents in many aircraft scenario ( $M$  across groups = 3.16) than in the few aircraft scenarios ( $M = 0.96$ ),  $F(1,$

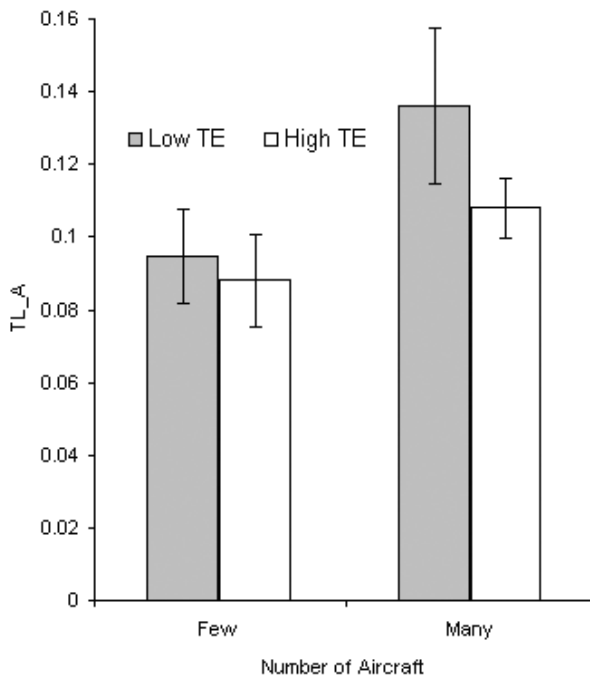
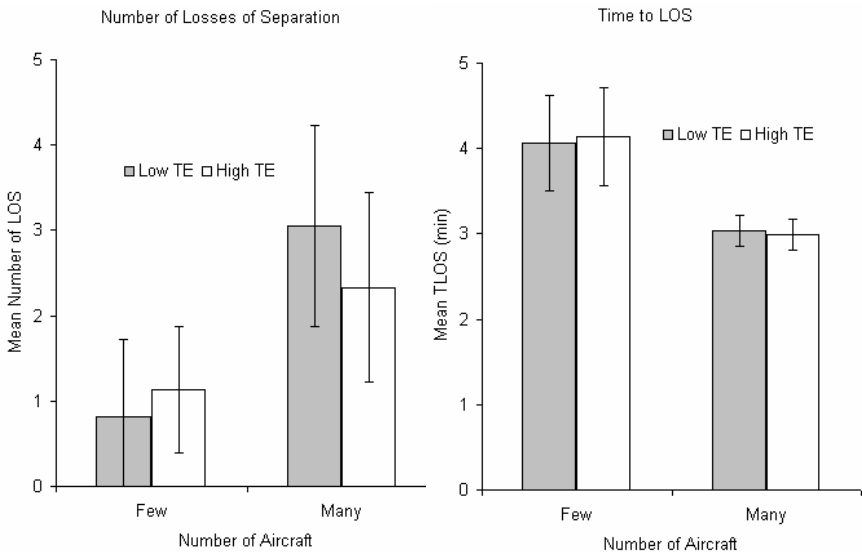


FIGURE 3 Actual taskload was significantly higher in the many aircraft scenario, attesting to successful taskload manipulation. Participants in the high time efficacy group were also able to manage their taskload better (keep it lower) than those who scored low in time efficacy, and this difference was greater in the busier scenarios. The error bars represent standard deviations of the means.

29) = 51.01,  $p < .001$ . No significant differences were found between the time efficacy groups ( $F < 1$ ). As to the time to loss of separation, the effect of aircraft count was again statistically significant,  $F(1, 29) = 95.38$ ,  $p < .001$ , and indicated that the participants resolved conflicts earlier in the few aircraft condition. Neither the effect of time efficacy group nor the interaction were significant ( $F < 1$ ), however. Figure 4 shows how ATC performance varied as a function of aircraft count and time efficacy group.

## DISCUSSION

This study showed that aircraft count significantly affected participants' SA, mental workload, and performance. Specifically, higher SA and lower workload ratings as well as better performance in terms of timelier resolution of conflicts and smaller numbers of loss of separation incidents were observed in the few aircraft condition. These results were, of course, expected, attesting to successful manipulation of traffic density as an independent variable, and in line with much previous



**FIGURE 4** There were significantly more loss of separation (LOS) incidents in the high traffic load condition than in the fewer aircraft scenario, attesting to the difficulty of the former. The low time efficacy group also incurred more LOS incidents than the high time efficacy group, but the difference was not significant. No between-group differences in the time to LOS (TLOS) were found, but all participants were significantly less timely in commanding conflict resolution maneuvers under high traffic load.

research (e.g., Kopardekar & Magyarits, 2003; Manning, Mills, Fox, Pfleider, & Mogilka, 2001). Some researchers have proposed that mental workload can be expressed in terms of the time required for information processing versus time available to do so (Hendy, Liao, & Milgram, 1997; Moray, Dessouky, Kijowski, & Adapathya, 1991). With the increase of aircraft count in the sector, the number of tasks, such as aircraft receive and handoff, and communication also increased. Moreover, the frequency and difficulty of conflict detection and resolution increased in the many aircraft condition. Therefore, time required under the many aircraft condition was far more than in the few aircraft condition, which was clearly reflected in the  $TL_A$  metric.

