Investigating the Influence of Ads on User Search Performance, Behaviour, and Experience during Information Seeking

Olivia Foulds
olivia.foulds@strath.ac.uk
University of Strathclyde
Glasgow, Scotland

Leif Azzopardi
leif.azzopardi@strath.ac.uk
University of Strathclyde
Glasgow, Scotland

Martin Halvey
martin.halvey@strath.ac.uk
University of Strathclyde
Glasgow, Scotland

ABSTRACT

The phenomenon of banner blindness explains that users can mentally ignore online advertisements (ads). However, eye-tracking studies have shown that users still fixate on ads, and even without direct gaze, ads still fall within a user's peripheral vision, which may negatively overload cognition. It is therefore unknown how blind, banner blindness, truly is, and what other effect ads may have on user's information seeking. To address this gap, a within-subjects design experiment was conducted with 37 participants who performed search tasks from the *TREC 2017 Common Core News Collection*, where 3 search tasks contained various types of ads, and one search task had no ads. Although our results showed that on average, participants retrieved similar amounts of relevant documents regardless of whether ads were present or absent, participants took significantly longer achieving this performance when ads were present. Furthermore, when ads were absent, participants reported less frustration, and not only believed they learned more, but a post-task recall test showed that participants actually did learn up to 38% more. Consequently, our findings suggest that banner blindness is more costly than just mere annoyance, and that the influence of ads on user's information retrieval recall may extend current theories of visual crowding.

CCS CONCEPTS

• **Information systems** → **Users and interactive retrieval**; • **Human-centered computing** → **HCI design and evaluation methods**.

KEYWORDS

Information retrieval; advertisements; banner blindness; clutter; crowding

1 INTRODUCTION

Online advertising has exponentially evolved since 1994 when the first banner advertisement (ad) appeared and limited numbers of people accessed a computer [60]. Now in the year 2020, roughly 88% of Western households own a digital device [5, 58], and the average internet user is exposed to roughly 200 ads daily [2]. As such, online ads have grown increasingly sophisticated and now cover a wide variety of formats including banners, click-bait, personalised ads, and newer designs such as advergames [43, 71]. A recent review that examined the last decade in online advertising research described how studies involving ads can be classified into six main themes: effectiveness, creative elements, context, personalization, mechanisms, and search advertising [43]. However these topics have all been driven from the perspective of the system or marketer, where the aim has been displaying ads that are more engaging, memorable, and clickable.

In parallel with the increase in online ad presence, users nowadays search for more and more information online. For example, E-Commerce has seen an unprecedented growth with a vast amount of users now engaging in online browsing to find desired products [32], 72% of "ordinary" users rely on the Internet to seek health-related information [12], and an estimated 3.6 billion users observe social information through social media accounts [20]. Yet in the field of Information Seeking and Retrieval (ISR) where one focus is to explicitly examine the retrieval of information online [21], research involving ads has concentrated on determining what ads should be displayed through a process called traffic allocation [19]. Similar to marketing objectives, traffic allocation aims at retrieving relevant ads to users that are most clickable and core to the business model of search engines [52]. However, the principles of ISR research state that research needs to move beyond the system and also consider the effects a system may have for its users [21].

Subjectively, many users report ads to be annoying [27] and options exist to remove their presence via ad blockers, browser reader modes, and/or paid subscriptions [36]. Alternatively, although some users continue to experience webpages with ads included, they report mentally skipping or filtering out the ads themselves – often resulting in a phenomenon dubbed *banner blindness* [17] or *ad avoidance* [42] – where ads are seemingly ignored, at least, from the advertiser's point of view. However eye-tracking studies have shown that users often continue to fixate on ads [62], and even without direct gaze ads would still lie in a user's peripheral vision, which may result in being implicitly processed [64]. Consequently, this paper aims to investigate whether so-called banner blindness is truly blind, and explores the main research questions of: ***How do ads affect: (1) user search performance; (2) search behaviours; and (3) user experiences, given an everyday search task.***

2 BACKGROUND AND MOTIVATION

To inform our experimental design, we draw on previous research across a diverse range of fields including advertising, marketing, information seeking and retrieval (ISR), and psychology. This has resulted in three main themes that can be broadly broken down into: 1) what aspects of an ad contribute to banner blindness; 2) how theories of visual attention suggest ads that lie in peripheral vision may create clutter and overload cognition; and 3) related works in ISR on visually complex webpages. As a result of these themes, a number of expectations regarding how ads may influence information seeking were formulated, which combined together then derive the present study's hypotheses.

2.1 Annoying Ads and Banner Blindness

When interviews were conducted about user information searches, webpages that were visually full with ads were subjectively identified as being distracting [51]. Annoying ads have therefore captured the attention of marketers to examine the economic and cognitive costs that annoying ads may produce. For example, summarised by Goldstein et al. [27, 28], annoying ads: interfere with user enjoyment of the webpage; cause users to worry about viruses and spyware; lower an advertiser's brand reputability; elicit trust issues regarding the publisher; and are less likely to be remembered. More recently, Alanazi et al. [3] performed an eye-tracking study on mobile devices, and found that users expressed lower satisfaction when ads were present. With many negative effects, it is unsurprising that users have developed strategies to avoid ads, through either physically blocking out their presence or just ignoring them.

How a user skips over an ad, either explicitly or implicitly, defines the concept of *Banner Blindness* or *Ad Avoidance*. In a study that examined user's attitudes towards ads, banner blindness was highlighted by many users, with one explaining: "I'm so used to seeing banner ads I tend to just ignore them" [67]. Previous research suggests that this concept is so common, that users rarely looked at ads and subsequently had very low recall of the ads that had been visible [17, 29, 37]. Factors that may affect banner blindness have been proposed such as the ad location [29, 31, 56], intrusiveness [10], or relevance to the task [13], to name a few. However the general consensus for why banner blindness exists appears to relate to how users deliberately avoid the ad to focus on the task at hand, especially when ads are task irrelevant [17, 37]. This is because during visual search, a user cannot attend to every visible element present, and therefore attention becomes selective to avoid information overload [37, 38, 55]. Thus under the concept of banner blindness, two expectations arise:

- *E1: During information seeking, a user's search behaviour and performance will be unaffected by ad presence.*
- *E2: Webpages containing ads will be perceived as more annoying than webpages without ads.*

2.2 Peripheral Visual Attention

Although users claim to ignore ads, findings from other studies would suggest that ads are still observed. Firstly, Tangmanee [62] showed that less than 10% of their users were able to correctly recall what ads had been viewed, despite nearly all fixating on at least one. Accordingly, users may *think* they ignore ads because they cannot remember them, but the ads *have* still been looked at. Secondly, Jahanian et al.[33] found that users who did not fixate on ads, but instead had to focus on the central point of a webpage for just 120 milliseconds, were successfully able to discriminate pages that contained ads from those that did not. Consequently, how the visual system processes peripheral information may suggest that ads, regardless of fixation, may covertly affect a user.

Research on visual attention more generally, irrespective of ads, has confirmed that even without direct gaze, distracting objects that lie in peripheral vision can still negatively overload cognition, a phenomenon called crowding [64, 68]. Crowding, otherwise known as clutter, has been known to affect many domains that involve search tasks, from safety-critical aviation operators who monitor multiple displays, to medical practitioners that scan through patient notes [1, 49]. For example, extra clutter in the form of added irrelevant information, resulted in the medical practitioner taking significantly longer to examine case information, while being less accurate and missing critical details that could impact upon diagnosis [49]. Yet, research is lacking for whether ads create negative clutter effects for everyday online information seeking.

