2005

Experimental Analysis of Task Prioritization Training for a Group of University Flight Technology Students

Amy L. Hoover

Kenneth H. Funk II

EXPERIMENTAL ANALYSIS OF TASK PRIORITIZATION TRAINING FOR A GROUP OF UNIVERSITY FLIGHT TECHNOLOGY STUDENTS

Amy L. Hoover
Central Washington University
Ellensburg, Washington

Kenneth H. Funk II
Oregon State University
Corvallis, Oregon

The purpose of this study was to evaluate changes in task prioritization performance between pilots who participated in a CTM training course and those who did not. A pretest-posttest control group design with random assignment was used. Pilots enrolled in the Central Washington University Flight Technology Program flew pretest and posttest simulated flights on a Frasca FTD. During a two week period between pretest and posttest simulated flights pilots in the experimental group participated in a CTM training course and pilots in the control group did not. Comparison of pre- and posttest error rates shows the experimental group had a 54% decrease in task prioritization errors and the control group had a 9% increase in errors.

Introduction

Pilots routinely perform multiple, concurrent tasks and the ability to effectively prioritize them for attention is a critical flying skill. Although pilots clearly understand this and generally practice concurrent task management (CTM) well, there are many instances in which failure to properly prioritize tasks or otherwise manage them effectively has led to a potentially dangerous incident or even a fatal accident (Chou et al, 1996).

Short-term memory appears to be a major limiting factor in CTM performance, so it is not surprising that a computational aid to augment human memory facilitated CTM performance in a low-fidelity flight simulator experiment (Funk and Braune, 1999). But technological limitations and other practical considerations strongly suggest that other means of improving CTM be explored, notably the training of CTM skills, including that of prioritizing tasks.

Bishara and Funk (2002) developed and evaluated a short (two-hour) CTM training module for general aviation pilots. In a pretest-posttest control group experiment, participants who received CTM training showed improvement in prospective memory performance. But results relating to task prioritization, a more general subskill, were ambiguous. This may have been due to several factors, including the quality of the training material (not developed by qualified flight instructors), the low fidelity of the simulator (Microsoft Flight Simulator was used), a small sample size (12), and the heterogeneity of the participants (reflecting a wide range of experience and skill). Although CTM performance is a significant factor in flight safety, the trainability of CTM, until now, has been in question.

Objective

The objective of this study was to carefully develop and evaluate CTM task prioritization training in a higher fidelity experimental environment using a more homogeneous population of participants.

Method

A pretest-posttest control group design was used. All participants flew a one hour simulated instrument flight on a Frasca 141 FTD (pretest) then flew another simulated flight two weeks later (posttest). The experimental group participated in a CTM training course during the two week interim and the control group did not.

Participants

Twenty-seven pilots enrolled in the Central Washington University Flight Technology Program participated in the experiment. Participants were randomly assigned to either the experimental or control group. All pilots had logged previous instrument time on the FTD used in the experiment. Regression analysis showed no correlation between participants' total flight time, instrument time, stage of training, total FTD time, and Frasca 141 FTD time with regards to CTM performance on the pretest, indicating the two groups were equivalent.

Flight Training Device

Two identical Frasca 141 FTDs were used and were configured as normally aspirated single engine fixed gear aircraft. The Avionics package included audio panel with marker beacons, dual VHF communication and navigation radios, DME, ADF, and a Garmin GNS430 IFR enroute and approach certified GPS/comm. The FTDs recorded all primary flight data including aircraft heading, altitude, airspeed, power settings, and position.

Procedure

Pre- and posttest simulated flights were conducted in a line oriented flight training (LOFT) format. The LOFT placed pilots in a high workload environment in Seattle Class B airspace and included radar vectors as well as pilot navigation, two precision instrument approaches, a multistage missed approach, and a holding procedure. Pilots conducted the simulated flights as per the CWU Standard Operating Procedures (SOP) manual; all checklists, flow checks, and callouts were the same used in their normal flight training.

Certified instrument flight instructors (CFIIs) were trained to administer the LOFT which was scripted with respect to air traffic control (ATC) communications and procedures. Flights were observed and scored in real time and again from videotape by a second scorer. Video cameras recorded a wide angle view that included the entire instrument panel, engine controls, yoke, rudder pedals, and pilots' hands and feet.

