Gaze Behavior in task- and non-task-related Web Interaction: Implications for Usability and UX

Fabian Engl[1], Erik Buchmann[2], Michael Burmester[3], and Jürgen Mottok[1]

[1] Software Engineering Laboratory for Safe and Secure Systems (LaS³), OTH Regensburg, Regensburg, Germany
{fabian.engl,juergen.mottok}@oth-regensburg.com

[2] Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI) Dresden/Leipzig, Leipzig University, Leipzig, Germany
erik.buchmann@uni-leipzig.de

[3] User Experience Research Lab (UXL), Stuttgart Media University (HdM), Stuttgart, Germany burmester@hdm-stuttgart.de

Abstract. Several studies show that eye movements differ between task-related search and non-task-related free web browsing. Yet, experiments typically rely on simplified single-page websites or screenshots. This paper extends that line of research by analyzing gaze recordings from 121 participants interacting with multi-page websites. We investigated whether core eye-tracking metrics systematically differ between both user states in complex real-world web interactions. The analysis revealed statistically significant effects for fixation duration, fixation distribution (x- and y-coordinates), and saccade duration, with medium to large Cohen's d effect sizes. Consequently, our findings highlight that task- and non-task-related user states persist in real-world web contexts and represent distinct interaction modes characterized by systematically different gaze patterns. This distinguishability has important implications for the design of usability- and UX-oriented eye-tracking studies, as both product qualities are defined by task-relatedness. Our results suggest that certain eye movements may be more usability- or UX-related depending on the user's interaction state.

Keywords: Eye-Tracking · HCI · Usability · UX

1 Introduction

Eye-tracking is widely used in behavioral and quantitative human-computer interaction (HCI) research, for example to evaluate how users perceive the usability and user experience (UX) of websites [7–9]. Conceptionally, **usability** eye-tracking studies assess **task-related** interaction aspects such as interface inconsistencies and errors [18], whereas **UX** studies are more concerned with **non-task-related** product qualities including design and aesthetics [6]. Such task- and state-dependent distinctions are also reflected in industry standards such as the ISO 9241-110 for *Ergonomics of human-system interaction* [3] or in theoretical HCI frameworks like the *pragmatic-hedonic* model [12] and the *components of user experience (CUE)* model [20].

This task- versus non-task-related differentiation is not addressed thoroughly in eye-tracking-based HCI research, particularly within realistic web contexts. Some eye-tracking studies have shown that participants exhibit different gaze patterns in both interaction states, however, most of these studies rely on simplified stimuli such as single-page websites or images [1, 5, 14, 21]. This leaves questions about whether identified gaze patterns generalize to natural web interactions. Two critical gaps remain for usability and UX research: Most notably, whether these state-dependent gaze patterns persist in complex multi-page web interactions remains an open question, as does how rapidly eye movement metrics shift following task completion.

To address this research gap, we build upon an existing eye-tracking dataset [4] — with permission of the authors — comprising recordings from 121 participants browsing six multi-page real-world websites. The study followed a two-stage design, where participants first performed assigned tasks and then continued browsing freely on the same website. Using this split, we conduct a quantitative comparison of task-related search versus non-task-related free browsing, alongside a qualitative temporal analysis of gaze behavior during transitions between the two states. Specifically, the quantitative analysis focuses on global differences in fixation- and saccade-based metrics across the two user states, while the qualitative analysis examines local temporal dynamics of gaze behavior in the transition period surrounding task completion. If consistent patterns are found here across websites, they may indicate a correspondence between theoretical distinctions in HCI and observable gaze behavior. Consistently different gaze characteristics across task- and non-task-related interactions would, drawing on the *pragmatic-hedonic* and *CUE* model [12, 20], open the possibility of differentiating between usability- and UX-related eye movement patterns.

The remainder of this paper is structured as follows: Section 2 reviews related work and introduces the research questions and hypotheses. Sections 3 and 4 describes the dataset and evaluation procedure, followed by results in Section 5. Finally, Sections 6, 7, and 8 discuss findings, implications, and conclusions.

2 Related Work

There are multiple studies which compare task- and non-task-related gaze patterns during web interaction, however, most originate outside of usability or UX research. Research objectives include distinguishing focal from global viewing [14], predicting different task types [2], or identifying common scanpaths [5]. Although most studies do not explicitly focus on usability or UX, they investigate a necessary empirical precondition for the current analysis: that eye-movements systematically differ between *task-* and *non-task-related* interaction.

Several studies [14–16] showed participants single-page websites, where they either had to search for images or browse freely. Results showed that during *task-related* interaction, users performed more fixations and saccades and made larger saccades compared to *non-task-related* browsing [14]. Additionally, very long fixations were observed immediately *pre-task-completion* [16].

Another study focused on scanpaths, represented through predefined areas of interest (AOIs) [5]. Here, *task-related* scanpaths overlapped by about 83%. In contrast, free browsing overlapped by only 53%, showing more variability between participants in *non-task-related* interaction [5]. AOIs, however, are stimulus-specific and don't generalize well, as websites differ in size, structure, and complexity. Thus, it is hard generalize these findings.

