2015

Modeling Task Prioritization Behaivors in a Time-Pressured Multitasking Environment

Takeaki Toma

Kenneth H. Funk II

Repository Citation

Toma, T., & Funk, K. H. (2015). Modeling Task Prioritization Behaivors in a Time-Pressured Multitasking Environment. *18th International Symposium on Aviation Psychology*, 446-451.
https://corescholar.libraries.wright.edu/isap_2015/31

MODELING TASK PRIORITIZATION BEHAIVORS IN A TIME-PRESSURED MULTITASKING ENVIRONMENT

Takeaki Toma
School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University
Corvallis, Oregon
Kenneth H. Funk II
School of Mechanical, Industrial, and Manufacturing Engineering, Oregon State University
Corvallis, Oregon

Cockpit task management (CTM) theory is structurally consistent with cognitive multitasking models. Based on the CTM framework, it is hypothesized that aviation task prioritization behavior in human multitasking may be influenced by importance, urgency, performance status, salience, and workload of tasks in a cockpit. A middle fidelity flight simulation study was conducted to test the above hypotheses. Questionnaire data indicated that the perceived task importance, the perceived task urgency and the perceived task salience had significant relationships with the perceived task priority after taking the individual difference and flight situational difference into account. The perceived task priority was related to the task execution time and task performance, but not correlated with task awareness level in the flight simulation.

Introduction

Human multitasking in transportation can be dangerous. A recent National Safety Council (NSC) white paper (2010) reported that 25% of all car crashes were caused by the use of a mobile device while driving, and an estimated 1.4 million crashes and 645,000 injuries were related to multitasking. Relatedly, many aviation incidents and accidents are reported to occur during multitasking. Chou, Madhavan, and Funk (1996) reviewed National Transportation Safety Board (NTSB) aircraft accident reports and NASA Aviation Safety Reporting System incident reports using cockpit task management (CTM) theory (Funk, 1991); they reported that 23% of aviation accidents and 49% of aviation incidents were rooted in CTM errors.

One reason for dangerous multitasking in transportation operations might be the difficulty of task prioritization. A well-known example is the crash of Eastern Air Lines Flight 401 on December 29, 1972 in the Florida Everglades. When the airplane was approaching the Miami airport, the pilots noticed that a landing gear indicator light did not turn on. The pilots communicated with the approach controller who gave a clearance to maintain an altitude of 2000 feet and hold west over the Everglades. The cockpit crew proceeded to put the plane in autopilot control and believed that the airplane was holding at an altitude of 2000 feet. However, the pilots did not notice that the autopilot setting had changed (probably due to a pilot mistakenly moving the yoke) and the airplane gradually descended towards the ground (NTSB, 1973). The root cause of the accident can be interpreted as a task prioritization error because the pilots wrongfully prioritized their attention to the landing gear problem instead of controlling the airplane; the pilots did not pay attention to the altitude because they focused too much on the landing gear problem (NTSB, 1973). This incident raises several questions: Why did the pilots prioritize diagnosing the landing gear problem higher than controlling the airplane? What is the mechanism behind task prioritization in aviation multitasking environments? We hypothesized that the aviation task prioritization process is influenced by five factors: task importance, task urgency, task status, task salience, and multitasking workload based on the antecedent study conducted by Colvin, Funk and Braune (2005). The objective of this research was to test the above hypotheses.

Potential Task Prioritization Factors

How should pilots manage multiple tasks in time-pressured and dynamic situations? How should pilots prioritize tasks in cockpits? Using a systems engineering approach, Funk (1991) developed Cockpit Task

Management (CTM) theory as “the process by which the flight crew manages an agenda of cockpit tasks” (p. 277). Funk’s (1991) CTM theory is structurally consistent with cognitive models, such as Wicken et al.’s (2013) human information processing stage model and Endsley’s (1995) situation awareness (SA) model. For example, CTM theory has situation awareness stages (CTM 2.a, 2.c, and 2.e) that correspond to Level-2 and Level-3 in the SA model. On the other hand, while the resource scope in CTM theory includes human resources (pilots) and equipment resources (autopilots, radios, displays and controls), the scope of SA theory and other human cognitive theories focus only on human sensory cognitive and motor resource. Chou et al. reported that at least two types of CTM errors (task initiation errors and task prioritization errors) occurred when the required cognitive resource was large and the number of concurrent tasks and task difficulty (flight path complexity) were high. Chou et al. (1996) concluded that high workload generated CTM errors and proposed the necessity for pilots to develop strategies to predict and handle high workload situations. Wilson (1998) and Funk, et al. (1999) reported that the automation of pilots’ tasks might increase task prioritization errors. For example, inappropriately designed automation may make it difficult for pilots to detect, diagnose and evaluate the consequence of automation failures.