However, advertisers and marketers *have* long been aware of the detrimental impact advertising clutter can have. Lee and Cho [70] showed that when webpages contained 4 extra ads, user's ad recognition significantly degraded. Similarly, when users perceived clutter to increase with increasing ad salience, a deterioration in ad memory was also found [37]. Marketers have therefore spent over 20 years attempting to find ways to cut through online clutter and engage the user to an ad [11]. Yet with more extreme ways being devised to grab a user's attention to ads, such as advergames that require user interaction [43], it is unknown whether this creates more visual clutter for a user. If ads create clutter and assimilate to theories in visual crowding, a contrasting expectation to E1 is:

- *E3: The presence of ads will cognitively overload users and negatively distract them from completing their information seeking goals, resulting in increased search time and worse search performance.*

2.3 Increasing Visible Webpage Elements

With awareness that webpages contain many elements that cannot be simultaneously processed [34], research in ISR that has increased visible elements during web search and monitored their effect on users provides clues for what might be expected to happen when ads are visible. Firstly, in multiple experiments that manipulated how many result snippets per page were presented, when fewer results were visible, users had more positive perceptions for the search task [53] and revealed the least self-reported workload and difficulty in finding relevant documents [35]. Secondly, in different research investigating aggregated search – where different types of media such as images or news are blended into the same search engine result page (SERP) – users rated the blended interface as significantly less usable [65]. Similarly, when web browser reader views have been implemented –which strips back webpage complexity through excluding items such as menus and images – users perceived the visual appeal of the page to significantly increase and were able to read pages 5% faster [41]. Although these three examples cover a range of visible elements during online search, they all

concluded similar results that increased elements on a webpage are associated with more negative perceptions. Oulasvirta et al. [53] explained these results using the concept of 'choice overload', where a user cannot exhaustively attend to everything they see, and this then degrades user satisfaction because the user worries they may have attended to the wrong thing. Alternatively, explanations may arise from theories in economic models of search, where users have to physically expend more effort to search through additional elements [9], and a higher effort then induces negative emotions [34]. Consequently, as the presence of ads also increases the amount of visible elements, this creates the following expectation:

- E4: *The presence of ads will result in users reporting more negative perceptions of the search task such as increased difficulty in finding relevant documents, and decreased visual appeal of web-pages.*

Other studies however, have identified that adding images can be beneficial for information seeking, and many ads contain images. For example, prior studies in aggregated search found that even although users had no reason to look for images, some images increased user interaction with the system, which resulted in greater accuracy at completing their search task of finding specific information embedded in a webpage [7]. Additionally, explorations in media communication identified that images stimulate engagement and interest in news stories [26]. This makes sense, as although it has already been identified that when a webpage is too visually complex users find it cognitively taxing to process [64], research since the 1970s has been aware that on the other end of the spectrum, if there are too few elements to process, then users may feel bored [Berlyne, 1970, cited in [54]]. Boredom may then lead users to distraction by other things around them, or cause them to abort their search all together. Therefore, an inverted U-Shape relationship has been proposed for visual stimuli [54] where users need exposure to a moderate amount of stimuli, as both too much, and too little, can be negative for different reasons: too much to process increases users' cognitive load, inducing frustration and negative emotions, which may result in a shorter duration of search [34]; yet equally, distractions can be beneficial by replenishing mental resources, resulting in lower reported workload and stress [44]. Thus given the right combination, adding images in the form of ads to webpages has the potential to offer the right balance of stimuli to keep users engaged, leading to the final expectation:

- E5: *The presence of ads will reduce boredom, increase user interaction, and result in higher search performance.*

2.4 Summary and Hypotheses

Overall, little research has explored how ads directly affect information seeking from a user's point of view. However studies that have examined visual attention, theories of crowding, and visually complex webpages during ISR have shown contrasting findings. On one hand, increasing visual complexity can result in users taking longer to complete tasks, missing important details, having greater difficulty finding relevant documents, while reporting less usability and workload. Yet equally, too little stimuli can elicit boredom, and so images may actually stimulate engagement and increase interaction, which results in greater task accuracy. Alternatively, the phenomenon of banner blindness may be so strong that regardless

of whether ads are present or absent, users may ignore them to focus on the task at hand. Consequently, the null hypothesis for the present study is:

- H0: *User search performance, behaviour, and experience will be unaffected by the presence of ads.*
- H1: *In contrast, the two-tailed non-directional hypothesis predicts that ads will affect user search performance, behaviour, and experience.*

3 METHOD

A within-user design was implemented where participants undertook four search tasks. One search task was the baseline condition where no ads were present, and the three other search tasks contained different kinds of ads visible. As previous studies have highlighted that relevant ads to the topic have elicited different attentional effects [37, 43], we created the three different ad conditions such that one showed ads that were topical to the task, another showed off-topic ads, and the remaining condition contained a mixture of both topical and off-topic ads. Due to space limitations and because the main focus of this paper was to analyse a general effect of ads regardless of type, we reserve exploration of the different ad conditions for future work. Consequently, to answer the main research questions of whether ads alter user's information seeking performances, behaviours, and experiences, the outcome measures for the three ad conditions were averaged to devise an overall measure of ads (Named from now on as 'AA' to denote 'All-Ads'). This operationalisation of averaging conditions aligned with the most recent literature we could find that compared the presence of one factor that had varying levels, against another condition where the main factor was absent [23]. Thus, two conditions are analysed in the present paper: **All-Ads (AA)** versus **No-Ads (NA)**.

3.1 Search Task

Each search task comprised of a simulated work task [14] that was situated in the context of a news-based retrieval system. This context was chosen to reflect a common scenario where many users retrieve news online [63], and has been used in prior ISR studies where users were given a topic and asked to gather as many relevant and different documents that were appropriate for learning about the given topic [35, 46, 47]. To quantify performance, all topics selected came from the *TREC Common Core 2017* collection *(CC2017)* [4] as this contains pre-assessed judgements for what is considered to be relevant for each given topic. CC2017 consists of over 1.8 million newspaper articles from the *New York Times (NYT)*, ranging from the period 1987 to 2006. Five topics were chosen that reportedly had similar levels of difficulties in other ISR studies [35, 46, 47]: Airport Security, Wildlife Extinction, Tropical Storms, Curbing Population Growth, and Piracy. All topics were randomly allocated to a condition for each participant, except Piracy, which remained as a stable practice task. For each task, participants were given a brief about what kinds of things would be relevant and irrelevant. For example, for the topic *Wildlife Extinction*, participants were explicitly told: "Find articles[1] that discuss efforts made by countries

---

[1]Note, from here on, articles will be referred to as documents to be in keeping with previous ISR studies
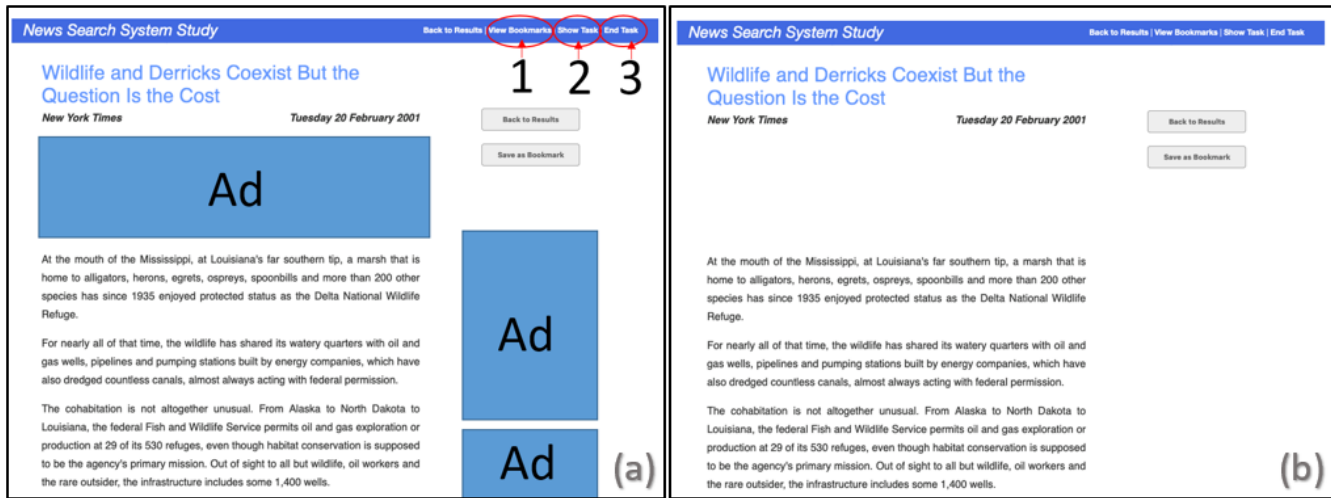
Figure 1: An example structure of the document view when (a) ads were present, and (b) ads were absent. Section 3.2 describes the annotations. Please note, for AA, additional ads were positioned on the right and bottom if a participant scrolled down.

other than the United States to prevent the extinction of wildlife species native to their countries."