Other studies take a more general approach and analyze how user shift attention across between tasks or across website regions [1, 21]. Instead of simple fixations- or saccades-based metrics, they used saliency maps [1] or grids [21] to track fixation shifts. Their results revealed that fixations were concentrated in the top-left-to-center screen region during *non-task-related* interaction and shifted more toward the left during *task-related* search [21]. Yet, these shifts were not consistent across all tasks [1].

Some publications also examined task- versus non-task-related user states in virtual reality [11, 13]. They also found *task-related* shifts in fixation distributions, with increased average y-coordinates [11]. In fact, similar spatial shifts are found across multiple in eye-tracking studies investigating these interaction modes [1, 21, 22]. In virtual reality, saccade amplitude and duration also varied between tasks but not consistently across all conditions [13].

Overall, multiple eye-tracking studies identify general patterns between task- and non-task-related behavior. These include shifts in fixation distributions [1, 13, 22], varying fixation durations [14], and changes in saccade characteristics like duration and amplitude [13, 14]. Nonetheless, existing studies exhibit two main limitations that restrict their transferability to real-world usability and UX contexts: First, many rely on static images or simplified single-page stimuli [1,5,14,21]. These simply do not capture the complexity of real-world website browsing, where users navigate across multiple pages driven by curiosity or perceived relevance. In fact, such stimulus simplifications have been criticized in eye-tracking-based web research, as they may not capture naturalistic browsing behavior [17]. Second, existing research typically builds upon a between-subject design, where participants either complete tasks **or** browse freely, but never transition between states within one website. This dichotomous design might overlook temporary changes in eye movements during these state transitions.

2.1 Research Questions and Hypotheses

Building on the identified shortcomings of existing literature, this section introduces the main research questions (RQ) and hypotheses (H) of the present work. Two RQs examine how task- and non-task-related eye movements differs in real-world web contexts: RQ1 first asks whether fixation- and saccade-based differences from prior eye-tracking studies replicate in a more complex, multi-page setting and across different websites. Following this, RQ2 focuses on the transition from task-related search to non-task-related browsing, examining gaze behavior immediately pre- and post-task-completion.

RQ1: To what extent do fixation distributions, fixation duration (FD), saccade duration (SD), and saccade amplitude (SA) differ in task-related versus non-task-related web interactions, and are these differences consistent across websites?

RQ2: How do gaze patterns, specifically fixation- and saccade-based metrics, change during the transition from pre- to post-task-completion?

H1: Fixation locations (x- and y-coordinates) shift, and FD, SD, and SA are higher during non-task-related free browsing compared to task-related searching, consistently across websites.

H2: Eye movements change immediately post-task-completion, resulting in a change in fixations locations (x- and y-coordinates) as well as duration-based metrics including FD and SD.

To address these RQs, we analyze the dataset introduced in [4], comprising gaze recordings from six real-world websites. The dataset is particularly well-suited to the present objectives, as it captures both user states — task-related and non-task-related browsing — including concrete timestamps of when users completed task. The following section briefly introduces the dataset and outlines the evaluation procedure of the present work.

3 Dataset

The dataset we build upon in this study [4] contains eye-tracking data from 121 participants interacting with six stimuli (real-world websites) in a within-subject design. Three stimuli were from German drinking water providers, three from the German Alpine Association (DAV). The study involved participants completing information-seeking tasks (e.g., finding dates or service details) on each website, with a randomized presentation order to reduce carryover effects. Participants initially had 30 seconds to familiarize themselves with each website followed by three minutes to complete the given task. Task completion was signaled by two consecutive right mouse clicks, after which participants were instructed to browse freely for the remaining time. This input ensured a clear separation between task- and non-task-related eye movements within each interaction.

The sample included 70 male and 51 female participants ($n = 121$), with 43 wearing glasses or contact lenses. The average age was 29.15 years ($SD = \pm 11,4$). Roughly half of the participants were students, while the rest were mainly junior and senior employees.

Eye movements were recorded at 250 Hz using Tobii Pro Fusion eye-trackers, which were attached to a 21-inch Full HD monitor. Participants had to meet a 0.75° calibration and validation threshold, recordings failing to do so were excluded. Recordings were excluded at website-level if the recording section exceeded 15% data loss. 54 website interactions were excluded, resulting in a total of 672 interactions (121 participants $\times$ 6 websites - 54 exclusions).

A comprehensive description of the study design, participant demographics, and technical setup can be found in [4].

4 Evaluation Procedure

To investigate whether fundamental differences in eye movements exist between both user states, our analysis focuses on the following eye-tracking metrics: fixation positions (x- and y-coordinates), FD, SD, and SA. Prior to statistical testing, we split the dataset both by website and interaction phase (pre- vs. post-task-completion), resulting in two user-state-specific subsets per website–metric combination and 30 subset pairs in total (6 websites $\times$ 5 metrics).

For each of these subsets we assessed normality using a Shapiro–Wilk test [19], but only a small proportion of subset pairs met the assumption ($p > 0.05$). Due to these violations, a non-parametric paired permutation test was used, which approximates the null distribution by repeatedly permuting group labels and recalculating the test statistic [10]. The p-value was computed as the proportion of equally or more extreme outcomes across 100,000 randomized permutations, with Cohen's d calculated to quantify effect strength. All analyses were conducted in Python, with the code and results linked in the acknowledgments under *Open Science and Reproducibility*.

