Predicting whether Users View Dynamic Content on the World Wide Web

CAROLINE JAY, ANDY BROWN, and SIMON HARPER, University of Manchester

Dynamic micro-content—interactive or updating widgets and features—is now widely used on the Web, but there is little understanding of how people allocate attention to it. In this article we present the results of an eye-tracking investigation examining how the nature of dynamic micro-content influences whether or not the user views it. We propose and validate the Dynamic Update Viewing-likelihood (DUV) model, a CHi-squared Automatic Interaction Detector (CHAID) model that predicts with around 80% accuracy whether users view dynamic updates as a function of how they are initiated, their size, and their duration. The model is constructed with data from live Web sites and does not rely on knowledge of the user's task to make its predictions, giving it a high level of external validity. We discuss one example of its application: informing how dynamic content should be presented in audio via assistive technology for people with visual impairments.

1. INTRODUCTION

One of the major technological changes behind "Web 2.0" is AJAX (Asynchronous Javascript And XML (Extensible Markup Language)), which enables parts of a Web page to change in isolation without needing to refresh the whole page. This allows pages to become more interactive, to the extent of appearing and behaving like traditional desktop programs. In addition to AJAX there are more traditional scripting technologies which can also alter the appearance of a page. Together these provide the Web designer with a number of ways to change sections of a page without changing the URL. We call these sections "dynamic micro-content," and the changes "dynamic updates." Examples of dynamic micro-content, which include auto-suggest

lists, tickers, slideshows, updating tables, tabs, and search facilities, can be seen on pages such as the Yahoo![1], and iGoogle[2] Web portals (see Figures 2 and 3).

Dynamic micro-content is now used very widely: of the 500 most popular Web sites worldwide, over 90% utilize Web technologies that can generate updates [Brown et al. 2010]. Despite this prevalence, little is known about how people allocate attention to dynamic updates, particularly on live Web sites in a naturalistic setting.

This article presents the Dynamic Update Viewing-likelihood (DUV) model predicting whether or not users view dynamic updates with a high degree of accuracy. The model is the culmination of a detailed eye-tracking investigation that examined people interacting with a wide variety of dynamic microcontent on live Web sites. More than 2000 dynamic updates were recorded across all the experimental sessions, and these were categorized according to a number of factors, before being analyzed with a CHAID decision tree algorithm. The analysis showed that the manner in which an update is initiated is the most important predictor of whether it is viewed, with click-activated updates most likely to be viewed, and automatic updates least likely. The size of click-activated updates, and the duration of automatic- and hover-, enter-key- or keystroke-activated updates also significantly predicts whether or not they are viewed.

Section 2 describes previous work looking at how people allocate attention to animated or updating content on Web pages. Section 3 describes the initial data collection and model creation process, and Section 4 explains how the model was validated. Section 5 discusses why the model takes the form it does, and what it says about how users interact with dynamic micro-content, and Section 6 describes an application of the model—informing the presentation of dynamic updates in audio for users with visual impairments—and considers how it compares to other models predicting user behavior on the Web. Finally, Section 7 summarizes the contribution and limitations of the investigation, and outlines directions for future research.

2. BACKGROUND AND RELATED WORK

Psychological studies indicate that both the appearance of new objects [Hillstrom and Yantis 1994] and the movement of existing ones [Franconeri and Simons 2003] are good at capturing attention, but the extent to which observers allocate their attention to dynamic data sources depends upon many factors, including color and movement direction [Saenz and Buracas 2003] and the presence of competing dynamic events within the display [Miller 1989]. Although movement appears to have a greater bottom-up salience than other visual cues [Carmi and Itti 2006], it is not clear that it truly captures attention in a stimulus-driven fashion, so whether or not someone views a moving or changing object is affected by top-down processes, such as the task he/she is performing at the time, and his/her anticipation of the change [Jingling and Yeh 2007; Pinto et al. 2008].

When people are viewing Web pages, it is implicit they have a task in mind (even if it is simply finding something interesting to look at), as well as numerous expectations of a page based on their previous experience of the Web [Habuchi et al. 2006], so there are many factors that may influence the allocation of attention in this visually complex medium.

A number of controlled studies have examined the effects of changing or moving Web content on task performance. Zhang [1999] found that animations were detrimental to performance in an information-seeking task, particularly if the task was simple and the animations were brightly colored, qualitatively similar to the information sought, and repeatedly flashed on and off. By contrast, Diaper and Waelend [2000], who criticized

---

[1]http://www.yahoo.com.
[2]http://www.google.com/ig.

Zhang's tasks for lacking external validity (participants searched for random strings or words amongst other random strings or words), found that unrelated graphics, whether animated or not, had little impact on participants' ability to search for an answer in a block of text. Similarly, van Schaik and Ling [2004] found that in an information retrieval task, peripheral animations did not affect performance, although they were reported to be more distracting than static graphics.

Maglio and Campbell [2000] examined the extent to which scrolling displays (tickers) either distracted participants or helped to provide them with desired information. Discrete motion displays, where a line of text appears and remains still, were preferable to continuous scrolling displays for two reasons: they were less likely to distract someone from his or her primary task (in this case, editing a block of text); and the start-stop motion helped to alert participants to the arrival of new information, enabling them to schedule their glance to just after it appears, rather than forcing them to constantly monitor the ticker.

Bartram et al. [2003] examined various types of display update, and found motion was more effective in a notification system than either color or shape change, and that different types of "moticon" (moving icon) varied in their ability to attract the participant's attention. Blinking moticons were the least distracting, followed by moticons that moved up in a sinusoidal motion then jumped back to their starting place, then moticons that zoomed smoothly to twice their starting size. Traveling moticons, which moved from side to side, or up and down, were the most distracting. When it came to the ease with which participants could detect different moticon notifications, however, the task they were doing had a significant impact: it appeared easier to detect movements that were different to those that occurred in their current task.

McCrickard et al. [2003] investigated the effect of various notification systems (tickers and alerts) on different types of task, finding that tickering text is best for supporting deeper comprehension, and fading best facilitates reaction. In line with Diaper and Waelend [2000] and van Schaik and Ling [2004], they also found that animated displays did not disrupt performance of the primary task, in this case, browsing. In a study investigating how to present help documentation on a secondary monitor, participants spent 18% of the time viewing "ambient" updating help when employed in an open-ended nontime-critical task, but only 3% of the time when doing something highly focused and time critical [Matejka et al. 2011].

Experiments examining the phenomenon of "banner blindness" (failure to attend to salient blocks of content, usually advertisements) have found that animation fails to improve either recognition or recall of such content [Bayles 2002], although memory for animated banners is better when users are browsing, rather than performing goal-directed tasks [Pagendarm and Schaumburg 2001]. Burke et al. [2005] found that the presence of animated banners does not impair performance in a visual search task (although people do report the task to be harder than when banners are static), and used eye tracking to show that people are no more likely to look at animated banner ads than static ones.

It is possible to draw some high-level conclusions from the collective results of these studies: people do appear to notice changes to Web pages, but can ignore them if they need to, and the extent to which they attract attention depends on both the nature of the change and the task the user is performing.

One issue affecting the studies reported before is that they all used specifically constructed, rather than "wild" Web pages. This is necessary when running a controlled study examining task performance, but it is at the expense of ecological validity; it is impossible to be sure that the results would apply to other kinds of Web use.

The investigation reported here addresses this issue, by studying people undertaking a variety of complex tasks on live Web sites. To keep interaction as naturalistic as

possible, people were not constrained in the way they used and viewed the sites (beyond being asked to complete some high-level tasks), and were not asked to comment on whether they noticed page updates, or found them useful or distracting. Instead, their gaze was monitored with an eye tracker, to determine if and when they viewed a change. Eye tracking has been used to investigate allocation of visual attention to Web stimuli in numerous situations (e.g., Djamasbi et al. [2010]; Granka et al. [2004, 2006]), but with the exception of Burke et al. [2005], it has not yet been used systematically to ascertain the circumstances under which people view updating content. Tracking eye movements is not a perfect way of determining where someone has directed his or her attention: fixated objects do not always register in conscious attention, and objects that are not fixated may still be attended to covertly [Beanland and Pammer 2010]. Nevertheless, its ability to reveal where a person is focusing, coupled with its unintrusive nature, makes eye tracking an excellent tool for answering the question of whether people view dynamic updates when they are using the Web.

3. EXPERIMENT 1

Previous studies have shown that the ability of certain types of animations, moticons, and tickers to attract or distract the user changes according to their visual characteristics and the user's task. These studies were all conducted in highly controlled environments, however, so the extent to which their results apply beyond the study setting is not clear.

The goal of experiment 1 was to ascertain whether it is possible to predict the likelihood that an update will be viewed in real-world Web interaction. There are potentially many factors that may affect whether or not an update is viewed, particularly those relating to its salience, such as the colors or levels of contrast it contains. Although "salience maps" have been produced for Web pages [Zheng et al. 2009; Jay et al. 2007], these apply only to static pages, and do not consider changes to content over time. Whether the change is an isolated instance or occurs over a longer period, and the proportion of the content within the HTML element that changes, may also impact on whether a user views it. These variables are difficult to accurately categorize, however, and as such were not considered in the study as independent variables. The relevance of the update to the user's task is also likely to affect whether it is viewed, but as the semantic nature of a task or goal is not known during real-world Web interaction, this was not considered as a factor either. Instead the study examined:

—the action used to trigger the update (none/hover/keystroke/enter key/click);
—the area of the HTML element containing the update in cm^2;
—whether the update replaced or added to the existing content on the page; and
—the duration of the update in seconds.

The study aim was to develop a predictive model of viewing behavior with a high level of external validity, so rather than systematically controlling the updates that appeared, users were set a series of tasks on live Web sites that were broadly designed to prompt them to interact with dynamic microcontent. Each update that occurred was then classified on a post hoc basis, and CHAID analysis was used to determine the parameters for categorizing each of the factors.

The hypotheses of the study can be summarized as follows.

—*H1.* The user will view updates s/he has initiated. It is not possible to predict the user's task in real-world Web interaction, but it is reasonable to assume that if the user triggers an update it is likely to be relevant to what s/he is doing, and s/he will view it.

—*H2.* The user will not view automatic updates. Previous research shows that users are not likely to view animated banners any more than static ones, so they may not view the majority of automatic updates.

—*H3.* The user is equally likely to fixate content that has been added to the page and content that replaces existing information on the page. As both types of update contain new information, it is not clear that one type of update would be more useful to the user than the other.

—*H4.* The larger an update region is, the more likely it is to be viewed. There is a greater probability that a fixation will land on a large region than a small one simply by chance, and changes to large areas may also be more salient and therefore more likely to be fixated.

—*H5.* The longer an update lasts, the more likely it is to be viewed. The longer an update remains on the screen, the greater the probability the user may "get around" to reading it, or fixate it by chance.

