Time Design: Analysing Human Temporal Control Behaviour in Dynamic Real-Time Systems [*]

Michael Hildebrandt, Karsten F. Loer and Michael D. Harrison
Human-Computer Interaction Group, Dept. of Computer Science
University of York, York YO10 5DD, United Kingdom
{hilde, loer, mdh}@cs.york.ac.uk

Abstract – *This paper presents work aimed at supporting the design of temporal aspects of socio-technical systems. Time Design is a framework for (a) analysing and representing temporal properties of the work domain, (b) generating design options that support timely, flexible and dependable function servicing, and (c) providing knowledge about the characteristics and biases of human temporal control behaviour. In support of the latter end, two microworld experiments that investigated temporal control decisions in a supervisory control task are presented. These experiments manipulated event rate, the duration of event rate blocks, the availability of online and offline event rate information, and the accuracy of this information. The studies identified conditions where attention to temporal information decreased and the use of conservative temporal control strategies increased.*

Keywords: Time, scheduling, decision-making, biases.

1   Introduction

Temporal reasoning, temporal judgment and temporal decision-making are ubiquitous components of control processes in dynamic real-time systems. A multitude of engineering approaches are available for designing and optimising temporal aspects of these systems, be it by supporting scheduling in multi-task environments or simply by increasing the speed of data analysis, decision-making or action implementation. However, while automation and optimisation are increasingly important in controlling dynamic systems, the human operator, in the role of supervisory controller, remains an indispensable element of the system. Designing such joint human-machine systems therefore requires knowledge about the characteristics of human temporal cognition. While it is tempting to simplify this design problem by modelling human temporal control decisions by the same normative models used for automatic components, research into human temporal decision-making [e.g. 6] has produced evidence for systematic temporal decision biases that lead to deviations from normative behaviour. The aim of this line of work is to produce, and make accessible to designers, knowledge about the characteristics of human temporal cognition and its interaction with features of the task and environment.

1.1   Time Design

Time Design [1] is an emerging research and development domain that draws on a diverse literature in cognitive psychology, psychophysics, sociology, computer science, engineering, Human Factors and HCI, as well a variety of representation, analysis and modelling techniques. Time Design is based on a functional view of time (as opposed to a descriptive view of time) and explores the ways in which human control behavior is sensitive to temporal information and temporal knowledge, what heuristics and biases occur in these temporal control decisions, and how temporal aspects of the system constitute degrees of freedom that the operator can use to make adaptive control decisions. Together, these notions of time begin to chart a design space that consists of, at least, the following dimensions:

- *Time as property of the machine or interface*: e.g. service rate, responsiveness, temporal validity of information.

- *Time as an aspects of user behaviour*: e.g. anticipative or reactive control mode, temporal reasoning biases, reaction to time stress, personal and social attitudes towards time.

- *Time as a task property*: e.g. interleavable, pre-emptable.

- *Time as a property of the environment*: e.g. predictability and regularity of task arrival, tempo and dynamism.

2   Previous research

The consequences of temporal errors can be severe, especially as work processes, from industrial processing to aviation, become increasingly fast-paced, synchronised and multi-tasked. While methods are available for analysing knowledge- and attention-based human error phenomena such as loss of situation awareness or mode

---

confusion, there is little support for the analysis of control processes where the input or output of a decision process is *temporal* (e.g. temporal knowledge and information, scheduling decisions or pacing). Our first contribution was therefore to provide a *conceptual model* that allows a system developer to reason about the temporal design options and requirements in the system (e.g. sequential rigidity or flexibility of functions, pace of work, display of temporal information) and to identify system-task configurations that may give rise to temporal errors. This *Dynamic Function Scheduling* (DFS) approach extends Dynamic Function Allocation (DFA) approaches by (a) adding a temporal dimension for allocation decisions, and (b) describing allocation as a process of value-based strategy selection [3]. DFS shares the core assumption of Human Factors work design approaches (such as DFA) that higher system dependability can be achieved by joint human-automation control than by either full automation or full manual operation. It extends such approaches by simultaneously considering the allocation of functions along the human-automation resource dimension *and* allocation in time (e.g. decision to postpone, interleave or drop functions). In doing so it takes advantage of the multitude of methods available in Operations Research for analysing and optimising temporal aspects of systems (e.g. queuing or scheduling models). This analysis is complemented by research in cognitive psychology and Human Factors that explores the causes and conditions of temporal error [5]. Temporal error (e.g. premature or late decisions, sequence errors, synchronization errors; c.f. [4]) has, for instance, been linked to conflicts among the temporal reference systems used by the operator [5], or to biases in temporal reasoning such as duration neglect [6].