Prioritization scheme used

A task prioritization scheme taught to pilots during primary and advanced flight training is the aviate, navigate, communicate (ANC) hierarchy (Chappell, 1998; FAA, 1999; Jeppesen, 2001, 2003a, 2003b; Kern, 1998; Kershner, 1998; Machado, 2001, 2003; Thom, 1991). For this study each task was defined based on pilot training manuals and literature as follows:

Aviate task: Included all items related to aircraft operation: airspeed, altitude, climb or descent rate, lift, thrust, and drag; e.g. primary aircraft control inputs (pitch, power, yaw, and roll), operation of lift and drag devices (flaps) and operation of primary engine systems.

Navigate task: Included items related to the current and future position of the aircraft, including vectors, course intercepting and tracking, identification of intersections and waypoints, and programming and operating the GPS and other navigation radios.

Communicate task: Included communications with ATC.

Definition of CTM Errors

Opportunities for twenty potential task prioritization errors were embedded at 14 challenge points throughout the one hour simulated flights. Challenge points were based on errors observed during a pilot study conducted prior to the experiments. Each challenge point provided an opportunity for the participant to divert his/her attention from a more important or more urgent task to a less urgent or less important task. Associated with each challenge point were specifications as to what actions would constitute which type of prioritization error. Types of prioritization errors included ignoring an aviate (flight control) task in order to navigate (aviate/navigate, 7 opportunities), aviate/communicate (7), navigate/communicate (5), and aviate/aviate (1) in which the pilot had to choose between two aviate tasks as to which was most critical to perform first.

Several of the challenge points were simply part of the LOFT scenario; they were embedded at a point where a pilot might make a task prioritization error and thus did not require any intervention. For example, challenge points were placed at locations in the flight scenario where there was potential for error if the pilot fixated on or became distracted by a navigate task at the expense of primary aviate tasks. Other challenge points required the CFII to act as ATC and call the pilot with information or instructions just before the pilot was leveling off or about to intercept course, or to cause a failure to a navigational facility or an aircraft system.

Performance criteria for determining if an error occurred was based on FAA-S-8081-4C Instrument Rating Practical Test Standards with respect to altitude, airspeed, heading, intercepting and tracking course, use of checklists, procedures, and ATC communications.

CTM Training Course

The training course followed standard practices and procedures common to the CWU training course outline (TCO) and university criteria for learning outcomes and assessment strategies. The course was taught by an FAA certified CFII and CWU flight technology professor. It consisted of two sessions 7 days apart that included reading, self-study, cooperative learning activities, guided discussion, and a reflective homework assignment. The course also emphasized procedural discipline with respect to task prioritization, including proper use of checklists, standard operating procedures, mnemonic aiding devices, situational awareness, and cockpit flow checks.

The first learning session consisted of a class discussion of selected materials related to aviation human factors, aeronautical decision making, situational awareness, workload management, and concurrent task management. Participants had prior knowledge of all those concepts from previous coursework and studies, thus the training did not introduce any new concepts but rather emphasized task prioritization as an important element of human factors and aeronautical decision making. Participants analyzed accident and incidents taken from the NTSB and NASA databases with respect to CTM errors and participated in class discussions of those data.

During the time between sessions participants were asked to reflect on at least one of their normal flights with respect to CTM concepts and how their awareness influenced their in-flight decision making. Students reflected in writing as well as through a verbal debriefing.

The second class session included an activity in which participants acted out role-playing scenarios designed to give insight into their reactions and behavior in the cockpit when confronted with CTM challenges. They also participated in a class discussion of strategies to improve pilot task prioritization performance and a guided discussion of the outcomes. A short quiz was given at the end of the second session to evaluate each pilot's progress and identify areas of improvement.

Results

CTM error data were recorded as a frequency distribution of raw scores and converted to a ratio score (number of errors: total number possible) for comparison. Table 1 and Figure 1 present CTM error scores for experimental and control groups. The control group showed a 9% increase in total CTM errors, and the experimental group showed a 54% decrease in total errors.