When the time available to the participants across conditions was constant, they should have higher actual taskload and experience higher workload in the many aircraft scenario, and this indeed was the case. Similarly, the participants rated their SA higher in the few aircraft condition. Lower mental workload and correspondingly higher SA in turn should result in better task performance in the simulated ATC scenarios. This relationship was borne out as well, evidenced by fewer conflicts and timely conflict resolution. Hence, although the theoretical relationships among taskload, workload, and SA might appear straightforward, it was nevertheless important that these relationships were also empirically demonstrated in realistic ATC simulations.

Most important, this study showed that high time efficacy, measured by the TMDI questionnaire (Huang & Zhang, 2001) and presumably predicting good time management practices, indeed resulted in significantly improved SA and lower experienced mental workload for the participating air traffic controller trainees. In ATC, concurrent tasks can unfold at the same time, resulting in high time pressure. Controllers need to use available time effectively. People with high time efficacy exhibit characteristically good time management practices, such as goal setting, task prioritization, time allocation, progress monitoring, and confidence in their ability to complete tasks in a given time. Good task awareness and effective workload management theoretically require such aptitude and skills. Our findings of significantly different SA and workload rating between the two time efficacy groups support our hypotheses. There were also significant differences in all three SART components between the two groups. Specifically, the high time efficacy group experienced less attention demand, more attention supply, and better situation comprehension in the scenarios.

Our results also indicated that although the effect of time efficacy on actual taskload was not statistically significant, an anticipated trend was nevertheless observable in that actual taskload was higher for the low time efficacy group than the high time efficacy group, suggesting that participants with high time efficacy and good time management skills managed their taskload more effectively than those low in time efficacy.

With respect to the time to loss of separation, no statistically significant differences were revealed between time efficacy groups. Less time theoretically indicates that participants take actions to resolve a potential conflict earlier, and low possibility of conflict occurrence. The putative relationship between time efficacy and timeliness of conflict resolution actions is not very clear, however, for there are no theoretically plausible reasons why time efficacy should influence controllers' performance in this task either way. Also, we do not have enough evidence to say whether small time to loss of separation reflects good or poor performance, but the fact that all participants solved conflicts in a timelier manner in the few aircraft than many aircraft scenarios suggest that it was indeed their preference, or that under lighter taskload when performance can be reasonably expected to be good, the participants solved conflicts sooner rather than later. Whether the participants changed their strategy under heavier taskload is a valid hypothesis, but one that we cannot test at this time given our data. We therefore conclude that individual preferences and control techniques not dependent on time management overrode any influence of time efficacy. These metrics are also rather coarse and have limited sensitivity in experimental simulations that include only a relatively small number of aircraft to be separated in scenarios of relatively short duration. However, the fact that participants in the high time efficacy group committed fewer conflicts in the many aircraft count condition than their low time efficacy counterparts suggested that their overall performance might have benefited from high time efficacy and better time management skills through better SA, lower experienced workload, and better taskload management.

Participants in our study had the very same level of ATC expertise due to their identical number of hours in training; furthermore, no psychological tests related to time management were administered to them when they were recruited as air traffic controller trainees and hence did not influence their selection into training. Therefore, the significant differences in SA and subjective mental workload ratings between the two groups can be attributed to the differences in time efficacy. The weaker results for actual taskload, conflict occurrences, and time to loss of separation might point to insufficient sensitivity of these metrics. Furthermore, as was stated earlier, available time management questionnaires, including the TMDI used in this research, were developed for and validated with college students, whose tasks were on a completely different time scale from those of controllers (Britton & Tesser, 1991; Huang & Zhang, 2001; Macan et al., 1990). Aptitudes and traits for time management in fast-paced tasks like ATC certainly warrant further research, as does development of better tools for their assessment among the general population. Although the TMDI time efficacy questionnaire might have been only partially applicable to fast-paced ATC tasks, the results nevertheless strongly suggest that time efficacy plays an important role in human performance, particularly in complex and dynamic task environments.