2.1 Microworld PaintShop1

The study reported in this paper follows up results of a previous microworld experiment that employed a similar supervisory control task (*PaintShop1*). The sections below outline the task and summarize the results obtained in the study (for details see [2]).

2.1.1 PaintShop1 supervisory control task

The PaintShop microworld simulates an industrial paint station in a factory where coats of paint are sprayed onto items. Each painted item earns the participant 1 pence. There is a steady stream of items passing through the station, but the event rate can change during the trial. In the first experiment, three event rate levels were used (*high*, *medium*, *low*, i.e. one item every 2, 3 or 4 seconds).

Event rates during each 3-minute trial could change in blocks of 1-minute duration. So for instance, a 3-minute trial might consist of 1-minute high event rate, followed by 1-minute low rate, followed by 1-minute medium rate. Half the participants received information about the arrival rate distribution prior to each trial, the

other half did not (between-subjects factor *availability of event rate information*). In half of the trials, participants could request a progress indicator by pressing the space bar (within-subjects factor *availability of progress indicator*). The progress indicator (visible in the centre of Fig. 1) showed the stage of the trial by means of a red time indicator that passed from left to right on a time scale that represented the 3-minute trial. For the group of participants *with* event rate information, the timeline display would also show the event rate information (tall grey block for high arrival rate, medium block for medium rate, small block for low rate). For the group *without* event rate information, the timeline would only provide visual markers at each 1-minute boundary, but no indication of the event rate level.

Items arrived on a single conveyor belt, but could be painted using two parallel paint stations. Items were distributed between these stations using a lift that delivered the item to whatever station was available, or would become available next (*automatic mode*). The operator was able to bias the delivery of items so that all items would either be delivered to the upper or to the lower station (*manual mode*).

Two modes were available for painting items, *automatic* and *manual*. In automatic mode, items would be processed as soon as they entered the station and no intervention by the operator was needed. Painting an item in automatic mode took 4 seconds. In manual mode, painting only took 2 seconds, but the operator would have to click the item with the mouse and keep the mouse button pressed for the whole of the paint cycle.

The main complication in this task was caused by faults in the form of blockages in the paint nozzles. Every 8-14 paint cycles, a station's nozzles would get blocked so that no further items could be processed. An imminent blockage was indicated by the pressure indicator above each station moving progressively from its normal position around the left-hand side of the display to the critical zone at the right of the display. It took 6 paint cycles for the nozzle to get from an initial blockage towards full blockage. The pressure indicator thus provided a visual warning of an imminent blockage, but intermediate states of blockage had no effect on the performance of the system (i.e. the paint durations). Two strategies, *repair* and *replace*, were available for servicing a fault.

Repair caused no financial cost, but put the station out of service for 24 seconds. As items keep arriving during repair periods, they will start to queue up in front of the lift unless the other station can process them quickly enough. The queue can have a maximum length of four items, after that each incoming item is removed from the belt and the money it could have earned is lost.

Replace put the station back into service immediately, but caused a financial cost of 6, 8 or 12 pence (between-subjects factor *repair cost*).

2.1.2 Control strategies

Based on the temporal characteristics of the system and the available temporal information, the participants' task was to develop control strategies that would maximize their financial payoff. To understand the trade-offs involved in these control decisions, the following implications need to be considered:

A *low* event rate can be handled by just one station on its own operating in *automatic* mode.

