# Manipulating Time Perception of Web Search Users

Cheng Luo[∗◦], Fan Zhang[∗□], Xue Li[∗◦], Yiqun Liu[∗◦], Min Zhang[∗◦], Shaoping Ma[∗◦], Delin Yang[∗+]
*Tsinghua National Laboratory for Information Science and Technology
[◦]Department of Computer Science and Technology, [□]Department of Physics,
[+]School of Economics and Management
Tsinghua University
Beijing 100084, China
luochengleo@gmail.com

## ABSTRACT

Time is an important factor in information retrieval studies including search evaluation, user behavior analysis and query understanding. In most of the previous works, time is usually an objective factor measured by timing devices. However, the time perceived by user seems more intuitive to describe the impact of time because search user's opinion is considered subjective. Psychological researches have reported that time perception can be affected by many physical and psychological factors. In this work, a laboratory study with 50 participants was adopted to investigate the impact of *Temporal Relevance*, e.g., the awareness of elapsed time, on time perception of Web search users. Experimental results show that participants in high temporal relevance environments tend to perceive significantly longer task durations than the actual ones. It shows that the perception of time can be manipulated in Web search scenario and reveals the necessity to take the factor of user perception into consideration in time-related Web search researches such as effort-based evaluation.

## CCS Concepts

•**General and reference** → **Evaluation;** *Performance;*
•**Information systems** → **Information retrieval;**

## Keywords

time perception, temporal relevance, search engine

## 1. INTRODUCTION

Time plays an essential role in multiple areas of Information Retrieval (IR) researches, such as search evaluation, user behavior analysis and search query understanding.

In a typical Web search scenario, a user issues a query to the search engine and hope to fulfill his/her information need with reasonable effort (search duration shouldn't be too long). From the system's perspective, the time reflects the effectiveness of the retrieval system. The users with slower response would have lower perceptions of system usability [1]. From the user's perspective, the user's expected and actual efforts spent on search task provide important implicit feedback of the experience in the search process, user variability and IR system evaluation [2].

In *search evaluation*, time-biased gain (TBG) [11] computes the benefits obtained by the users and accounts for user effort taken to achieve those benefits in terms of time. As a more recent instantiation of TBG, time well spent (TWS) [3], expressed as the total time spent consuming relevant material v.s. the total time spent searching, measures both benefits and effort in meaningful units. In *user behavior analysis*, click dwell time is one of the most important features used to identify satisfied clicks. The clicks with dwell time longer than 30s are usually assumed to be more likely from users who are reading relevant results than short clicks (shorter than 15s) [6]. In *query understanding*, the time factor should also be taken into consideration. First, queries with temporal intents should be identified so that search engines can decide whether it should rank results based on timeliness or not. Second, search engine should be able to identify situations where users have urgent information needs in the context of an acute problem [9].

To the best of our knowledge, time mentioned in previous works is the objective time, measured by timing devices in experiments or systems. However, the perceived time seems more intuitive to describe the impact of time on IR evaluation where user behavior is considered subjective.

Time perception, referring to the subjective experience of time, is a construction of the brain that is manipulable and distortable under certain circumstances. In the field of psychology, substantial studies indicated that many aspects of our cognitive and behavioral functioning are based on processing temporal information to some extent [10]. Psychologists have found that the perception of the passing of time is influenced by many subjective factors, such as body temperature, emotion and etc [7].

As a factor that has an impact on interval length estimation in psychology, *Temporal Relevance* (TR) was defined by Zakay [13] as the *"level of relevancy and importance of time dimension in a specific state required for the optimal adaptation to the external environment"*. TR is one of the main determinants of the level of temporal awareness [8].