Section 3.1 next describes the procedure used for collecting the data, and Section 3.2 describes the process of data analysis.

3.1. Method

Participants started each task on a live Web page that was rich in dynamic micro-content, and it was anticipated that they may use some of this content in the process of completing it. To ensure external validity, people were free to approach the tasks however they wished. In addition, participants spent several minutes browsing, with no specific instructions.

Participants' eye movements were recorded as they completed the tasks. The recordings were then reviewed, and every update encountered by a participant logged and categorized. A CHAID decision tree algorithm was used to ascertain which features of a dynamic update reliably predict whether it is fixated by a user.

3.1.1. Stimuli. The full range and scope of dynamic micro-content usage is currently unknown. In the experiment, we wished to document interaction with popular, live Web sites that contained a significant amount of dynamically updating content. Participants completed tasks on, or starting from, the home pages of the following sites.

—*Kayak* (Figure 1). This is a travel site incorporating a flight search engine.
—*Yahoo!* (Figure 2). This is a Web portal that contains links to news and information from Yahoo! and other sources.
—*iGoogle* (Figure 3). This is a Web portal that contains links to news and information from Google and other sources.
—*Google Suggest* (Figure 4). This is a Google search page that provides an auto-suggest list for the term one types in the search box, based on common searches.

In all cases, interaction was documented with live Web sites, which meant that every participant experienced slightly different content.

3.1.2. Tasks. As the aim of the study was to investigate how people allocate attention to dynamic updates, the directed tasks used in the experiment were designed to prompt interaction with various items of dynamic micro-content. People were not obliged to use this content, however; they were simply asked to achieve a goal, and were free to accomplish this however they chose. Participants also completed two browsing tasks where no particular goal was specified, and they were free to interact (or not interact) with dynamic micro-content as they wished. The instructions given to participants for each task are listed next, with the dynamic micro-content they might be expected to use.

Fig. 1. The Kayak travel site homepage. (©Kayak.com.)

Fig. 2. The Yahoo! Web portal home page. (©Yahoo!)

Fig. 3.   The iGoogle Web portal home page. (©Google.)

Fig. 4.   The Google Suggest home page. (©Google.)

Fig. 5. A participant views the auto-suggest list when selecting the destination airport in task Kayak Search. (©Kayak.com.)
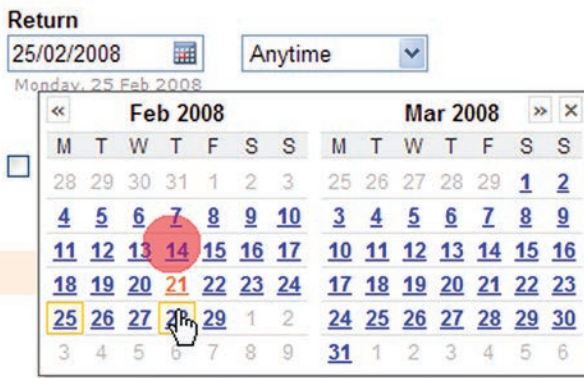
Fig. 6. A participant selects the departure date from the calendar in task Kayak Search. The calendar appears when the participant edits the date field, or clicks the symbol to the right of the date field. (©Kayak.com.)

—*Task Kayak Search*. Search for a flight for one person from Manchester to New York, leaving on <date> and returning a week later.

> When the user types in the "From" and "To" text fields, a list of suggestions (auto-suggest list) matching his/her input appears (see Figure 5[3]). A calendar appears when the focus enters the date input boxes, showing a month formatted as a table, and allowing users to click on a date to select it (see Figure 6). After submitting the search form, a new page is loaded, which presents results as they appear. A progress bar in the top left of the page indicates how much of the process has completed, below which is an updating list of the airlines that have been searched so far. Initially the center of the page contains a message telling the user that a search is in progress; after a few seconds this is replaced with a table of flights, which updates as results are received (see Figure 7).

————
[3]In Figures 5 to 15, the red dots indicate user fixations, and the lines saccades.
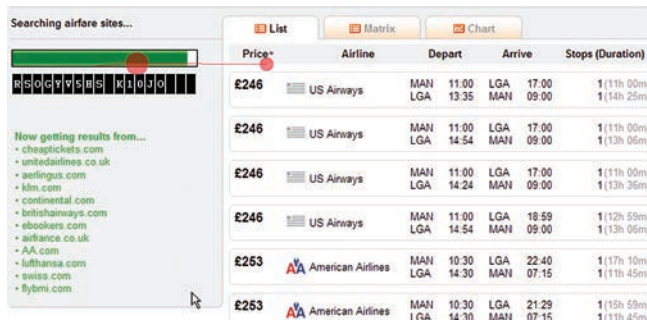
Fig. 7. A participant's gaze switches between the progress bar and the updating table of flights as the search progresses in task Kayak Search. (©Kayak.com.)
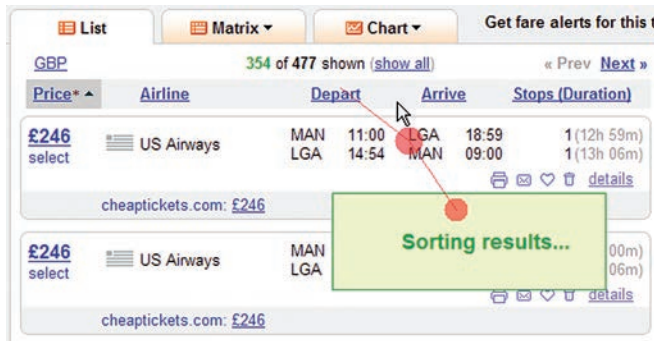
Fig. 8. A participant views the "Sorting results..." box after clicking the "Depart" column heading, when asked to name the last flight of the day in task Kayak Latest. (©Kayak.com.)

—*Task Kayak Cheapest*. What is the least expensive flight available?

> The results table may be reordered dynamically by clicking on the headings, for example, clicking the "Price" column heading reorders the rows according to price. A box containing the message "Sorting results..." appears while the rows are being reordered (see Figure 8).

—*Task Kayak Latest*. When does the last flight of the day leave?

> Users may encounter the reordering table, as in task Kayak Cheapest.

—*Task Yahoo! Headlines*. Could you browse the headlines in the center of the page, and find one about a story that interests you? When you have found a headline, could you read it out?

> The Yahoo! page has news stories arranged by category into tabs. The default view is "In the News"; clicking one of the other tabs ("World", "Local", or "Finance") replaces the headlines with a set relevant to that particular category (see Figure 9).

—*Task Yahoo! Weather*. What will the weather be like in New York today?

> The top-right area of the Yahoo! page has six buttons ("Mail", "Messenger", "Radio", "Weather", "Local", "Horoscopes"). When the mouse pointer hovers over one of these, an area below it expands to show relevant content, for example, a search box for defining a weather forecast location (see Figure 10).
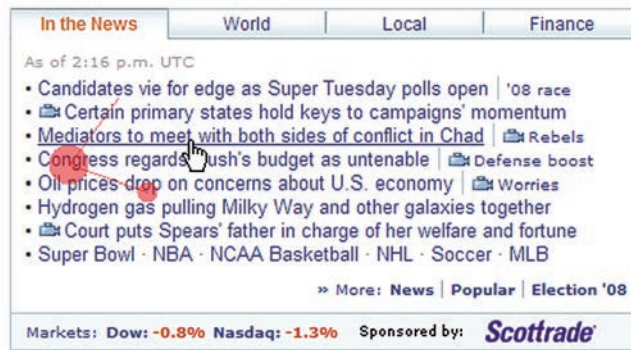
Fig. 9. A participant views the headlines in the center section of the Yahoo! page in task Yahoo! Headlines.
(©Yahoo!)
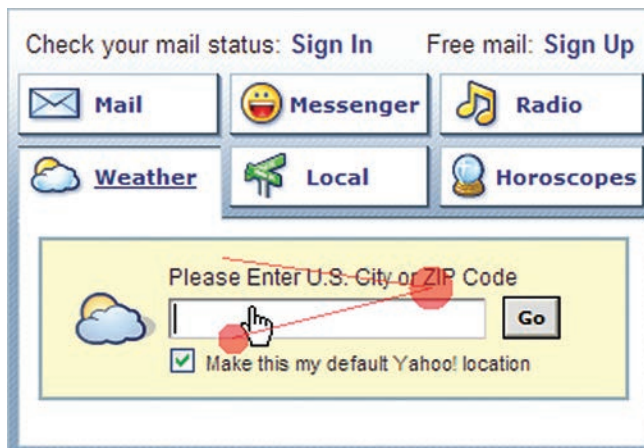
Fig. 10. A participant views the search box produced by the "Weather" button in task Yahoo! Weather.
(©Yahoo!)

—*Task Yahoo! Radio*. Could you locate a link to a radio station that interests you?

> The hover-initiated "Radio" button expands in the same way as the "Weather"
> button, revealing an image and brief description of a radio station. Arrows to the
> left and right of this description allow the user to dynamically change the infor-
> mation to scroll backwards or forwards through a set of different radio stations
> (see Figure 11). This type of content is often known as a slideshow or carousel.[4]

—*Task Yahoo! Sports-Entertainment*. Find a sports or entertainment story that inter-
ests you, and read out the first sentence.

> "Sports" and "Entertainment" stories are presented in tabs, as described in task
> Yahoo! Headlines. (See Figure 12(a)).

─────────
[4]http://developer.yahoo.com/yui/carousel/.
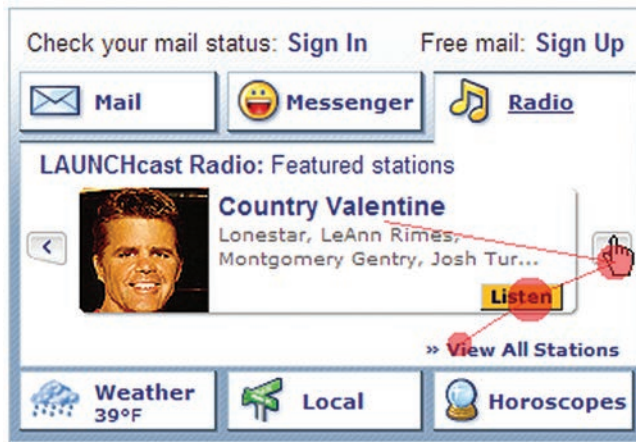
Fig. 11.   A participant views the carousel produced by the "Radio" button in task Yahoo! Radio. (©Yahoo!)

—*Task iGoogle CNN*. There is a section on the page that contains stories from CNN.com. Could you take a look at these, and read out the one you find most interesting?

> The iGoogle page has a section with three news headlines from CNN. To the left of each is a "+" sign, which, if clicked, causes an area below it to expand, revealing more detail about the story (see Figure 13).

—*Task iGoogle How To*. There is a section on the page entitled "How To of the day". Could you take a look at this, and read out the "How To" you are most interested in?