A *medium* rate can be handled by one station in *manual* mode or two stations in *automatic* mode. Both options create slack time.

A *high* event rate can be handled by one station in *manual* mode or two stations in *automatic* mode.

The second piece of important temporal information is the amount of items that get lost when stations are unavailable (e.g. while being under repair). During training, participants were given the example of a situation where both stations were unavailable simultaneously. If both stations were put under repair mode, they would be unavailable for 24 seconds. During that time, 12 items would be lost under *high* event rate, 8 items under *medium* event rate, and 6 under *low* event rate. Thus depending on the current event rate, the state of the other station, the painting mode used on the other station (manual or automatic), and the replace cost (6, 8 or 12 pence), *repair* or *replace* might be the fault servicing decision that produces the greatest benefit.

Of particular interest were participants' use of temporal information and the effect of availability of temporal information (in the form of event rate knowledge and availability of a progress indicator) on control strategies. Online event rate information allowed participants to optimise control strategies in two respects:

(a) It allowed them to judge the periods when both stations were required (e.g. under high event rate) and when only one station on its own could handle the workload (e.g. under low event rate). When only a single station was required, all items could be routed to one station. The other station could either be rested or repaired *prematurely* (i.e. before it was broken) so that a fresh station was available in case the other one broke or a high event rate period started.

(b) When a station *had* broken, the choice of an appropriate fault servicing decision benefited from knowledge of any event rate changes that may occur during the course of the 24-second repair period. For instance, when a switch to a low event rate was imminent, *repair* may suddenly become the preferred option, while with an upcoming switch to high workload, *replace* may produce the higher pay-off.

Participants who received both event rate information *and* had an online progress indicator available were therefore able, at least potentially, to make accurate fault servicing decisions. When the online progress indicator was not available for these participants (as it was on half of the trials), they still had the benefit of knowledge of event rate information, but would have to use their own time perception abilities if they wanted to take event rate boundaries into account.

Participants without event rate information could only judge event rate based on current knowledge (i.e. if event rate had just switched to low, they knew this rate would persist for the next minute). In those trials where the online progress indicator was available, they could use the markers at the one-minute intervals to judge when an event rate change would occur. For instance, if they currently operated under low workload, they may want to rest or repair one of the stations so that it would be available for the next block, which might well bring a higher event rate.

That participants did indeed develop sophisticated temporal control strategies is illustrated by the following quotes from the feedback forms participants completed after each trial (for more quotes, see [2]):

6p repair cost, no event rate information, no timescale: "I wanted to set both machines to repair half way through the medium section so they were both ready for the heavy period, but not having the timeline made that very difficult to judge."

12p repair cost, no event rate information, timescale available: "The timeline is very useful to check how far through you are and also if the event rate is about to change so you can plan ahead. [...] But it is only really useful in medium or low periods when you have time to look at it."

8p repair cost, no event rate information, no timescale: "Definitely works best having one machine being repaired all the time and using the other in manual mode. Usually seem to lose about 6 or 7 during a run so better than replacing the nozzles."

2.1.3 Results of PaintShop1

Of the results obtained in the first study, only those relevant to the current study shall be reported here (see [2] for more results).

Event rate information. Based on the considerations above, the most accurate fault servicing decisions would be expected in trials where both event rate information and the online progress indicator were available. Analysis of the number of lost items confirmed the benefit of event rate information, but showed no additional benefit of the online progress indicator.

Use of progress indicator. Results also showed a significant increase in use of the progress indicator as event rate decreased (c.f. second user comment above).