3.2 Search System

To run the experiment, a standard webpage search interface was created where participants could issue queries, 10 result snippets per page would then become visible, and participants could click on any link to view the entire document, where they could then bookmark it if thought relevant. Three additional buttons were positioned at the top of the webpage (See Figure 1): **1)** 'View bookmarks' created a pop-up where participants could see the documents they had already saved; **2)** 'Show task' allowed participants to remind themselves of the specific task, as it has previously been found that information workers struggle with their memory for the exact task to be completed [45]; and **3)** 'End task' allowed participants to move onto the next section when they felt they had found enough.

The Whoosh Information Retrieval (IR) toolkit[2] with the *BM25* retrieval algorithm ($\beta = 0.75$) was used as the underlying retrieval system. The $P@10$ values were computed for every query, which scores how many relevant results were among the top 10 results presented to a participant. To double check **NA** was similar to **AA**, the $P@10$ values for every query in each condition were compared using a t-test which returned a non-significant result, confirming that generally, participants in each condition saw similar levels of relevant results ($P@10$ mean in **NA**: 0.33, and **AA**: 0.31).

For conditions where ads were displayed, a banner ad was located at the top, bottom, and four ads in the right rail of the webpage, on both the SERP and document pages, as done in previous advertising research [6]. The ads were always randomly selected from a pool of ads on each page load depending on the condition – with the pool consisting of topical, off-topic, or a mixture of both ads. If an ad was clicked, this would yield a popup window displaying a larger version of the ad. Note that in our study, we only observed

four clicks on ads over all participants, meaning that few ads were actively engaged with. For the no-ads condition, blank space was left to ensure that the content information in the webpages was always in the same location and that there was no bias in presenting the information higher up.

3.3 Advertisements

As identified in Section 1 above, ads come in many formats, and so to minimise potential confounds of factors such as animations, personalised-ads, or other interactive ads affecting performance, we chose to focus on static banner ads as these appear to be the most commonly researched in the literature [43]. Static banners were sourced from the *Ads of the World* database[3] where, for each topic, a selection of topical and off-topic ads were selected. In line with Buscher et al. [18], topical ads were defined by their appropriateness to the search task. Three volunteers manually examined all ads for their appropriateness to each topic. All raters had to agree on an advert's appropriateness to be considered either topical or off-topic. Inconclusive ads were discarded. This created 6 databases where each topic had its own selection of 40 topical ads, and then one large database comprised of 200 ads that were off-topic for all topics.

3.4 Outcome Measures

Dependent variables for this study were split into three main categories: *search performance*, considering how well participants performed; *search behaviours*, considering participants' interactions with the system; and *search experiences*, considering what participants thought and felt about the task, system, and personally.
Measuring Search Performance. Using the TREC CC2017 relevance judgements [4], we were able to estimate participant search performance through counting how many documents that participants had saved were considered TREC-relevant for the given topic.

The *accuracy* of a participant's searching ability was further calculated by: 1) the proportion of TREC-relevant documents saved over the total number saved (**Relevant-Saved/ Total-Saved**); 2) how many TREC-relevant documents had been saved in relation to how many TREC-relevant[4] documents a participant had hovered over in the SERP (**Relevant-Saved/ Relevant-Hovered**); and *3)* had actually been clicked on (**Relevant- Saved/ Relevant-Clicked**).

Additionally, immediately post-task, to give an indication of how much participants had learned from their search, participants had to recall as many concepts that they had previously found (**Concepts Recalled**). For example, in the topic *Wildlife extinction*, participants were asked to recall the names of species and the country that had taken preventative measures to save the species. To analyse whether the concepts recalled had just been learned or were already prior knowledge, two checking measures were implemented. Firstly, before each task, every participant completed a brief survey where they indicated how much they knew about the topic on a Likert-type scale ranging from *1 - Nothing* to *5 - A lot*. Secondly, each concept recalled was checked against the documents that they had identified, and only counted as correct if the content matched what a participant had interacted with in their search.

Measuring Search Behaviour. To provide exploratory insights into participant's search behaviours, various interactions with the search system were logged for each topic, using behavioural measures that have been widely used in previous ISR studies [15, 16, 35, 65] such as: the number of queries issued; average query length; documents hovered over and clicked on (including those which had been pre-assessed as relevant). From the log, we also computed a series of time-based measures, including: time taken until a first click; time until first saved document; total time spent querying per topic; total time spent on SERPs; total time spent examining documents overall; and total session time per topic.

Measuring User Search Experience. Participant subjective search experiences were analysed using multiple surveys after each condition. The surveys were split into three sections, with 5-point Likert-type items adapted from various studies [8, 22, 27, 44, 48]:

(1) Task-Focused Survey: Participants focused on their perception of the task over three statements with various scales. Questions included: *(a)* how much did you learn? *1- Nothing* to *5- A lot*; *(b)* how interesting was the topic? *1- Not at all* to *5- Very*; and *(c)* how difficult was the task? *1- Very easy* to *5- Very difficult*.

(2) User-Focused Survey: Using a scale of *1- Strongly disagree* to *5- Strongly agree*, participants expressed how the search made them feel, considering their: *frustration*; *confidence*; *enjoyment*; *satisfaction* (with their decisions); and *tiredness*.

(3) System-Focused Survey: Participants rated their perception of the system over statements with the scale of *1-Strongly disagree* to *5-Strongly agree*. Questions considered how: *aesthetically appealing*, *boring*; *annoying*; *confusing*; and *engaging* the system was.

3.5 Demographics

38 participants completed the study: 23 males; 14 females; and 1 did not disclose their demographic details. Ages ranged from 18 to 58, with a mean of 32 years old. One male participant, aged 22, was considered an extreme outlier and removed from analysis as their

logs demonstrated figures more than 20 times larger than the other participant logs. All 37 remaining participants were native English speakers with a range of educational backgrounds, as self reported highest level of education achieved included: 4 post-graduates, 24 college graduates, and 8 high school graduates. For taking part, participants were compensated with the equivalent of US$13.

3.6 Experimental Procedure

The experiment was conducted on the online platform Prolific[5] where participants must have had a minimum screen resolution of 1024×768, and disabled any ad-blockers. Programmatic checks ensured these requirements were complied with. In accordance with the ethical approval sought from our University department's Ethics Committee, each participant was provided with instructions about what they would be required to do and then gave informed consent if they were happy to proceed. After completing a short demographics survey, the participants undertook the practice task using the *Piracy* topic, to familiarise themselves with the system and learn how to query, browse, and save documents. Before each main search task, participants completed a brief survey about their knowledge for the topic and then continued onto the task. When participants felt they had saved enough relevant documents, they were to press the button labelled 'End task'. Alternatively, to ensure the overall experiment did not overrun, the system would automatically move onto the next part after eight minutes. Eight minutes was chosen as a similar experimental setup that also used the TREC Common Core 2017 collection found that users spent approximately seven minutes per task [47]. Immediately afterwards, three user search experience questionnaires were given (described previously in detail in Section 3.4). Then, a post-task questionnaire was given to assess how many concepts participants could recall from their search on that topic. This process continued for the remaining 3 search tasks. To ensure that topic and ordering effects were minimised, a fully factorial design was implemented where the ad-type was always randomly rotated between topics, and the order of topics always varied for each participant.

