The Impact of System Response Time on Navigating through Hypertext

Herbert A. Meyer and Michael Hildebrandt

Department of Psychology, University of Kassel

Abstract. Two experiments investigated the effects of system response time (SRT) on navigating through hypertext. SRT duration was varied systematically and manipulated between subjects. The dependent variables were residence time, emotional strain, and memory performance. All three variables were influenced by SRT. As a prominent result, a synchronization between computer's and human's response time was observed.

In descriptive terms, the distinguishing feature of browsing a hypertext database is that the user accesses the linked elements in an order of his own choice. To formalize this process, two temporal components can be identified. First, there is the technology-induced time of the transmitting a requested document from the server to the user's workstation. This time span is referred to as system response time (SRT). Second, there is the user-determined period of perception and engagement with the requested document. This period, during which a multitude of actions and decisions may occur, is collectively called residence time.

It could be assumed that residence time and SRT are independent of one another. However, research in software engineering and ergonomics clearly contradicts such an assumption. It turns out that not only too long SRT, but also too short SRT has serious effects on human-computer interaction. On the one hand, work productivity and satisfaction decreases as SRT increases. On the other hand, there is also evidence that very short delays may induce an agitated work style, leading to a reduction in performance accuracy (Barber, 1983; Boucsein, 1987; Kohlisch & Kuhmann, 1997; Miller, 1968). Shneiderman (1984) explained the agitated work style by the user's tendency to adopt the work speed of the system.

Two experiments manipulated SRT levels to investigate if these phenomena occur in browsing a hypertext interface to a large and homogeneous database consisting of entertaining pictorial material.

Experiment 1

Method

Materials. The hypertext database consisted of 2604 fashion photographs (1304 thumbnails, 93x142 pixel, and 1304 full-size pictures, 390x595 pixel) of 12 well-known designers, shot during fashion shows of the 1996 season. The control

structure supported navigation on three levels. Level (1) gave the choice of 12 designers. Choosing a designer led to the preview level showing eight thumbnails from the collection. There was a navigation panel consisting of three buttons for "previous eight pictures", "up to designer level", "next eight pictures" (represented as icons by left-, up-, or right-pointed triangles). Clicking any thumbnail led to the picture level where a large version of the respective photo was shown. There were again three navigation buttons for previous, up (to the thumbnail level), next. The materials were presented in a familiar World Wide Web browser (Netscape Communicator 4.5, opened to full screen size, with the Navigation-, Location- and Personal toolbar hidden). JavaScript was used to record navigation at 10ms accuracy and to simulate between-subject SRT of 0.5 s, 1.5 s, 2.5 s, and 3.5 s. When a link was activated, the page unloaded and an empty white background remained visible until the SRT had elapsed and the chosen page was displayed. The study was run on a Windows NT 4.0 workstation (Pentium II, 300 MHz, 64MB RAM), using a 17" color monitor at 1024x768 screen resolution, 85Hz refresh rate, 65536 colors. All input was by a Microsoft serial mouse.

Procedure. The experimenter demonstrated all features of the site using an alternative version of the database, which was set to the same SRT level as was used in the subsequent trial. When the participants had understood all three levels and the various navigation buttons, they were instructed to navigate through the site freely and on their own for some time until the experimenter returned to the room. They were told they participated in an experiment on time perception and should afterwards report how long the session had lasted, and give their impression of the site on a questionnaire. No mentioning was made of the SRT manipulation. The program terminated the session after 60 clicks. Immediately before and after the session, the participants were given a mood rating scale (SES; Hampel, 1977).

Participants. Participants were 44 students (23 woman, 21 men) between 20 and 39 years of age (mean 25.3) as paid volunteers. They were randomly assigned to the four experimental groups, tested individually and de-briefed at the end of the experiment.

Measures. To assess the effect of SRT on residence time, we obtained two dependent variables. The recorded residence times were classified as *horizontal navigation* (progressing through a sequence of either thumbnails or pictures, i.e. clicking "next-next"), or *vertical navigation* (terminating a sequence, changing direction, e.g. "next-up"). To test the influence of SRT on emotional strain we calculated the increase in annoyance while browsing the site (post/pre-difference on the mood scale). To control for the assumption that women are more interested in fashion than men, sex was entered as a factor in all analyses of variance

(ANOVA). Because there was no significant interaction between sex and SRT, we will not report results concerning this factor.

Results and Discussion

In the upper part of Figure 1, residence times for both horizontal and vertical navigation are plotted against SRT levels. The curves run almost in parallel and ANOVA confirmed a main effect in each case [F (3,36) = 4.45, p < .01, resp. F (3,36) = 8.29, p < .01]. In addition, SRT duration had a significant effect on emotional strain [F (3,36) = 3.47, p <.05]. The lower part of Figure 1 shows that beyond three seconds, SRT increased annoyance.

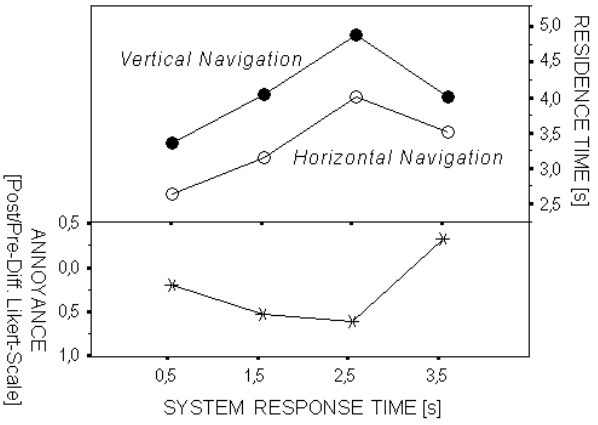

Figure 1. Results of Experiment 1 (N = 44)

Interpreting the above pattern, it seems as if the computer's speed paces the user's speed. This relationship brakes down only when SRT causes annoyance. A second experiment was carried out to investigate the cognitive correlates of this effect.