3   PaintShop2

The study reported in this paper was designed to address some of the question raised by the PaintShop1 study. Of particular interest was the effect of event rate information that did not seem to be mediated by the availability of an online progress indicator.
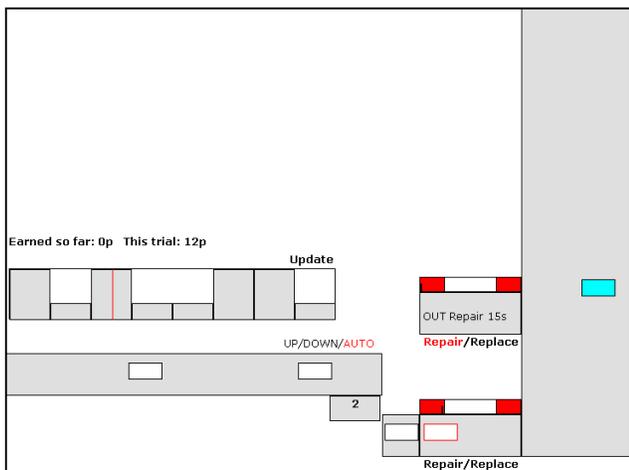

Figure 1. Microworld simulation PaintShop2

The current experiment used the same basic supervisory control task as PaintShop1, but focused it on the temporal aspects of the fault servicing decision by reducing the degrees of freedom in the control strategy, reducing the event rate levels, increasing the salience of temporal information and introducing deviations from the provided event rate information. In particular, the following changes to the PaintShop1 task (c.f. section 2.1.1) were made:

- The option to paint manually was removed. All items were painted automatically (duration: 4 seconds).

- The pile-up queue was shortened. As soon as the lift was occupied, any incoming items were removed.

- There was only one replace cost level, 6 pence.

- There were only two event rate levels, *high* (1 item every 2 seconds) and *low* (1 item every 4 seconds).

- In PaintShop1, event rate blocks had a duration of 1 minute. In this study, three different block durations were used: *short* (12 seconds), *medium* (24 seconds) and *long* (48 seconds). Fig. 1 shows a timescale with medium block duration.

- Event rate information was provided for *all* participants before each trial (e.g. for the trial in Fig. 1, the participant would have been told that it consisted of 24-second blocks and a sequence of high-low-high-low-low-high-high-low).

- For half the participants the event rate information was also visible in the timeline display (as in Fig. 1), for the other half the timeline only showed markers at block boundaries (i.e. every 12, 24 or 48 seconds).

- Participants had to work through 18 trials of about 3 minutes duration. In 6 of these trials, the event rate and block duration corresponded to the actual information participants were given before the trial. In 6 of the trials, there was a switch in the block structure, so that the actual event rate and block length information did no longer correspond to the information given before the trial or displayed in the timescale (within-subjects factor *schedule switch*). In these situations, both the block duration and the event rate distribution would change, never just one of these on its own. A switch would be noticeable either by a change in event rate that did not occur at an assumed event rate boundary (e.g. the event rate might change somewhere in the middle of a 48-second block if the actual schedule had changed to 12-second blocks), or by event rate changes that did not go in the expected direction (e.g. if a change from high to low event rate is expected in a 24-second schedule but instead the event rate remains high because the schedule has changed to 48-second blocks). In 6 of the trials, two of these changes occurred. Participants could request an updated version of the schedule. If the schedule had actually changed, the new, correct version would be displayed. If it had not changed, participants were informed of this fact. They could make a maximum of three update requests per trial.

- The timeline was available in half of the trials and unavailable in the other (within-Ss factor *timescale*).

4   Method

The control elements and factor structure were described in the previous section. The experiment was conducted as a fully computer-based task with the simulation implemented in JavaScript and running on a Windows 2000 PC. Participants received a 20-minute online tutorial that introduced them to all aspects of the system. This was followed by the 18 experimental trials.

After every trial, participants could provide feedback about the control strategies they used and about any problems that might have occurred during the trial. They also provided NASA Task Load Index ratings. As compensation, participants received the money they earned in the experiment, plus a flat rate of £5 (£10-15 in total). The participant with the highest score received an additional £20. 18 participants completed the experiment.