4 RESULTS

To answer the main research questions (as described in Section 1), using the user metrics derived from search tasks completed under the conditions of No-Ads (**NA**) and All-Ads (**AA**), the following results will be separated into: (1) User Performance; (2) User behaviour; and (3) User experience.

For performance and behavioural measures, Paired Sample T-Tests administered using Pingouin [66] determined whether there were any significant differences between **NA** and **AA**. Before each t-test was administered, the difference between **NA** and **AA** for each dependent variable was analysed for normal distribution using both visual inspection and the Shapiro-Wilk Test. Out of 19 variables, six breached normality *(Relevant-Saved, Relevant-Saved/Relevant-Clicked, Number of queries, Average query length, Average time taken until first click, and Time session Overall)*. For these six variables, main effects were examined with a more conservative alpha value of $\alpha = 0.01$, as recommended when data breaches the assumption of

---

[4]Note, from here on, TREC-relevant documents will just be referred to as relevant.

Table 1: The performance measures (mean ± standard deviation) for No-Ads (NA) and All-Ads (AA). For a detailed description of each performance measure, refer to Section 3.4. Bold signifies significant differences between NA and AA.

Performance	No-ads (NA)	All-ads (AA)
P@10	0.33 ± 0.21	0.31 ± 0.09
Total documents saved	4.62 ± 3.02	4.84 ± 3.12
Relevant-Saved	3.08 ± 2.25	2.78 ± 1.24
Relevant-Saved / Total-Saved (%)	67 ± 29	61 ± 15
Relevant-Saved / Relevant-Hovered (%)	**43 ± 31**	**33 ± 16**
Relevant-Saved / Relevant-Clicked (%)	77 ± 25	78 ± 19
Concepts Recalled	**3.78 ± 1.97**	**2.74 ± 1.28**

Table 2: The behavioural measures (mean ± standard deviation) based on No-Ads (NA) and All-Ads (AA). Measures of time are in seconds. Bold signifies significant differences.

Behaviour	No-ads (NA)	All-ads (AA)
Number of queries	3.27 ± 2.28	3.95 ± 1.81
Average query length	29.16 ± 14.13	28.41 ± 8.11
Total document hover count	31.68 ± 20.23	36.46 ± 19.57
Total relevant document hover count	10.11 ± 7.62	10.21 ± 6.02
Total document click count	7.24 ± 3.93	7.85 ± 3.55
Total relevant document click count	4.03 ± 2.71	3.83 ± 1.74
Average time taken until first click	11.93 ± 17	10.70 ± 6.21
Time till first saved document	65.6 ± 52.75	88.16 ± 55.88
Total time on documents	**204.37 ± 108.19**	**236.92 ± 90.78**
Total time on SERP	105.04 ± 55.83	114.83 ± 44.09
Total time spent querying	**22.30 ± 15.36**	**29.58 ± 15.94**
Time session overall	**357.44 ± 131.99**	**402.45 ± 105.24**

normality [61]. The remaining 13 dependent variables maintained the standard alpha value of $\alpha = 0.05$ to assess significance.

4.1 User Performance

Table 1 reports the mean (± standard deviation) for the observed performance measures for the conditions **NA** (no ads, baseline) and **AA** (average over all ad conditions). In terms of accuracy during the search task, participants saved similar numbers of documents (**NA**:4.62, **AA**:4.84) and relevant documents (**NA**:3.08, **AA**:2.78) in both conditions. Additionally, the number of saved relevant documents out of how many relevant documents had been clicked on, was almost identical between conditions (**NA**:77%, **AA**:78%). However, a trend did emerge that participants in **NA** achieved a higher percentage accuracy for how many documents they had saved were classed as relevant (**NA**:67%, **AA**:61%). Similarly, a significant difference emerged that showed participants achieved a higher percentage accuracy for how many relevant documents they had hovered over and then saved, during **NA** (**NA**:43%, **AA**:33%), $t(36) = -2.08, p=.044$.

When analysing post-task performance, a measure of knowledge-gain demonstrated a significant effect that participants in **NA** could generally recall an entire extra concept learned from their search (3.78) compared to participants in **AA**(2.74), $t(35) = -2.81, p=.008$. This score did not appear to relate to user prior knowledge, as out of the 152 search tasks completed throughout this study, only 5 of the tasks consisted of participants answering that they "Knew details" about the topic in our pre-task knowledge test. Furthermore, each concept listed was checked by the lead researcher to see if it coincided with a document that had been viewed, and only counted as a correct concept listing if it matched.

4.2 User Behaviour

As well as differences occurring between conditions for *performance* during the search task, both conditions also elicited variations in *behaviour* in order to achieve their performance. In terms of the

amount of time spent completing the task, Table 2 shows that participants undertaking the **NA** condition took 357.44 seconds on average, but with **AA**, they took 402.45 seconds on average – roughly 13% longer. This difference was significant, $t(36) = 2.98, p=.005$. This finding appears to be attributable to participants spending significantly more time both querying (**NA**:22.30, **AA**:29.58, $t(36) = 2.35$, $p=.024$) and examining documents (**NA**:204.37s, **AA**: 236.92s, $t(36) =2.73, p=.01$) when ads were present. Although participants also spent longer examining the SERP during **AA**, this difference did not reach significance (**NA**:105.04s, **AA**:114.83s). Furthermore, although participants took a similar amount of time until their first click in both conditions (**NA**:11.93s, **AA**:10.70s), participants in **NA** took much less time – roughly 26% less – until they saved their first document (**NA**:65.6s, **AA**:88.16s).

Despite large variations in time, other measures of search behaviour between conditions did not reach significance. For example, the number of queries issued (**NA**:3.27, **AA**:3.95) and average query length (**NA**:29.16, **AA**:28.41 characters) remained similar for both conditions. Furthermore, participants hovered over (**NA**:10.11, **AA**:10.21), and clicked on (**NA**:4.03, **AA**:3.83), the same amount of relevant documents irrespective of condition when rounding to the nearest whole number. Yet interestingly, although not significant, trends did emerge that participants hovered over roughly 4 more documents in **AA** (**NA**: 31.78, **AA**: 36.46) and also clicked on roughly one extra document in this condition (**NA**:7.24, **AA**:7.85). Thus, in **AA**, participants hovered over and clicked on more documents, yet they did not hover over or click on more *relevant* documents overall. This might imply that ads somehow distracted participants to hover over and click on more irrelevant documents.

4.3 User Experience

To analyse the Likert-type items for each survey question, Wilcoxon signed-rank tests compared the no ads condition (**NA**) against the

average of all ads (**AA**). Main effects were then examined with $\alpha = 0.05$. To establish where the significance levels lay, the median score of each condition was the key indicator. The mean of each condition was also analysed, as done in similar research that used this kind of Likert-type data [51], and results are reported in Table 3.

Task-Focused Survey. Examining Table 3, we can start to build up a picture of the subjective differences reported by participants for each survey. In terms of the user's perceptions of the task they had just completed, after no ads had been visible, significantly more participants agreed that the task was easier to complete (**NA**:2, **AA**:3, *p=0.012*) and that they believed they had learned significantly more (**NA**:4, **AA**:3, *p=0.024*). These differences were not a result of how interesting the task was perceived to be, as no significant differences occurred for this question, *p=0.246*.

