QWERTY: The Effects of Typing on Web Search Behavior

Kevin Ong
RMIT University
kevin.ong@rmit.edu.au

Kalervo Järvelin
University of Tampere
kalervo.jarvelin@uta.fi

Mark Sanderson
RMIT University
mark.sanderson@rmit.edu.au

Falk Scholer
RMIT University
falk.scholer@rmit.edu.au

ABSTRACT

Typing is a common form of query input for search engines and other information retrieval systems; we therefore investigate the relationship between typing behavior and search interactions. The search process is interactive and typically requires entering one or more queries, and assessing both summaries from Search Engine Result Pages and the underlying documents, to ultimately satisfy some information need. Under the Search Economic Theory (SET) model of interactive information retrieval, differences in query costs will result in search behavior changes. We investigate how differences in query inputs themselves may relate to Search Economic Theory by conducting a lab-based experiment to observe how text entries influence subsequent search interactions. Our results indicate that for faster typing speeds, more queries are entered in a session, while both query lengths and assessment times are lower.

CCS CONCEPTS

• **Information systems → Users and interactive retrieval**; • **Human-centered computing → Text input**;

KEYWORDS

Search Economic Theory, Text Input, Web Search Behavior

1 INTRODUCTION

Interactive Information Retrieval (IIR) researchers seek to understand the interactions between an information seeker and an information retrieval system [7]. Modeling IIR interactions is a non-trivial process as many factors can contribute to differences in the search behavior. The search process itself is an iterative sequence of actions that often requires multiple queries to acquire the desired amount of relevant information for an underlying information need to be resolved. In IIR, the focus is on the interaction between a human and a search system. The interaction can take place over three modalities, such as search by using/viewing images, speaking [17] or traditionally by typing into a search box. In this paper we focus on typing.

Nowadays, searchers expect rapid responses from search engines [15]. However, the amount of useful information gathered during the search process is also dependent on how well searchers are able to translate their search inputs (such as typing), into useful information output [2]. IIR is a collaborative effort between the searcher and the system to answer an information need. However, research efforts have mainly focused on the changes in human search behavior as a response to system response delays [4, 12], rather than on the effort exerted by both humans and systems as part of the search process. Maxwell and Azzopardi [12] observed that as total time on queries increases, then more time will be spent on assessing the returned Search Engine Result Pages (SERPs). Query cost is inevitably tied to searcher's own input interactions on the search interface itself.

Typing behavior has received renewed research attention recently [6, 10, 11] but not within the IIR community. In this paper, we study the relationship between typing speed and search behavior, and how an individual's search and typing behavior may be used as a factor for explaining IIR behavior.

2 RESEARCH QUESTIONS

The main research question studied in this work is:

"What is the relationship between a user's typing speed and their search behavior?"

The search process arises in order to satisfy some information needs, and SET suggests that when query costs increase, more time will be spent on assessments. We investigate if typing speed can be used as an indicator to understand the differences that might contribute to different query-assessment trade-offs, by considering the following specific research questions.

RQ1, ***What is the relationship between typing speed and the number of queries and assessment time per query***? Azzopardi et al. [2] postulated that user search interactions can be modeled by an economic theory. We seek to understand how the cost of typing can be incorporated into such a model.

RQ2, ***Will a faster searcher type more characters per query***? As the effort to type is lower for faster searchers, we seek to understand if they will choose to issue queries that are longer.

RQ3, ***Do differences in typing speed for different topics reflect a user's topic interest and familiarity***? Past work found that changes in searchers' behavior can be measured from their emotion and typing behavior [5, 10]. Informed by Kelly and Cool [9], we seek to understand how topic interest and familiarity may be associated with differences in the typing behaviors of searchers.

3 BACKGROUND

Past work on SET and typing behavior revealed how different search costs may give rise to different search behaviors. Using the Interaction-Cost hypothesis [1], we propose that query cost based on typing effort can be used as a factor to explain the trade-offs between alternative search strategies in IIR.