Table 1. *Task prioritization error rates for each group. Mean scores are shown with standard deviation in parentheses.*

Group	Pretest	Posttest
Experimental	0.24 (0.12)	0.11 (0.08)
control	0.23 (0.15)	0.25 (0.10)

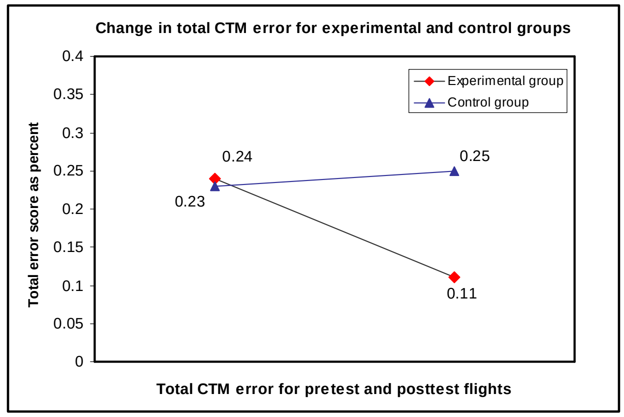

Figure 1. *Graph showing the change in total CTM error scores for each group expressed as a percent of total possible errors.*

There were 14 in the experimental group and 13 in the control group. An F-test for homoscedasticity found the samples had equivalent variance and K-S test and Q-Q plot showed they were normally distributed, so an independent samples t-test was used to compare the two groups. Because data showed a posttest reduction in CTM errors for the experimental group compared to the control group a one-tailed test was used yielding a t = 2.67 at p = 0.007 (Table 2).

Table 2. *Independent samples t-test*

t-test for equality of means				
		t (t critical)	*df*	Significance (1-tailed)
Posttest-pretest differen ce	Equivalent variances assumed	2.67 (1.71)	25	0.007
	Equivalent variances not assumed	2.68 (1.71)	24.3	0.006

Discussion

Results show the control group made the same or more prioritization errors overall in the posttest flight compared to the pretest; individual pilots showed an increase, a decrease, or no change in errors. Such a distribution would be expected from randomly sampling a group of pilots during two discreet flights. If there were no effect from the CTM training course then the experimental group should show a similar distribution of pretest and posttest scores. However, the experimental group had a much larger decrease in total CTM errors between pretest and posttest flights compared to the control group.

It seems reasonable that any well designed training course would show some effect during the short term, but a major question that arises is; how long will it take that effect to disappear, or to drop below acceptable performance standards? The answers to those questions would need to be assessed by testing the same participants at a later date, as well as controlling for effects of extraneous variables that might affect their performance.

The amount of time between the pretest and posttest simulated flights (2 weeks) represented a trade-off between internal and external validity. The time period was kept short enough to reduce history effects, but that did not allow the study to comment on longer term effects of the training. For a longer period of time, control for extraneous variables, including further training in human factors and additional flight experience, might be difficult. However, pretest data indicated no correlation between this particular group of participants' total flight time, instrument time, or FTD time and their CTM performance, so controlling for the influence of such extraneous variables might be a reasonable possibility.

A related question is whether or not any learning actually took place; pilots who received training showed a decrease in CTM errors and an improvement in performance over a two week period of time, but it is not known from this study whether they actually retained the new information or learned new behaviors that will endure.

Since all pilots in the experiment had previously studied concepts of prioritization and task management during their regular flight training, it is possible that the reduction in CTM errors by the experimental group might represent a sensitization effect; the only difference between the two groups could have been that the experimental group was focused on those concepts during the short term and did not actually code the information into their long term memory.

The issue of whether learning occurred is a critical one and also difficult to resolve because a teacher or instructor often does not have the ability to evaluate students after they leave the learning environment. More follow-up studies are needed to comment on the long term effects of the training. Additionally, a qualitative response from participants at some future time might also reflect on whether or not they felt learning occurred.

Pilots in the experimental group who showed the greatest reduction in CTM error scores were the ones that originally made the most errors. Thus it could be that the reduction in errors might simply represent a regression toward the mean for those pilots. However, the fact that several pilots in the control group also scored a large number of errors in the pretest without a corresponding reduction in errors for the posttest indicates that regression was probably not the cause for that trend in the experimental group.

What the data does suggest is that pilots who performed the worst seemed to benefit more from the training than those who initially made a low number of errors. Alternatively, pilots who made only one or two errors in the pretest and posttest were not able to be evaluated with respect to a training effect since there were only a fixed number of challenge points and it was not possible to show a large improvement in error scores for those pilots.

One error that more than half the pilots in both groups made involved a missed approach procedure (MAP) that called for the pilot to climb via the localizer course to 2000 feet, then to identify a specific intersection as the point to commence climb to 5000 feet while continuing to track the localizer course. Many of the pilots became fixated on the task of either programming the GPS for the waypoint or tuning and identifying the VOR to identify the cross radial for the intersection and either strayed off course, deviated from altitude, or both, while attempting to identify the fix. In several cases the video tapes showed pilots were not even looking at their flight instruments while operating the GPS unit. A few pilots were off altitude by as much as 500 feet and off the localizer course by a full needle deflection as a result of their fixation on the navigate task.