User-Centred Survey. When asked to consider how they felt after completing each task, in both conditions, participants reported similar levels of (a) tiredness (**NA**:2, **AA**:2) and (b) confidence in their decisions (**NA**:4, **AA**:4). However, for the amount of enjoyment experienced completing the task, although the median was higher after no ads had been visible (**NA**:4, **AA**:3), indicating that more participants agreed they enjoyed completing the task when ads were absent, this difference failed to reach significance, *p=0.065*. Nonetheless, significant differences between conditions did occur for levels of frustration (*p=0.003*) and satisfaction (*p=0.002*): more participants agreed they were frustrated when ads were present (**NA**:3, **AA**:2); and more participants agreed they felt more satisfied with their search performance when ads were absent (**NA**:4, **AA**:3).

System-Focused Survey. From Table 3, it can seen that the presence of ads did not result in participants believing the system was more confusing or differently aesthetically appealing, as no significant differences between conditions occurred, and both the median and means between **NA** and **AA** were very similar for those questions. However, although the medians were identical between conditions for the questions that asked participants about how engaging, boring, or annoying the system was, Wilcoxon tests still returned significant results. It has previously been identified that this is possible because Wilcoxon does not just test the null hypothesis that the median of a distribution is equal to some value, but rather is a ranked sum test [25]. Thus to analyse where the differences lay for these significant questions, the counts for how many participants selected each option were compared and displayed as percentages of the whole sample in Figure 2. This showed that when asked about how boring the system was, more participants disagreed that the system was boring in **AA**, and more participants agreed it was boring in **NA**, *p=0.02*. Similarly, more participants agreed that the system was engaging in **AA**, compared to more strongly disagreeing that the system was engaging during **NA**, *p=0.002*. However, despite participants appearing to find the system with ads present as more engaging and less boring, they equally found it to be more annoying, with more participants disagreeing that it was annoying in **NA**, and more agreeing it was annoying in **AA**, *p=0.041*.

5 DISCUSSION

In response to the main research questions of how ads affect information seeking in terms of user search performance, behaviour,

and user experience, many significant results were found. Similar to previous research where users identified ads to be annoying [27, 28], our participants also believed the search system was significantly more annoying to use when ads were present, confirming our initial expectation E2 (See Section 2). When ads have previously been reported as annoying, users have claimed to ignore them resulting in Banner Blindness, supposedly allowing the user greater focus on the task at hand [17, 37]. Under this assumption, it was expected that if ads are ignored, then they will not affect a user's information seeking (E1). The present study only partly supports this expectation, as although participants in both conditions retrieved similar amounts of relevant documents, participants had slightly higher accuracy when no ads were visible for: the percentage of how many saved documents were actually relevant; and how many relevant documents had been saved in relation to how many had been hovered over. Furthermore, measures of behaviour also differed where participants took significantly longer completing the task amid ads, which reaffirmed E3. Additionally, although the presence of ads reduced user-reported boredom as predicted by E5, ad presence was also associated with greater levels of frustration, annoyance, and less user-reported satisfaction with their search performance. These negative feelings then translated into participants finding the task more difficult to complete in comparison to when ads were absent, which corresponds to E4. Finally, participants not only believed they learned less after ads had been visible, but post-task tests revealed that participants recalled significantly less about their information seeking during the conditions amongst ads, which again supported E3. Taken together, the null hypothesis, which stated that ads will not affect user information seeking (See Section 2), can be rejected.

Crowding. Explanations for why participants take significantly less time completing the search task in **AA** would appear consistent with previous research on visual crowding that identified individuals perform search tasks with longer response times when visual stimuli increases [39, 49, 50]. Specifically, crowding arises when the impact of nearby contours interfere with and reduce visual discrimination when trying to focus on a target [39]. This occurs because distraction is innate: humans are predisposed to be aware of everything that is going on around them [57]. But as human brains are limited [59], an excess of visual stimuli (either directly fixated on or present in peripheral vision) exceeds cognitive attentional resources and short-term memory ability, which results in a bottleneck that impairs object perception [39]. Furthermore, as participants in the present study also demonstrated poorer recall post-task in **AA**, this reaffirms the concept of crowding, where short-term memory capabilities are exceeded [30]. Previously, crowding of ads has only been identified in studies where too many visible ads have impaired recognition for what ads had been viewed [37, 70]. The present study therefore furthers this understanding of ad crowding and provides the first evidence (to the best of our awareness) that the negative effects of ad crowding extend onto (a) increased response times; and (b) reduced retention of what users learn during a search task- at least, for short-term memory.

Negative user experience. Previous literature was inconclusive about what effect ads might have on user experience. Some studies showed that extra stimuli created worse perceptions in one scenario [53], yet replenished mental resources resulting in lower

Table 3: The survey questions and associated median scores and mean (± standard deviation). Questions highlighted in bold and ending in * signify that Wilcoxon tests were significant between the No-ads condition (NA) and All-ads (AA).

Survey	Question	Scale	MEDIAN		MEAN + STD	
			NA	AA	NA	AA
Task	**How much did you learn about this topic?***	*1: Nothing → 5: A lot*	4	3	3.68 ± 0.82	3.11 ± 0.57
	How interesting was this topic?	*1: Not at all → 5: Very*	4	3	3.59 ± 0.98	3.43 ± 0.65
	How difficult was this task to complete?*	*1: Very easy → 5: Very Difficult*	2	3	2.46 ± 1.04	2.92 ± 1.01
User	**I felt frustrated while doing the task***		2	3	2.22 ± 0.92	2.86 ± 1.06
	I was confident in my decisions		4	4	3.76 ± 0.83	3.54 ± 0.87
	I enjoyed completing this task	*1: Strongly disagree →*	4	3	3.62 ± 0.95	3.35 ± 0.72
	I was satisfied with my search performance*	*5: Strongly agree*	4	3	3.70 ± 0.91	3.22 ± 0.95
	I felt tired when completing this task		2	2	2.00 ± 0.88	2.22 ± 0.92
System	The system was aesthetically appealing		3	3	3.00 ± 1.18	3.05 ± 1.15
	The system was boring*		3	3	3.14 ± 0.92	2.78 ± 1.08
	The system was annoying*	*1: Strongly disagree →*	2	2	2.14 ± 0.75	2.51 ± 0.93
	The system was confusing	*5: Strongly agree*	2	2	1.97 ± 0.83	2.14 ± 0.71
	The system was engaging*		3	3	2.54 ± 0.90	2.97 ± 0.83
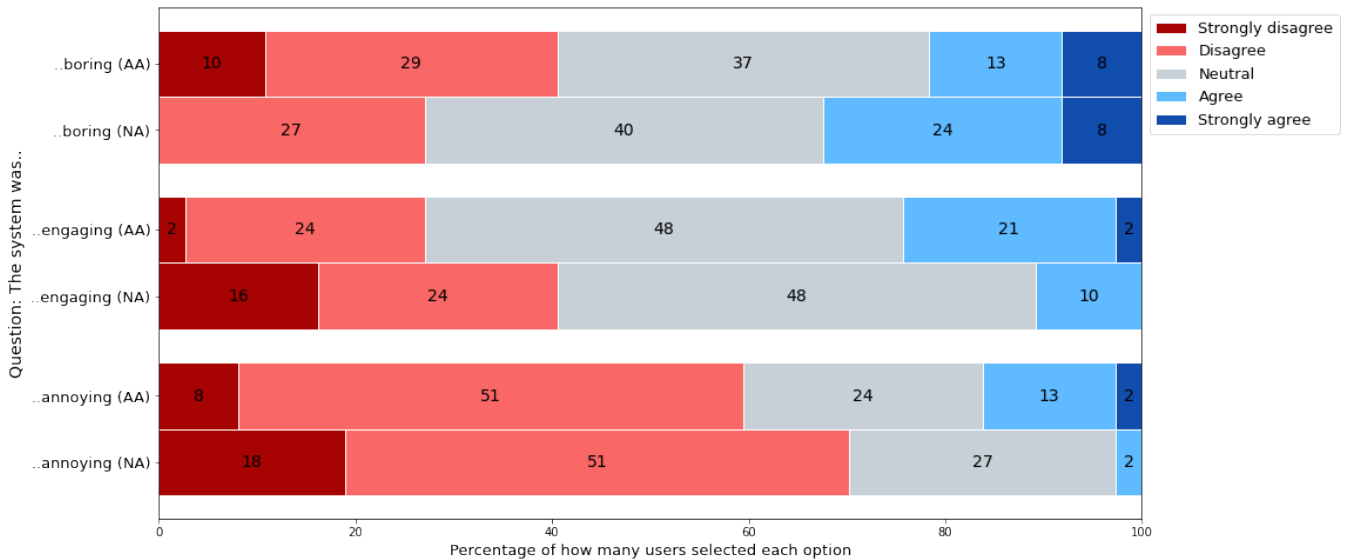
Figure 2: The percentages for how many participants responded with each option on the Likert-type scale for three system-focused survey questions in both No-ads (NA) and All-ads(AA). Specifically, 'boring', 'engaging' and 'annoying' are displayed as these questions returned significant Wilcoxon test results between NA and AA despite having equivalent medians.

