2005

# Decision Making During an Airline Rescheduling Task: a Contextual Control Model Description

Karen M. Feigh

Amy R. Pritchett

Julie A. Jacko

Tina Denq

# DECISION MAKING DURING AN AIRLINE RESCHEDULING TASK: A CONTEXTUAL CONTROL MODEL DESCRIPTION

**Karen M. Feigh**, **Amy R. Pritchett**, **Julie A. Jacko**, **Tina Denq**
Georgia Institute of Technology
Atlanta, Georgia

This paper examines decision makers' selection of contextual control modes as described by Hollnagel's Contextual Control Model, and evaluates real-time, unobtrusive measures of a decision maker's immediate mode. In a two-part experiment, participants performed airline rescheduling tasks. The first portion varied task time limits, the second introduced a sudden change in the task. Participants reported operating in, and transitioning between, different contextual control modes in response to time limits and task changes. Computer interaction did not correlate to contextual control modes. Contextual control modes did not correlate with TLX ratings of demand and effort, but did correlate with TLX-frustration and TLX-performance ratings. The results suggest that decision making performance may be determined by use of context-appropriate contextual control modes, and imply that the design of decision aids should work to support those modes.

## Introduction

Airline managers of a typical large U.S. airline are responsible for the daily operation of large regions or fleets of aircraft, often with 40-50 flights departing every hour. They oversee daily operations that are often disrupted by weather, ATC delays and maintenance problems. and are responsible for implementing flight delays, cancellations, "aircraft swaps" and the use of reserve crews to minimize the impact of such disruptions relative to the nominal flight schedule. Decisions must often be made quickly, frequently based on uncertain information. Many elements must be requested from other personnel (e.g., the maintenance department's estimate of a repair time). Other information must be retrieved from cumbersome text-based interfaces presenting data about hundreds of flights.

Our own observations have revealed that managers' approaches to this task can vary wildly. On a day with few disruptions the manager may consider many possible alternatives to minimize flight delays. Alternatively, on a busy travel day with major disruptions, the manager may resort to sweeping measures such as operating the entire fleet an hour behind schedule. This study hypothesized that these changes in decision making behavior may be described by different contextual control modes (CCM).

A large number of decision models which view decision making as the cognitive task of selecting from a set of alternatives. One accounting for some of the multiple decision models has recognized the tendency for human decision makers to "select" or "switch" cognitive strategies as a coping strategy in the face of stressors. Strategy switches include speed/accuracy trade-offs, task shedding, and the use of simpler strategies (e.g. Svenson, et al., 1993; Maule, 1997; Orasnau, 1997), which are not always explained simply as methods to reduce workload. While the selection of a strategy is often modeled as a cost-benefit activity (Maule, 1997), studies have also described cases where decision makers chose to increase their effort to maintain performance under perceived time constraints (e.g. Todd, et al., 1994; Kerstholt, 1996).

Hollnagel contends that the "the degree of control a person will have over a situation can vary. It seems reasonable to think of control as a continuous dimension where at one end there will be a high degree of control and at the other there will be little or no control" (Hollnagel, 1993). To better describe this continuum of control, Hollnagel has developed a classification of four contextual control modes:

• "*Scrambled control* denotes the case where the choice of next action is completely unpredictable or random." (Hollnagel, 1993, pp. 168)

• "*Opportunistic control* corresponds to the case when the next action is chosen from the current context alone, and mainly based on the salient features." (Hollnagel, 1993, pp. 169-170)

• "*Tactical control* is characteristic of situations where the person's event horizon goes beyond the dominant needs of the present, but the possible actions considered are still very much related to the immediate extrapolations from the context." (Hollnagel, 1993, pp. 170)

• "Strategic control means that the person is using a wider event horizon and looking ahead at higher level goals... The strategic control mode should provide a more efficient and robust performance, and thus be the ideal to strive for."(Hollnagel, 1993, pp. 170)

An important aspect of Hollnagel's contextual control model (COCOM) is the idea that individuals will transition between CCM to maintain control over a changing situation (Stanton, et al., 2001; Jobidon, et al., 2004). Hollnagel states that, "The change between control modes is determined by a combination of situational and person (or internal) conditions – in other words by the existing context…," (Hollnagel, 1993 pp. 194). Thus, the control mode must be appropriate to the context. An erroneous assessment of context, such as an incorrect subjective assessment of available time, may lead to use of a CCM that will not result in the best performance possible in the available time. For example, the impact of time pressure has been experimentally linked to CCMs in dynamic tasks, (e.g., Jobidon, et al., 2004) who concluded that time pressure and corresponding 'worse' CCMs lead to poorer performance.