The issue of fixation has become an area of great concern in the flight training industry in recent years; over the past 5 years general aviation cockpits have incorporated more sophisticated IFR certified GPS units, and in the past 3 years flat panel primary flight displays (PFDs) and multifunction displays (MFDs) have been installed in training aircraft.

Wilson (1998) found that as the level of sophisticated instruments and automation increases on airline flight decks the potential for CTM errors also increases. Also, in a more general meta-study of airline flight deck human factors issues, Funk et al (1999) found the attentional demands of automation to be problematic. It is likely that the same potential exists for increased sophistication in general aviation cockpits, including training aircraft. Pilots who preprogrammed the GPS while still on the ground at did not make that fixation error.

Conclusions and Recommendations

Experimental analysis showed that the group of university flight students who participated in the CTM training course improved their task prioritization performance over a two week period of time. The decrease in task prioritization errors for pilots in the experimental group did not seem to be a result of regression toward the mean. It is not certain whether that performance increase had a longer lasting effect.

Pilots who did not participate in the CTM short course did not markedly improve in their prioritization performance; they either showed an increase, decrease, or no change in performance.

One particular error that emerged was that of pilots fixating on the GPS display to the exclusion of aircraft control, sometimes showing dangerously large deviations in altitude and course. Fixation errors are of critical importance in the current flight training environment as modern cockpits utilize more sophisticated displays and avionics.

Based on the findings from this study, the following recommendations for future research are presented:

- The same experiments could be conducted with students at a different flight school to enhance external validity.

- The experiment could be conducted with a longer time period between pretest and posttest flights and controlled for extraneous variables to test for long term training effects.

- A time-series design could be used to determine longer term training effects.

- A regression-discontinuity design might describe training effects for pilots who initially performed lower and those that performed higher.

- A qualitative study using responses from participants could comment on the extent of learning that occurred.

- A larger sample size could be used to enhance external validity.

- A study could be designed to test pilots in cockpits with various levels of complexity, for example using one of the many new general aviation flat panel PFD/MFD or virtual 3D displays installed in many new aircraft.

- A cost to benefits analysis could be conducted to determine if task prioritization training should be incorporated as a component of a training course outline.

References

Bishara, S. and K. Funk (2002). Training pilots to prioritize tasks, *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, Baltimore Maryland, September 30 – October 4, 2002, 96-100

Chappell, S. L. (1998). Disciplined attention, the next best thing to a crystal ball. In T. Kern, *Flight discipline*. New York: McGraw-Hill.

Chou, C.D., D. Madhavan, and K.H. Funk (1996). Studies of cockpit task management errors, *International Journal of Aviation Psychology*, Vol. 6, No. 4, 307-320.

Federal Aviation Administration (1999). *Aviation instructor's handbook* (FAA-H-8083-9). Washington D.C.: U.S. Government Printing Office.

Funk, K., B. Lyall, J. Wilson, R. Vint, M. Niemczyk, C. Suroteguh, and G. Owen (1999). Flight deck automation issues, *International Journal of Aviation Psychology*, Vol. 9, No. 2, 109-123.

Funk, K. and R. Braune (1999). *The AgendaManager: A knowledge-based system to facilitate the management of flight deck activities*, SAE 1999-01-5536, SAE International, Warrendale, PA, presented at 1999 World Aviation Congress, San Francisco, CA, October 19-21, 1999.

Jeppesen (2002). *Flight instructor manual*. Englewood, CO: Author.

Jeppesen (2003a). *Instrument commercial manual*. Englewood, CO: Author.

Jeppesen (2003b). *Private pilot manual*. Englewood, CO: Author.

Kern, T. (1998). *Flight discipline*. New York: McGraw-Hill.

Kershner, W. K. (1998). *The instrument flight manual* (5th ed.). Ames, IA: Iowa State University Press.

Machado, R. (2001). *Rod Machado's Private Pilot Handbook.* San Clemente, CA: The Aviation Speaker's Bureau.

Machado, R. (2003) *Rod Machado's instrument pilot's survival manual*. San Clemente, CA: The Aviation Speaker's Bureau.

Thom, T. (1991). *The pilot's manual: volume 2 flight operations.* New York: Center for Aviation Theory.

Wilson, J. R. (1998). *The effect of automation on the frequency oft task prioritization errors on commercial aircraft flight decks: An ASRS incident report study*. Unpublished master's thesis, Oregon State University, Corvallis.