Then what is the limit of human ability to compute task priority? Shakeri and Funk (2007) tested how people can calculate the tradeoff among their CTM task prioritization criteria (importance, urgency and status of tasks) in multitasking with a juggler’s paradigm. A participant monitored six hypothetical tasks on a computer screen. The status and urgency of each task was displayed with a bar similar to a battery icon charging level bar. Each task had a different level of importance, urgency, and status property to be taken into consideration for task prioritization, and reported these four main findings. First, the participants were not able to achieve perfect task-prioritization (they scored 71% to 87% in task prioritization performance to perfect task-prioritization of a near-optimal algorithm). Second, the participants were more aware of the importance of tasks that were static and displayed on the screen and failed to recognize the dynamically changing urgency or status of tasks that required mental computation and prediction. Third, the participants overemphasized the penalty score of the task-prioritization decision, which indicated the difficulty participants had ignoring the salient task; and fourth, participants “learned” which tasks should be prioritized more highly than others (strategic task management).

What are the most relevant and useful factors for the rational design of the aircraft cockpit? Colvin, Funk and Braune (2005) conducted a flight simulator study in which they asked pilot participants to report the six primal candidate factors they used for task prioritization. The first factor reported was the perceived salience of stimuli that relates to a task. Colvin et al., hypothesized that “the priority of a task is directly proportional to its salience” . Shakeri and Funk (2007) reported that people could not ignore salient stimuli in task prioritization. Furthermore, if task-related stimuli are not salient, inattentional/change blindness phenomena may occur (Simons & Chabris, 1999; Simons and Levin 1998; Strayer, Drews & Johnston, 2003). The salience of task-related stimuli are also a factor in visual attention prioritization in the SEEV model (Wickens, Helleberg, Horry & Talleur, 2003). The second factor participants reported in Colvin et al.’s study was the perceived importance (or value) of a task. They hypothesized that “the priority of a task is directly proportional to its importance” (Colvin et al., p334). The importance (or value) criterion is often used for tradeoffs in multi-attribute utility decisions between pros (importance or value) against costs. For example, Kushleyeva, Dario, Salvucci, and Frank (2005) constructed a task-prioritization model that trades off a cost factor and the value factor multiplied by their probabilities, and Wickens et al. (2003) and Funk (1991) used importance / value factor for a visual attention prioritization model and the CTM theory, respectively. The third factor reported by Colvin et al.’s participants was the perceived performance status of a task. Colvin et al. hypothesized that “the priority of a task is directly proportional to its importance” (p335). The perceived status of a task may influence the task prioritization decision, because understanding of the current status is situation awareness Level-2, which is a foundation of sound decision-making (Endsley, 1995). Wickens (2003) noted that pilots need to have situation awareness of task status for sound decision-making along with spatial awareness and system awareness. Furthermore, Altman and Trafton (2002) reported that people tend to forget to recall, resume, and execute tasks in a suspended status. For example, the pilots of Spanair Flight 5022 forgot to complete a flaps checklist item while taxiing, resulting in its crash. Thus even high priority tasks may be forgotten and observers (e.g., accident investigators) may regard it as an "inappropriate task prioritization decision". Furthermore, the status of tasks (CTM 2.c) is used as a task prioritization factor in CTM theory (Funk, 1991). The fourth potential task prioritization factor is the perceived urgency of a task. Colvin et al. hypothesized that "the priority of a task is directly proportional to its urgency" (p335). The perceived urgency of task may be defined as the buffer time, or time remaining until the deadline of the task (Wickens, et al., 2013), which reflects the projected situation awareness of Level-3 SA (Endsley, 1995). The lack or inappropriate perception of task urgency may lead to fatal aviation accidents (e.g., In the Flight 401 accident case, the pilots did not notice the urgency of the task). The Threaded Cognition Multitasking model uses the task urgency factor for task prioritization (Salvucci & Taatgen, 2011; Salvucci, Taatgen and Borst, 2009), and Funk (1991) also use it as the task prioritization factor in CTM. The fifth potential factor is the expectation of a task. Colvin et al. (2005) hypothesized that "the priority of a task is directly proportional to its consistency with procedure or with other pilot expectations" (Colvin et al., p335). Simons and Chabris (1999), and Simons and Levin (1997) showed that lacking the expectation of stimuli may generate inattention/change blindness cognitive phenomena. Endsley (1995) argued that an expected mental model (i.e., expectation) affects situation awareness that will influence where attention is directed and how perceived information is interpreted in a top-down cognitive process. Furthermore, the projected future situation of the environment (SA level-3) will become an expectation that affects the above points (Endsley, 1995). Thus, the expectation factor is used in prioritization models (e.g., Wickens et al., 2003). The sixth potential factor is the perceived cost or effort of a task or its workload. Colvin et al. (2005) stated that "the priority of a task is proportional to the time/effort required to perform it" (Colvin et al., p336). Chou et al. (1996) reported that high workload adversely affects multitasking performance. Furthermore, high workload or high switching costs of tasks (e.g., Allport 1994) will delay the execution of prioritized tasks in time (e.g., Lee, McGehee, Brown, & Reyes, 2002). When people cannot make adequate progress in concurrent multitasking they may suspend one or more tasks for later resumption (e.g., Altman & Trafton, 2002; Salvucci, Taatgen & Borst, 2009). As mentioned before, outside observers (e.g., accident investigators) may potentially regard delayed execution as inappropriate task prioritization decisions in incidents or accidents (e.g., Flight 401; NTSB, 1973). Generally, cost (effort or workload) is considered as a tradeoff against value (i.e., importance) in multi-attribute utility decisions, and several human multitasking models utilize "cost" as a task prioritization factor (e.g., Wickens et al., 2013, Kushleyeva et al. 2005).