workload and stress in another [44]. Our results provide a mixture of both good *and* bad experiences and feelings for the user engaged in information seeking. Beyond participants reporting that the system containing ads was more annoying, other negative perceptions of the search task were also found. For example, similar to research that found increased visual elements in the form of more result snippets visible on a SERP made it more difficult for users to find relevant documents [35], participants in the present study also believed that the task was significantly more difficult to complete when more visual elements (in the form of ads) were present. Similarly, after completing the task when ads were visible, participants reported feeling more frustrated and less satisfied with their search performance. As participants in both conditions ultimately retrieved the same amount of relevant documents, it is interesting that they simultaneously reported such different user experiences. However, our results also showed that participants undertaking **AA** hovered over more documents overall- yet no more relevant documents. As the P@10 values demonstrated no significant differences between conditions for the amount of relevant documents physically being shown to participants in either **NA** or **AA**, this might imply that participants were somehow more distracted and hovered over more incorrect documents during **AA**. This makes sense, as when users have 'choice overload', they are predisposed to distraction [39, 57], which then induces negative emotions [34, 53]. Therefore, from a user's perspective, the presence of banner ads should be minimised during information seeking.

Positive user experience. However, removing all ads may actually be negative for users, because for some aspects of user experience, ads were found to be positive. For example, in line with previous research which stated that images can stimulate engagement [26], and too little visual stimuli to process can result in boredom [54], the present study also found that when ads were absent, participants rated the system as significantly less engaging and more boring. Thus in other contexts where time or memory is not a critical factor, for example e-commerce or browsing social media, the presence of ads may provide a more enjoyable information seeking experience, through greater engagement and less boredom. This would correspond with previous research on ad attention which found that during a reading task, only 50% of users reported paying attention to ads, which was in comparison to 80% of users reportedly attending to ads when freely browsing online [37]. Consequently, if a user is undertaking information seeking where time is not of the essence, then ads should be kept to improve user experience. Conversely, when a goal-orientated task must be completed, physically removing ads through the use of an ad-blocker may be more beneficial to a user. Future research could thus focus on developing adaptive systems that can tailor the visual presentation of ads, depending on the task being conducted.

Contrasts with Previous Findings. Furthermore, the survey questions that returned similar results between conditions also raise various points for discussion. When prior work examined browser reader views where ads were removed (as well as other visual elements such as images), users perceived the visual appeal of the webpage to significantly increase [41]. Yet, the present study found that participants rated the system with and without ads as having the same levels of aesthetic appeal. Additionally, previous research has suggested that increased cognitive load may induce greater confusion or tiredness [39, 50]. Nevertheless, participants in the present study reported no difference in tiredness levels or confusion between conditions. Summarising these points together, although these results are at odds with previous studies, they are still unsurprising as only static banner ads were used in the present study. Animated ads are known to distract users more[24, 37]. Thus, it may be possible that more visually intrusive ads might mirror previous studies to also find that ads do increase confusion, tiredness levels, or alter webpage visual appeal. Alternatively, the contrasting results found may be attributable to the fact that ads in the present study were always positioned in consistent locations with equivalent sizing. Alternatively, previous studies have found that ad placement can be unpredictable [17], and unpredictable ads receive less attention [18]. Thus it could be possible that participants in the current study were able to get used to the ad presence, and as it was consistent, this resulted in factors such as confusion and tiredness not differing between conditions. Nonetheless, the current findings hold promise for current designers that banner ads are not all bad. If ways to minimise search time and improve recall are explored, future research could hopefully find a balance of ad displayed that allows both a positive user experience *and* performance.

Limitations and Future Work. In addition to some observations identified above such as only investigating static ads with consistent sizing and locations, a few other limitations can be noted. Firstly, the present findings are based on a small sample of only 37 users and in the context of ads in a news search environment. This

limits the power to generalise results: with new users with varying individual differences, or other contexts, such as entertainment or navigational search, different results may be found. Secondly, the present study only displayed graphical banner ads that differed in terms of how topical they were to the task. Thus, other types such as paid search-based textual ads, where users may be unable to distinguish their presence to organic results [40, 69], may alter user attention leading to different performance and experience outcomes. Similarly, it could then be argued that the current baseline condition of **NA** was too visually obvious from the **AA** condition, and that this made participants more aware of ad presence. However we believe this was ecologically and experimentally valid because: (a) it simulated the use of an ad-blocker; and (b) it ensured that any results found were in fact due to ad presence. Finally, a fixed number of ads were used on every page that a participant viewed. However, advertising clutter can be more formally defined as the 'ratio of promotional to non-promotional material in a medium' [Speck & Elliott, 1997, cited in [42]]. Consequently, an open question remains as to whether fewer amounts of ads that are visible to a user at any given glance, may ameliorate some of the negative effects found in the current study, such as increased time spent searching, reduced recall for what has been learned, or enhanced user satisfaction. Further research is therefore needed to examine the best type, location, and quantity of ads that can be visible to a user without interrupting their main experience.

6 CONCLUSION

In this paper, a search task in the context of news articles was devised to examine how the presence of banner ads affect user search performance, behaviour and experience during information seeking. While it has previously been recognised that ads are annoying distractions which users often ignore – resulting in so-called *banner blindness*– this study shows that banner blindness is more complex, as users take significantly longer completing the task when ads are present, and report more negative search experiences such as significantly higher levels of frustration, annoyance, and difficulty completing the task. Furthermore, we show that ads can significantly impede user's abilities in retaining the information they had just retrieved, with lower recall of relevant concepts post-task. Consequently, this study provides the first evidence that the negative effects of clutter –where an excess of visual stimuli overloads cognition– extend onto information seeking when ads are present. This opens up future research for better understanding: what the impact of ads, and more generally visual clutter, is on user's information seeking; and how ads can be displayed that create the most beneficial search outcomes and positive experience for the user.

REFERENCES

[1] Stephen H. Adamo, Matthew S. Cain, and Stephen R. Mitroff. 2015. Targets need their own personal space: Effects of clutter on multiple-target search accuracy. *Perception* 44, 10 (2015), 1203–1214. https://doi.org/10.1177/0301006615594921

[2] AdBlock. 2017. How to Make the Most of AdBlock. https://blog.getadblock.com/how-to-make-the-most-of-adblock-304f0e9dc1bc

[3] Afrah Olayan Alanazi, Mark Sanderson, and Jaewon Kim. 2020. The Impact of Ad Quality and Position on Mobile SERPs. (2020), 3–6.