SET proposed that differences in query costs will result in changes in search behavior [1]. The search process is a combination of inputs (**Q**, **A**) per search session, where

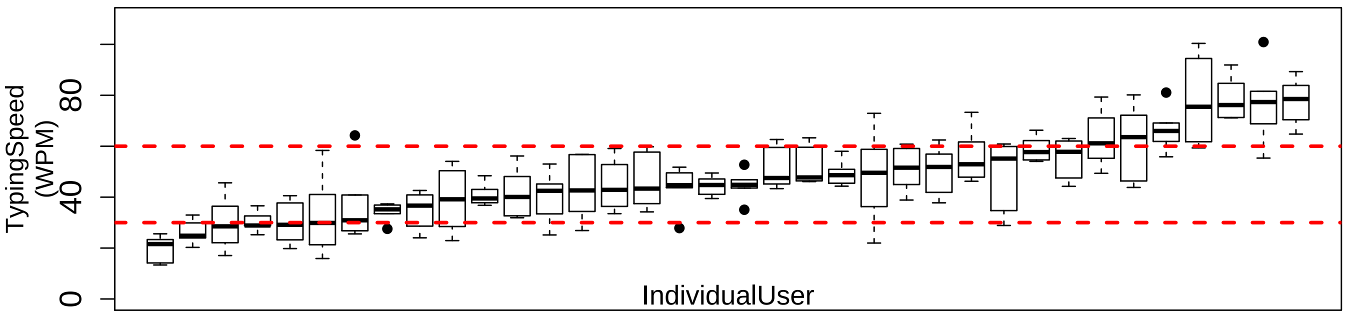

Figure 1: Typing speed of each individual user across their search tasks

- **Q** is the number queries that a user will issue,
- **A** is the number of assessments per query.

A refers to both assessments of SERPs and web documents. The combination of (**Q**, **A**) will be dependent on the relative cost of a query against the cost of assessing a web page. The cost estimations can either be interaction-based [2] or time-based [12]. Under the interaction-based approach [2], the combination of (Q, A) produces a cumulative gain (CG), which can be measured as

$$g(\mathbf{Q}, \mathbf{A}) = \alpha * \mathbf{Q}^{\beta} * \mathbf{A}^{(1-\beta)}$$

where the parameter α represents how well a searcher can convert their interactions into finding relevant documents, and β is the relative cost of **Q** to **A**.

Maxwell and Azzopardi [12] studied the effect of query cost on search behavior by subjecting users to two forms of delays: five seconds SERP response delays, or five seconds document download delays. They found that when total querying time increased, document assessments increased. Similarly, when both SERP response and document download delays were introduced, users spent more time examining documents.

Typing Salthouse [14] studied the effects of age and skills on typing behavior, and found that experience but not age influenced typing speed. A skilled typist is formally trained on typing on keyboards. In subsequent work, John [8] suggested a baseline of 30–60 Words Per Minute (WPM) as an average typing speed on a standard keyboard. There has been renewed interest in typing [6, 10, 11]. Lim et al. [10] studied the effects of emotion on typing behavior and found that highly stressed users spent longer on tasks and clicked more rapidly but typed more slowly and were more error prone. Feit et al. [6] studied 30 participants aged 20 – 55 and found that non-skilled typists can attain the same level of typing speed and accuracy compared to skilled typists (58 vs 59 WPM). Logan et al. [11] found that skilled typists typed at 45 WPM when composing text. In general, previous research [6, 11] found non-skilled typists typed as well as skilled typists, unlike earlier works [8, 14]. This suggests that the growing computer literacy over the decades might have narrowed the gaps between these two groups of typists.

4 METHOD

To investigate the relationship between typing and search behavior, a user study was carried out.

Searchers: 36 searchers (18 males and 18 females) aged from 18–47 with a mean age of 26 were recruited via opportunistic sampling at a local campus library. While 13 participants reported that English was not their first language, all participants were confident in reading and typing in English. A 15" Macbook laptop was provided for the experiment, and participants were asked if they were familiar with the type of keyboard; half reported that they were unfamiliar with the keyboard provided. The participants were told they had 45 minutes in total to complete six topics, and were compensated at the end of the experiment with a $20 voucher for participation.

Data: Data were collected as part of a larger experiment where individual search interactions were collected [13]. Participants were given two attempts at an initial typing test. The test allowed the participants to familiarize themselves with the keyboard. They were then asked to conduct searches for six provided informational search tasks. Unknown to the participants, the initial SERPs were controlled by varying quality of search results on the first page [13].[1] Search results for any query reformulations per task were obtained from a live commercial search engine. Participants were asked to carry out the tasks as they would on a normal web search engine. Task order was rotated and counterbalanced with different types of initial starting conditions. They were able to enter as many queries as they liked, and were asked to find as many relevant search results as they needed to satisfy the search task. Participants were free to distribute their time as they saw fit and were not pressed to complete each task within a set time frame. Total time remaining was provided to the participants if requested, but was otherwise not indicated.