Research Questions and Methodology

Based on important insight from the above literature, the following research questions were raised about task prioritization behavior in aviation human multitasking. **Research question-1**: Can perceived task priority be explained by the following five factors? 1. perceived importance of tasks, 2. perceived urgency of tasks, 3. perceived status of tasks, 4. perceived salience of tasks, 5. perceived workload. **Research question-2**: Is the perceived priority of a task consistent in the level of task awareness and task execution performance?

Sixteen pilots (15 males and 1 female) were recruited for a flight simulation experiment. All the pilots possessed a private pilot license with an average of 4,508 hours of total flying experience (minimum 65 hours, maximum 31,000 hours) and an average of 2,274 hours of single pilot hours. Their ages were between 24 and 82 and average age was 49.3. They were compensated for 2 hours of data collection (duration of the experiment) with a $25 gift card. The experiment was conducted at the Human Factors Engineering Laboratory at Oregon State University. Subjects flew a Cessna 172 RG airplane in a X-plane®-based general aviation flight simulator. In order to run the experiment and collect multitasking behavioral data, the following instruments were prepared: a computer-synthesized voice for the ATC communication (operated by the experimenter), a flight checklist, and flight charts. Each participant practiced the X-plane flight simulator, Air Traffic Control (ATC) communication system, the flight plan, and aircraft checklists to become familiar with the system before the start of the simulation. In this practice session, the participant was allowed to fly in the simulator in a similar condition to the experiment. After the practice was completed, the participant could ask for clarification regarding simulator operation. The second session was data collection using the flight simulator. Flight data (e.g., headings, altitudes, airspeeds, engine parameters, radio frequencies, and flight control movements) were automatically recorded by the simulator. Every few minutes the flight simulation was frozen and the participant was asked to rate his or her perception of the importance of tasks, urgency of tasks, status of tasks, salience of tasks, workload of tasks, and perceived priority of tasks at that moment. Here, we followed Endsley (1995b)'s query guideline such that the timing of each freeze for query was randomly determined at each experimental block. Because of its unpredictable interruption, it was assumed that the participants could not anticipate or prepare for queries beforehand, which could provide unbiased estimates of the participant's task-prioritization decisions. Furthermore, behavioral audio and video data were recorded throughout the flight. A simple but challenging flight scenario was prepared. After becoming familiar with the flight simulation in the practice session, pilots conducted a simulated flight scenario using two VORs (VHF Omni Directional Radio Range). Pilot participants communicated with the experimenter who played the role of the ATC controller with the computer-synthesized voice based on a predetermined communication script. The pilot was reminded that in the simulation, flight safety was the ultimate goal, and it was assumed that the participating pilots prioritized among four tasks: Aviate (vertical control), Navigate (lateral control), Communicate, and Manage Systems. In each of eight situations, expected and unexpected flight instrument problems challenged the participants. Repeating problems included those problems that pilots could expect to happen because they reoccurred multiple times in the scenario. Those problems were Pitot tube clog that caused a airspeed indicator malfunction in situations 2, 4, and 7; a low fuel problem occurring in situations 6, 7, and 8; and an altimeter malfunction in situations 4 and 8. Non-repeating problems included those problems that pilots could not easily expect to happen because each problem occurred only once. Those problems were artificial horizontal indicator malfunction, vacuum pump indicator, vertical speed indicator, and navigation instrument malfunctions. No information was given to pilots about which problem(s) might occur or repeat in the scenario.

