

# Time Pressure and System Delays in Information Search

Anita Crescenzi  
School of Information and  
Library Science, University of  
North Carolina  
Chapel Hill, NC, USA  
amcc@email.unc.edu

Diane Kelly  
School of Information and  
Library Science, University of  
North Carolina  
Chapel Hill, NC, USA  
dianek@email.unc.edu

Leif Azzopardi  
School of Computing Science,  
University of Glasgow  
Glasgow, United Kingdom  
leif@dcs.gla.ac.uk

## ABSTRACT

We report preliminary results of the impact of time pressure and system delays on search behavior from a laboratory study with forty-three participants. To induce time pressure, we randomly assigned half of our study participants to a treatment condition where they were only allowed five minutes to search for each of four ad-hoc search topics. The other half of the participants were given no task time limits. For half of participants' search tasks ( $n=2$ ), five second delays were introduced after queries were submitted and SERP results were clicked. Results showed that participants in the time pressure condition queried at a significantly higher rate, viewed significantly fewer documents per query, had significantly shallower hover and view depths, and spent significantly less time examining documents and SERPs. We found few significant differences in search behavior for system delay or interaction effects between time pressure and system delay. These initial results show time pressure has a significant impact on search behavior and suggest the design of search interfaces and features that support people who are searching under time pressure.

## Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—*Search Process*

## Keywords

Search Behavior; Time Pressure; System Delays

## 1. INTRODUCTION

Recently, increased attention has been paid to how system delays impact search behavior and the user experience in the context of interactive information retrieval (IIR) [1, 8, 15, 16]. Users of search systems with slower query response times have been found to have lower perceptions of system usability and helpfulness [1]. Studies examining system delays have found users are tolerant of search delays up

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or to publish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

to a “tipping point” beyond which behaviors and perceptions change [15]. User sensitivity to delay varies across individuals, the average speed of the system used [1], and how the information is presented [15]. It has also been found that searchers who experience both query and document response delays view fewer documents, but mark a higher percentage as relevant, although the delay had no impact on the number of queries issued, the number of documents marked relevant, the depth of results inspected or search accuracy [8].

Although there has been a resurgence of interest in the impact of search system delay on user experience, a long line of research precedes this interest. In 1968, Miller [9] described system response times exceeding two seconds as changing the character of “man-computer conversational transactions” and provided a taxonomy of acceptable response times depending on the type and complexity of system interactions and the extent to which the response delay would disrupt the user's task. Since this early work, researchers have found that perceived wait time and the impact of delays can be reduced by informing users of the system response status [11], incrementally loading pages [2], or displaying filler text or images [6] while the system is processing its response.

While system delays have been investigated in IIR, one variable that has not been investigated much is time pressure. Research on time pressured decision-making has found that people adapt under time pressure: they may spend less time in each decision stage, spend less time processing individual pieces of information, accelerate the pace of the task or information processing, or even satisfice in their decisions [12]. Time has been shown to be an important contextual factor in search [13]; however, the impact of time pressure on search behaviors and perceptions is less well understood. Recent studies have shown when searchers have less time to complete tasks, they have lower perceptions of their search performance [7] and task completion [17], less satisfaction with results [7, 17] and less knowledge gained [7]. Searchers also report lower pre-task search confidence [7] and more negative affect [7] and stress [17]. Time pressure has also been found to impact search satisfaction even when the same amount of time was given to complete tasks [4]. Finally, Mishra et al. [10] found dissatisfaction among people who searched for time-critical information.

In this paper, we report preliminary results of a study investigating the impact of time pressure, system delays and interaction effects on search behaviors and user experience. We present the results of our analyses of interaction measures as well as time pressure and delay manipulation checks.

## 2. METHOD

We conducted a laboratory experiment where we manipulated task time limit and system response time. To manipulate task time limit, half of the participants were given an unexpected task time limit just before they started the experimental search tasks. This time limit was presented via pop-up message after they completed the practice task and informed them they had only 5 minutes to complete each experimental search task (Figure 1). These instructions were intended to simulate a situation in which an urgent notification changes one’s expectations about the time available to complete a task. The 5 minute time limit represented approximately half of the mean task completion times from a prior study [5] where different participants used the same system to complete three of the same tasks without time limits. The system tracked the time, and after participants had been searching for 5 minutes, popped-up another message indicating time was up.