5 Results

To address RQ1, we partitioned the eye-tracking dataset by website and user state (pre- versus post-task-completion). Task completion timestamps were derived from user-provided inputs, which we matched to the eye-tracker timestamps. For 12 of the 672 recorded web interactions (1.8%), user inputs were ambiguous or otherwise unreliable. Consequently, we excluded the affected web interactions from all subsequent analyses to avoid misclassification errors. For all remaining website interactions, we labeled pre-task-completion eye movements as *task-related* and post-task-completion eye movements as *non-task-related*. We further excluded all participants who failed to complete the task within the three-minute time limit, as the paired permutation test required both task- and non-task-related gaze data. The number of users who successfully completed the task and their corresponding *task-* and *non-task-related* eye movements are outlined in Table 1. Finally, we averaged fixation- and saccade-based metrics per participant to represent overall trends. Since stimulus order was randomized across participants, order effects can be ruled out. The numbering in the diagrams therefore bears no relation to presentation order.

Table 1. Table showing the number of users per website (W) retained for analysis and the number of task- and non-task-related eye movements (fixations and saccades).

	W1	**W2**	**W3**	**W4**	**W5**	**W6**
Included Web Interactions	64	109	50	107	86	106
Task-related Eye Movements	41.408	16.886	34.807	44.203	36.976	38.594
Non-task-related Eye Movements	90.841	111.387	108.066	85.420	102.576	92.218

Results for RQ1 are presented in box plots, which allow for a direct visual comparison of the distribution of eye movement metrics across task- and non-task-related interactions. Outliers exceeding ± 2.5 SD are represented as gray dots. To prevent y-axis distortion, we removed extreme outliers from individual plots and indicated this with red squares at the diagram borders.

We also checked for participants with extreme eye movement frequencies using a Z-test at the stimulus level. Except for two participants (P045 and P074), deviations beyond $|Z| \geq 2$ were scattered across participants and stimuli, with no indication of systematic outliers. We further performed a sensitivity analysis to examine how the two participants with atypical eye movement counts influenced the results, but excluding them had no effect on statistical significance. Therefore, both were† retained for the evaluation. The statistical results, including p-values and Cohen's d effect sizes, are provided in Table 2.

Table 2. Table comparing mean eye-tracking metrics between task-related search and non-task-related free browsing, alongside permutation test results including observed differences, p-values, and Cohen's d effect sizes.

Eye-Tracking Metric	**Website**	**During Task (task-related)**	**Free Browsing (non-task-related)**	**Observed Permutation Difference**	***p*-value**	**Cohen's *d***
Mean Fixation Duration (FD)	W1	248.9 ms	267.3 ms	+18.4 ms	**<0.001**	0.529
	W2	279.6 ms	265.4 ms	-15.2 ms	**<0.001**	-0.337
	W3	248.4 ms	264.4 ms	+16.0 ms	**0.003**	0.441
	W4	248.3 ms	264.1 ms	+15.9 ms	**<0.001**	0.495
	W5	242.2 ms	260.2 ms	+18.0 ms	**<0.001**	0.414
	W6	273.0 ms	254.7 ms	-18.4 ms	**<0.001**	-0.458
Mean x-coordinate	W1	783.9 px	789.3 px	+5 px	0.392	0.108
	W2	1024.0 px	928.3 px	-96 px	**<0.001**	-1.162
	W3	851.3 px	886.7 px	+36 px	**0.027**	0.324
	W4	834.3 px	886.3 px	+52 px	**<0.001**	0.838
	W5	876.4 px	884.8 px	+9 px	0.449	0.082
	W6	829.3 px	912.3 px	+83 px	**<0.001**	0.757
Mean y-coordinate	W1	591.2 px	531.1 px	-60.2 px	**<0.001**	-0.820
	W2	464.3 px	506.5 px	+42.1 px	**<0.001**	0.419
	W3	538.8 px	464.6 px	-74.3 px	**<0.001**	-1.163
	W4	652.6 px	562.4 px	-90.2 px	**<0.001**	-1.178
	W5	531.1 px	518.8 px	-12.3 px	0.313	-0.109
	W6	500.9 px	497.1 px	-3.9 px	0.602	-0.052
Mean Saccade Duration (SD)	W1	37.4 ms	42.4 ms	+4.9 ms	**<0.001**	0.824
	W2	36.9 ms	44.4 ms	+7.5 ms	**<0.001**	0.656
	W3	38.7 ms	42.7 ms	+4.0 ms	**<0.001**	0.516
	W4	37.2 ms	43.3 ms	+6.0 ms	**<0.001**	0.776
	W5	39.8 ms	42.3 ms	+2.6 ms	**0.027**	0.228
	W6	40.3 ms	41.9 ms	+1.6 ms	0.273	0.114
Mean Saccade Amplitude (SA)	W1	137.3 px	151.2 px	+13.9 px	**<0.001**	0.489
	W2	157.9 px	175.9 px	+18.0 px	**<0.001**	0.439
	W3	181.3 px	186.0 px	+4.7 px	0.576	0.090
	W4	171.6 px	174.3 px	+2.7 px	0.328	0.095
	W5	208.5 px	204.1 px	-4.5 px	0.371	-0.097
	W6	182.1 px	177.6 px	-4.5 px	0.465	-0.076

Starting with fixation-based metrics, Figure 1 and Table 2 show statistical significant changes in task- versus non-task-related gaze behavior in terms of mean FD. The paired permutation test revealed statistically significant differences across all six websites ($p < 0.001$ to $p = 0.003$). However, the direction of these effects varied. On W2 and W6, mean FD decreased by roughly 15 and 18 milliseconds (ms) during free browsing compared to task-related search, whereas on the remaining websites FD increased by a similar magnitude, approximately 16 to 18 ms. The Cohen's d analysis shows moderate effects ranging from $d = \pm 0.337$ to $d = \pm 0.529$ for all websites.