To answer the first research question, each pilot's perceptions of task priority and five hypothesized prioritization criteria were obtained by asking the probe questions on Table 1. The obtained perceived task priority, all response variables and explanatory variables were numerically coded for statistical hypotheses tests. The perceived task priority score was modeled using five explanatory variables with regression models, and each factor's effect was estimated with the corresponding coefficient in mixed models. Here, each of the five task-prioritization decision criteria was used as the fixed effect. Intercepts for subjects and by-subjects slopes were used as random effects in the model. A likelihood ratio test was used for testing a linear relationship between the perceived priority task priority and each of the five task prioritization decision criteria.

Table 1.
Variables and Probe Questions in the questionnaires.

Variables	**Probe Questions**
Perceived Task Priority Score (Y)	Which task did you prioritize at this moment? (pair-wise comparison)
Perceived Task Importance Score ($X1$)	Based on your comprehension of the current situation, which task was more important? (pair-wise comparison)
Perceived Task Urgency Score ($X2$)	Based on your projection of the future status, rate the urgency of each task by its "buffer time"; the amount of time you could delay the task before it requires your attention to maintain safe flight.
Perceived Task Performance Score ($X3$)	Based on your comprehension of the current situation, rate the performance of task, how successful you believe in accomplishing the goal of the task set by yourself?
Perceived Task Salience Score ($X4$)	Based on your current perception, which task was more salient and drew your attention at the moment? (pairwise comparison)
Perceived Task Workload Score ($X5$)	NASA's TLX workload questionnaire applied to each task singly

To answer the second research question regarding the relationships between perceived task priority and task awareness and task execution, subjects were instructed to verbally report as soon as they noticed any problems or abnormal situations during the flight simulation. The unnoticed time period was measured from the recorded video for estimating the level of task related problem awareness. The directional deviation and altitude deviation from the target path were obtained from the flight recorder.

Results

There were clear linear relationships between the perceived task priority and the perceived task importance score, the perceived buffer time (i.e., task urgency), and the perceived salience score. There was not enough evidence to conclude the existence of linear relationships on the perceived performance status score and the perceived workload score Table 2 summarizes the linear relationships between the four perceived task priority scores and the five potential prioritization criteria scores. Each P-value shows the linearity test result between the perceived task priority and the perceived task prioritization criterion. The relative weights methodology revealed that the perceived importance and the perceived salience scores explained most of the variance of the perceived task priority scores.

Table 2.
P-Values of five potential task prioritization decision criteria for four tasks.

Y: Priority of each task	**X1: Perceived Importance**	**X2: Perceived Urgency**	**X3: Perceived Performance Status**	**X4: Perceived Salience of tasks**	**X5: Perceived Workload**
Aviate task	P-Value<0.001 $R^2 = 0.70$	P-value<0.03 $R^2 = 0.55$	P-Value=0.88 $R^2 = 0.47$	P-Value<0.001 $R^2 = 0.60$	P-Value=0.03 $R^2 = 0.48$
Navigate task	P-Value<0.001 $R^2 = 0.58$	P-Value=0.07 $R^2 = 0.29$	P-Value<0.06 $R^2 = 0.34$	P-Value<0.001 $R^2 = 0.40$	P-Value=0.29 $R^2 = 0.26$
Communicate	P-Value<0.005 $R^2 = 0.66$	P-Value=0.02 $R^2 = 0.41$	P-Value=0.66 $R^2 = 0.34$	P-Value<0.001 $R^2 = 52$	P-Value=0.23 $R^2 = 0.17$
Manage Systems	P-Value=0.59 $R^2 = 0.58$	P-Value=0.002 $R^2 = 0.43$	P-Value=0.11 $R^2 = 0.40$	P-Value<0.001 $R^2 = 0.50$	P-Value=0.22 $R^2 = 0.33$

The second research question result was such that the perceived task priority score had a linear relationship with the actual task execution time and the task execution performance while the task awareness level was not improved by the higher perceived task priority score. For example, the higher perceived Navigate task priority (P=0.06) improved the directional deviation, and it was worsened as the number of concurrent tasks increased (P=0.13). On the other hand, the time period of inattention/change blindness was not improved by the perceived task priority score (P-Value=0.68), and it was worsened as the number of concurrent tasks increased (P-Value=0.02).