However, the degradation in performance may not directly relate to choosing a 'worse' CCM. Inappropriate use of a higher control mode may also result in lower performance. For example, empirical studies by Oransanu et al. (1993) and Johnson et al. (2002) described how mismatches between context and decision strategies could have detrimental effects on performance. Unexpectedly, these mismatches can occur with reductions in workload, suggesting that CCMs and their appropriateness to the context can be better predictors of decision making performance than workload measures alone.

Decision support tools may be tailored for specific decision modes (Niwa, et al., 2002; Johnson, et al., 2002). However, very little work has been done to investigate measures which would allow real-time identification of an individuals'. Therefore, this research also investigates potential easily observable indicators of a decision maker's immediate CCM.

## Objectives

The objectives of this study were twofold. First, we endeavored to verify the impact of time constraints and changes in task demands on human cognitive behavior as described by CCMs. Second, we sought to identify measures of CCMs including measures of information seeking behavior and a self-assessment.

## Method

*Participants*

Participants in this experiment were undergraduate students. Data from 16 participants (12 males and 5 females) will be discussed here. The participants had a mean age of 22 years (ranging from 18 to 34 years), and had no previous airline scheduling experience. No selection criteria were used to qualify or disqualify participants.

*Experiment Task and Procedure*

Participants were asked to assume the role of airline manager for a small airline (4 airports, 4 aircraft and 12-16 flights). In the first part of the experiment the participants were presented with a disruption to an established flight schedule. Disruptions included weather and unexpected maintenance issues. They were instructed to strand as few passengers as possible while following some basic rules (e.g., all flights must terminate by midnight), and asked to find the best solution possible within a given time limit.

In the second part of the experiment, in addition to a the up-front disruption, a change in context was suddenly introduced part way into the task by telling participants that an aircraft had just announced they needed to divert to an airport due to a bomb threat, creating a further disruption. At the end of each run, participants were asked to record their solutions and the number of passengers it stranded. They were also asked to provide a self-assessment of workload and CCM.

The participants had access to complementary computer based and non-computer based information about the flight schedules. The information external to the computer mimicked information which is normally requested from a person who is not in the immediate vicinity, and thus carried a time cost. This external information represents information beneficial but not necessary for the completion of the task; by assigning a time cost to this supplementary information, its access suggests adequate subjectively available time for a tactical or strategic CCM.

Each participant conducted six runs. The first, a training run, had a simplified task to introduce the task, computer interface, and information available. The following five runs asked participants to find the best solution possible for a specified disruption in the time provided.

*Apparatus*

The experiment was conducted at a standard computer terminal with keyboard and mouse. The display was 17in. flat panel display set to a resolution of 1280 by 1024 pixels. Participants were also given a piece of paper and a pencil. The interface approximated the text-based terminal windows used by airline sector managers, with command buttons substituted for text-based commands.

*Experiment Design*

The two independent variables were time limit and the introduction of contextual change. In the first part of the experiment four time limits were used: 18, 13, 8, and 3 minutes. The final run (i.e., the second part of the experiment) introduced contextual change two minutes into the.

The scenario order, time limits and run order were balanced using a Latin square to minimize order, learning, and scenario effects. In the second part, the time limit was fixed at eight minutes, contained the same scenario task, and was always given last so that participants would not anticipate such a disruption in subsequent runs.

*Dependent Measures*

The data of interest were categorized into the following six groups:

*Computer Interaction* Key logging and mouse tracking software automatically recorded the frequency of requests for information from the computer and delete key hits.

*Interaction External to the Computer* External interaction was measured by the number of times the participant sought external information.

*NASA Modified Task Load Index (TLX)* Workload ratings were collected after each run via the six NASA TLX subjective rating sub-scales: mental demand, physical demand, temporal demand, performance, effort, and frustration.