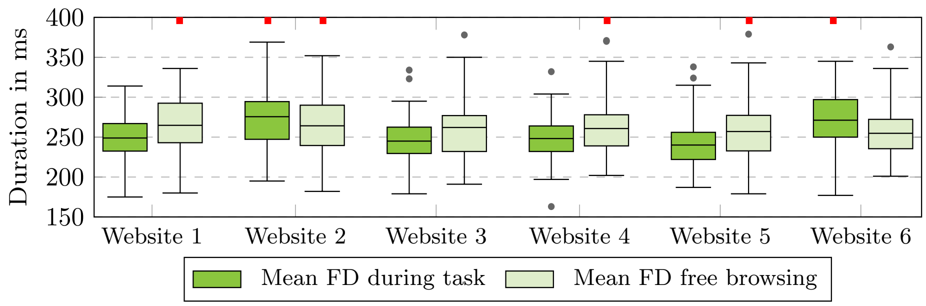

Fig. 1. Boxplot illustrating differences in participants' mean fixation duration (FD) between task-related search and non-task-related free browsing.

Stronger effects were observed for fixation distribution. As shown in Figure 2, four websites (W2, W3, W4, and W6) exhibited statistically significant horizontal shifts in mean fixation coordinates ($p < 0.001$ to $p = 0.027$). W1 through W4 additionally showed statistically significant vertical shifts ($p < 0.001$). On average, mean x-coordinates increased between 36 and 83 pixels (px) and mean y-coordinates decreased between 60 and 90 px during non-task-related free browsing (see Table 2). W2 is the only exception to this pattern. Despite some non-significant results, effect sizes ranging from $d = \pm 0.324$ to $d = \pm 1.178$ indicate more pronounced changes in terms of fixation location compared to FD.

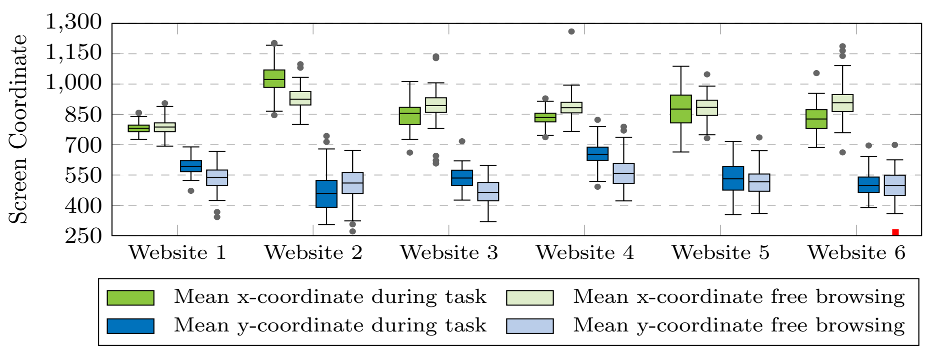

Fig. 2. Boxplot illustrating differences in participants' horizontal and vertical fixation distribution (x- and y-coordinates) between task-related search and non-task-related free browsing.

Turning to saccade-related metrics, Figure 3 shows statistically significant changes for SD during non-task-related free browsing across W1 to W5 ($p <$ 0.001 to $p = 0.027$), with small-to-large effect sizes ranging from $d = 0.228$ to $d = 0.824$. Across these five websites, participants exhibited on average 4 to 7.5 ms longer saccades during non-task-related free browsing compared to task-related searching. W6 shows a similar tendency, however, the mean SD increase of only 1.6 ms did not reach statistical significance.

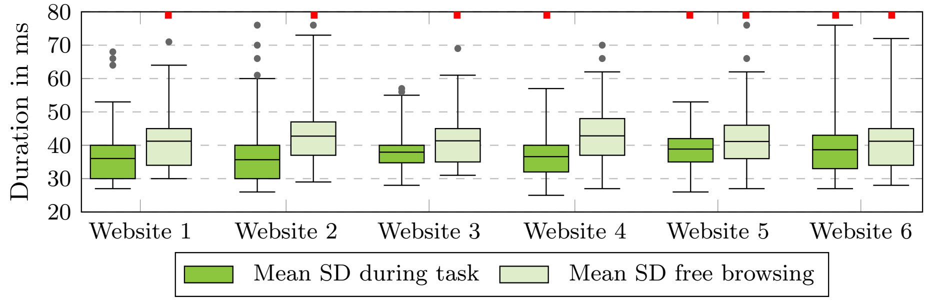

Fig. 3. Boxplot illustrating differences in participants' mean saccade duration (SD) between task-related search and non-task-related free browsing

Despite a consistent increase in mean SD, Figure 4 shows that there is only limited variation in terms of SA between task-related search and non-task-related free browsing. Significant increases of 14 and 18 px were observed only for W1 and W2 ($p < 0.001$). Both websites show medium effect sizes $d = 0.489$ and $d = 0.439$, respectively, highlighting a similar level of change between both user interaction states. Overall, SA differences between task- and non-task-related interactions are thus only partially consistent across websites.