> The "How To" section has headlines which may be expanded, in the same way as the CNN headlines.

—*Task iGoogle Weather*. What will the weather be like in Manchester today and tomorrow?

> The "Weather" section on iGoogle offers a search box for users to enter a destination for weather forecasts. When the search has completed, the search box is replaced with text and images giving the forecast (Figure 14).

—*Task Google Suggest Manchester/Browse*. Search for the University of Manchester/anything you wish.

> These tasks start from the Google Suggest page, which contains a search box. Suggestions are presented below the input box in an auto-suggest list, updating (but not necessarily changing) after each keystroke (see Figure 15).

—*Task Yahoo!/iGoogle Browse*. Starting on the Yahoo!/iGoogle home page, look at whatever you wish. You will be browsing for two minutes: the investigator will tell you when the time is up.

> These tasks are unstructured and the user may encounter a wide variety of dynamic microcontent.

(a) the participant clicks the "Entertainment" tab and views the new content

(b) the participant clicks on a summary

(c) the summary replaces the main content

Fig. 12.   A participant uses the tab facility to find a story in task Yahoo! Sports-Entertainment. (©Yahoo!)

3.1.3. Participants. Seventeen male and 13 female volunteers, aged between 18 and 34, took part in the study. All participants used the Internet on a daily basis. Participants were recruited through the University of Manchester research volunteering notice board[5], and as such were either staff or students at the University, or friends of staff or students.

––––––––
[5]http://www.studentnet.manchester.ac.uk/volunteer/.

Fig. 13. A participant clicks on the "+" sign to view a summary in task iGoogle CNN. (©Google.)
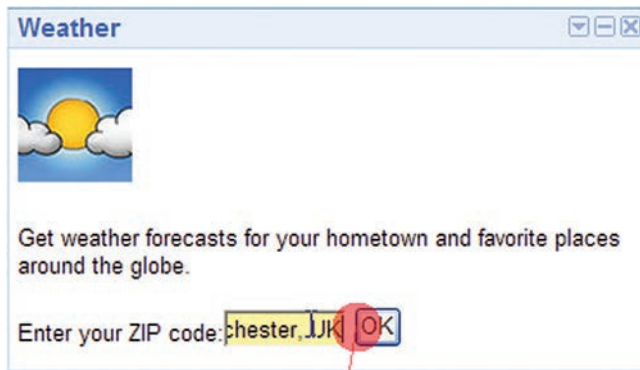
(a) the participant searches for Manchester

(b) the participant views the weather

Fig. 14. A participant uses the search box in the "Weather" section to find the forecast for Manchester in task iGoogle Weather. (©Google.)

3.1.4. Procedure. Participants sat in front of a 17" monitor with a built-in Tobii 1750 eye tracker. The page data and eye movements were recorded on a SONY VAIO VGN-FS315S laptop using Tobii Studio Professional Edition software. Participants started the experiment on an index page which contained links to the sites. They visited the Web sites in a random order, but performed all of the tasks on the site (again, in a random order, with the exception of the Kayak tasks, which were completed

Fig. 15. A participant views the auto-suggest list in task Google Suggest Manchester. (©Google.)

in sequence due to the stepwise nature of the interaction required by the site) before moving on to the next. They completed the directed tasks first, and then the browsing tasks, finally completing a short questionnaire asking how frequently they used the Web and whether they had encountered the sites used in the evaluation previously. All the experimental materials (consent form, information sheet, and questionnaire) are contained in the Web Ergonomics Lab ePrints data repository, along with a detailed description of how people completed the tasks, the Tobii Studio Project file containing the eye-tracking data, and an SPSS file containing the data entered into the analysis [Jay and Brown 2008].

3.2. Results

The questionnaire results showed that 22 participants had previously used at least one of the Web sites featured in the study (or a similar site). All participants, including the eight who had not encountered the sites before, used the Internet on a daily basis, however, and as such are highly likely to have encountered dynamic micro content previously [Brown et al. 2010]. Observation of gaze replays showed that all participants were indeed able to fully engage in their tasks using the dynamic micro-content.

Participants completed the tasks in a wide variety of ways, encountering various types of dynamic content. As the focus of this article is understanding whether people view dynamic updates as a function of their characteristics, but irrespective of user task, a detailed qualitative description of how people completed the tasks is not provided here, but can be found in the study's technical report [Jay and Brown 2008]. The sections that follow describe the procedure (Section 3.2.1) and results (Section 3.2.2) of the decision tree analysis used to determine the factors that affect whether or not a user views a dynamic update.

3.2.1. Analysis Procedure. There is no software presently available that can automatically map gaze data to updates occurring on Web pages, so instances of dynamic content were logged manually. Each Web page containing dynamic micro-content was defined as a "scene" in Tobii Studio, and each item of dynamic micro-content within that page as an "Area Of Interest" (AOI), with its boundaries determined by the outside edge of the HTML element in which the change occurred. An update was defined as any visible change[6] that occurred on a Web page, regardless of the underlying technology that caused it.

---

[6]Changes caused by manipulating the view in the browser—such as scrolling or resizing the page—were not categorized as updates.

The start and end times of every update were identified by watching the gaze replay slowed down by a factor of 5 (the lowest setting) and pausing it and recording the time stamp at the relevant points. As this was done manually, the recorded time is slightly later than the actual time (comparing the investigator's response to events such as mouse clicks with the time recorded in the Tobii log file suggests the delay is in the region of 100–200 ms), but we have not applied any adjustment as the precise fixation time is not crucial to the current study, and we did not wish to add further noise to the data. The start time was logged as the instant at which the update started to occur visibly. An update was said to end due to one of the following occurrences: the user navigated to a new page; the user scrolled the page; the dynamic micro-content was otherwise moved or replaced (usually by another update). A script was used to compare the update instances with the gaze data log files, and calculate the time and duration of fixations for each update.

Whether the content was viewed (yes/no) constituted the dependent variable in the analysis. A single fixation occurring at any point during the lifetime of the content resulted in a "yes" classification, as it was assumed this would provide sufficient information to either dismiss the content or focus on it further. The independent variables were: the action used to trigger the update (none/hover/keystroke/enter key/click); the area of the HTML element containing the update in cm^2; whether the update replaced or added to the existing content on the page; and the duration of the update in seconds. As participants experienced varying types and numbers of dynamic updates, and it is possible that some people are more likely to view updates than others, an identifier for the participant who encountered the update was also entered as an independent variable. Web site was not used as a predictor, as sites were live and their content and layout varied over time. Action, whether or not content was replaced or added to, and participant identifier were entered as categorical variables and duration and area were entered as continuous variables.

To confirm the existence of a relationship between the characteristics of the update, and the likelihood that it was viewed, the data were analyzed in SPSS 15.0 using the CHi-squared Automatic Interaction Detector (CHAID) classification tree procedure. CHAID was chosen for analysis of the dataset due to its ability to automatically detect and clearly illustrate interactions between variables. It takes both categorical and continuous variables as input, and produces a tree illustrating the relative contribution of each variable in predicting a particular outcome. While categorical variables are treated as discrete groups, continuous variables are partitioned as part of the CHAID analysis process into optimally predictive categories. This is useful for defining the parameters at which to classify an update as "viewed" or "not viewed".[7]

The procedure was first carried out using filtered eye-tracking data that only included fixations occurring within a 40 degree radius for at least 100 ms. This was chosen as a threshold because people are known to be able to formulate an impression of a scene in this time [Potter and Levy 1969] and it is recommended by Tobii as the appropriate threshold to use when people are viewing mixed content (i.e., text and images)[8]. As responses to dynamic updates on Web pages have not previously been studied and it is therefore unclear precisely where the parameters for a "meaningful" fixation should be set, the analysis was repeated with the raw data, which includes all fixations that are greater than 20 ms. This process, reported in Appendix A, confirmed that the filter applied to the data was appropriate. The results of the analysis conducted with the filtered data are reported in the following section.

---

[7]The raw data used for the analysis are available at http://wel-eprints.cs.manchester.ac.uk/49/.

[8]http://www.slideshare.net/AcuityETS/tobii-studio-quick-start-guide.

Table I. The Instances of Dynamic Microcontent Encountered by Participants
during the study, Categorized according to Content Type and Action that
Triggered Them

Content Type	Keystroke	Click	Enter	Hover	None
Additional form field	0	1	0	0	0
Animated advertisement	0	0	0	0	272
Animation	0	6	11	0	40
Auto-suggest list	467	0	0	0	0
Button text change	0	28	1	0	0
Calendar	0	51	0	0	0
Carousel	0	16	0	0	0
Description	0	0	0	5	0
Drop-down list	0	6	0	0	0
Drop-down menu	0	17	0	13	0
Expansion box	0	6	0	0	1
Interactive advertisement	0	1	0	0	0
List box	0	9	12	0	0
Message box	0	1	0	0	0
Interactive content	0	0	0	14	0
Pop-up box	32	33	0	12	0
Pop-up box/animation	0	14	1	0	0
Progress bar	0	31	0	0	0
Reordered table	0	43	0	0	0
Search results	0	17	14	1	0
Slideshow	0	0	0	0	14
Status text	0	1	0	0	0
Tabbed content	0	50	0	0	0
Text field selection	26	100	10	0	0
Ticker	0	0	0	0	37
Updating list	0	59	0	0	0
Video	0	11	0	0	0
Video replacement	2	0	0	0	0
Total	**527**	**501**	**49**	**45**	**364**

3.2.2. Analysis Results. Specifically, 1486 updates were documented, of which 364 occurred automatically and 1122 occurred in response to a user action. Automatic updates included animations, movies, and tickers. User-initiated updates included auto-suggest lists, tabbed content, tables, movies, progress bars, pop-up boxes, and drop-down menus. Some types of dynamic micro-content, such as slideshows, were both automatic and user-initiated. Table I categorizes the instances of dynamic micro-content that appeared on the Web sites, as a function of how their type and how they were initiated.

The results of the analysis are shown in Figure 16. The model is able to predict whether or not an update is viewed with 78% accuracy (see Table II). Participant was not a significant predictor in the model, indicating that the likelihood that a dynamic update was viewed did not vary according to the individual who encountered it. Whether the update added to or replaced content on the page did not affect whether it was viewed either[9].