### Before you begin



You just received a message that your research meeting has been just rescheduled for tomorrow and you only have 20 minutes today to create your lists of articles on each topic.

Since you only have 20 minutes, you decide to limit yourself to search for 5 minutes per topic. The system will let you know when you have searched for 5 minutes.

Click "Next" to start the first search task.

Figure 1: Task time limit instructions

Participants who were not given a task time limit were not provided with any instructions regarding the amount of time to spend on each task, although all participants were aware that the experimental session would last 1hr and 15min. If these participants were still searching after 1hr and 5min (i.e., with approximately 10 minutes remaining in the experimental session), we stopped them so they could complete the exit questionnaire and interview. Based on our previous experiences, we did not expect we would need to do this; however, this occurred in seven cases. To minimize time pressure caused by researcher presence [14], the researcher was stationed in an adjacent room while participants completed the study.

System response time was a within-subjects variable with two levels: delay and no delay. Participants completed two tasks in which we introduced 5 second delays after they submitted queries and clicked on SERP results and two tasks in which the system was not delayed. We selected a 5 second delay based on prior research which indicated that a delay of 4 seconds changed behaviors but not attitudes and is usually considered acceptable to users (delays of 7-10 seconds have been found to be unacceptable by users [15]). We presented delays in 2 orders: half of the participants experienced delays during their first and third tasks while the other half experienced delays during their second and fourth tasks.

The basic procedure was as follows. Participants were first shown the experiment instructions on screen including the simulated work task scenario and then completed a practice task to become familiar with the system; we did not impose a time limit or system delay during the practice task. Once participants completed the practice task, those with a task time limit were shown the instructions in Figure 1 before being presented with their first search task. For each search task, participants were shown the topic, completed

a pre-search questionnaire, conducted their search and then completed a post-search questionnaire. After conducting all four searches, participants completed an exit questionnaire. Following this, the researcher returned to the room for a final interview and debriefing.

### 2.1 Scenario and Tasks

Following Borlund’s [3] guidelines, search tasks were situated in the context of a project in which a research team needs to compare news coverage about four topics during the (consecutive) U.S. presidential terms of Clinton, Bush and Obama. Participants were told their job was to examine news during Clinton’s second term, 1996-2000, which corresponded to the date range of the test collection [18]. We wanted the scenario to provide participants with a rationale for searching articles from this time period and to make this characteristic of the collection salient to help ground participants’ relevance judgments. The scenario indicated that participants should find 8-12 articles per topic.

We selected topics from the collection with contemporary relevance where prior studies have reported that participants found the tasks at least moderately interesting and difficult [5]. We aimed to minimize potential variability in interest and difficulty in order to examine the impact of our independent variables. Four topics were used: wildlife extinction (347), journalist risks (354), piracy (367) and population growth (435). Participants were presented with the topic descriptions before each search.

To minimize order effects, we created a set of 24 topic rotations. We assigned the topic rotations to time limit and delay order combinations such that each topic rotation was used twice: once in each time limit condition and once in each delay order. We randomly assigned participants to time limit, delay order and topic rotation.

### 2.2 Search Questionnaires

Our search questionnaires measured several constructs including system usability, task difficulty, metacognitive actions and satisfaction. Several items functioned as manipulation checks to see if participants felt time pressure and noticed system delays (see Table 1). Participants responded to these items with a 7-point scale (1 = strongly disagree, 7 = strongly agree). In this paper, we only present items related to the manipulation checks and will analyze the additional measures in future work.