5 Results

The experiment manipulated availability of event rate information (*evRateInfo*) as between-subjects independent variable, and availability of online progress indicator (*timescale*) and number of schedule switches (*switches*; none, 1, 2) as within-subjects variables. For some of the analyses, data was aggregated by within-trial variables event rate (*evRate*; high, low) and block duration (*blkDuration*; 12s, 24s, 48s). Performance-related dependent variables were the amount of money earned per trial and the number of painted items. Strategy-related dependent variables were the number of repair and replace decisions, timescale requests, changes in lift direction, and requests for schedule updates.

Analysis of variance (ANOVA) of the performance-related dependent variables with the between-subjects factor *evRateInfo* and within-subjects factors *timescale, switches, evRate* and *blkDuration* showed no significant results. ANOVA of the strategy-related variables with the independent variable mentioned above showed the following significant results.

5.1 Requests for schedule updates

The frequency of participants' request for an updated version of the schedule was affected by the number of schedule switches ($F_{(2,32)}$=4.92, p<.05; Fig. 2).
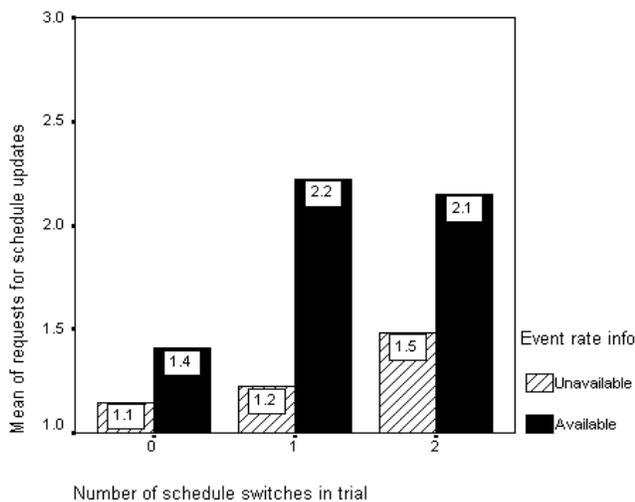
Figure 2. Effect of *switches* on schedule update requests

5.2 Timeline requests

There was a highly significant effect of *switches* on frequency of timeline requests ($F_{(2,30)}$=39.76, p<.001) and a marginally significant interaction of *switches* and *evRateInfo* ($F_{(2,30)}$=3.04, p<.065 ; Fig 3).
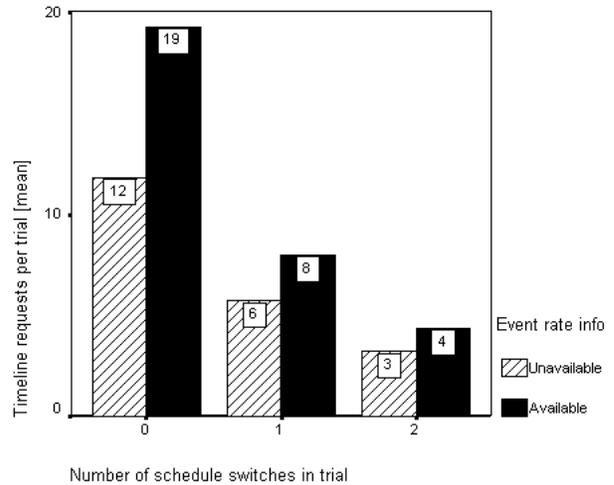
Figure 3. Interaction effect of *switches* and *evRateInfo* on timeline requests

5.3 Changes in lift direction

There was a highly significant effect of *switches* on frequency of changes in lift direction ($F_{(2,16)}$=91.29, p<.001; Fig. 4).
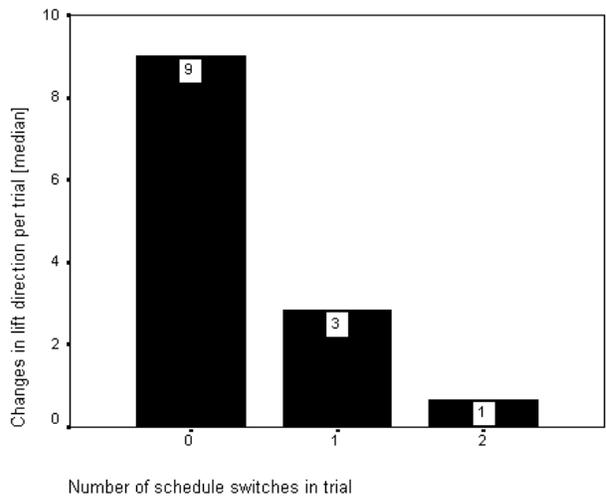
Figure 4. Effect of *switches* on changes in lift direction