---

[9]Figure 16 shows the model providing the best fit. Models considering only action and duration, or action and area, also predicted whether updates were viewed, albeit less accurately. These models are shown in Appendices B and C, respectively.
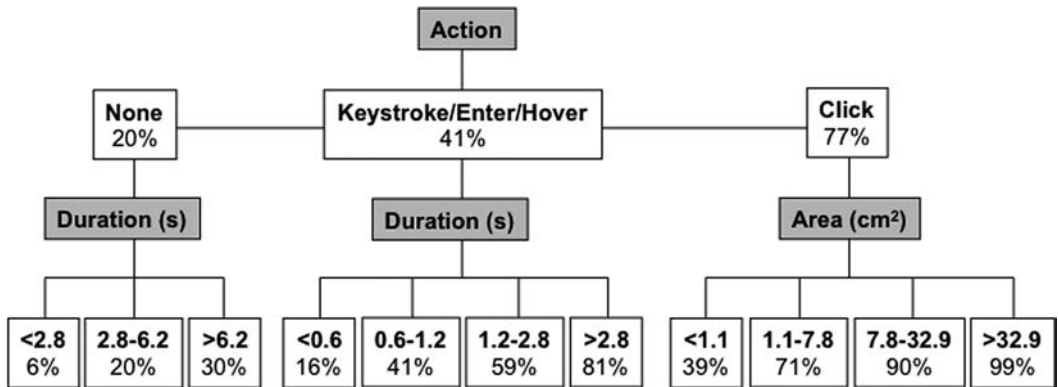
Fig. 16.   The CHAID model showing the likelihood an update was viewed as a function of the action that triggered it, its duration in seconds, and its area in cm^2.

Table II. Classification Table for the CHAID Model

	Observed	(Predicted)	% Correct
Viewed	507	(710)	84%
Not Viewed	653	(776)	71%
Overall	1160	(1486)	78%

The most important predictor was the way that the update was initiated. In all, 77% of updates activated by a click were viewed, followed by 41% of hover-, enter-key-, or keystroke-activated updates (which were entered into the analysis separately, but were not significantly different) and 20% of automatic updates ($\chi^2(2, n = 1486) = 292$, $p < 0.001$).

The duration of an automatic update significantly affected the likelihood it would be viewed. Only 6% of updates shorter than 2.8 sec were viewed, compared with 20% of updates that lasted between 2.8 and 6.2 sec, and 30% of updates longer than 6.2 sec ($\chi^2(2, n = 363) = 23.1$, $p < 0.001$). Duration also affected how likely hover-, enter- key-, or keystroke-activated updates were to be viewed. Only 16% of updates shorter than 0.6 sec were viewed, followed by 41% of those between 0.6 and 1.2 sec, 59% of those between 1.2 and 2.8 sec, and 81% of those longer than 2.8 sec ($\chi^2(3, n = 619) = 157$, $p < 0.001$).

Size, rather than duration, affected the likelihood that click-activated updates would be viewed. Specifically, 39% of those smaller than 1.1 cm^2 were viewed, 71% of those between 1.1 and 7.8 cm^2, 90% of those between 7.8 and 32.9 cm^2, and 99% of those greater than 32.9 cm^2 ($\chi^2(3, n = 504) = 133.4$, $p < 0.001$).

3.3. Discussion of Experiment 1

The experiment was designed to explore how people allocate attention to dynamic updates that appear in the course of normal Web use. Participants completed a number of high-level tasks on live Web sites, and experienced a large number of updates (50 per participant on average). These updates were categorized according to key and measurable features, and a CHAID decision tree procedure was used to ascertain whether it is possible to predict whether an update is viewed as a function of these features.

The resulting model is able to predict whether or not an update is viewed with 78% accuracy. How an update is initiated is by far the best predictor of whether it is viewed: 77% of click-activated updates, 41% of hover-/enter-key-/keystroke-activated updates, and just 20% of automatic updates were viewed overall. Area is the best predictor
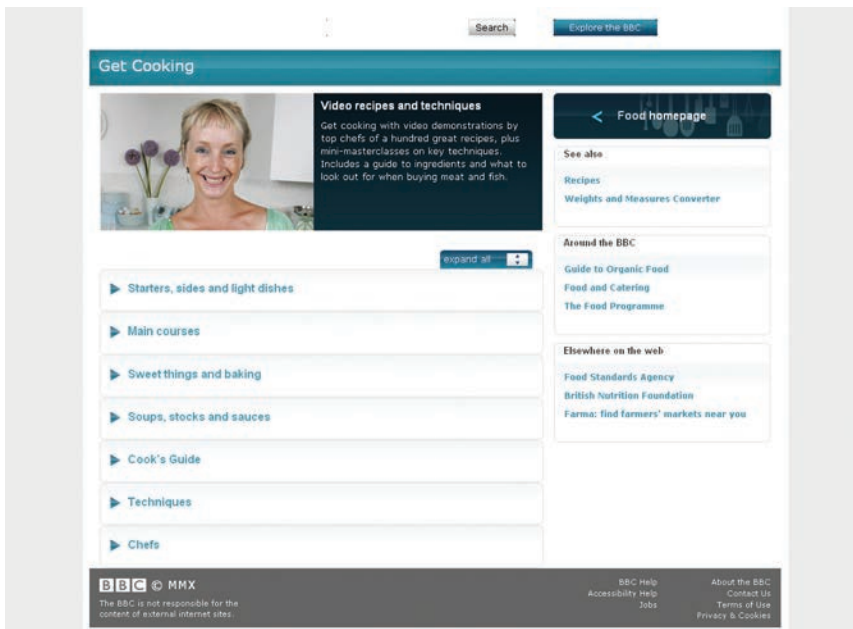
Fig. 17.   The BBC Get Cooking home page. (©BBC.)

of whether click-activated updates are viewed, and duration is the best predictor for automatic-, keystroke-, enter-key-, or hover-activated updates.

Although the model appears to provide a good fit, additional data are necessary to validate it. The original input data were collected under circumstances designed to produce naturalistic Web interaction, to ensure the model had external validity. As the participants were still instructed to perform particular tasks, however, it cannot be assumed the model will apply to other types of Web use. To confirm the model's validity, it is necessary to collect further data from different people using different Web sites in different ways.

4. EXPERIMENT 2: EVALUATING THE MODEL

A second study was carried out approximately two years later, specifically to collect data with which to validate the model.

4.1. Method

Experiment 2 was designed to determine the external validity of the model. Data were collected and analyzed as described in Section 3.1.

4.1.1. Participants. Five female and seven male participants aged between 18 and 34 were recruited through the University of Manchester research volunteering notice board. All participants used the Internet on a daily basis. None had participated in experiment 1.

4.1.2. Procedure. As the aim of the study was to validate the model with data that were as close to "normal" Web use as possible, participants were not required to complete any particular tasks. They spent three minutes on each of the following sites, chosen because they had a large amount of dynamic content: BBC Get Cooking[10] (see Figure 17);

---

[10]Now offline.

Fig. 18.   The GMTV home page. (©ITV plc.)
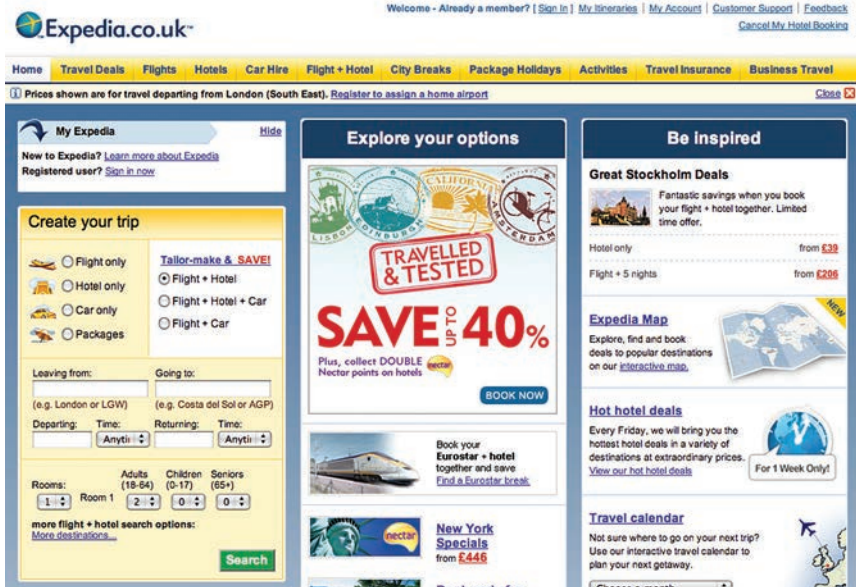
Fig. 19.   The Expedia home page. (©Expedia, Inc.)

GMTV[11] (see Figure 18), and Expedia[12] (see Figure 19). They also spent five minutes
visiting whichever sites they wished, starting from the Google home page. Participants

---

[11]cms.gm.tv.

[12]www.expedia.com.

Table III. The Instances of Dynamic Micro-content Encountered by Participants during the Validation Study, Categorized according to Content Type and Action that Triggered Them

| Content type	Keystroke	Click	Enter	Hover	None
Animated advertisement	0	0	0	0	200
Animation	0	0	0	0	2
Auto-suggest list	93	0	0	0	0
Calendar	0	41	0	0	0
Slideshow	0	0	0	28	1
Drop-down menu	0	23	0	51	0
Expansion box	0	25	0	0	0
Pop-up box	0	1	0	5	1
Status text	0	2	0	0	1
Carousel	0	16	0	0	0
Progress bar	0	9	0	0	0
Search results	0	2	0	0	0
Slideshow advertisement	0	0	0	0	14
Tabbed content	0	1	0	0	0
Text field selection	3	34	0	0	0
Ticker	0	0	0	0	28
Video	0	4	0	0	0
Total	**96**	**158**	**0**	**84**	**247**

were free to browse and use the sites however they wanted. The order in which they visited the sites varied according to a Latin square.

4.2. Results

Inadequate data was obtained from two of the participants (both were wearing glasses, and it appeared that the lens type prevented the eye tracker from adequately picking up their gaze), so the results are described for data from 10 participants (four female and six male).

The data were logged and categorized as specified in Section 3.2.2, and were entered into the CHAID procedure as test data with which to validate the existing model. Data were recorded for 247 automatic and 338 user-activated updates. Table III provides a summary of the main types of update participants experienced.

The validation data fit the model slightly better than the training data, with whether or not 83% of updates were fixated being accurately predicted, compared to 78% (see Figure 20 and Table IV). Fewer updates of all types were fixated—11% of automatic, 38% of hover-/enter-key-/keystroke-activated, and 68% of click-activated ($\chi^2(2, n = 585) = 292$, $p < 0.001$)—but the relative proportions of updates fixated in each category are very similar to those seen with the training data. Of the automatic updates, 7% of those shorter than 2.8s, 9% of those between 2.8 and 6.2 sec, and 19% of those longer than 6.2s were viewed ($\chi^2(2, n = 246) = 23.1$, $p < 0.001$). Of click-activated updates, 25% of those smaller than 1.1 cm^2, 83% of those between 1.1 and 7.8 cm^2, 79% of those between 7.8 and 32.9 cm^2, and 100% of those larger than 32.9 cm^2 were viewed ($\chi^2(3, n = 158) = 162.7$, $p < 0.001$). Of hover-/enter-key-/keystroke-activated updates, 20% of those less than 0.6 s, 35% of those between 0.6 and 1.2 s, 78% of those between 1.2 and 2.8 sec, and 64% of those longer than 2.8s were fixated ($\chi^2(3, n = 181) = 157$, $p < 0.001$).