### 2.3 Search System

All participants used the same search system which resembled a standard search engine. The Whoosh IR Toolkit was used as the core of the retrieval system, with BM25 as the retrieval algorithm, using standard parameters, but with an implicit ANDing of query terms to restrict the retrieval set. Only a portion of the collection was indexed to ensure that the natural response time of the system was quick and constant. This was to control the effects of both the natural and experimental delays as much as possible. The documents included in the index consisted of all the TREC pooled documents for the given topics, along with the top 100 documents retrieved from the full index given a set of user generated queries from a past study (approximately 100-200 queries per topic). In total, over 200,000 documents were indexed. In all conditions, the search system displayed a spinning wheel when it was busy.

## 2.4 Participants

Forty-five people from a major research university participated including undergraduate students ( $n=13$ ), graduate students ( $n=8$ ) and staff ( $n=22$ ) representing many different majors and occupations. Participants ranged in age from 18-59 years ( $M=32.1$ ,  $SD=12.8$ ) and 31 were female. Participants were compensated \$20 USD.

## 3. RESULTS

We analyzed 163 tasks completed by 43 participants using a mixed ANOVA. We excluded data from 2 participants with no task time limit who spent most of the time completing a single search task, 7 tasks where participants with no task time limit were stopped in their third task or final task, and 2 tasks where participants accidentally ended the tasks by pressing the wrong button.

### 3.1 Time and Delay

We conducted a series of manipulation checks to make sure our experimental manipulations were successful. We found those with the task time limit reported feeling significantly more time pressure, a greater need to work fast and more rushed (Table 1). We found significant differences in perceived quickness of SERP and document display between the delay conditions. We found no interaction effects, but note a significant effect of delay on perceived time pressure.

We examined task completion times and found significant effects of task time limit and system delay on the total task time as shown in Table 1. The introduction of time limits and system delays complicates comparison of interaction measures as the total delay experienced by any one person is a function of the number of queries issued and documents viewed by this person. For this reason, we also examined adjusted task times where we subtracted the cumulative delay experienced by each participant for each task from the total task time; we found significant main effects of time limit and no main effect of delay. Interaction effects were significant ( $F(2,118)=7.03$ ,  $p<.01$ ).

To check the delay manipulation, we examined query and document processing times (in seconds). As expected, we found no significant differences in organic system response times and significant differences in the mean total processing time (organic + delay) according to delay condition. It took longer for the system to return requests for tasks with delay condition (query:  $M=5.92$ ,  $SD=0.96$ ; document:  $M=5.20$ ,  $SD=1.16$ ) than for tasks without delay (query:  $M=0.94$ ,  $SD=0.84$ ; document:  $M=0.29$ ,  $SD=0.34$ ). We found no main effect for time limit or interaction effects.

### 3.2 Search Behaviors

Table 1 shows the means and standard deviations of participants' search behaviors by time limit and response time condition. With one exception, there were no significant interaction effects, so we only report main effects. We normalized time-based measures to facilitate comparison. The adjusted rate-based measures reflect the number of queries, document views and documents marked relevant per minute using adjusted times; that is, total time minus time spent in delay as in Taylor, Dennis & Cummings [15]. While participants who experienced a task time limit had higher rates of querying, document viewing and marking documents as relevant, only the difference for query rate was significant. There were no significant differences according to whether or

not participants experienced delays; in fact, these measures were very similar in these two conditions.

When we examine actions per query, we see those who experienced the task time limit examined search results more shallowly. They viewed fewer SERPs and documents and also did not navigate as deeply in the search results list. All of these differences were statistically significant. Participants with a task time limit also spent significantly less time viewing documents and SERPs. Although the trend was for participants to view fewer SERPs and documents per query when they experienced delays and not go as deeply in the search results list, these differences were not significant. Participants spent significantly more time viewing documents for delayed tasks, but they spent roughly equal amounts of time viewing SERPs for their delayed and non-delayed tasks.

With respect to performance, there were no differences in participants' save rates (% viewed marked relevant) according to time limit or response time conditions. Although we have not examined the content of the documents participants saved, we examined whether participants met the scenario goal of finding 8-12 relevant articles. This goal was achieved more often for tasks where no task time limit was present (96% vs. 84% of tasks) and for tasks where no delays were present (92% vs. 86% of tasks). Only 76% of tasks with both a delay and a task time limit had at least 8 pages marked relevant compared to 100% of tasks with no task time limit and delay.