General Discussion

Following is our preliminary interpretation of those results. Scrutinizing the perceived urgency criteria data revealed two possible mechanisms of task awareness (i.e., inattention/change blindness). The first possible mechanism was a task-prioritization decision problem in sequential multitasking. Some pilots rated larger task buffer times (i.e., not urgent), suggesting that the high priority task was postponed; deterministically it took a longer time to notice the task related problem. Thus, task importance, task salience, and task urgency criteria should be addressed in cockpit design and pilot training for better task prioritization decisions. The second possible mechanism was the cognitive resource interference problem during concurrent multitasking. Subjects worked hard to conduct concurrent multitasking because they reported multiple tasks were simultaneously urgent. When pilots perceived multiple tasks to have high urgency and high priority, they would start concurrent multitasking. During concurrent multitasking, the task awareness level dropped in a stochastic way. When the instrument panel malfunction signals were neither expected nor salient, the mean time to notice the task signal was 137.8 seconds and no significant effect of number of concurrent tasks was observed (P-value=0.9). When the instrument panel malfunction signals were expected but without salient stimuli, the mean time to notice the task signal was 122.5 seconds but it increased as the number of concurrent multitasks increased (P-value=0.01). When the instrument panel malfunction produced a salient signal but it was not expected, the mean time to notice the task signal was 24.8 seconds, but it increased as the number of concurrent multitask increased (slope P-value=0.03). When a problem was expected and had a salient signal, the mean time to notice the task signal was 23.9 seconds and it did not increase with more concurrent multitasks (slope P-value=0.7). This indicate that the bottom-up factors (signal salience, signal expectation), and the number of concurrent multitask should be taken into consideration in cockpit design and pilot training to increase awareness of tasks related information.

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks.

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. Cognitive science, 26(1), 39-83.

Chou, C. C., Madhavan, D., & Funk, K. (1996). Studies of cockpit task management errors. The International Journal of Aviation Psychology, 6(4), 307-320.

Colvin, K., Funk, K., & Braune, R. (2005). Task Prioritization Factors: Two Part-Task Simulator Studies. The International Journal of Aviation Psychology, 15(4), 321-338.

Endsley, M. R. (1995a). Toward a theory of situation awareness in dynamic systems. Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(1), 32-64.

Funk, K. (1991). Cockpit task management: Preliminary definitions, normative theory, error taxonomy, and design recommendations. The International Journal of Aviation Psychology, 1(4), 271-285.

Funk, K., Suroteguh, C., Wilson, J., & Lyall, B. (1998, October). Flight deck automation and task management. In IEEE INTERNATIONAL CONFERENCE ON SYSTEMS MAN AND CYBERNETICS (Vol. 1, pp. 863-868). INSTITUTE OF ELECTRICAL ENGINEERS INC (IEEE).

Kushleyeva, Yelena, Dario D. Salvucci, and Frank J. Lee. "Deciding when to switch tasks in time-critical multitasking." Cognitive Systems Research 6.1 (2005): 41-49.

Lee, J. D., McGehee, D. V., Brown, T. L., & Reyes, M. L. (2002). Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. Human Factors: The Journal of the Human Factors and Ergonomics Society, 44(2), 314-334.

National Safety Council White Paper. (2010). Understanding the distracted brain: Why driving while using hands-free cell phones is risky behavior. (on-line publication)

NTSB. (1973). Aircraft accident report. Eastern Air Lines, Incorporated, L-1011, N310EA, Miami, Florida, December 29, 1972. Report No. NTSB-AAR-73-14. Washington, DC: National Transportation Safety Board

NTSB (1987). Aircraft accident report. Northwest Airlines, Incorporated, MD-82, Detroit, Michigan, August 16, 1987. Report No. Washington, DC: National Transportation Safety Board

Salvucci, D. D., Taatgen, N. A., & Borst, J. P. (2009). Toward a unified theory of the multitasking continuum: From concurrent performance to task switching, interruption, and resumption.

Salvucci, D. D., & Taatgen, N. A. (2010). The multitasking mind. Oxford University Press.

Shakeri, S., & Funk, K. (2007). A comparison of human and near-optimal task management behavior. Human Factors: The Journal of the Human Factors and Ergonomics Society, 49(3), 400-416.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception-London, 28(9), 1059-1074.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. Psychonomic Bulletin & Review, 5(4), 644-649.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. Human Factors: The Journal of the Human Factors and Ergonomics Society, 45(3), 360-380.

Wickens, C. D. & Hollands, J. G. (2013). Engineering psychology and human performance: Prentice Hall New Jersey.

Wilson, J. R. (1998). The effect of automation on the frequency of task prioritization errors on commercial aircraft decks: An ASRS incident report study.