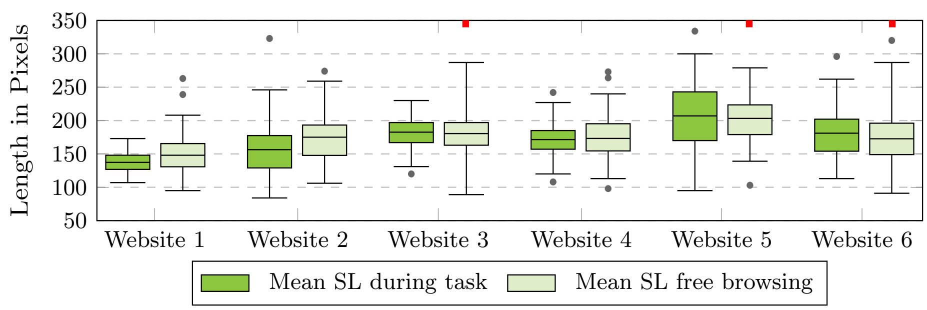

Fig. 4. Boxplot illustrating differences in participants' mean saccade amplitude (SA) between task-related search and non-task-related free browsing

Overall, both fixation- and saccade-based metrics often showed statistically significant differences between task-related and non-task-related gaze behavior. For instance, fixation distributions shifted upwards and towards the left side of the screen during free browsing, which aligns with existing research [21]. SD increased significantly for five out of six websites, with the remaining site exhibiting a similar change. Contrary to existing findings [14], SA did not increase alongside SD. Across metrics, differences between task- and non-task-related web interaction thus vary considerably, showing that findings from existing literature only partially generalize to more complex real-world websites.

Despite frequent significant differences at the individual website-level, these patterns were rarely consistent across all six websites, with significant results typically limited to a subset of stimuli, as can be seen in Table 3. As our analysis revealed no uniform changes in eye-tracking metrics across all websites, **H1 is rejected**. Nevertheless, significant differences at the individual website-level recurrently indicate that task- and non-task-related interactions elicit distinct gaze patterns, albeit not universally. Notably, where significant differences were observed, Cohen's d effect sizes were comparable in magnitude across the affected websites, suggesting that these effects are similarly strong when they occur.

Table 3. Table showing how mean saccade- and fixation-based eye-tracking metrics changed during non-task-related free browsing compared to prior task-related searching (↑ increase, ↓ decrease, → no statistical significant change).

Eye-Tracking Metric	W1	W2	W3	W4	W5	W6
Fixation Duration (FD)	↑	↓	↑	↑	↑	↓
Fixation x-coordinate	→	↓	↑	↑	→	↑
Fixation y-coordinate	↓	↑	↓	↓	→	→
Saccade Duration (SD)	↑	↑	↑	↑	↑	→
Saccade Amplitude (SA)	↑	↑	→	→	→	→

For **RQ2**, we shifted the analysis to a narrower time window around task completion. Each eye movement was assigned a time offset relative to the task-completion event ($t = 0$). Only eye movements within a 10-second window pre- ($t = -10$) and post-task-completion ($t = 10$) were considered. For RQ2, the analysis focuses only on fixation distribution (x- and y-coordinates), FD, and SD as they showed the most consistent statistically significant effects across websites in RQ1. We excluded SA as the prior analysis found significant results for only two of the six websites.

For SD — where we found consistent differences in eye movements between pre- and post-task-completion across the majority of websites — Figure 5 shows that this increase occurs immediately post-task-completion. Roughly two seconds after completing the task ($t = 2$), SD increases from around 37.5 to 44.5 ms on average. This trend is universal across all six websites.

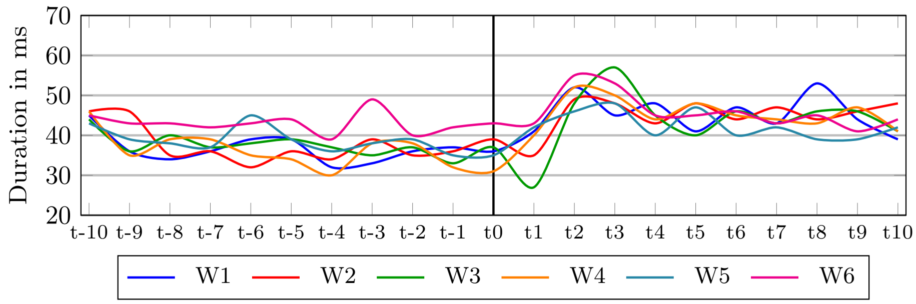

Fig. 5. Line chart illustrating temporal changes in mean saccade duration (SD) across all websites (W), averaged across participants, within a 20-second window spanning 10 seconds pre- and post-task completion (t=0).