Experiment 2

Method

Materials and procedure were the same as in Experiment 1. SRT levels were set to 0.75 s, 1.25 s, 1.75 s, 2.25 s, 2.75 s, 3.25 s, and 3.75 s. An additional memory test was administered after the second mood rating. Without prior instruction, participants were given max. 12 pictures they had seen during browsing (*old*), mixed with max. 12 distractor pictures (*new*). The task was to decide if a picture was old or new. As a measure of recognition, discrimination indices *Pr* were

computed (Snodgrass & Corwin, 1988). Participants were 140 students (80 female, 60 male) aged 18 to 42 (mean 24.8) as paid volunteers.

Results and Discussion

The influence of SRT on annoyance could not be confirmed ($F < 1$), while the main effects for both residence time measures (vertical and horizontal) was replicated [F (6,126) = 5.57, $p < .01$, resp. F (6,124) = 4.71, $p < .01$]. The uppermost diagram of Figure 2 reports peak recognition performance at a SRT of 1.75 s [nearly significant; F (6,89) = 2.14, $p = .06$]. At both longer and shorter SRT, performance is worse.

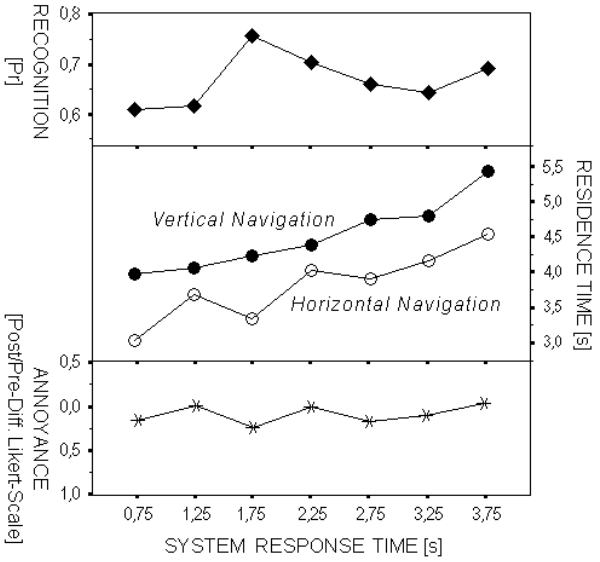

Figure 2. Results of Experiment 2 (N=140)

Importantly, there was again evidence for the pacing effect, indicated by the linear trend in the SRT/residence time plot (middle diagram of Figure 1). Both experiments are consistent in their results on the relationship between SRT and mood: Experiment 1 showed a drop in residence time only when SRT caused significant annoyance. There was no drop in residence time in Experiment 2, but no increased annoyance was observed either. The data concerning memory performance agrees well with findings in software engineering: very short SRT induces an agitated, careless work style, resulting in low accuracy. Long SRT impedes or disrupts the work flow by generating cognitive overhead.

Conclusions

Our results clearly confirm that the computer's response speed is an important characteristic of hypertext applications. We showed effects of experimentally manipulated SRT on emotional strain and memory performance, and especially on residence time. In a subtle way, the computer acted as a "pacemaker" for the user's hypertext exploration. Our results indicate that SRT should not *always* be reduced as much as technically possible, but that its duration can be optimized to support cognitive processing.

The negative effects on mood and memory obtained at the upper end of the SRT spectrum, i.e. beyond two or three seconds, should be taken very seriously by designers of hypertext applications. They are particularly relevant for developing learning applications by suggesting that memory performance can be optimized at a certain SRT level. Designers must have possible bandwidth restrictions of their target audience in mind and should, in many instances, trade graphical opulence for response rapidity.

To summarize, proper time design has to become an integral part of the design and evaluation process of hypertext applications. Further research is needed to develop inconspicuous interaction-tracking tools, and to set standards for the interpretation of the parameters obtained.

References

Barber, R.E & Lucas, H.C. (1983). System response time, operator productivity, and job satisfaction. *Communications of the ACM, 26*, 972-986.

Boucsein, W. (1987). Psychophysiological investigation of stress induced by temporal factors in human-computer interaction. In M. Frese, E. Ulich & W. Dzida (Eds.), *Psychological issues of human-computer interaction in the work place*, (163-181). Amsterdam: Elsevier.

Hampel, R. (1977). Adjektiv-Skalen zur Einschätzung der Stimmung (SES). *Diagnostica, 23*, 43-60.

Kohlisch, O., & Kuhmann, W. (1997). System response time and readiness for task execution: The optimum duration of inter-task delays. *Ergonomics, 40*, 265-280.

Miller, R.B. (1968). Response time in man-computer conversational transactions. In *AFIPS Conference Proceedings, 1968 Fall Joint Computer Conference* (Vol. 33, 267-277). Montvale, NJ: AFIPS Press.

Shneiderman, B. (1984). Response time and display rate in human performance with computers. *Computing Surveys, 16*, 265-285.

Snodgrass, J.G., and Corwin, J. (1988). Pragmatics of measuring recognition memory: Application to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34-50.