*Self-Assessment of Contextual Control Mode* At the start of the experiment, subjects were briefed about the CCMs using Hollnagel's description for each. Then, on the questionnaire administered at the end of each run participants indicated the CCM they used during most of the task on a scale of 1-10, where the four CCMs were equally arranged and explicitly labeled at the 1 (scrambled), 4 (opportunistic), 7 (tactical) and 10 (strategic) marks. Additionally, participants were asked to state if they felt that they had transitioned from one CCM to another during the course of the task.

*Performance* Each scenario was designed to have at least four valid solutions. To standardize across all scenarios, the solutions were ranked according to the number of passengers stranded and the number of flights cancelled or delayed. The four best solutions were ranked one through four. All other valid solutions were given a rank of five. All invalid or incomplete solutions were assigned a rank of six.

## Results

*Experiment Part 1*

A general linear model of the self-assessed CCM. This model indicated main effects due to scenarios (F=3.989, p=0.024) and time limit (F=5.348, p=0.008). Pairwise comparisons found differences between two scenarios (p=0.017). Time limit differences were found between 3min-13min (p=0.017), and 3min-18min (p=0.007) levels, as shown in Figure 1.

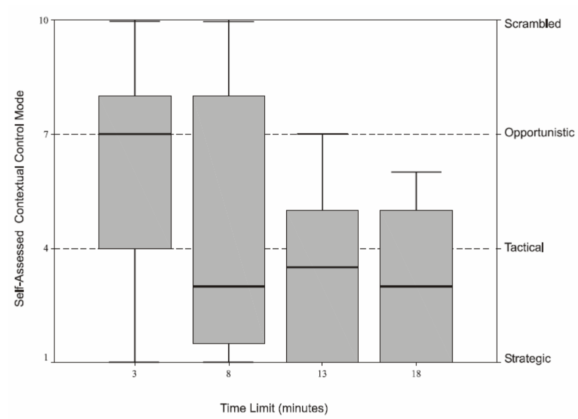


**Figure 1.** *Self-assessed CCM as a Factor of the Time Limit Imposed.*

A linear regression was performed on the reported CCM to examine the impact of observable indicators. The full model included the average time between mouse clicks, time limit, and the percentage of external information used. The model was found to be significant (F=4.656, p=0.003), however the average time between mouse clicks did not significantly contribute. There was a significant correlation between the percentage of external information used and time limit (r=0.653, p < 0.001). Likewise, a general linear model evaluated the six raw TLX subscale scores. Time limit was found to be a significant source of variance only in the TLX-temporal measures (F=10.208, p<0.001). Pairwise comparisons revealed that there were significant differences between the three minute level and all other levels (p<0.05).

A linear regression was performed on the raw TLX subscales. The full model included those measures which could be available for a real time assessment of CCM: the average time between mouse clicks, time limit, and the percentage of external information used. The model was found to be significant for the TLX-temporal subscale (F=9.736, p<0.001). Reduced models were found to be significant for the TLX-physical and TLX-frustration subscales. The reduced model for the TLX-temporal subscale only included time limit (F=28.976, p<0.001). The reduced model for the TLX-physical subscale included both time limit and the average time between mouse clicks (F=3.206, p=0.047), whereas the reduced model for the TLX-frustration subscale only included the average time between mouse clicks (F=6.111, p=0.016).

To compare self-assessed CCMs and workload, a linear regression was performed on the self-assessed CCMs, where the model included all six TLX subscales, average time between mouse clicks and the percentage of external information used (see Figure 2).

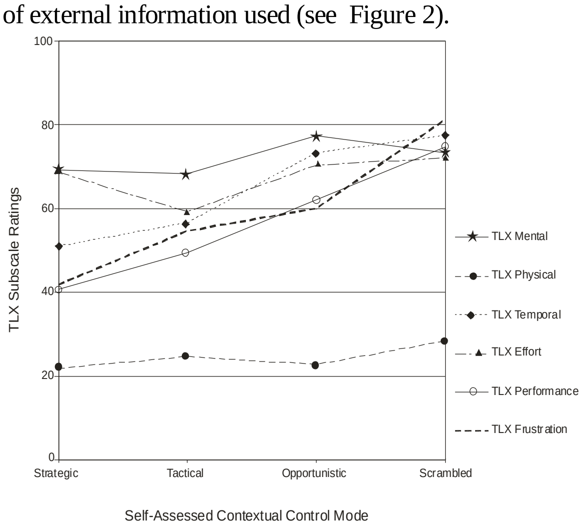


**Figure 2.** *TLX Subscale Scores by Self-Assessed Transition Direction in CCM*

The model was found to be significant (F=5.108, p<0.001). However, only the percentage of external information used, TLX-frustration, TLX-performance, and TLX-temporal subscales were found to significantly contribute to the model.