Turning to fixation-based metrics, Figure 6 shows a sharp increase in FD immediately before completing the task ($t = -1$), followed by a return to baseline levels about one to two seconds later. While this pattern is consistent across all websites, it is less pronounced for W1. Additionally, W2 and W6 showed a slight decrease in FD post-task-completion. The remaining websites, show no sustained difference between pre- and post-task-completion.

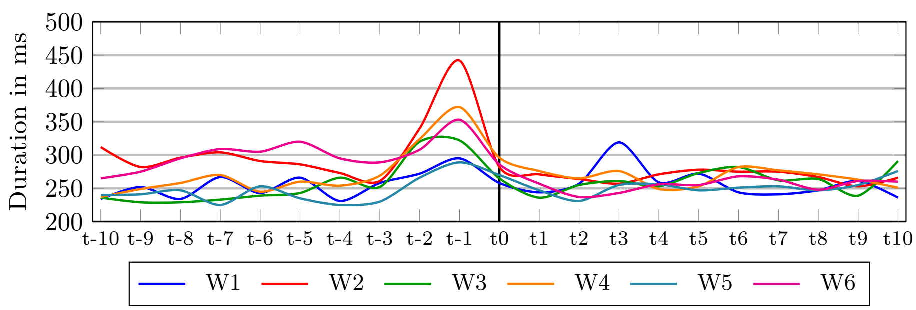

Fig. 6. Line chart illustrating temporal changes in mean fixation duration (FD) across all websites (W), averaged across participants, within a 20-second window spanning 10 seconds pre- and post-task completion (t=0).

Finally, Figure 7 reveals systematic gaze shifts both horizontally and vertically around the point of task completion ($t = 0 \pm 2$). Prior to completing a task, participants progressively scanned downward, reflected by a steady increase in mean y-coordinates, followed by an immediate decrease upon transitioning from pre- to post-task-completion. In terms of mean x-coordinates, W2, W4, and W6 exhibited a rightward shift pre-task-completion, followed by a leftward shift post-task-completion. W1, W3, and W5 show less changes in horizontal fixation distribution.

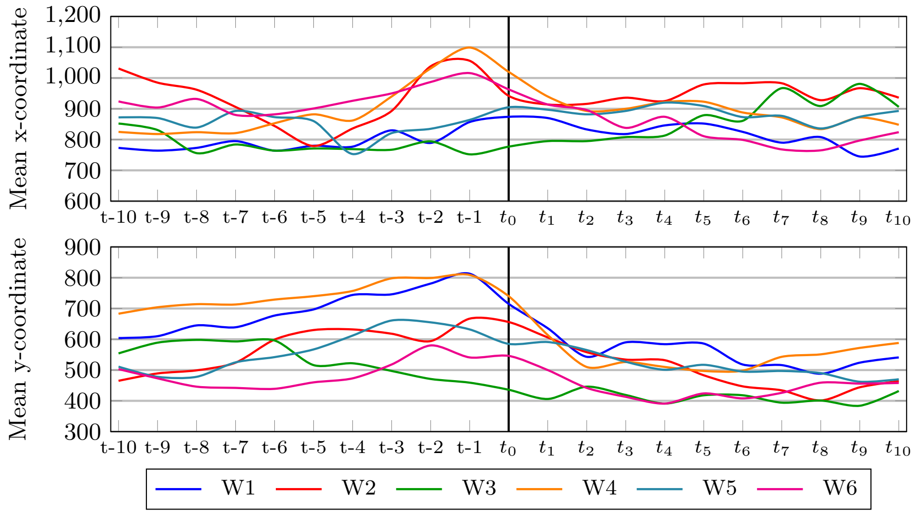

Fig. 7. Line chart illustrating temporal changes in mean fixation x- and y-coordinates across all websites (W), averaged across participants, within a 20-second window spanning 10 seconds pre- and post-task completion (t=0).

In summary, the transition-based analysis supports and extends several findings from RQ1. First, the observed increase in SD during non-task-related free browsing was abrupt rather than gradual, appearing at the onset of the post-task-completion phase. Second, the transient spike in FD immediately before completing the task is consistent with existing literature [16]. Additionally, FD decreased directly post-task-completion only for W2 and W6, which were also the only websites showing an overall decline in mean FD during free browsing in RQ1. Third, fixation distributions shifted systematically around task completion, with participants directing attention towards the top of the screen immediately post-task — also consistent with prior findings [11]. Taken together, these findings indicate that eye-tracking metrics differ between interaction states and can change immediately at the moment of transition, reflecting a clear behavioral shift between user states. Consequently, **H2 is accepted**.

6 Discussion and Limitations

The present findings show that eye movements differ between task-related search and non-task-related free browsing on realistic, multi-page websites. This extends — and partially aligns with — existing results from simpler single-page settings. In terms of fixation-related metrics, FD changed significantly between both interaction states across all six websites and fixation distributions shifted systematically across four of the six websites. This shows that users direct their gaze to distinct screen regions depending on whether they are involved in a task-related search or non-task-related browsing. Identified shifts are consistent across stimuli, despite different website structures and content. Particularly the vertical fixation shifts, are consistent with prior work on single-page websites [1,11,21,22] and highlight that users experience an attentional shift when transitioning from task- to non-task-related interaction. That said, the direction of FD change was not uniform: four websites showed a significant increase during free browsing with the remaining two showing a decrease, indicating that currently unknown website-specific characteristics also influence this effect.