## CONCLUSIONS

Despite the discrepancies between the TMDI questionnaire used to assess time efficacy and the pace and time scale of ATC tasks, our results clearly demonstrated that time efficacy had an effect on air traffic controllers' subjective SA and mental workload, as well as on objectively measured performance (taskload and number of losses of separation). These results have three major implications. First, it might be useful to develop explicit training modules in controller training programs that teach how to improve time efficacy and time management skills for enhanced SA, decreased experienced workload, and increased performance. Our results also suggest that it might be possible to select individuals with high time efficacy as air traffic controllers. Establishing the predictive validity of time efficacy questionnaires for screening purposes naturally requires much more work, possibly involving development of a time efficacy scale based on ATC work behavior. Further analyses of work sample behaviors might also provide better measures of ATC time management and be helpful for establishing appropriate time management training modules. Finally, our results provide a strong impetus for investigation of time efficacy in general, individual aptitudes for it, and whether and how it could be trained. Such work would be generalizable to a wide variety of supervisory control tasks and jobs, where prioritization of tasks and timing of actions are critical to the system performance and safety.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under collaborative Agreement No. 30700237/C010805. All views and conclusions presented here are those of the authors and do not necessarily represent the positions of the sponsor.

## REFERENCES

- Bandura, A. (1999). Human agency in social cognitive theory. *American Psychologist*, *44*, 1175–1184.
- Britton, B. K., & Tesser, A. (1991). Effects of time-management practices on college grades. *Journal of Education Psychology*, *83*, 405–410.
- Bunce, D., & Sisa, L. (2002). Age differences in perceived workload across a short vigil. *Ergonomics*, *45*, 949–960.
- Calabresi, R., & Cohen, J. (1968). Personality and time attitudes. *Journal of Abnormal Psychology*, *73*, 431–439.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*, 32–64.

- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLS (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North-Holland Elsevier Science.
- Hendy, K. C., Liao, J., & Milgram, P. (1997). Combining time and intensity effects in assessing operator information-processing load. *Human Factors*, 39, 30–47.
- Hopkin, V. D. (1995). *Human factors in air traffic control*. London: Taylor & Francis.
- Huang, X. T., & Zhang, Z. J. (2001). The compiling of adolescence time management disposition inventory. *Acta Psychologica Sinica*, 33, 338–343.
- Kopardekar, P., & Magyarits, S. (2003). Measurement and prediction of dynamic density. In *Paper presented at the 5th USA/Europe Air Traffic Management R & D Seminar*, Budapest, Hungary. Available at [http://www.tc.faa.gov/acb300/techreports/DD\\_ATM2003\\_7-03.pdf](http://www.tc.faa.gov/acb300/techreports/DD_ATM2003_7-03.pdf). (Accessed on March 6, 2008).
- Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlation with academic performance and stress. *Journal of Education Psychology*, 82, 760–768.
- Manning, C. A., Mills, S. H., Fox, C., Pfleider, E., & Mogilka, H. J. (2001). *Investigating the validity of performance and objective workload evaluation research (POWER)* (DOT/FAA/AM-01/10). Washington, DC: Federal Aviation Administration.
- Moray, N., Dessouky, M. I., Kijowski, B. A., & Adapathya, R. (1991). Strategic behavior, workload, and performance in task scheduling. *Human Factors*, 33, 607–629.
- Ramos, R. A., Heil, M. C., & Manning, C. A. (2001a). *Documentation of validity for the AT-SAT computerized test battery: Volume I* (DOT/FAA/AM-01/5). Washington, DC: Office of Aviation Medicine.
- Ramos, R. A., Heil, M. C., & Manning, C. A. (2001b). *Documentation of validity for the AT-SAT computerized test battery: Volume II* (DOT/FAA/AM-01/6). Washington, DC: Office of Aviation Medicine.
- Rantanen, E. M., & Nunes, A. (2005). Hierarchical conflict detection in air traffic control. *International Journal of Aviation Psychology*, 15, 339–362.
- Sperandio, J. C. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics*, 21, 195–202.
- Stein, E. S. (1985). *Air traffic controller workload: An examination of a workload probe* (DOT/FAA/CT-TN84/24). Atlantic City, NJ: Federal Aviation Administration.
- Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness in aerospace operations (AGARD-CP-478)* (pp. 3/1–3/17). Neuilly Sur Seine, France: Advisory Group for Aerospace Research and Development.
- Wessman, A. E. (1973). Personality and the subjective experience of time. *Journal of Personality Assessment*, 37, 103–114.
- Wickens, C. D. (2002). Situation awareness and workload in aviation. *Current Directions in Psychological Science*, 11(4), 128–133.

Manuscript first received: May 2008