The effect of time limit, observable indicators, CCMs and TLX subscales on performance were then examined. A Kruskall-Wallis mean rank comparison found a marginally significant effect of time limit on participant performance ($\chi^2$=6.333, p=0.096), as shown in Figure 3. Paired comparisons found a significant difference between performance in the 8 and 13 time limit levels (Z=-2.104, p=0.035).

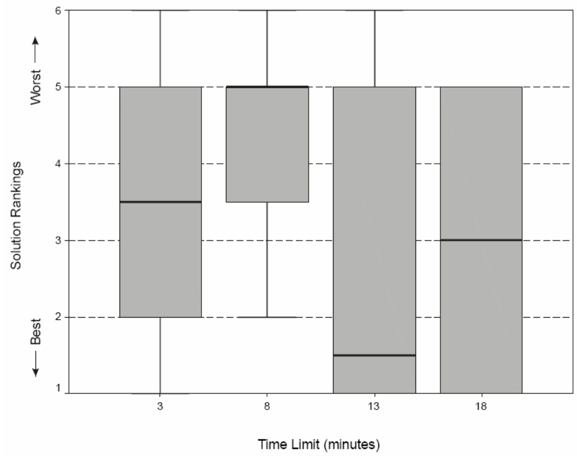


**Figure 3:** *Performance as a Factor of Time Limit*

A further Kruskall-Wallis mean rank comparison did not find differences in participant performance based on their self-assessed CCM. However, when individual paired comparisons were conducted a significant difference between participant performance was found between the opportunistic and the scrambled levels (p=0.033). A linear regression was performed on participant performance where the full model included all six TLX subscales. Neither the full model nor any of the individual TLX subscales were found to be statistically significant.

*Experiment Part 2*

In 10 of 16 participants (63% of the runs) there was a self-reported transition due to the contextual change of unexpectedly announcing (to the participant) that an aircraft was diverting to another airport, further disrupting the flight schedule. A general linear model was used to evaluate whether the inclusion of a contextual change affected the average time between mouse clicks, the TLX subscales, self-

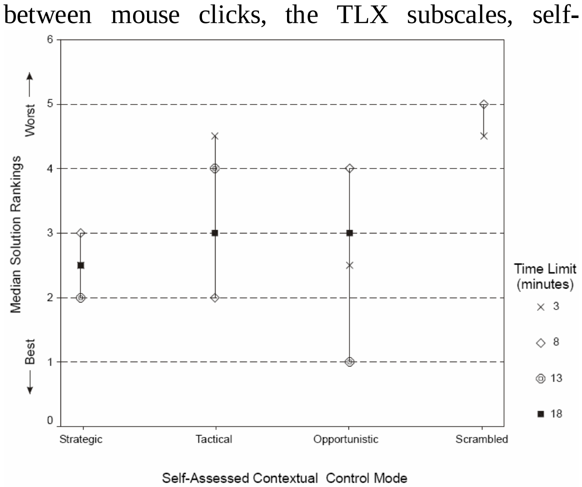


**Figure 4.** *Median Solution Performance by Self-Assessed CCM*

assessed CCM, or direction of CCM transition. Analysis of the model indicated that contextual change did not affect the average time between mouse clicks, the CCM or the CCM transition amount in a statistically significant manner. The analysis also indicated that the contextual change did affect the TLX-mental (F=11.309, p = 0.001), TLX-temporal (F=13.153, p=0.001) and TLX-frustration (F=4.681, p=0.034) subscales. Kruskall-Wallis mean rank comparisons found no significant effects due to the contextual change in performance, percentage of external information used or rule violations.

Adding the impact of a contextual change to the model generated in the first part of the experiment, which included time limit, the percentage of external information used, and contextual change, found that contextual change is also a statistical predictor of CCMs. The new model was significant (F=5.900, p<0.001.