[4] James Allan, Donna Harman, Evangelos Kanoulas, Dan Li, Christophe Van Gysel, and Ellen M Voorhees. 2017. TREC 2017 Common Core Track Overview. In *TREC*.

[5] Thomas Alsop. 2020. UK households: ownership of home computers 1985-2018. *Statista* (Feb 2020). https://www.statista.com/statistics/289191/household-penetration-of-home-computers-in-the-uk/

[6] Ioannis Arapakis and Luis A. Leiva. 2020. Learning Efficient Representations of Mouse Movements to Predict User Attention. (2020), 1309–1318. arXiv:2006.01644 http://arxiv.org/abs/2006.01644

[7] Jaime. Arguello and Robert. Capra. 2012. The Effect of Aggregated Search Coherence on Search Behavior Categories and Subject Descriptors. *Proc. 21st ACM CIKM* (2012), 1293–1302.

[8] Jaime Arguello and Bogeum Choi. 2019. The effects of working memory, perceptual speed, and inhibition in aggregated search. *ACM Transactions on Information Systems (TOIS)* 37, 3 (2019), 1–34.

[9] Leif Azzopardi. 2011. The economics in interactive information retrieval. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*. 15–24.

[10] Hyejin Bang, Jooyoung Kim, and Dongwon Choi. 2018. Exploring the effects of ad-task relevance and ad salience on ad avoidance: The moderating role of internet use motivation. *Computers in Human Behavior* 89, April (2018), 70–78. https://doi.org/10.1016/j.chb.2018.07.020

[11] Karen L Becker-Olsen. 2003. And now, a word from our sponsor–a look at the effects of sponsored content and banner advertising. *Journal of Advertising* 32, 2 (2003), 17–32.

[12] Sima Bhowmik, Md Main Uddin Rony, Md Mahfuzul Haque, Kristen Alley Swain, and Naeemul Hassan. 2019. Examining the Role of Clickbait Headlines to Engage Readers with Reliable Health-related Information. *arXiv preprint arXiv:1911.11214* (2019).

[13] Benjamin Boeuf, François A. Carrillat, and Alain D'Astous. 2018. Interference effects in competitive sponsorship clutter. *Psychology and Marketing* 35, 12 (2018), 968–979. https://doi.org/10.1002/mar.21149

[14] Pia Borlund and Peter Ingwersen. 1997. The development of a method for the evaluation of interactive information retrieval systems. *Journal of documentation* (1997).

[15] Horatiu Bota, Ke Zhou, and Joemon M Jose. 2016. Playing your cards right: The effect of entity cards on search behaviour and workload. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*. 131–140.

[16] Kathy Brennan, Diane Kelly, and Jaime Arguello. 2014. The effect of cognitive abilities on information search for tasks of varying levels of complexity. In *Proceedings of the 5th Information Interaction in Context Symposium*. 165–174.

[17] Moira Burke, Anthony Hornof, Erik Nilsen, and Nicholas Gorman. 2005. High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Transactions on Computer-Human Interaction (TOCHI)* 12, 4 (2005), 423–445.

[18] Georg Buscher, Susan T Dumais, and Edward Cutrell. 2010. The good, the bad, and the random: an eye-tracking study of ad quality in web search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. 42–49.

[19] Dagui Chen, Junqi Jin, Weinan Zhang, Fei Pan, Lvyin Niu, Chuan Yu, Jun Wang, Han Li, Jian Xu, and Kun Gai. 2019. Learning to Advertise for Organic Traffic Maximization in E-Commerce Product Feeds. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 2527–2535.

[20] J. Clement. 2020. Number of social media users worldwide. *Statista* (Jul 2020). https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

[21] Colleen Cool and Nicholas J Belkin. 2011. Interactive information retrieval: history and background.

[22] Anita Crescenzi, Diane Kelly, and Leif Azzopardi. 2016. Impacts of time constraints and system delays on user experience. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*. 141–150.

[23] Jonathan Daume, Sebastian Graetz, Thomas Gruber, Andreas K Engel, and Uwe Friese. 2017. Cognitive control during audiovisual working memory engages frontotemporal theta-band interactions. *Scientific reports* 7, 1 (2017), 1–13.

[24] Fangfang Diao and S. Shyam Sundar. 2004. Orienting response and memory for web advertisements: Exploring effects of pop-up window and animation. *Communication Research* 31, 5 (2004), 537–567. https://doi.org/10.1177/0093650204267932

[25] UCLA Institute for Digital Research & Education Statistical Consulting. 2019. Why is the Mann-Whitney significant when the medians are equal? https://stats.idre.ucla.edu/other/mult-pkg/faq/general/faq-why-is-the-mann-whitney-significant-when-the-medians-are-equal/

[26] Sylwia Frankowska-Takhari, Andrew MacFarlane, Ayşe Göker, and Simone Stumpf. 2017. Selecting and tailoring of images for visual impact in online journalism. *Information Research* 22, 1 (2017), 1619.

[27] Daniel G Goldstein, R Preston McAfee, and Siddharth Suri. 2013. The cost of annoying ads. In *Proceedings of the 22nd international conference on World Wide Web*. 459–470.

[28] Daniel G Goldstein, Siddharth Suri, R Preston McAfee, Matthew Ekstrand-Abueg, and Fernando Diaz. 2014. The economic and cognitive costs of annoying display advertisements. *Journal of Marketing Research* 51, 6 (2014), 742–752.

[29] Hacer Güner and Yavuz İnal. 2015. The effect of banner location on banner recognition in a Turkish government website: an eye tracking study. (2015), 65–72.

[30] Jay Hegdé, Serena K. Thompson, Mark Brady, and Daniel Kersten. 2012. Object recognition in clutter: cortical responses depend on the type of learning. *Frontiers in Human Neuroscience* 6, June (2012), 1–15. https://doi.org/10.3389/fnhum.2012.00170

[31] Yi Ting Huang. 2018. *The female gaze: Content composition and slot position in personalized banner ads, and how they influence visual attention in online shoppers.* Vol. 82. Elsevier B.V. 1–15 pages. https://doi.org/10.1016/j.chb.2017.12.038

[32] Rahul Radhakrishnan Iyer, Rohan Kohli, and Shrimai Prabhumoye. 2020. Modeling Product Search Relevance in e-Commerce. *arXiv preprint arXiv:2001.04980* (2020).

[33] Ali Jahanian, Shaiyan Keshvari, and Ruth Rosenholtz. 2018. Web pages: What can you see in a single fixation? *Cognitive Research: Principles and Implications* 3, 1 (2018). https://doi.org/10.1186/s41235-018-0099-2

[34] Jarosław Jankowski, Juho Hamari, and Jarosław Wątróbski. 2019. A gradual approach for maximising user conversion without compromising experience with high visual intensity website elements. *Internet Research* 29, 1 (2019), 194–217. https://doi.org/10.1108/IntR-09-2016-0271

[35] Diane Kelly and Leif Azzopardi. 2015. How Many Results per Page? A Study of SERP Size, Search Behavior and User Experience. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '15)*. Association for Computing Machinery, New York, NY, USA, 183–192. https://doi.org/10.1145/2766462.2767732

[36] Martin Kratky-Katz. 2015. *The cost of ad blocking. 2015 Ad Blocking Report.* PageFair and Adobe. https://blockthrough.com/2015/08/10/ad-blocking-report/

[37] Jarmo Kuisma. 2015. *Consumer Perception of Online Advertising.* Vol. 157. 175 pages. https://aaltodoc.aalto.fi/bitstream/handle/123456789/18246/isbn9789526064291.pdf?sequence=2{%}0Ahttp://urn.fi/URN:ISBN:978-952-60-6430-7

[38] Laure Léger and Aline Chevalier. 2017. Location and orientation of panel on the screen as a structural visual element to highlight text displayed. *New Review of Hypermedia and Multimedia* 23, 3 (2017), 207–227. https://doi.org/10.1080/13614568.2017.1399468

[39] Dennis M. Levi. 2008. Crowding-An essential bottleneck for object recognition: A mini-review. *Vision Research* 48, 5 (2008), 635–654. https://doi.org/10.1016/j.visres.2007.12.009

[40] Dirk Lewandowski. 2017. Users' understanding of search engine advertisements. *Journal of Information Science Theory and Practice* 5, 4 (2017), 6–25.