4.3. Discussion of Experiment 2

The model provides an excellent fit to the validation data, accurately predicting whether or not 83% of updates are viewed. The model may be a better fit to this
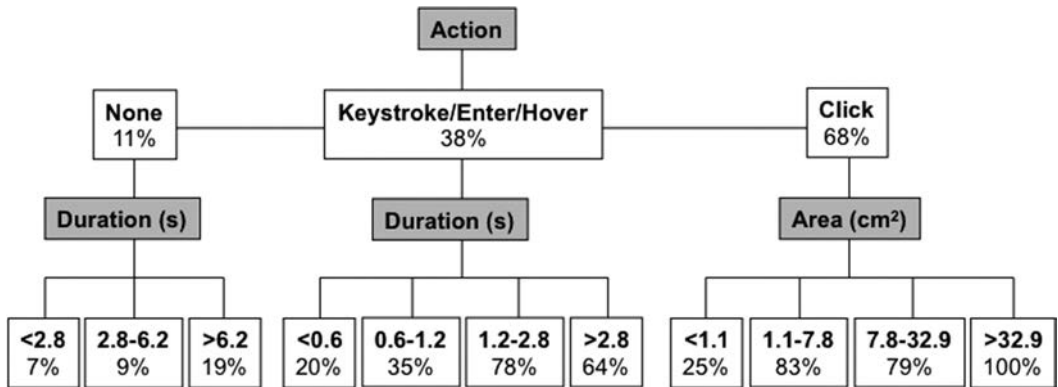
Fig. 20. The CHAID model created with the validation data, showing the likelihood an update was viewed as a function of the action that triggered it, its duration in seconds, and its area in cm^2.

Table IV. Classification Table for the Validation Data
CHAID Model

	Observed	(Predicted)	% Correct
Viewed	128	(202)	63%
Not Viewed	358	(383)	94%
Overall	486	(585)	83%

dataset because it is more likely to correctly predict when an update is not viewed than when it is, and the validation data contained a greater proportion of automatic updates. It should also be noted that participants were from the same demographic group (young, English-speaking adults recruited through the University research volunteer notice board) in both studies, and data from other participants (for example, older users) may produce different results.

5. DISCUSSION

Previous studies have shown that updating Web content can distract or aid users performing particular tasks on artificially constructed Web pages, but so far none has investigated how users respond to the rich and complex dynamic micro-content frequently encountered during normal Web use. This investigation had the aim of understanding how people allocate attention to dynamic updates on real Web pages. Although there are potentially many factors that may influence this process, the results of a CHAID analysis show that it is possible to predict the likelihood that someone will view a dynamic update with an accuracy of around 80%, by considering a few measurable characteristics of the update.

The best predictor of whether an update is viewed is the way in which it is triggered, with participants most likely to view click-activated updates and least likely to view automatic updates. *H2* is supported by the data, as only 20% of automatic updates are fixated. *H1* is only partially supported, however. While people view 77% of updates initiated by a click, they view only 41% of those activated by a mouse hover or keystroke. Clicking is an accepted means of requesting information and previous research shows that eye and mouse movements are often coordinated [Rodden et al. 2008]. For example, when a user clicks the "Go" button after entering a location in the weather section of the Yahoo! Web site, he or she is specifically asking for a summary of the weather, and is therefore likely to view the resulting content. An interesting question with regard to

click-activated updates may be why participants do not fixate all of them. The answer may lie in the fact that although many such updates will be crucial to the task at hand, a number will not be. The weather summary is a key item of information essential to the task of finding out the weather forecast, and as such is anticipated by the user. It is not the only update activated by the click, however: a small rotating arrow icon also appears on the right of the text field after the request has been submitted. This indicates that the request has gone to the server, and a response should be expected shortly. It informs the user that something is happening, but it does not provide information that is essential to the task. The fact that the proportion of click-activated updates viewed increases with size indicates that salience plays a role in whether such updates are fixated: only 39% of those smaller than 1.1 cm^2 are viewed, compared to 71% of those greater than 1.1 cm^2.

Why are the majority of mouse-hover-, enter-key-, or keystroke-activated updates not viewed, even when they have been initiated by the user? In the case of hover-activated updates, it may be because they were not initiated deliberately. Whereas a click is a decisive action, hovering occurs continually, and as such hover-activated updates sometimes appear when the user does not intend them to. If this happens while the mouse is moving from one location to another, the relevant information (such as a menu or "tool tip") will appear only fleetingly. If it is unexpected and irrelevant to the user's current activity, such an update is unlikely to be viewed. Longer updates are more likely to result from an intentional hover: the user is requesting the information, and as such is more likely to view it. This explanation is supported by the model: only 16% of hover- and keystroke-activated updates shorter than 0.6s are viewed, but this proportion rises considerably as duration increases, to 41% of those between 0.6 and 1.2s, 59% of those between 1.2 and 2.8s, and 81% of those longer than 2.8s.

This relationship between duration and user intention may apply to keystroke-activated updates too, which occur almost entirely during the use of auto-suggest lists. Although participants do not look at every change to the list (it is usually updated with each keystroke and many people look at the keyboard to type), suggestion lists are glanced at—and selected from—by the majority of participants at some point during their lifetime [Brown et al. 2009]. The longer duration of a keystroke-activated update may therefore arise as the user pauses typing in order to view the list.

Although pressing the enter key is in one sense a keystroke, it performs a different function to other keystrokes: activating a control rather than entering text. As it often has the same role as a click, it is perhaps surprising that they are not classified in the same way by the model. While the click requires visual attention, the enter keystroke does not. Updates caused by pressing a button often appear close to that button, and this may be why they are more likely to be viewed when the user has clicked the colocated control, rather than pressing enter.

H3. is supported by the data, as whether the update replaced or added to the content on the page did not predict whether it was fixated. *H4* and *H5* are only partially supported, however, as duration is not a significant predictor of whether a click-activated update is fixated, and size does not affect the likelihood that automatic-, hover-, enter-key-, or keystroke-activated updates are fixated. The reason that the duration of click-activated updates does not affect the likelihood they are viewed may be because they are generally relevant to the current activity in the user's workflow, and are therefore attended to quickly if at all. That size does not predict the likelihood that automatic-, keystroke-, or hover-activated updates are viewed is interesting, as this provides evidence that larger—and potentially more salient—updates are no better at attracting attention than smaller ones when they are irrelevant to the user's task.

6. APPLICATIONS OF THE MODEL

The DUV model has potential applications in a number of domains. It has value from a design perspective; showing, for example, that automatic updates never have more than a 30% chance of being viewed, however long they remain on the screen, and that while salience (as indicated by size) affects the likelihood that click-activated, task-relevant updates are viewed, it has no impact for any other type of update. The model also describes how one could expect the user to behave on the "average" Web page containing dynamic content. This would be useful in a usability context as a benchmark for determining whether people are viewing updates more or less frequently than expected.

Shortly, two further contributions are discussed in detail: how the DUV model can be used to guide the design of screen-reading assistive technology; and how it adds to the domain of modeling Web-based interaction.

6.1. User Agent Design Recommendations

The DUV model provides a means of predicting how likely an update is to be viewed, and illuminating how users interact with different categories of dynamic micro-content.

A greater understanding of what causes a user to attend to—or ignore—an update would be of particular value when dealing with situations where bandwidth is limited, and decisions must be made about how and when to present updating information to the user. This issue arises with small-screen or mobile devices, where only part of a Web page can be viewed at one time, but is perhaps most pertinent to screen-reading assistive technology devices (used by people with visual impairments), which present Web pages in a serial audio stream. As only a single item of information can be presented at once, knowing whether to relay an update—potentially interrupting the information flow and causing confusion—is crucial. The following recommendations show how the model may be applied to improve screen-reader presentation of dynamic updates, but several would also apply to presentation via mobile or screen magnification devices.

Often, Web pages contain competing streams of content, where updates occur either simultaneously or in quick succession. In this situation, the challenge is to present the right content to the user at the right time. According to the model, participants view the majority of click-activated updates larger than 1.1 cm^2, so presenting these should be a priority. Although those smaller than 1.1 cm^2 are only fixated 39% of the time, they are still viewed more frequently than even the longest automatic updates. As it is likely they bear some relevance to the task, they may be useful to the user. The appropriate rule may therefore be the following.

> Attempt to present all click-activated updates, prioritizing the largest.

The longer a hover update is, the more likely it is to be fixated. This may be because shorter updates are activated by mistake (due to the mouse passing over it while the user completes another task); the longer duration of a hover-activated update is a signal the user intends to view the information it contains. Similarly, the longer duration of a keystroke-activated update occurs when the user pauses typing to view an auto-suggest list. The majority of hover- or keystroke-activated updates lasting longer than 1.2 sec are viewed by users, so this could be viewed as the threshold after which to present them. Waiting to see whether the updates reach this length before presenting them may cause an irritating delay, however. A better rule for this type of update may therefore be the following.

> Present hover-activated updates when there is a pause in mouse movement[13], and keystroke-activated updates when there is a pause in typing.

What exactly constitutes a pause is open to debate, but as an example, if a user with visual impairments is typing in a search field, the screen reader should start reading suggestions whenever it has finished echoing the last key the user has pressed.

Automatic updates are only ever viewed the minority of the time, regardless of how big they are or how long they last. For this reason, presenting them should not be a priority. The fact that they are fixated some of the time means they are not completely ignored by users, however, and there may be times when they are interesting or useful. For this reason, it may be useful to notify users that automatic updates are occurring. A rule for automatic updates may therefore be as follows.

> Alert users to the occurrence of automatic updates, so they can view or listen to them if they wish.

On some Web sites (such as those containing many animated advertisements) automatic updates occur very frequently, and in this case even a discrete notification could be annoying. Automatic updates are rarely directly relevant to the user's task, so he or she may wish to ignore them completely, at the small risk of missing information of interest. An additional rule could therefore be as follows.

> Provide an option allowing users to ignore all automatic updates.

These recommendations are not exhaustive, but demonstrate how the model may be applied to improve user agent design. At present, screen readers deal very badly with dynamic updates—often not informing the user they have occurred—so an applied understanding of how sighted users interact with dynamic micro-content could significantly improve the available technology.

6.2. Understanding and Predicting Web Interaction

There are a number of existing models that explain and predict how users interact with the Web, focusing on how they make choices about navigation [tat Fu and Pirolli 2007; Miller and Remington 2004; Kitajima et al. 2005; Teo and John 2010], and how they search for information within a page [Halverson and Hornof 2007].