Kruskall-Wallis mean rank comparisons were performed to see if the time limit, TLX subscales, or contextual change affected self-assessed CCM transitions. Of these, the only significant predictor of self-assessed CCMs is the TLX-frustration subscale. As shown in Figure 5, the TLX-frustration subscale was significantly affected by reported CCM transitions ( $^{2}$=6.948, p=0.008), with a higher frustration level when participants reported a transition in either direction.

## Discussion

The first part of the experiment examined the impact of time limits on human cognitive behavior as described by CCMs. The analysis revealed that, while there is a general trend for the self-assessed CCM to increase (become more strategic) with decreased time pressure, a linear trend is not strictly observed. Similarly, participants' performance did not linearly correlate with the self-assessed CCM. Many of the poorer performing data points correspond to self-assessments of 'opportunistic' control modes in the eight minute time limit condition and to 'tactical' control in the three and thirteen minute conditions, in addition to the conditions where the participants self-assessed their control mode as 'scrambled'.

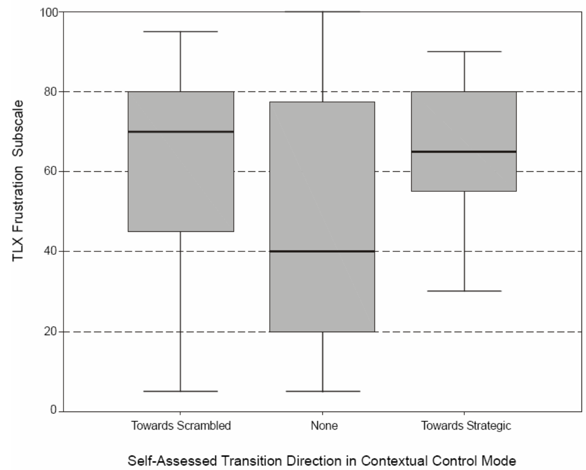


**Figure 5.** *TLX Frustration Scores by Self-Assessed Transition Direction in CCM*

These two findings may together correspond to the findings of the study by Johnson et al. (2002) in which participants sometimes appeared to ambitiously switch to inappropriate modes of behavior which could not generate high levels of performance within the time provided. These effects may correspond to poor assessments of subjectively available time in relation to the demands of the task.

The results also indicated that participants felt that their behaviors were more closely related to performance (and to frustration, defined in the TLX description as difficulties in achieving desired performance) than to measures of load and effort. As seen in Figure 2, only TLX workload ratings for performance and frustration were found to be statistical predictors of ratings of CCMs; TLX measures of demand and effort were not significant.

CCMs and workload differed in significant ways. Self-assessments of CCMs correlated with actual performance, whereas TLX ratings of perceived workload, including self-assessed performance, did not. Likewise, CCMs and the TLX subscales were predicted by different factors. For three of the TLX subscales, the observable indicators (average time between mouse clicks, amount of external information used, time limit) tested here were statistical predictors of TLX temporal by time limit, TLX physical by time limit and average time between mouse clicks, and TLX frustration by average time between mouse clicks. In contrast, CCM, while statistically predicted by the TLX performance and TLX frustration subscales, was not statistically predicted by any of the observable indicators.

In addition to the ‘overall’ CCM within each run, participants also reported transitioning between modes, with the transitions not statistically predicted by any of the observable indicators. Likewise, TLX-frustration was the only statistical predictor of transitions between CCMs during a run, albeit a comparatively weak predictor, as seen in Figure 5.

## Conclusion

Participants in this study were able to provide a self-assessment of CCM. These self-assessments yielded a significant relationship to decision making performance and to contextual factors generally thought to impact performance, such as information sought, and self-assessed temporal demand, performance and frustration. These results support Hollnagel’s representation of CCMs as involving more than a direct consequence of workload.

From the perspective of CCMs, the best performance within a given context (including time limit) will be attained when the decision maker applies the most appropriate CCM. Conversely, poor performance in this experiment corresponded not only to severe time limits demanding a ‘lower’ CCM, but also to perhaps over-optimistic attempts at ‘higher’ CCMs when sufficient time did not exist to carry them through. This perspective explains the results of earlier studies in which more time available sometimes led to a decrease in decision making performance.