5.4 Repair decisions

The number of repair decisions was significantly affected by *event rate* ($F_{(1,11)}$=25.17, p<.001), an interaction between *event rate* and *block duration* ($F_{(2,22)}$=8.99, p<.001) and an interaction between *timescale* and *block duration* ($F_{(2,22)}$=4.08, p<.05; Fig. 5).
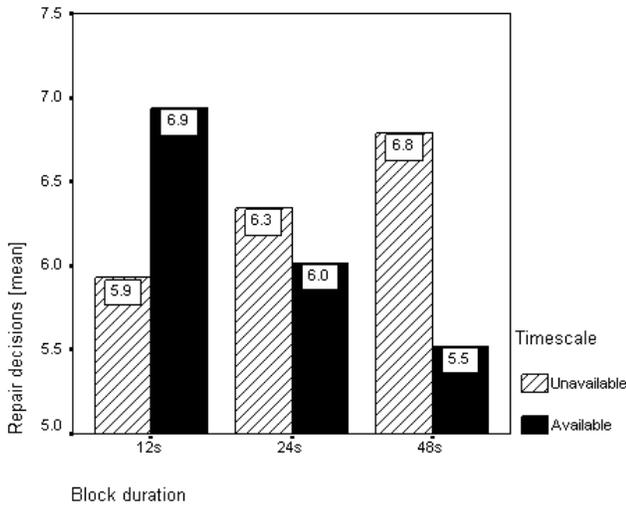
Figure 5. Interaction between *availability of timescale* and *block duration* on number of repair decisions

5.5 Replace decisions

The number of replace decisions was significantly affected by *event rate* ($F_{(1,5)}$=27.48, p<.005), an interaction between *timescale* and *switch* ($F_{(2,12)}$=5.88, p<.05; Fig 6), and an interaction between *timescale, switch* and *event rate* ($F_{(2,12)}$= 8.41, p<.005).
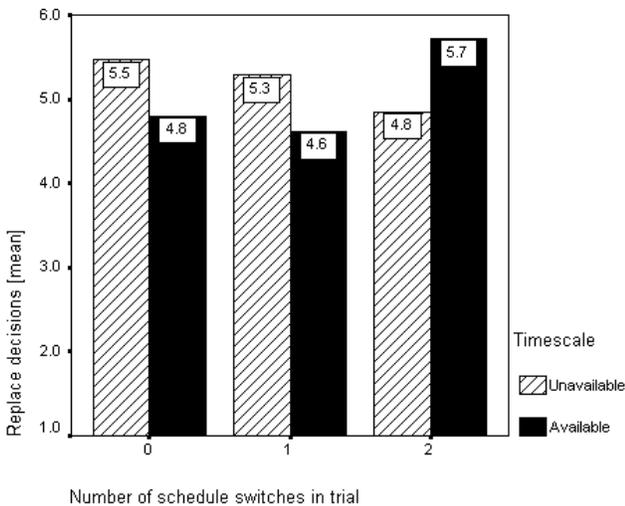

Figure 6. Interaction between *availability of timescale* and *switches* on number of replace decisions

6 Discussion

The main difference between this and the previous experiment was the introduction of "surprise" trials where the actual event rate schedule and block length would deviate from the expected one. Although this intervention had no direct effect on the performance measures, it had a significant effect on the strategy indicators. Fig. 4 suggests that under conditions of scheduling surprises (1 or 2 switches), participants abandon strategies that adapt

dynamically and anticipatively to the changes in workload levels, such as routing all items to one station and keeping the other one as backup. Fig. 3 shows further evidence that participants' interest in temporal information decreases with the occurrence of schedule deviations: timeline usage decreases dramatically when schedule switches are introduced, especially for the group whose timeline showed event rate information (as opposed to the group whose timeline only showed block boundary markers).