The DUV model contributes to our understanding of and ability to predict user behavior on the Web in a number of ways. First, it considers dynamic displays, predicting whether people will view changes to content rather than estimating which link they will click or object they will fixate within a static display (phenomena that have been investigated in previous research). To the authors' knowledge, this is the first time interaction with dynamic Web content has been modeled in detail. Second, it contributes by showing it is not necessary to directly consider the semantics of the user's task or goal to make an accurate prediction about his or her behavior. Modeling user behavior as a function of a known goal is a powerful way of improving our understanding of how people make decisions on the Web, but a limitation of this approach is that it is impossible to precisely determine a user's goal during real-life Web interaction. Here we show that to predict whether users will view a dynamic update it is not necessary to explicitly understand their goal a priori.

---

[13]Not all users with visual impairments use a mouse (although many do with the ZoomText screen reader), but screen readers such as JAWS have a hover keystroke function instead.

Although the values of the factors used in the model can be measured objectively, the DUV model does not make its prediction completely blind to the user's task. The manner in which an update is activated is a property of the Web content, but it reflects the user's intention; for example, if content is triggered by a click, it is likely to have been requested by the user and is therefore important to his or her goal. In the case of keystroke- and hover-activated updates, it is less clear whether they are important to the user, which is why it is necessary to consider duration. The length of time these updates stay on the screen indicates how relevant they are to the user's task: if they appear for only a short time, they are likely to be irrelevant; if they persist for a longer period, it is likely they are being used. Together, these measures serve as a proxy for the user's goal that is sufficient to enable a prediction of his/her subsequent viewing behavior.

7. CONCLUSIONS AND FUTURE WORK

The DUV model represents the first successful attempt to model the likelihood that users will view dynamic micro content on the Web. It is able to predict, with around 80% accuracy, whether someone will view a dynamic update as a function of the action used to trigger it, its duration, and its size. As it was constructed and validated using data from live Web sites, it has a high level of external validity.

The model has potential applications in a number of domains, including Web design, user agent design, and usability. One of the most significant contributions of the investigation is that it shows that data describing how the user interacts with the Web page can serve as a proxy for task, such that behavior can be predicted without knowledge of any semantic details of the user's goal. The DUV model demonstrates this for the particular case of viewing dynamic updates; it would be interesting to see whether this applies to other forms of Web interaction as well.

The model alone does not provide a complete picture of how people interact with dynamic micro-content. As a person may focus on something but not register it in conscious attention, or may attend to something covertly without looking at it directly [Beanland and Pammer 2010], it may not always reflect how users allocate their attention—as opposed to their gaze—to dynamic content. It is also important to consider that it describes how people use dynamic micro-content at present, and that as the Web evolves, the model may need to evolve as well. The validation data were collected two years after the initial data but still fit the model well, indicating potential longevity, but only time will tell whether it continues to predict behavior accurately. Google now updates not only suggestions, but search results with every keystroke. According to the model, duration, rather than area, is key to determining whether keystroke updates are fixated; it would be interesting to see whether this rule applies to Google's new functionality.

In spite of these issues, the work documented here is significant. It represents the first study to investigate whether people view dynamic updates on live Web sites, and the first successful attempt to predict Web interaction without directly considering the user's goal. The result is a comprehensive and accurate model with a high degree of external validity which has a number of applications.

APPENDIXES

A. RAW DATA ANALYSIS

The model constructed from the raw gaze data (which considered an update to have been viewed if it received a fixation at least 20ms in length) is shown in Figure 22 and Table V. It is slightly weaker than the filtered data model (accurately predicting whether 76%, rather than 78% of updates were fixated), but qualitatively very similar:

Table V. Classification Table for the Raw Gaze Data
CHAID Model

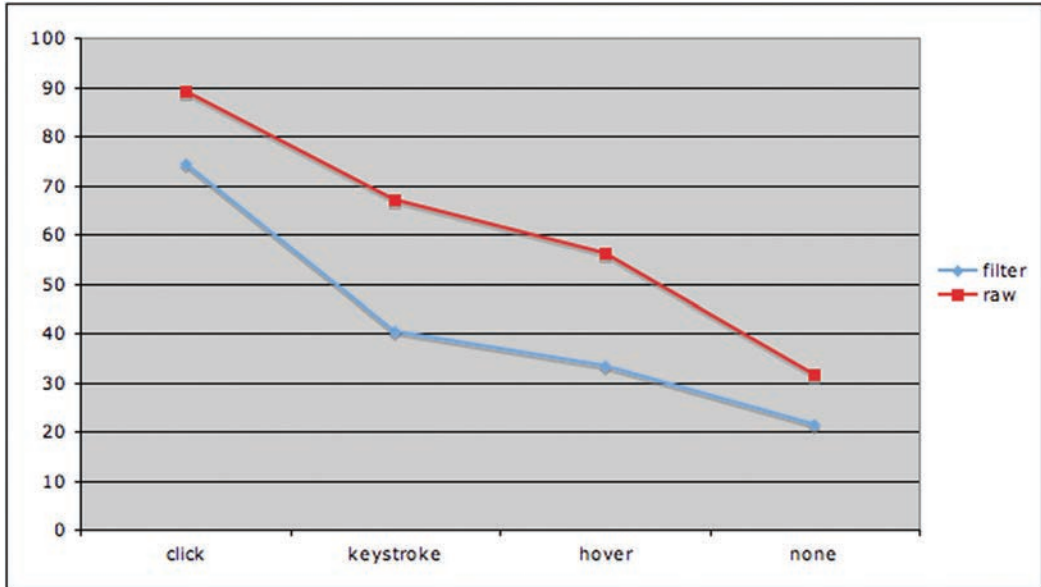|	Observed	(Predicted)	% Correct
Viewed	732	(972)	75%
Not Viewed	394	(514)	77%
Overall	1126	(1486)	76%

Fig. 21. The percentage of updates viewed according to how they were initiated (click, hover, keystroke or automatic) for the raw and filtered data.

the strongest predictor was how the update was initiated, with 90% of click-activated, 66% of hover-, enter-key-, or keystroke-activated, and 30% of automatic updates viewed ($\chi^2(2, n = 1486) = 342, p < 0.001$). Duration affected the likelihood that automatic and hover/enter-key/keystroke-activated updates would be viewed: 15% of automatic updates shorter than 2.8 sec were viewed, 31% of those between 2.8 and 9.6 sec, and 44% of those longer than 9.6 sec (($\chi^2(2, n = 363) = 19.5, p < 0.001$); of the hover/enter-key/keystrokes updates, 48% of those less than 0.6 sec, 64% of those between 0.6 and 1.2 sec, and 88% of those longer than 1.2 sec were viewed ($\chi^2(2, n = 619) = 91, p < 0.001$). Again, the likelihood that click-activated updates were viewed varied according to size: 76% of those smaller than 1.1 cm^2, 87% of those between 1.1 and 7.8 cm^2, and 98% of those larger than 7.8 cm^2 were viewed ($\chi^2(2, n = 504) = 40.7, p < 0.001$). The only qualitative difference between this and the filtered data model is the presence of a further branch of the tree for automatic updates, which shows that for those lasting between 2.8 and 9.6 sec, 18% of those smaller than 2.2 cm^2 and 38% of those larger than 2.2 cm^2 are viewed ($\chi^2(1, n = 166) = 7.3, p < 0.05$).

   Figure 21 shows the percentage of updates fixated as a function of how they were initiated. A GLM repeated measures procedure with the Greenhouse-Geisser adjustment applied shows that as one might expect, the percentage is significantly higher when the threshold for updates is set at 20ms (the raw data) than at 100ms ($F_{1,17} = 71.4$, $p < 0.001$). There also appears to be an interaction effect, which shows that the difference between the proportion of updates fixated according to the raw and filtered
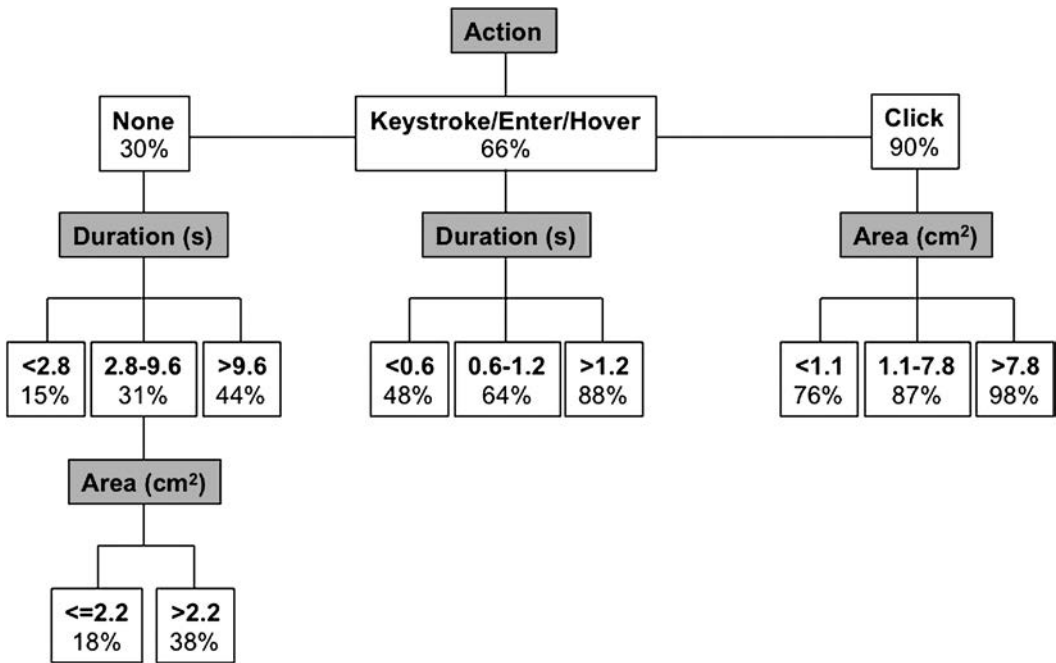
Fig. 22. The CHAID model created with the raw eye-tracking data, showing the likelihood an update was viewed as a function of the action that triggered it, its duration in seconds, and its area in cm^2.

Table VI. Percentage of Updates Fixated when Considering the Filtered Data and the Raw Data, and the Difference between the Two

Trigger	Click	Keystroke	Hover	None
Filtered data %	74	40	33	21
Raw data %	89	67	56	32
% difference	15	27	23	10

data measures varies according to how they are initiated: while an additional 10% of automatic updates are fixated when considering the raw rather than the filtered data, this rises to 15% for click-activated updates, 23% for hover-activated updates, and 27% for keystroke-activated updates (see Table VI). This is not found to be significant at the 5% level ($F_{1.8,30.3} = 3.2$, p < 0.06), but it is worth running this analysis again with the data collected for validation, to ensure a Type-II error has not occurred.