Participants could mark a search result as relevant within the SERP itself by clicking a checkbox, without having to open the underlying document (which they could do, should they choose to). Participants were also given both a pre-task questionnaire on topic interest and familiarity, and post-task questionnaire on topic difficulty for each task.

Logging: To capture typing behavior, a search engine interface was created to log the typing speed of each query using JavaScript. The following interactions were logged:

- The number of queries issued.

[1]The initial starting conditions did not influence the final number of documents saved ($F = 0.2$, $p = 0.95$).

- The average task typing speed, measured in WPM. This is the number of keystrokes divided by five, between the first and last key presses within the time taken, as described by John [8].[2] It is inclusive of all non-character keystrokes, such as backspaces and other non-characters.
- The number of key press errors: backspaces and other non-characters.
- The number of documents clicked and marked by the participants as relevant.
- The amount of time spent typing and assessing SERPs and documents.

Task-searcher pairs (i.e. search interactions performed by each searcher on each task) are classified into three groups based on their average on-task WPM. The following thresholds were proposed by John [8]: fast (>60 WPM), average (30–60 WPM) and slow (<30 WPM).

5 RESULTS

The data from the user study was analyzed based on individual task-searcher combinations (i.e. search interactions performed by each searcher on each task) because we conjectured that typing speed might also vary with both topic familiarity and interest [9, 16]. There were 213 task responses in total. Three tasks from two searchers were removed due to non-logging issues.

Figure 1 shows the average task typing speed on each topic across all six topics, for all 36 participants . Two horizontal, red dashed lines mark the boundary at 30 and 60 WPM. Some typists had consistent typing speeds (shorter boxes) while others varied substantially (taller boxes). This figure shows that typing speed varies significantly across individuals ($F = 8.1, p < .0001$) and that an individual's typing speed is dependent on the topic that the task-searcher is working on. The mean speed across all task-searchers was 48 WPM with a SD of 17.5.

Time-based and interaction-based measurements are recorded in Table 1. Chi-Square (X^2) or ANOVA (F-score) values are recorded for each relevant variable. A follow-up Tukey's honest significance difference (TukeyHSD) test is conducted at $p < 0.05$ for values analyzed with ANOVA. Values that are significantly different from two other categories using TukeyHSD pairwise tests are bold.

Across three speed categories, there were 38 slow, 128 average, and 47 fast task-searchers ($X^2 = 69.2, p < .001$). For those who were familiar with the laptop keyboard, there were 14 slow, 63 average, and 30 fast task-searchers ($X^2 = 35.1, p < .001$). For task-searchers who were not, there were 24 slow, 65 average, and 17 fast task-searchers ($X^2 = 38.1, p < .001$). Regardless of familiarity with the keyboard, more than 60% were average task-searchers. The differences in rates of yes/no familiarity responses within each speed group were not significant (slow: $X^2 = 2.63, p = 0.1$; average: $X^2 = 0.03, p = 0.9$; fast: $X^2 = 3.60, p = 0.06$).

For time-based response measures, fast task-searchers spent significantly less time issuing queries ($F = 77.1, p < .001$) and assessing documents and SERPs ($F = 4.9, p < .001$) than the other types of task-searchers. On the interaction-based response measures, the slow, average and fast task-searchers typed 24.3, 45.9 and 72.8 WPM respectively ($F = 370.0, p < .001$). Fast task-searchers

[2]The division is for the average number of characters in an English word.

Table 1: Mean or median ($\widetilde{x}$) (and standard deviations, SD) of searchers' interactions per task. Significant differences (ANOVA) denoted by: *p < 0.05, **p < 0.01, ***p < 0.001. Values for pairwise comparisons (TukeyHSD; p < .05) are bold.