## 4. DISCUSSION AND CONCLUSIONS

We investigated the impact of time pressure and system delays on search behaviors. Results showed participants who were given task time limits queried at a significantly higher rate, viewed significantly fewer documents per query, had significantly shallower hover and view depths, and spent significantly less time SERPs. They also spent significantly less time examining documents, and presumably, making judgments about which documents to save.

We did not find significant differences in search behavior with respect to system delays beyond the time spent completing the task (total and adjusted) and the time spent per document. While the delay was noticed by most participants, it did not significantly impact their search behaviors. These findings are consistent with other studies [1, 8]. Of course, it is likely that larger delay times would produce different results: Taylor et al. [15] found a tipping point between 7 and 11 seconds for behavioral impacts but some changes to satisfaction were observed at lower thresholds. Whether the delay is consistent or intermittent also seems to play a role [1]. Further analysis of the questionnaires and debriefings will likely provide additional insight about how time pressure and system delays impacted the user experience.

Interestingly, we found considerable variability in the task completion times for those without task time limits and two participants, whose data were excluded, used nearly the entire experimental session to search for a single task. For those tasks included in analyses, task completion times ranged from 1.76min to 36.27min. For 10 tasks, the total task completion time exceeded 15min. These variations surprised us because they differ considerably from what we have observed in previous laboratory studies, where the general challenge is often to create tasks that will require sustained interaction.

	Time Limit (main effects)			Response Time (main effects)		
	Session	Task	F	No Delay	Delay	F
<b>Manipulation checks</b>						
felt time pressure	3.43 (1.62)	5.25 (1.62)	15.27***	4.29 (1.90)	4.61 (1.80)	4.12*
needed to work fast	3.56 (1.56)	5.62 (1.50)	27.30***	4.70 (1.82)	4.71 (1.86)	0.00
felt hurried or rushed	3.03 (1.49)	4.91 (1.75)	17.93***	4.03 (1.96)	4.11 (1.82)	0.18
displayed search results quickly	4.31 (1.67)	4.65 (1.78)	0.99	4.96 (1.58)	4.01 (1.76)	19.73***
displayed articles quickly	4.39 (1.54)	4.68 (1.92)	0.63	5.00 (1.68)	4.09 (1.74)	17.72***
<b>Interaction measures</b>						
total task time (m)	9.04 (6.12)	4.60 (0.98)	21.82****	6.13 (4.00)	7.01 (5.27)	7.61**
adjusted task time (m)	8.25 (5.63)	4.08 (1.02)	21.29****	6.13 (4.00)	5.70 (4.67)	0.03
adjusted query rate	0.61 (0.44)	1.04 (0.61)	16.72***	0.85 (0.61)	0.85 (0.54)	0.03
adjusted view rate	2.28 (0.93)	2.63 (1.24)	2.55	2.52 (1.18)	2.42 (1.07)	1.60
adjusted mark relevant rate	1.64 (0.98)	2.05 (1.24)	3.22	1.85 (1.19)	1.89 (1.12)	0.00
SERPs viewed per q	2.25 (1.67)	1.64 (0.79)	5.36*	1.98 (1.54)	1.83 (0.96)	0.90
docs viewed per q	6.09 (5.75)	3.90 (3.47)	4.81*	5.08 (5.04)	4.65 (4.41)	0.71
hover depth per q	17.52 (16.70)	11.34 (8.78)	5.23*	14.77 (15.46)	13.34 (10.43)	0.77
view depth per q	14.88 (15.95)	8.89 (7.98)	6.51*	11.91 (14.46)	11.15 (10.11)	0.20
total time per doc (s)	14.05 (6.50)	8.82 (5.46)	17.26***	10.22 (5.54)	12.07 (7.23)	10.63***
total SERP time per q (s)	61.53 (60.35)	34.86 (23.44)	13.18***	46.73 (52.80)	46.55 (36.98)	0.00
save rate	0.70 (0.18)	0.73 (0.21)	0.64	0.69 (0.20)	0.74 (0.20)	2.30