The raw validation data fit the model slightly better than the training data, with whether or not 79% of updates were fixated being correctly predicted, compared to 76% (see Figure 23). As with the filtered data, the relative proportions of updates fixated in the validation and training data were very similar. At the first level of the tree, 18% of automatic, 57% of keystroke/enter-key/hover-activated, and 89% of click-activated updates were fixated ($\chi^2(2, n = 585) = 342$, p < 0.001). Of the automatic updates, 11% of those shorter than 2.8s, 18% of those between 2.8 and 9.6s and 48% of those longer than 9.6s were fixated ($\chi^2(2, n = 246) = 19.5$, p < 0.005). Of those automatic updates that were between 2.8 and 9.6s, 14% of those smaller and 25% of those larger than 2.2 cm^2 were fixated ($\chi^2(1, n = 25) = 7.3$, p < 0.05). Moreover, 40% of keystroke/enter-key/hover-activated updates shorter than 0.6s were viewed, 63% of those between 0.6 and 1.2s, and 81% of those longer than 1.2 s ($\chi^2(2, n = 181) = 80.9$, p < 0.001). Of the click-activated updates, 71% of those smaller than 1.1 cm^2, 92% of those between
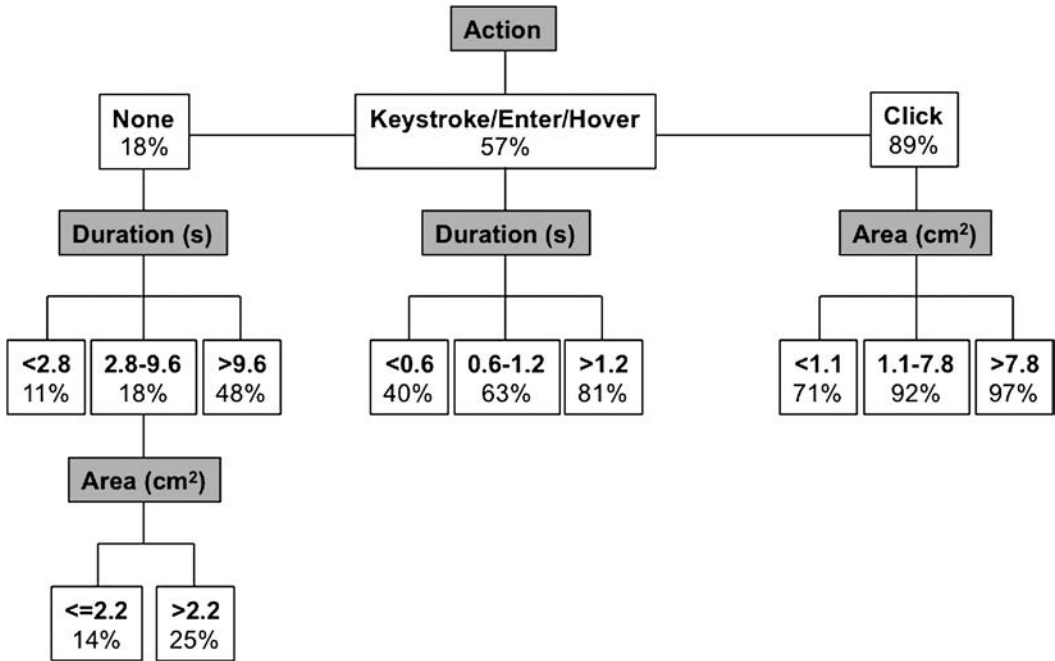
Fig. 23. The CHAID model created with the raw validation eye-tracking data, showing the likelihood an update was viewed as a function of the action that triggered it, its duration in seconds, and its area in cm^2.

Table VII. Classification Table for the Raw Validation
Gaze Data CHAID Model

	Observed	(Predicted)	% Correct
Viewed	211	(288)	73%
Not Viewed	250	(297)	84%
Overall	461	(585)	79%

Table VIII. Classification Table for the Action and Duration
Filtered Gaze Data CHAID model

	Observed	(Predicted)	% Correct
Viewed	546	(710)	77%
Not Viewed	586	(776)	76%
Overall	1132	(1486)	76%

1.1 and 7.8 cm^2, and 97% of those greater than 7.8 cm^2 were fixated ($\chi^2(2, n = 158) = 61.6$, $p < 0.001$).

Figure 24 shows the percentage of validation data updates fixated as a function of how they were initiated. Again the percentage is significantly higher when the threshold for updates is set at 20ms (the raw data) than at 100ms ($F_{1,5} = 27.2$, $p < 0.005$), but there is no significant trigger by measure interaction ($p > 0.2$). This indicates that including fixations shorter than 100ms serves to add noise to the model, rather than increase its accuracy. Section 5 therefore discusses the filtered rather than the raw data model.

B. MODEL FOR ACTION AND DURATION ONLY

The model considering action and duration only is able to predict whether an update is fixated with an accuracy of 76% in the case of the training data (see Figure 25 and
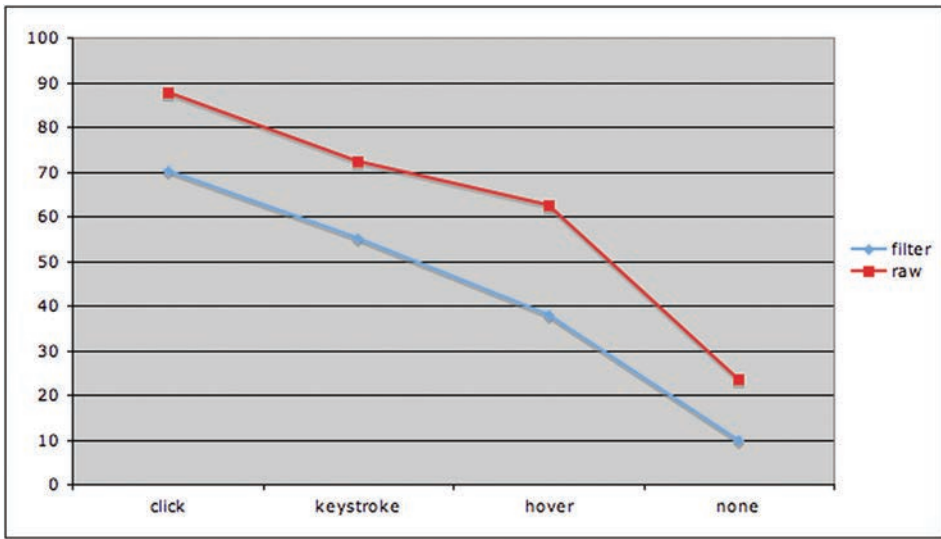
Fig. 24. The percentage of updates viewed according to how they were initiated (click, hover, keystroke or automatic) for the raw and filtered validation data.
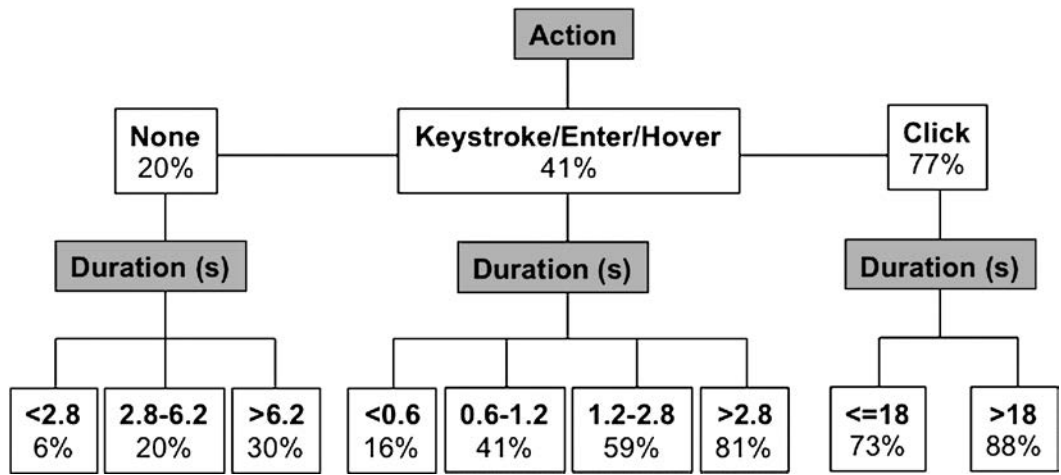

Fig. 25. The CHAID model created with the filtered eye-tracking data, showing the likelihood an update was viewed as a function of the action that triggered it and its duration in seconds.

Table IX. Classification Table for the Action and Duration
Filtered Gaze Data CHAID Model

	Observed	(Predicted)	% Correct
Viewed	140	(202)	69%
Not Viewed	322	(383)	84%
Overall	462	(585)	79%

Table IX) and 79% in the case of the validation data (see Figure 26 and Table X). The only difference between this model and that reported in the main article is the effect of duration for click updates; in the case of the training data, 73% of those longer than 18 sec are fixated compared to 88% of those longer ($\chi^2(1, n = 386) = 9.9$, $p = 0.015$). In
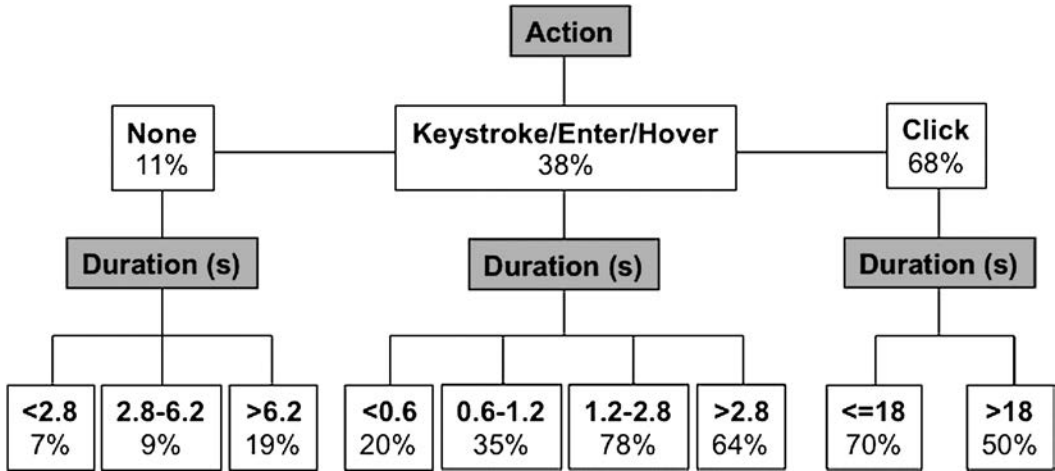
Fig. 26.   The CHAID model created with the filtered validation eye-tracking data, showing the likelihood an update was viewed as a function of the action that triggered it and its duration in seconds.
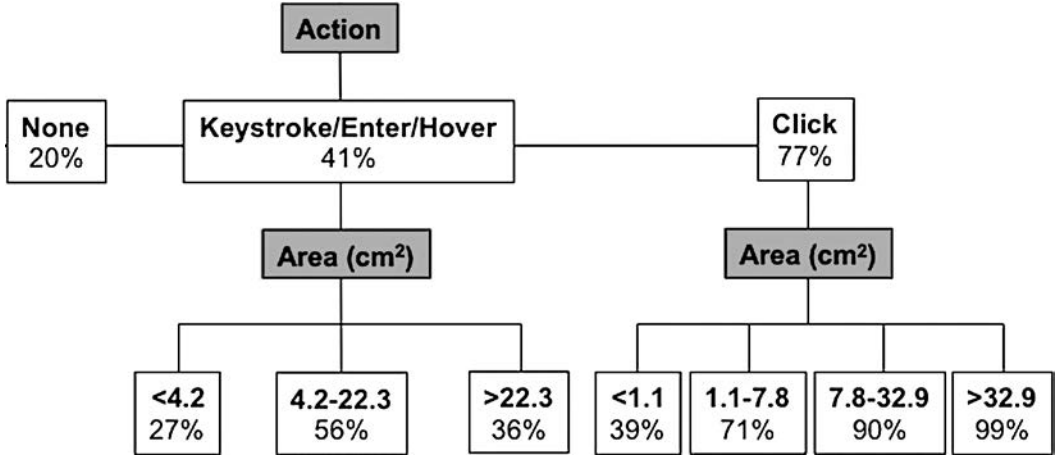

Fig. 27.   The CHAID model created with the filtered eye-tracking data, showing the likelihood an update was viewed as a function of the action that triggered it and its area in cm^2.