For saccade-based metrics, the present findings diverge more from existing literature. While SD increased consistently across five of the six websites in task- versus non-task-related interaction, SA increased for only two — contrary to prior findings of uniformly longer saccades [14]. Two factors likely contribute to this discrepancy. First, the task structure itself is not directly comparable: the reference study involved a visual search for images or icons, while the present study required locating text-based information. These two search modes may produce inherently different baseline scanning patterns that compare differently to free browsing. Second, unlike single-page stimuli where content is finite, the multi-page websites offer participants additional pages to explore, potentially leading to continued, more relaxed, scanning once a task is completed. Nonetheless, our findings and prior literature both identify differences in fixation and saccade metrics between task- and non-task-related website interaction.

The temporal analysis around task completion provides further insight into not only how both interaction states differ, but also when these differences emerge. Consistent with prior research [16], spikes in FD occur immediately pre-task-completion. This elevation, however, does not persist post-task-completion. A sustained post-task-completion decline in FD was observed only on websites associated with lower mean FD during non-task-related interaction, suggesting that — unlike decreases — FD increases gradually rather than abruptly. The opposite is true for SD, which increases right after completing the task and remains elevated. A similar effect is observed for fixation locations. Here participants shift their gaze toward the top of the screen directly post-task-completion — consistent with findings from single-page studies [13, 21]. This suggests that during real-world web interaction, saccade behavior and fixation distributions can shift abruptly post-task-completion.

Beyond metric-level differences, the contrast between interaction states may also reflect a broader shift in attentional control and direction. Task-related interaction is likely more influenced by top-down attention, where gaze is directed towards goal-relevant content. In contrast, the absence of a specific goal during free browsing may increase susceptibility to bottom-up, stimulus-driven influences. Future work could investigate this relationship between top-down and bottom-up attention and task-related gaze behavior more directly.

Finally, several limitations should be acknowledged. First, the observed effects are not consistent across all stimuli, indicating that website-specific characteristics influence gaze behavior in each state. Second, the analysis is based on data aggregated across participants and does not account for individual differences in interaction behavior. This is particularly relevant for non-task-related viewing, where gaze patterns are shown to diverge more strongly between participants [5]. A participant-level analysis could uncover more distinct interaction strategies, demographic influences such as age, or differences between participants who successfully completed the task and those who did not. Third, search- or website-specific characteristics such as task difficulty or website structure may influence eye-movement behavior, visual search strategies and consequently the underlying metrics discussed here. These factors should be examined more systematically in future work. Finally, this analysis focuses strongly on detectable changes in eye-tracking metrics around task-completion. Future research could employ more advances techniques such as machine learning to see how strongly these patterns are and how accurately such transitions can be detected - particularly when the interaction is not reduced to the time frame of task-completion.

7 Implications for Usability and UX Research

The results of this analysis indicate that task- and non-task-related web interaction constitute distinct behavioral states, reflected in measurable differences in eye-movement patterns. Although effects vary across metrics and websites, patterns often span across multiple stimuli, suggesting that user states are behaviorally distinguishable even in complex, multi-page web environments.

These findings have important implications for eye-tracking-based HCI research. First, distinguishing between both states is critical for usability and UX research, as pooling data across interaction states may obscure differences in eye-movement metrics. Researchers aiming to assess usability should therefore ensure that data is collected within clearly defined task-related contexts and exclude post-task-completion data, as eye movements can change rapidly once task constraints are removed. Including post-task-completion interactions may obscure usability-relevant eye movements, particularly in short sessions. Conversely, studies focusing on UX should explicitly incorporate phases of non-task-related interaction. Even loosely defined tasks — like selecting a product of choice — may induce task-related gaze behavior and unwanted bias.

More broadly, these findings resonate with the pragmatic–hedonic model [12], which distinguishes between task-related (pragmatic) and non-task-related (hedonic) qualities of digital products. The observed gaze differences may directly reflect this distinction, with task-related patterns capturing pragmatic and non-task-related patterns capturing hedonic behavior. If this mapping holds, certain eye movements may be more usability- or UX-related depending on the user's interaction state. This potential link between eye movements and HCI theory remains hypothetical and needs to be verified through further studies. Either way, HCI-based eye-tracking studies should explicitly account for task- and non-task-related interaction states during the study design and interpretation of results. Ignoring this distinction risks conflating distinct cognitive processes and may lead to misleading conclusions about perceived usability or UX.

8 Conclusion

This study investigates gaze differences between task- and non-task-related interactions across six real-world multi-page websites. The results demonstrate that both interaction states are associated with systematically different gaze patterns, most notably in fixation distributions, saccade duration (SD), and fixation duration (FD). A temporal analysis around task completion further revealed a transient spike in FD immediately prior to completing a task, alongside a sustained increase in SD post-task-completion, indicating that gaze behavior shifts rapidly once task constraints are removed. These findings underscore the importance of accounting for task- and non-task-related interaction states in eye-tracking-based usability and UX research. Thus, HCI-based eye-tracking studies might consider excluding task-related segments when analyzing UX, and think about removing non-task-related segments when analyzing usability.