[41] Qisheng Li, Meredith Ringel Morris, Adam Fourney, Kevin Larson, and Katharina Reinecke. 2019. The Impact of Web Browser Reader Views on Reading Speed and User Experience. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[42] Yujie Li. 2019. User Perception Affects Search Engine Advertising Avoidance: Moderating Role of User Characteristics. *Social Behavior and Personality: an international journal* 47, 4 (2019), 1–12.

[43] Yuping Liu-Thompkins. 2019. A decade of online advertising research: What we learned and what we need to know. *Journal of advertising* 48, 1 (2019), 1–13.

[44] Gloria Mark, Shamsi Iqbal, and Mary Czerwinski. 2017. How blocking distractions affects workplace focus and productivity. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 928–934.

[45] Tara Matthews, Mary Czerwinski, George Robertson, and Desney Tan. 2006. Clipping lists and change borders: Improving multitasking efficiency with peripheral information design. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2. 989–998. https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/chi2006-clippinglists.pdf

[46] David Maxwell, Leif Azzopardi, and Yashar Moshfeghi. 2017. A study of snippet length and informativeness behaviour, performance and user experience. In *SIGIR 2017 - Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Association for Computing Machinery, Inc, 135–144. https://doi.org/10.1145/3077136.3080824

[47] David Maxwell, Leif Azzopardi, and Yashar Moshfeghi. 2019. The impact of result diversification on search behaviour and performance. *Information Retrieval Journal* 22, 5 (2019), 422–446.

[48] Lori McCay-Peet, Mounia Lalmas, and Vidhya Navalpakkam. 2012. On saliency, affect and focused attention. In *Proceedings of the SIGCHI Conference on Human*

Factors in Computing Systems. 541–550.

[49] Nadine Moacdieh and Nadine Sarter. 2015. Clutter in Electronic Medical Records. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 4 (2015), 591–606. https://doi.org/10.1177/0018720814564594

[50] Nadine M Moacdieh and Nadine Sarter. 2017. Using eye tracking to detect the effects of clutter on visual search in real time. *IEEE Transactions on Human-Machine Systems* 47, 6 (2017), 896–902.

[51] Meredith Ringel Morris, Adam Fourney, Abdullah Ali, and Laura Vonessen. 2018. Understanding the needs of searchers with dyslexia. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2018-April. Association for Computing Machinery. https://doi.org/10.1145/3173574.3173609

[52] Rita Moura and Beatriz Casais. 2020. Search Engine Marketing to Attract International Digital Traffic. In *Digital Marketing Strategies and Models for Competitive Business*. IGI Global, 74–102.

[53] Antti Oulasvirta, Janne P Hukkinen, and Barry Schwartz. 2009. When more is less: the paradox of choice in search engine use. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*. 516–523.

[54] Supavich (Fone) Pengnate, Rathindra Sarathy, and Jin Kyu Lee. 2018. The Engagement of Website Initial Aesthetic Impressions: An Experimental Investigation. *International Journal of Human-Computer Interaction* 35, 16 (2018), 1517–1531. https://doi.org/10.1080/10447318.2018.1554319

[55] Daniel J. Rea, Stela H. Seo, Neil Bruce, and James E. Young. 2017. Movers, Shakers, and Those Who Stand Still: Visual Attention-grabbing Techniques in Robot Teleoperation. *ACM/IEEE International Conference on Human-Robot Interaction* Part F1271 (2017), 398–407. https://doi.org/10.1145/2909824.3020246

[56] Marc Resnick and William Albert. 2014. The impact of advertising location and user task on the emergence of banner ad blindness: An eye-tracking study. *International Journal of Human-Computer Interaction* 30, 3 (2014), 206–219.

[57] Marc Resnick, William Albert, and Yunzhi Huang. 2016. The attention grabbing salience of viscerally engaging images. *Proceedings of the Human Factors and Ergonomics Society* 2001 (2016), 744–748. https://doi.org/10.1177/1541931213601171

[58] Camille L Ryan and Jamie M Lewis. 2017. *Computer and internet use in the United States: 2015*. US Department of Commerce, Economics and Statistics Administration. https://www.census.gov/library/publications/2017/acs/acs-37.html

[59] Stoo Sepp, Steven J Howard, Sharon Tindall-Ford, Shirley Agostinho, and Fred Paas. 2019. Cognitive load theory and human movement: Towards an integrated model of working memory. *Educational Psychology Review* (2019), 1–25.

[60] Ryan Singel. 2010. Oct. 27, 1994: Web Gives Birth to Banner Ads. https://www.wired.com/2010/10/1027hotwired-banner-ads/

[61] Statistics Solutions. 2013. Normality [WWW Document]. https://www.statisticssolutions.com/normality/

[62] Chatpong Tangmanee. 2016. Fixation and recall of YouTube ad banners: An eye-tracking study. *International Journal of Electronic Commerce Studies* 7, 1 (2016), 49–76. https://doi.org/10.7903/ijecs.1404

[63] Daniel Trielli and Nicholas Diakopoulos. 2019. Search as news curator: The role of Google in shaping attention to news information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.

[64] Junichi Tsurukawa, Mohammed Al-Sada, and Tatsuo Nakajima. 2015. Filtering visual information for reducing visual cognitive load. *UbiComp and ISWC 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the Proceedings of the 2015 ACM International Symposium on Wearable Computers* (2015), 33–36. https://doi.org/10.1145/2800835.2800852

[65] Lauren Turpin, Diane Kelly, and Jaime Arguello. 2016. To Blend or Not to Blend? Perceptual Speed, Visual Memory and Aggregated Search. *Sigir 2016* (2016), 1021–1024. https://doi.org/10.1145/2911451.2914809

[66] Raphael Vallat. 2018. Pingouin: statistics in Python. *Journal of Open Source Software* 3, 31 (2018), 1026.

[67] Aku Visuri, Simo Hosio, and Denzil Ferreira. 2017. Exploring mobile ad formats to increase brand recollection and enhance user experience. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. 311–319.

[68] Maarten W.A. Wijntjes and Ruth Rosenholtz. 2018. Context mitigates crowding: Peripheral object recognition in real-world images. *Cognition* 180, July (2018), 158–164. https://doi.org/10.1016/j.cognition.2018.06.015

[69] Bartosz W Wojdynski and Nathaniel J Evans. 2020. The covert advertising recognition and effects (CARE) model: Processes of persuasion in native advertising and other masked formats. *International Journal of Advertising* 39, 1 (2020), 4–31.

[70] Sang Yeal Lee and Yong-suk Cho. 2010. Do Web Users Care About Banner Ads Anymore? The Effects of Frequency and Clutter in Web Advertising. *Journal of Promotion Management* 16 (2010), 288–302. https://doi.org/10.1080/10496490903582594

[71] Ke Zhou, Miriam Redi, Andy Haines, and Mounia Lalmas. 2016. Predicting pre-click quality for native advertisements. *25th International World Wide Web Conference, WWW 2016* (2016), 299–310. https://doi.org/10.1145/2872427.2883053