Table X. Classification Table for the Action and Area
Filtered Gaze Data CHAID Model

	Observed	(Predicted)	% Correct
Viewed	444	(710)	63%
Not Viewed	645	(776)	83%
Overall	1089	(1486)	73%

the case of the validation data, 70% of those shorter than 18 sec are fixated, but only 50% of those longer ($\chi^2(1, n = 108) = 9.9, p < 0.015$).

C. MODEL FOR ACTION AND AREA ONLY

The model considering only action and area predicts the likelihood an update will be viewed with 73% accuracy in the case of the training data (see Figure 27 and Table X)
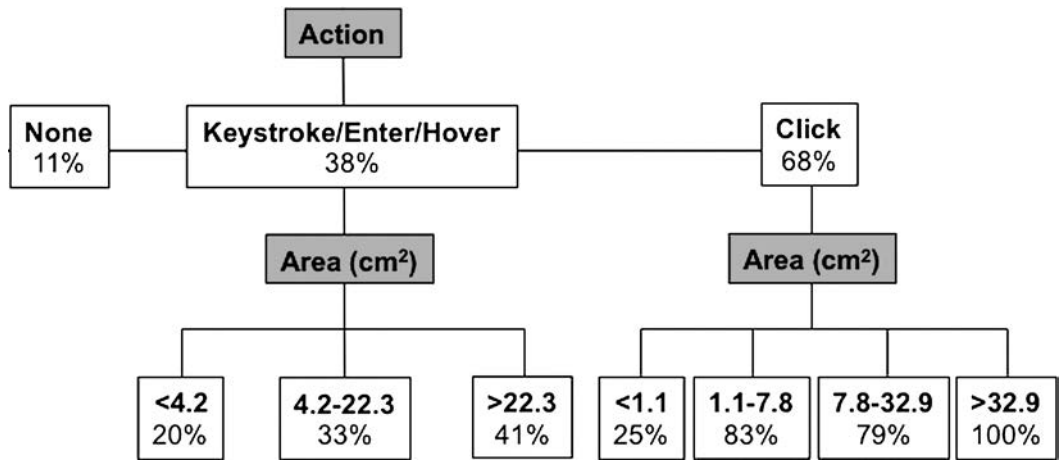
Fig. 28. The CHAID model created with the filtered validation eye-tracking data, showing the likelihood an update was viewed as a function of the action that triggered it and its area in cm^2.

Table XI. Classification Table for the Action and Area
Filtered Validation Gaze Data CHAID Model

	Observed	(Predicted)	% Correct
Viewed	115	(202)	57%
Not Viewed	331	(383)	86%
Overall	446	(585)	76%

and 76% in the case of the validation data (see Figure 28 and Table XI). The size of an automatic update has no impact on whether it is viewed. In the case of keystroke-, enter-key-, or hover-activated update training data, 27% of those smaller than 4.2 cm^2 are viewed, Moreover, 56% of those between 4.2 and 22.3 cm^2 are viewed, and 36% of those greater than 22.3 cm^2 are viewed ($\chi^2(2, n = 251) = 27.4$, $p < 0.001$). Of the validation data, 20% of those smaller than 4.2 cm^2, 33 % of those between 4.2 and 22.3 cm^2, and 41% of those greater than 22.3 cm^2 are viewed ($\chi^2(2, n = 68) = 27.4$, $p < 0.001$). It may be that larger updates of this type are slightly less likely to be viewed because they are often auto-suggest lists, and the longer lists tend to occur at the start of typing, before users look at them.

REFERENCES

BARTRAM, L., WARE, C., AND CALVERT, T. 2003. Moticons: Detection, distraction and task. *Int. J. Hum.-Comput. Stud. 58*, 5, 515–545.

BAYLES, M. 2002. Designing online banner advertisements: Should we animate? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 363–368.

BEANLAND, V. AND PAMMER, K. 2010. Looking without seeing or seeing without looking? Eye movements in sustained inattentional blindness. *Vis. Res. 50*, 977–988.

BROWN, A., JAY, C., CHEN, A. Q., AND HARPER, S. 2010. Web 2.0: An accessibility-oriented survey of dynamic updates. *Univers. Access Inf. Soc. 1,* 2.

BROWN, A., JAY, C., AND HARPER, S. 2009. Audio representation of auto suggest lists. In *Proceedings of the Cross-Disciplinary Conference on Web Accessibility (W4A'09)*. 58–61.

BURKE, M., HORNOF, A., NILSEN, E., AND GORMAN, N. 2005. High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Trans. Comput.-Hum. Interact. 12*, 423–445.

CARMI, R. AND ITTI, L. 2006. Visual causes versus correlates of attention selection in dynamic scenes. *Vis. Res. 46*, 4333–4345.

DIAPER, D. AND WAELEND, P. 2000. World wide web working whilst ignoring graphics: Good news for web page designers. *Interact. Comput. 13*, 2, 163–181.

DJAMASBI, S., SIEGEL, M., AND TULLIS, T. 2010. Generation y, web design, and eye tracking. *Int. J. Hum.-Comput. Stud. 68*, 307–323.

FRANCONERI, S. AND SIMONS, D. 2003. Moving and looming stimuli capture attention. *Percept. Psychophys. 67*, 999–1010.

GRANKA, L., HEMBROOKE, H., AND GAY, G. 2006. Location location location: Viewing patterns on www pages. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA'06)*. ACM Press, New York, 43–43.

GRANKA, L. A., JOACHIMS, T., AND GAY, G. 2004. Eye-tracking analysis of user behavior in www search. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'04)*. ACM Press, New York, 478–479.

HABUCHI, Y., TAKEUCHI, H., AND KITAJIMA, M. 2006. The influence of web browsing experience on web-viewing behavior. In *Proceedings of the Symposium on Eye Tracking Research Applications (ETRA'06)*. ACM Press, New York, 47–47.

HALVERSON, T. AND HORNOF, A. J. 2007. A minimal model for predicting visual search in human-computer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*. ACM Press, New York, 431–434.

HILLSTROM, A. AND YANTIS, S. 1994. Visual attention and motion capture. *Percept. Psychophys. 55*, 4, 399–411.

JAY, C. AND BROWN, A. 2008. User review document: Results of initial sighted user investigations. Tech. rep., University of Manchester. http://hcw-eprints.cs.man.ac.uk/49/.

JAY, C., STEVENS, R., GLENCROSS, M., CHALMERS, A., AND YANG, C. 2007. How people use presentation to search for a link: Expanding the understanding of accessibility on the web. *Univers. Access Inf. Soc. 6*, 3, 307–320.

JINGLING, L. AND YEH, S.-L. 2007. New objects do not capture attention without a top-down setting: Evidence from an inattentional blindness task. *Vis. Cogn. 15*, 6, 661–684.

KITAJIMA, M., BLACKMON, M. H., AND POLSON, P. G. 2005. Cognitive architecture for website design and usability evaluation: Comprehension and information scent in performing by exploration. In *Proceedings of the 11th International Conference on Human Computer Interaction*. 343–373.

MAGLIO, P. P. AND CAMPBELL, C. S. 2000. Tradeoffs in displaying peripheral information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'00)*. ACM Press, New York, 241–248.

MATEJKA, J., LI, W., GROSSMAN, T., AND FITZMAURICE, G. 2011. Ambient help. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*. ACM Press, 2751–2760.

MCCRICKARD, D. S., CATRAMBONE, R., CHEWAR, C. M., AND STASKO, J. T. 2003. Establishing tradeoffs that leverage attention for utility: Empirically evaluating information display in notification systems. *Int. J. Hum.-Comput. Stud. 58*, 5, 547–582.

MILLER, C. S. AND REMINGTON, R. W. 2004. Modeling information navigation: Implications for information architecture. *Hum.-Comput. Interact. 19*.

MILLER, J. 1989. The control of visual attention by abrupt visual onsets and offsets. *Percept. Psychopys. 45*, 567–571.

PAGENDARM, M. AND SCHAUMBURG, H. 2001. Why are users banner-blind? The impact of navigation style on the perception of web banners. *J. Digital Inf. 2*.

PINTO, Y., OLIVERS, C. N. L., AND THEEUWES, J. 2008. Static items are automatically prioritised in a dynamic environment. *Vis. Cogn. 17*.

POTTER, M. C. AND LEVY, E. I. 1969. Recognition memory for a rapid sequence of pictures. *J. Experiment. Psychol. 81*, 10–15.

RODDEN, K., FU, X., AULA, A., AND SPIRO, I. 2008. Eye-mouse coordination patterns on web search pages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'08)*. ACM Press, Springer, 2997–3002.

SAENZ, M. AND BURACAS, G. T. AND BOYNTON, G. 2003. Global feature-based attention for motion and color. *Vis. Res. 43*, 6, 629–637.

TAT FU, W. AND PIROLLI, P. 2007. Snif-act: A cognitive model of user navigation on the world wide web. *Hum.-Comput. Interact. 22*, 355–412.

TEO, L. AND JOHN, B. E. 2010. The evolution of a goal-directed exploration model: Effects of information scent and goback utility on successful exploration. Wiley Online Library 3. 253–259.

VAN SCHAIK, P. AND LING, J. 2004. The effects of graphical display and screen ratio on information retrieval in web pages. *Comput. Hum. Behav. 22*, 870–884.

ZHANG, P. 1999. *Doing Business on the Internet: Opportunities and Pitfalls*. Springer.

ZHENG, X. S., CHAKRABORTY, I., LIN, J. J.-W., AND RAUSCHENBERGER, R. 2009. Correlating low-level image statistics with users -Rapid aesthetic and affective judgments of web pages. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI'09)*. ACM Press, New York, 1–10.