This effect is surprising because it cannot simply be attributed to the reduced utility of the timeline when it showed inaccurate information, as participants had the option to update the timeline so that it would show the current, valid schedule. That participants were indeed sensitive to schedule switches is suggested by Fig. 2, which shows the number of update requests made per trial. The group whose timeline display also showed event rate information managed, on average, to spot both changes in trials that included two switches. In trials with none or only one switch they made, on average, one unnecessary update request. The group whose timeline only showed block boundary information under-recognized schedule switches at least in trials that involved two switches.

That schedule switches not only had an effect on information gathering but also on strategy selection is documented in Fig. 6. During trials where the timeline was available, the number of replace decisions increased with the number of schedule switches. Based on the available data it is not clear whether this effect is due to erroneous fault servicing decisions during periods where participants worked under incorrect scheduling information, or to a genuine strategy shift as a reaction to conditions of temporal uncertainty (note that replace is a conservative temporal strategy as it makes the station available immediately).

Repair decisions, on the other hand, imply a higher degree of temporal commitment, as the station will be unavailable for 24 seconds. Fig. 5 suggests that in the absence of an online progress indicator, participants are more inclined to accept this temporal commitment under the relatively predictable conditions of 48-second blocks than under the frequent event rate changes associated with shorter blocks. The data pattern is reversed when the timescale is available; here the most repair decisions occur during 12-second blocks. As this condition has the most block boundaries (15, vs. 7 for 24-second blocks and 3 for 48-second blocks), participants may perceive the mere proximity to an event rate boundary as an opportunity for choosing a strategy that involves a temporal commitment. In situations where fault servicing decisions occur near an event rate boundary (from *high* to *low* or vice versa), participants may over-estimate the utility of a repair decision by either underestimating the

duration of the remaining *high* period or by overestimating the duration of the remaining *low* period.

7 Conclusions

The two studies investigated supervisory control decisions in rich temporal environments. Results suggest that increasing workload, changes in the expected event rate distribution, and longer durations of constant event rate may lead to a decrease in monitoring of temporal information and an increase in the use of time-conservative control strategies. Further studies are needed to disambiguate these results by dissociating event rate switches from block duration switches and by creating conditions where participants do not have the option to update scheduling information.

Work is currently under way to build normative utility models of this dynamic task against which the empirical data can be compared. These models will also provide a representation whereby the empirical results can be made accessible to system developers. Integrated into design-level approaches for exploring the functional aspects of time in socio-technical systems (c.f. [1]), these representations will support developers in analysing potential temporal decision biases and in providing design solutions to mitigate them.

References

[1]   M. Hildebrandt, A. Dix, and H. A. Meyer. Time design. *Extended Abstracts of CHI 2004,* pp. 1737-1738.

[2]   M. Hildebrandt, and M.D. Harrison. PaintShop: A Microworld Experiment Investigating Temporal Decisions in a Supervisory Control Task. *Proc. HFES 2004*, in print.

[3]   M. Hildebrandt, and M.D. Harrison. Putting time (back) into Dynamic Function Allocation. *Proc. HFES 2003*, pp. 488-492.

[4]   E. Hollnagel. The phenotype of erroneous actions: Implications for HCI design. In G.R.S. Weir and J.L. Alty (Eds.), *Human-Computer Interaction and Complex Systems*, pp. 73-121. Academic Press, London, 1991.

[5]   V. De Keyser. Time in ergonomics research. *Ergonomics*, 38, pp. 1639-1660, 1995.

[6]   C. Varey, and D. Kahneman. Experiences extended across time: Evaluation of moments and episodes. *Journal of Behavioral Decision Making*, 5, pp. 169-185, 1992.