From this definition, we can see that different tasks may lead to different levels of TR in Web search scenario. For



*CHIIR '16, March 13 - 17, 2016, Carrboro, NC, USA*



ISBN 978-1-4503-3751-9/16/03…$15.00
DOI: http://dx.doi.org/10.1145/2854946.2854994

example, users who are supposed to meet urgent information needs (e.g. collecting information for an unfamiliar topic in a few minutes as described in the cases from [4]) should have high TR in their tasks. However, it remains uninvestigated whether TR will affect the time perception of Web search users and further have influence on the perception of system effectiveness.

In this paper, we report preliminary results of a study investigating the impact of TR on time perceiving and user experience in Web search scenario. Especially, we focus on the impact of TR on time perception in completing search tasks. We conduct an experimental study by manipulating temporal relevance for different groups of participants and measure their time perceptions while completing search tasks. Experimental results show that participants in high TR scenario tend to perceive longer durations than they actually spent in tasks, which suggests that we could make a better inference of user experience by considering the scenario the user involves in. This result also accords with existing findings in more general settings [12], which shows the probability of extending more findings in human time perception to Web search related studies.

## 2. EXPERIMENT SETUP

We conducted a laboratory study where TR is manipulated by specific signals in search interface. Users were asked to estimate the time they have spent in each search task under different manipulation settings.

There were 16 informational search tasks adopted in the experiment. For each task, we first randomly selected a medium-frequency (1,000 to 10,000 monthly) query from the log of a commercial search engine. Then a backstory was created as suggested in [2] to explain the information need of the task. Three annotators worked together to create the backstories based on the original query and corresponding clicked results. The backstories were carefully discussed until agreements were reached. In our experiment, the backstories were read and recorded by one of the authors and the same recordings were played to all the participants to ensure that they receive the same task information.

To encourage the participants to engage more into the tasks, the assessors organized a question for each task which required the participants to summarize the information gained in the search process. Participants had to answer the question correctly to get the payment to guarantee quality. An example of our tasks is shown in Table 1.

**Table 1: An example of search tasks**

| | |
|---|---|
| Query | Miranda Warning |
| Backstory | There is a famous concept called Miranda Warning in the law of the U.S. Please search for the explanation and history. |
| Question | Briefly explain what is Miranda Warning with your own words. |

We built an Web-based experimental search system to provide modified search results from a popular commercial search engine. All ads and sponsors' links were removed. We also removed the vertical results and query suggestions to reduce possible behavior biases during searching. Besides these changes, the search system looked just like a traditional commercial search engine. The users could issue a query, click results and switch to the landing pages in the Web page as they wish. Query reformulation was not allowed in our system to ensure that users get the same results from the search engine and TR would not be affected by other factors such as result quality.

We randomly assigned half of our participants to a treatment condition where they were shown a "timing block", a flashing colored block, on top-right corner of the search result page (SERP). The Web-based search system is represented on an Dell workstation with a 17" display running at 1366*768. The timing block was 200*20 pixels in size. A coloration and flashing scheme was applied to better visually inform the user about elapsed time: the colored block was initially in green and flash slowly. In the search process, as time goes on, the bar would subsequentially turn to orange, red and it would flash more and more frequently. An example of user interface is shown in Figure 1(a). The other half of participants were given no suggestion about the purpose of the experiment and the timing block was not shown on their SERPs. We did not use the progress bar to prevent the participant estimate the duration based on the distance of progress bar. According to their findings, the coloration and flashing scheme adopted in our work would strengthen the time pressure during search process significantly.