Acknowledgements This research was funded by the EDIH *Digital Innovation Ostbayern (DInO)*, funded by the European Union and European Funds for Regional Development (ERDF) (References: 101083427 & 20-3092.10-THD-105).

Open Science and Reproducibility In support of open science, all data is published on Zenodo: https://doi.org/10.5281/zenodo.20764567. The archive further contains detailed statistical results at both the participant and stimulus level as well as a link to the GitHub repository with the Python code for the statistical implementation.

AI Usage Disclosure We acknowledge the use of LLMs including DeepL Write, ChatGPT, Perplexity, and Claude to assist in language refinement or during the coding processes, not to generate content or ideas. Those originated solely from the authors or are based on the cited literature.

References

1. Betz, T., Kietzmann, T.C., Wilming, N., Koenig, P.: Investigating task-dependent top-down effects on overt visual attention. Journal of vision **10**(3), 15–15 (2010)
2. Castilla, D., Del Tejo Catalá, O., Pons, P., Signol, F., Rey, B., Suso-Ribera, C., Perez-Cortes, J.C.: Improving the understanding of web user behaviors through machine learning analysis of eye-tracking data: D. castilla et al. User Modeling and User-Adapted Interaction **34**(2), 293–322 (2024)
3. DIN EN ISO: Ergonomics of human-system interaction — part 110: Interaction principles (2020), iSO 9241-110:2020
4. Engl, F., Mottok, J., Burmester, M.: Predicting usability and ux based on eye movements: Identifying cross-stimuli interaction patterns with machine learning. In: Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems. pp. 1–20 (2026)

5. Eraslan, S., Yesilada, Y.: Patterns in eyetracking scanpaths and the affecting factors. Journal of Web Engineering pp. 363–385 (2015)
6. García, M., Cano, S.: Eye tracking to evaluate the user experience (ux): Literature review. In: International Conference on Human-Computer Interaction. pp. 134–145. Springer (2022)
7. Goldberg, J.H., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. International journal of industrial ergonomics **24**(6), 631–645 (1999)
8. Goldberg, J.H., Stimson, M.J., Lewenstein, M., Scott, N., Wichansky, A.M.: Eye tracking in web search tasks: design implications. In: Proceedings of the 2002 symposium on Eye tracking research & applications. pp. 51–58. Association for Computing Machinery, New York, NY, USA (2002)
9. Goldberg, J.H., Wichansky, A.M.: Eye tracking in usability evaluation: A practitioner's guide. In: the Mind's Eye, pp. 493–516. Elsevier, Amsterdam (2003)
10. Good, P.: Permutation, parametric and bootstrap tests of hypotheses. Springer (2005)
11. Hadnett-Hunter, J., Nicolaou, G., O'Neill, E., Proulx, M.: The effect of task on visual attention in interactive virtual environments. ACM Transactions on Applied Perception (TAP) **16**(3), 1–17 (2019)
12. Hassenzahl, M., Platz, A., Burmester, M., Lehner, K.: Hedonic and ergonomic quality aspects determine a software's appeal. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems. pp. 201–208. ACM, The Hague, Netherlands (Apr 2000)
13. Malpica, S., Martin, D., Serrano, A., Gutierrez, D., Masia, B.: Task-dependent visual behavior in immersive environments: A comparative study of free exploration, memory and visual search. IEEE transactions on visualization and computer graphics **29**(11), 4417–4425 (2023)
14. Milisavljevic, A., Abate, F., Le Bras, T., Gosselin, B., Mancas, M., Doré-Mazars, K.: Similarities and differences between eye and mouse dynamics during web pages exploration. Frontiers in Psychology **12**, 554595 (2021)
15. Milisavljevic, A., Le Bras, T., Abate, F., Gosselin, B., Petermann, C., Mancas, M., Doré-Mazars, K.: Different visual explorations between free-viewing and target finding tasks in websites: evidence from temporal analyses of ambient and focal modes. In: 20th European Conference on Eye Movements. vol. 12, p. 390 (2019)
16. Milisavljevic, A., Le Bras, T., Mancas, M., Petermann, C., Gosselin, B., Doré-Mazars, K.: A dynamic approach of searching behaviour in webpages. In: 41st European Conference on Visual Perception (2018)
17. Nielsen, J., Pernice, K.: Eyetracking web usability. New Riders (2010)
18. Novák, J.Š., Masner, J., Benda, P., Šimek, P., Merunka, V.: Eye tracking, usability, and user experience: A systematic review. International Journal of Human–Computer Interaction **40**(17), 4484–4500 (2024)
19. Shaphiro, S., Wilk, M., et al.: An analysis of variance test for normality. Biometrika **52**(3), 591–611 (1965)
20. Thüring, M., Mahlke, S.: Usability, aesthetics and emotions in human–technology interaction. International journal of psychology **42**(4), 253–264 (2007)
21. Vidyapu, S., Vedula, V.S., Bhattacharya, S.: Investigating and modeling the web elements' visual feature influence on free-viewing attention. ACM Transactions on the Web (TWEB) **15**(1), 1–27 (2020)
22. Zhou, X., Viola, I., Rossi, S., Cesar, P.: Comparison of visual saliency for dynamic point clouds: Task-free vs. task-dependent. IEEE Transactions on Visualization and Computer Graphics (2025)