	Categories of task-searchers			X^2
	Slow (<30 WPM)	**Average (30-60 WPM)**	**Fast (>60 WPM)**	
Number of task-searchers	38	128	47	69.2***
Familiarity with keyboard provided	Yes: 14 No: 24	Yes: 63 No: 65	Yes: 30 No: 17	35.1*** 38.1***
Time-based (per query)				**F**
Typing Time / Query (in sec)	**22.2 (10.6)**	**10.9 (5.7)**	**5.2 (2.1)**	77.1***
Assessment Time / Query (in sec)	79.4 (40.3)	73.2 (44.3)	**54.3 (28.1)**	4.9***
Total Time / Query (in sec)	101.6 41.9	84.1 44.9	**59.6 28.3**	11.4***
Interaction-based (per query)				**F**
On-Topic Typing / Query (WPM)	**24.3 (4.2)**	**45.9 (8.2)**	**72.8 (10.7)**	370.0***
Typed Characters / Query	42.5 (20.1)	37.3 (16.5)	**29.2 (9.9)**	7.7***
Error Characters / Query	11.0 (13.2)	7.2 (7.4)	**3.6 (4.6)**	12.0***
Output Characters / Query	31.6 (10.7)	30.2 (12.9)	25.7 (7.9)	3.3*
Documents Marked / Query	3.4 (2.2)	3.9 (2.9)	4.3 (3.3)	0.9
Documents Clicked / Query	1.1 (1.3)	1.0 (1.2)	0.6 (1.0)	2.9
Total Interaction counts				**F**
Number of Queries	**1.7 (0.9)**	**2.4 (1.4)**	**3.3 (2.1)**	13.1***
Documents Marked (in total)	5.4 (3.8)	8.6 (8.0)	**12.5 (10.1)**	8.5***
Documents Clicked (in total)	1.6 (1.8)	2.4 (2.9)	1.7 (2.8)	1.7
Self report				**F**
Age	26.9 (5.6)	26.5 (7.8)	**23.1 (4.3)**	4.9**
Topic Interest	$\widetilde{x} = 7$ (2.4)	$\widetilde{x} = 6$ (2.7)	$\widetilde{x} = 6$ (2.2)	0.5
Topic Familiarity	$\widetilde{x} = 3$ (2.4)	$\widetilde{x} = 4$ (2.4)	$\widetilde{x} = 3$ (2.2)	0.07
Topic Difficulty	$\widetilde{x} = 4$ (2.3)	$\widetilde{x} = 4$ (2.4)	$\widetilde{x} = 4$ (2.3)	2.4

also typed significantly shorter queries ($F = 7.7, p < .001$) and were also less error-prone ($F = 12.0, p < .001$) than both average and slow task-searchers. Fast task-searchers entered significantly greater number of queries ($F = 11.1, p < .001$), while the average numbers of documents marked did not vary significantly between groups ($F = 0.9, p = 0.435$). However, fast task-searchers marked more documents than both average and slow task-searchers in total ($F = 8.5, p < .001$). Regarding demographic data, fast task-searchers are significantly younger at 23.1 years on average, compared to 26.5 and 26.9 for average and slow task-searchers ($F = 4.9, p < .01$).

Topic interest and topic familiarity were established through pre-task surveys, with participants asked to rate each factor on a scale of 1 – 10. The median score ($\widetilde{x}$) is reported. Categorizing task-searchers based on their typing speed, there were no significant differences between different groups for topic interest ($F = 0.5, p = 0.6$), topic familiarity ($F = 0.1, p = 0.93$) or topic difficulty ($F = 2.4, p = 0.10$).

Limitations. In this study, we made an assumption that all documents saved were relevant - similar to the Click-through Hypothesis and did not record additional details about the documents saved. Auto-corrections and query suggestions were also excluded in the experimental design.

6 DISCUSSION AND CONCLUSIONS

Overall, our study participants demonstrated a range of typing speeds. On average, their speeds were comparable to that observed by Logan et al. [11] (48 vs 45 WPM). Our experimental data also showed that age may have contributed to differences in typing speed, which is a factor to consider for future experiments.

Framed within SET, we posit that higher typing cost is associated with lower typing speed and higher typing time. Regarding **RQ1**, as typing speed increased, more queries were issued and less time was spent per query, while slow task-searchers spent more time on both typing and assessments. Typing cost can be considered as one of the factors of total query cost, and is shown to support the cost-interaction hypothesis of SET. However, when typing cost increased, the total number of queries decreased, while the number of assessments per query did not change significantly. Allowing task-searchers to mark documents as relevant, without requiring document click-throughs, may be the cause of this artifact. On the other hand, for the time-based measures, fast task-searchers spent significantly less time on both queries and assessments.

For **RQ2**, fast task-searchers, as opposed to the average and slow ones, typed shorter queries, and entered a higher number of queries. We expected that as typing cost is reduced, then more characters would have been typed as a result; however, the results indicated the opposite. We conjecture that the shorter query lengths were balanced by the increased number of queries. In future work, we plan to study how different task-searcher types adapt their search strategies to changes in search system and document download delays.

For **RQ3**, while a relationship between topic familiarity, interest and typing speed was expected, we were not able to conclude that there is a difference in our study, based on typing speed. Given our sample size, there was not enough data for typing speed to reveal the differences in topic familiarity and interest.