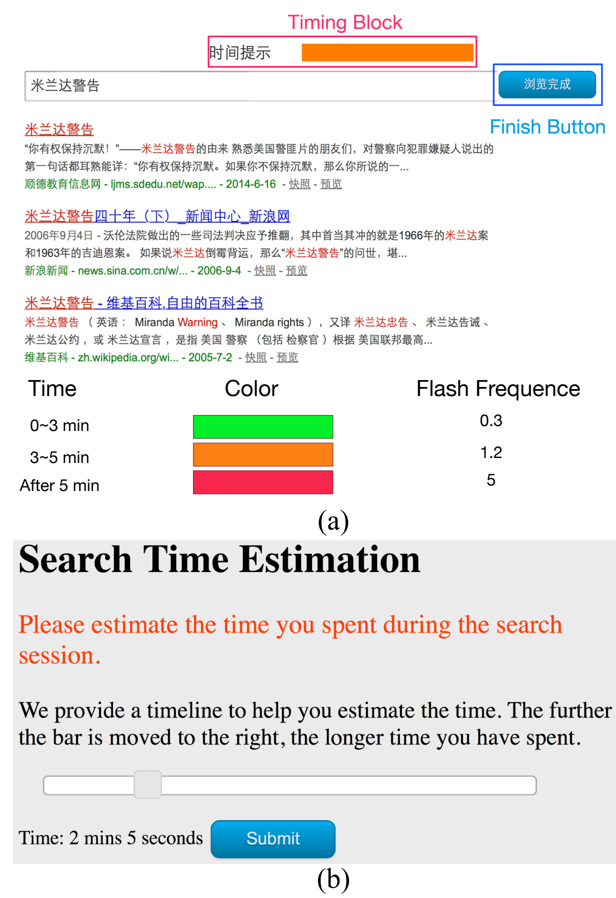


**Figure 1: User interface for the treatment group (The length of the timeline is 10 mins because each task was expected to be finished in 10 mins, but not had to be finished in 10 mins)**

Before the experiments began, we first asked the participants to take off their watches or other timing devices such as cellphones, tablets and music players. The clock on computer desktop was also removed. During the experiment, the participants were not allowed to acquire time from external environments. Then they were shown a video instruction of the experiments. The only difference in the instruction video between the treatment and control groups was the introduction to the timing block on SERPs. For both groups of participants, we had no time limit for the tasks. For each task, first the backstory was read to them via a voice element on the Web page. The backstory could be repeated for multiple times until the participant understood the intent. Then the participant could search for relevant information with the given query and corresponding SERP. He/she could finish the search by clicking the finish button on the SERP. After searching, the participant was required to estimate the time spent during the task in seconds, e.g., the duration of the whole search session as shown in Figure 1(b). Since they never knew how their perceived time varied from actual time, their ability of estimating time wouldn't improve. Then he/she was asked to answer the predefined question and annotate his/her satisfaction in five-point scale. After 6 tasks, he/she was asked to take a short break in case of fatigue. After the completion of all the tasks, an interview was performed to investigate their experience in the experiment.

We recruited 50 undergraduate students (23 females and 27 males) from a university. All of them reported that they were familiar with basic usage of search engines. For each participant, he/she was randomly assigned to the control/treatment group and required to finished 12 of the 16 tasks. We dynamically selected tasks for them to make sure all the tasks were finished in a balance way. A Latin-Square arrangement was used to minimize the order effects. All the interactions such as mouse clicks and mouse movements were recorded by the experimental system for further analysis. Each participant was compensated $20 USD.

# 3. RESULTS AND ANALYSIS

We first look into the average of objective dwell time (dtime) and perceived time (ptime) in both groups. Both dtime and ptime are the time spent during each task in seconds. The results in Figure 2 show that the average dwell time in treatment group is shorter ($p$-$value$ $< 0.01$) than that in control group. A potential reason is that the participants in high TR environment would perceive more time pressure and spend less time processing individual information pieces [5].

The user perceived time (ptime) in the treatment group is significantly longer (10.36% with $p$-$value$ $< 0.01$) than the dwell time while no significant difference is observed between dwell time and perceived time in the control group, which means that in high TR situation, human beings are more sensitive to the time effort and they attend to perceive longer durations than usual. This result also accords with previous psychological experiments in general settings [12].

We also compare the difference between average dwell time and user perceived time of each task in Figure 3. For 15 of the 16 tasks, the average perceived time in treatment group is longer than the average dwell time. While in the control group, the perceived time is very close to the dwell time. In both control group and treatment group, the variance of dwell time and perceived time between users is very great, which means that the actual dwell time and perceive time on each task varies significantly across users and tasks.