These insights imply several design considerations. Decision makers operating within a fairly stable context might benefit from decision aids streamlined to support information-seeking, decision and action-taking behaviors which support the CCM most appropriate to that context. Keeping the context stable maybe seen as an important aspect of workload management. Evidence of this can be seen in standard ATC operating procedures. Controllers maintain focus on the near term and could be hypothesized as using tactical CCM, whereas the traffic flow managers are responsible for more strategic decisions and can by hypothesized to use tactical and strategic CCMs. When a controller is no longer able to manager the volume of traffic they are paired up with a D-side controller. This can be viewed as a controller no longer being able to operate at a tactical CCM, i.e. with out the additional controller they would be forced to operate at an opportunistic CCM due to traffic.

However, in many other aviation situations the decision maker’s context can vary from hour to hour and from day to day, such as the airline rescheduling task examined here and other aviation related jobs. In these cases, the decision aid may need to be capable of supporting several different CCMs. This may be achieved through one large interface which centrally emphasizes the most salient information needed in opportunistic CCMs while also supporting the information seeking and explorative behaviors corresponding to tactical and strategic CCMs.

One could argue that the differences in assessment of how much information to give a pilot or a controller stems from CCM. Depending on which “level” the pilot or controller is operating at will greatly influence how much information and which types of displays would be most helpful. At an extreme, an aid may be envisioned with separate interfaces for each of the CCMs potentially employed by its user. Such an aid could, in theory, switch automatically between interfaces in response to its user’s transitions between CCMs, i.e., an “adaptive decision aid” equivalent to “adaptive automation.” However, as the real-time indicators examined in this study were not able to statistically predict CCM, some other indicators or methods of assessing the user’s control behavior would be required. Participants’ ability to self-assess their CCM suggests that decision makers may be able to manually switch between interfaces to obtain the level of support they require, i.e., an “adaptable decision aid” may be a better approach to support pilots and controllers by allowing them to chose how much information they need. With experience, interface switching may itself be another component of an expert’s adaptation to the operating environment. Before such expertise is developed, however, another potential role of the interface may also be to present contextual factors that allow the controller, pilots, and airline operations managers to better select the CCM most appropriate to their immediate situation.

## Acknowledgements

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship, and the experiment was supported in part through grants awarded to Julie A. Jacko by the Intel Corporation and the National Science Foundation (BES-9896304)

## References

Hollnagel, E. (1993). Human reliability analysis: Context and control. London, UK: Academic Press.

Hollnagel, E. (2002). Time and Time Again. Vol. 3(2), 143-158.

Hutchins, E. (1995). Cognition in the Wild.

Jobidon, M.-E., Rousseau, R., & Breton, R. (2004). Time in the Control of a Dynamic Environment. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 557-561. HFES.

Johnson, K. E., Kuchar, J. K., & Oman, C. M.Experimental Study of Automation to Support Time-Critical Replanning Decisions. Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, 5. 2002: HFES.

Kerstholt, J. H. (1996). The Effect of Information Costs on Strategy Selection in Dynamic Tasks. Vol. 94, 273-290.

Maule, J. A. (1997). Strategies for Adapting to Time Pressure. R. Flin, E. Salas, M. Strub, & L. Martin, Decision Making Under Stress, 271-280. Aldershot, UK: Ashgate Publishing Ltd.

Niwa, Y., & Hollnagel, E. (2002). Principles of Performance Monitoring in Coupled Human-Machine Systems. Johannsen, IFAC Analysis, Design and Evaluation of Human-Machine Systems, 303-307.

Orasnau, J. (1997). Stress and Naturalistic Decision Making: Strengthening the Weak Links. R. Flin, E. Salas, M. Strub, & L. Martin, Decision Making Under Stress, 43-66. Aldershot, UK: Ashgate Publishing Ltd.

Svenson, O., & Edland, A. (1993). On Judgment and Decision Making Under Time Pressure and the Control Process Industries. IEEE International Conference on Systems, Man & Cybernetics, 367-375. IEEE.

Todd, P. A., & Benbasat, I. (1994). The Influence of Decision Aids on Choice Strategies Under Conditions of High Cognitive Load. Vol. 24(4), 537-547.