**Table 1: Manipulation check and interaction signals. Means (sd) and F-test results by time constraint (df=1, 41) and response time (df=1,118), \*p<0.05; \*\*p<0.01; \*\*\*p<0.001; \*\*\*\*p<0.0001**

While preliminary, our initial results open a new line of inquiry into how time pressure impacts search behaviors and how search tools might be designed to support people who are searching under time pressure. For example, interface features that make it easier for people to query, assess relevance and monitor progress might be especially beneficial. Displaying fewer results per SERP might also help keep people who are searching under time pressure stay focused.

## 5. REFERENCES

- [1] I. Arapakis, X. Bai, and B. B. Cambazoglu. Impact of response latency on user behavior in web search. In *Proc. of the 37th ACM SIGIR conference*, pages 103–112, 2014.
- [2] N. Bhatti, A. Bouch, and A. Kuchinsky. Integrating user-perceived quality into Web server design. *Computer Networks*, 33(1-6):1–16, 2000.
- [3] P. Borlund. The IIR evaluation model: A framework for evaluation of interactive information retrieval systems. *Information Research*, 8(3):1–34, 2003.
- [4] A. Crescenzi, R. Capra, and J. Arguello. Time pressure, user satisfaction and task difficulty. *Proc. of the American Society for Info. Sci. and Tech.*, 50(1):1–4, 2013.
- [5] A. Edwards, D. Kelly, and L. Azzopardi. The impact of query interface design on stress, workload and performance. In *Proc. of the 37th ECIR Conference*, pages 691–702, Vienna, Austria, 2015.
- [6] Y. Lee, A. N. K. Chen, and V. Ilie. Can online wait be managed? The effect of filler interfaces and presentation modes on perceived waiting time online. *MIS Quarterly*, 36(2):365–394, 2012.
- [7] C. Liu, F. Yang, Y. Zhao, Q. Jiang, and L. Zhang. What does time constraint mean to information searchers? In *Proc. of the 5th IIX Conference*, pages 227–230, 2014.
- [8] D. Maxwell and L. Azzopardi. Stuck in traffic: How temporal delays affect search behavior. In *Proc. of the 5th IIX Conference*, pages 155–164, 2014.
- [9] R. B. Miller. Response time in man-computer conversational transactions. pages 267–277. 1968.
- [10] N. Mishra, R. W. White, S. Jeong, and E. Horvitz. Time-critical search. In *Proc. of the 37th ACM SIGIR conference*, pages 747–756, 2014.
- [11] R. Molich and J. Nielsen. Improving a human-computer dialogue. *Comm. of the ACM*, 33(3):338–348, 1990.
- [12] J. W. Payne, J. R. Bettman, and E. J. Johnson. *The Adaptive Decision Maker*. Cambridge Uni. Press, 1993.
- [13] R. Savolainen. Time as a context of information seeking. *Library & Info. Sci. Research*, 28(1):110–127, Mar. 2006.
- [14] D. N. Stone and K. Kadous. The joint effects of task-related negative affect and task difficulty in multiattribute choice. *Organizational Behavior and Human Decision Processes*, 70(2):159–174, May 1997.
- [15] N. J. Taylor, A. R. Dennis, and J. W. Cummings. Situation normality and the shape of search: The effects of time delays and information presentation on search behavior. *J. of the A. Soc. for Info. Sci. and Tech.*, 64(5):909–928, 2013.
- [16] J. Teevan, K. Collins-Thompson, R. W. White, S. T. Dumais, and Y. Kim. Slow search: Information retrieval without time constraints. In *Proc. of the 7th HCIR Symposium*, pages 1–10, 2013.
- [17] A. Tombros, I. Ruthven, and J. M. Jose. How users assess web pages for information seeking. *J. of the American Society for Info. Sci. and Tech.*, 56(4):327–344, 2005.
- [18] E. M. Voorhees. Overview of the trec 2005 robust retrieval track. In *Proceedings of TREC-14*, 2006.