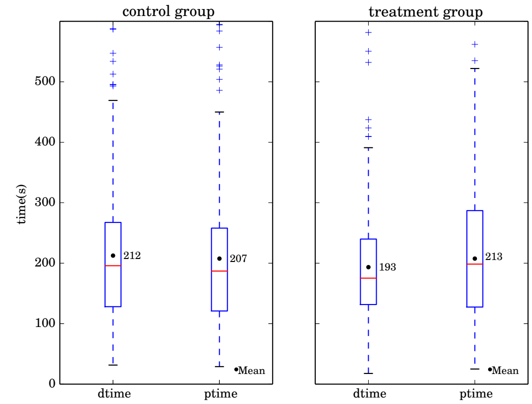


**Figure 2: The comparison between average dtime and average ptime**

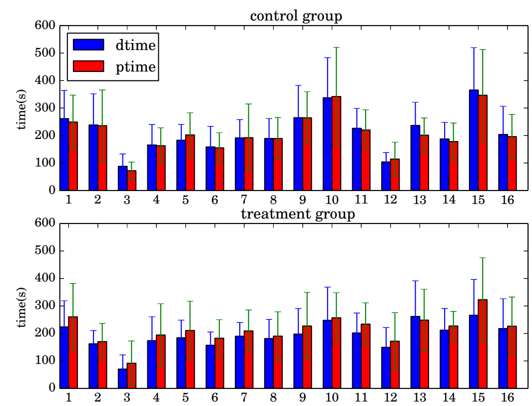


**Figure 3: The comparison between dtime and ptime on each task**

For each individual participant, we calculate the *average estimation offset* (AEO) on all the tasks finished by the participant. Figure 4 shows that there are more participants whose AEO are positive in the treatment group than in the control group. The difference of a user's time perception in different TR settings might be helpful for us to understand the ability and variability of the time perception. We would like to leave this problem in our future work.

$$AEO = \sum_{t \in tasks} \frac{ptime_t - dtime_t}{dtime_t} \qquad (1)$$

For each individual task, we count the number of users whose perceived times longer or shorter than dwell time in the control/treatment group. The results are represented in Figure 5. In the treatment group, for 14 of the 16 tasks, there are more users who have longer perceived time than dwell time. While in the control group, the participants in two conditions (dtime>ptime, dtime<ptime) follow a more balanced distribution.