We conjecture that the fast task-searchers were able to mark more documents as relevant because of the higher number of query reformulations they performed. This is because the average number of documents marked did not differ significantly across different task-searcher types. In terms of propensity of typing errors, we observed that fast task-searchers made fewer typing errors than average or slow task-searchers (12% vs 19% and 26% error rates, respectively). For future work, we plan to investigate what contributed to differences in error rate, for example by studying the other aspects of typing, such as consistency and rhythm. We also plan to investigate how typing speed can be used as a signal to support different types of task-searchers, such as showing different forms of snippets for different task-searchers [3]. These findings support observations by Azzopardi et al. [2] and Maxwell and Azzopardi [12] that revision to SET is needed.

In conclusion, we show the relationship between task typing speed and search behavior. Fast task-searchers were younger. They also entered more queries with shorter lengths and marked more documents in total while slow and average task-searchers were more error-prone and spent more time per query. We did not observe a relationship between topic interest and familiarity with task typing speed, given our sample size.

7 ACKNOWLEDGEMENTS

This project is funded by ARC Discovery Grant, ref: DP140102655 and an APA scholarship. Travel funding is also provided by ACM SIGIR for the lead author to attend the conference.

REFERENCES

[1] Leif Azzopardi. 2011. The Economics in Interactive Information Retrieval. In *Proceedings of SIGIR*. 15–24.
[2] Leif Azzopardi, Diane Kelly, and Kathy Brennan. 2013. How Query Cost Affects Search Behavior. In *Proceedings of SIGIR*. 23–32.
[3] Ruey-Cheng Chen, J. Shane Culpepper, Tadele T. Damessie, Timothy Jones, Ahmed Mourad, Kevin Ong, Falk Scholer, and Evi Yulianti. 2015. RMIT at the TREC 2015 LiveQA Track. In *Proceedings of TREC*.
[4] Anita Crescenzi, Diane Kelly, and Leif Azzopardi. 2015. Time Pressure and System Delays in Information Search. In *Proceedings of CHIIR*. 767–770.
[5] Ashlee Edwards and Diane Kelly. 2017. Engaged or Frustrated?: Disambiguating Emotional State in Search. In *Proceedings of SIGIR*. 125–134.
[6] Anna Maria Feit, Daryl Weir, and Antti Oulasvirta. 2016. How We Type: Movement Strategies and Performance in Everyday Typing. In *Proceedings of CHI*. 4262–4273.
[7] Peter Ingwersen and Kalervo Järvelin. 2005. Information Retrieval in Context: IRiX. In *ACM SIGIR Forum*, Vol. 39. 31–39.
[8] Bonnie E. John. 1996. TYPIST: A Theory of Performance in Skilled Typing. *Human–Computer Interaction* 11, 4 (1996), 321–355.
[9] Diane Kelly and Colleen Cool. 2002. The Effects of Topic Familiarity on Information Search Behavior. In *Proceedings of JCDL*. 74–75.
[10] Yee Mei Lim, Aladdin Ayesh, and Martin Stacey. 2015. The Effects of Typing Demand on Emotional Stress, Mouse and Keystroke Behaviours. In *Intelligent Systems in Science and Information 2014*. 209–225.
[11] Gordon D. Logan, Jana E. Ulrich, and Dakota R. B. Lindsey. 2016. Different (Key)Strokes for Different Folks: How Standard and Nonstandard Typists Balance Fitts' Law and Hick's Law. *Journal of Experimental Psychology: Human Perception and Performance* 42, 12 (2016), 2084–2102.
[12] David Maxwell and Leif Azzopardi. 2014. Stuck in Traffic: How Temporal Delays Affect Search Behaviour. In *Proceedings of IIiX*. 155–164.
[13] Kevin Ong, Kalervo Järvelin, Mark Sanderson, and Falk Scholer. 2017. Using Information Scent to Understand Mobile and Desktop Web Search Behavior. In *Proceedings of SIGIR*. 295–304.
[14] Timothy Salthouse. 1984. Effects of Age and Skill in Typing. *Journal of Experimental Psychology: General* 113 (1984), 345–371.
[15] Eric Schurman and Jake Brutlag. 2009. Performance Related Changes and Their User Impacts. In *Velocity: Web Performance and Operations Conference*.
[16] Ben Shneiderman. 1984. Response Time and Display Rate in Human Performance with Computers. *Comput. Surveys* 16, 3 (1984), 265–285.
[17] Johanne R. Trippas, Damiano Spina, Mark Sanderson, and Lawrence Cavedon. 2015. Towards Understanding the Impact of Length in Web Search Result Summaries over a Speech-Only Communication Channel. In *Proceedings of SIGIR*. 991–994.