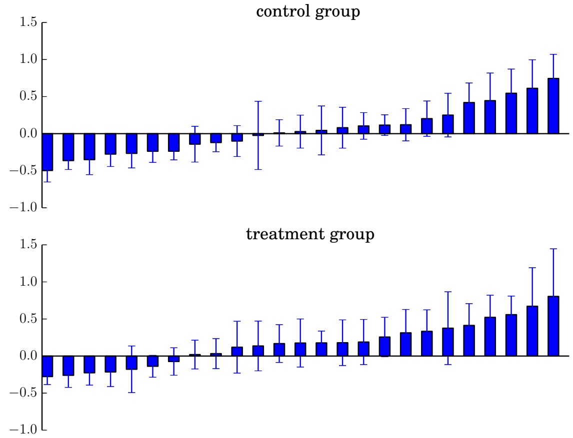


**Figure 4: The Average Estimation Offset (AEO) of participants from control/treatment groups**

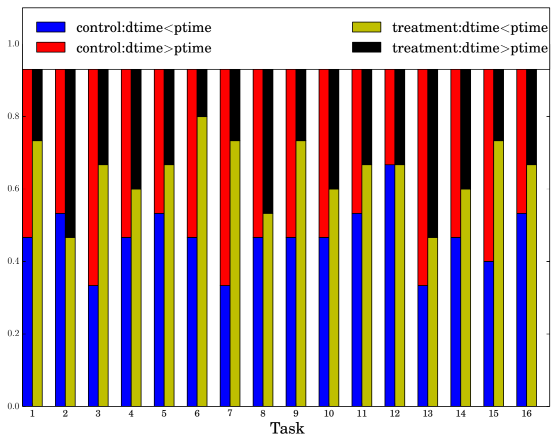


**Figure 5: The distribution of users whose perceived times are longer/shorter than dwell times**

Based on these results, we can conclude that the TR, e.g., the awareness of elapsed time, would have a significant effect on time perception in Web search environment. The users in high TR settings tend to perceive longer time than the participants in low TR settings. This finding may present a potential way to improve time based evaluation frameworks (such as TBG, etc.) by replacing the objective measured time with the user perceived time.

# 4. CONCLUSION AND FUTURE WORK

In this work, we investigate the impact of TR on users' time perception in Web search environment. Results showed that the participants who were in a high TR settings had longer perceived time than the ones in low TR settings, which accords with the previous findings in psychology, which means that we could make better inference of user experience by considering the temporal features of search tasks and the scenario the users involved in. In the future work, we plan to explore the relationship between TR, time perception and user satisfaction and improve the time-based measurements by incorporating the user perception information into existing frameworks.


# 5. ACKNOWLEDGMENTS

This work was supported by Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation, National Key Basic Research Program (2015CB358700) and Natural Science Foundation (61532011, 61472206) of China. Part of the work has been done at the Tsinghua-NUS NExT Search Centre, which is supported by the Singapore National Research Foundation & Interactive Digital Media R&D Program Office, MDA under research grant (WBS:R-252-300-001-490).


# 6. REFERENCES


[1] I. Arapakis, X. Bai, and B. B. Cambazoglu. Impact of response latency on user behavior in web search. In *Proceedings of the 37th international ACM SIGIR'14*, pages 103–112. ACM, 2014.

[2] P. Bailey, A. Moffat, F. Scholer, and P. Thomas. User variability and ir system evaluation. In *Proceedings of the 38th International ACM SIGIR'15*, pages 625–634, New York, NY, USA, 2015. ACM.

[3] C. L. A. Clarke and M. D. Smucker. Time well spent. In *Proceedings of the 5th Information Interaction in Context Symposium*, IIiX '14, pages 205–214, New York, NY, USA, 2014. ACM.

[4] A. Crescenzi, R. Capra, and J. Arguello. Time pressure, user satisfaction and task difficulty. *Proceedings of the American Society for Information Science and Technology*, 50(1):1–4, 2013.

[5] A. Crescenzi, D. Kelly, and L. Azzopardi. Time pressure and system delays in information search. In *SIGIR'15*.

[6] S. Fox, K. Karnawat, M. Mydland, S. Dumais, and T. White. Evaluating implicit measures to improve web search. *ACM Trans. Inf. Syst.*, 23(2):147–168, Apr. 2005.

[7] S. Grondin. Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention*, 2010.

[8] F. Macar, V. Pouthas, and W. J. Friedman. *Time, action and cognition: Towards bridging the gap.* Number 66. Springer Science & Business Media, 1992.

[9] N. Mishra, R. W. White, S. Ieong, and E. Horvitz. Time-critical search. In *Proceedings of the 37th International ACM SIGIR'14*, pages 747–756, New York, NY, USA, 2014. ACM.

[10] E. Pöppel. Lost in time: historical foundations and the 3-second-window of temporal integration. *Acta Neurobiol Exp (Wars)*, 64:295–301, 2004.

[11] M. D. Smucker and C. L. Clarke. Time-based calibration of effectiveness measures. In *Proceedings of the 35th International ACM SIGIR'12*, pages 95–104, New York, NY, USA, 2012. ACM.

[12] M. Sucala, B. Scheckner, and D. David. Psychological time: interval length judgments and subjective passage of time judgments. *Current psychology letters. Behaviour, brain & cognition*, 26(2, 2010), 2011.

[13] D. Zakay. *On prospective time estimation, temporal relevance and temporal uncertainty.* Springer, 1992.