Journal of Eye Movement Research
3(4):5, 1-20

Repeated Web Page Visits and the Scanpath Theory: A Recurrent Pattern Detection Approach

Michael Burmester
Stuttgart Media University

Marcus Mast
Stuttgart Media University

This paper investigates the eye movement sequences of users visiting web pages repeatedly. We are interested in potential habituation due to repeated exposure. The scanpath theory posits that every person learns an idiosyncratic gaze sequence on first exposure to a stimulus and re-applies it on subsequent exposures. Josephson and Holmes (2002) tested the applicability of this hypothesis to web page revisitation but results were inconclusive. With a recurrent temporal pattern detection technique, we examine additional aspects and expose scanpaths. Results do not suggest direct applicability of the scanpath theory. While repetitive scan patterns occurred and were individually distinctive, their occurrence was variable, there were often several different patterns per person, and patterns were not primarily formed on the first exposure. However, extensive patterning occurred for some participants yet not for others which deserves further study into its determinants.

Keywords: web pages, revisitation, eye movements, scanpath theory, sequence comparison, T-patterns

Introduction

A high number of web page visits are revisits. The recurrence rate, i.e. the percentage of revisits among web page visits has been reported in several long-term studies of web navigation behavior as 58% (Tauscher & Greenberg, 1997), 81% (Cockburn & McKenzie, 2001), and more recently with an improved measurement technique as 46% (Obendorf, Weinreich, Herder, & Mayer, 2007). With such high recurrence rates, the question arises how user behavior is characterized on repeated visits. Is user behavior different on each visit or do users develop habits? Habitual, repetitive behavior would likely show in the sequence of eye movements the user follows to scan a page.

In the context of repeated stimulus exposure, Noton and Stark (1971a, 1971b) proposed that every person forms his or her own typical sequence of initial eye movements on first exposure to a stimulus and re-applies it, with some variation, on subsequent exposures as a mechanism of recognition. The associated model of visual recognition became known as the "scanpath theory" and was sometimes controversially discussed (Henderson, 2003; Walker-Smith, Gale, & Findlay, 1977) but has been persistently taken up in various fields such as advertising (Pieters, Rosbergen, & Wedel, 1999) or scene perception (Foulsham & Underwood, 2008). The model predicts gaze sequences to be similar within a person across exposures to the same stimulus but dissimilar between persons.

Josephson and Holmes (2002) tested the scanpath theory's prediction of within-person gaze sequence stability for applicability to web page revisitation. They presented participants several times with identical web pages and measured gaze sequence similarity with the string-edit technique (Wagner & Fischer, 1974). Since only pairwise comparisons can be made with string-editing, multidimensional scaling and clustering techniques were applied subsequently to identify groups of similar sequences. Results were not sufficiently clear to allow for a definite conclusion. There seemed to be a tendency for individually similar gaze sequences but there was also evidence for similarity between persons in many cases.

In this paper, we build on the work of Josephson and Holmes (2002) but apply T-pattern detection (Magnusson, 2000) to overcome some limitations of the string-edit technique. T-pattern detection originates from the

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

analysis of social interaction and extracts recurrent (spatio-) temporal patterns within and across sequences. It allows to additionally test a second central prediction of the scanpath theory: that scanpaths are formed on first encounter with the stimulus. There are no studies known to us having addressed this question. Also, T-pattern detection can provide quantitative measures of scanpath repetition through pattern length and quantity comparisons. The main advantage of the T-pattern approach however is that repeated elements (in this case regions of interest, ROIs) are exposed as T-patterns along with the occurrence times of the pattern. This reveals the actual scanpath and provides insight into the process of its formation.

We employ a paradigm where participants view web pages with partially updated content from visit to visit. We considered this a more common revisitation activity than repeated visits to exactly identical pages as in Josephson and Holmes (2002). The scanpath theory has rarely been tested with stimuli that change from viewing to viewing (two exceptions being Privitera, 2006; Stark et al., 2001). We consider this aspect in the discussion.

The Scanpath Theory

Building on the work of Yarbus (1967) and Jeannerod, Gerin, and Pernier (1968) who had found evidence for cyclic scanning during picture viewing, Noton and Stark (1971a, 1971b) examined the eye movement of persons repeatedly viewing the same line drawing. They found that the initial eye movements on first stimulus exposure were often repeated on subsequent exposures. Gaze sequences of a given person for a given stimulus resembled each other. Gaze sequences were dissimilar however between persons for a given stimulus and between stimuli for a given person. Noton and Stark argued that the within-person similarity results from checking a stored mental model for fit with the currently observed image. They proposed that the representation of visual information in memory is an alternating sequence of sensory (representing a stimulus feature) and motor (representing a saccadic vector to the next fixation location) memory traces. This sequence of memory traces is laid down during initial viewing and run as a control program on subsequent exposures to facilitate recognition. A person's idiosyncratic, recurrent scan pattern was termed a scanpath. Subsequent research under different experimental paradigms (e.g., Brandt & Stark, 1997; Parker, 1978; Pieters et al., 1999; Stark & Ellis, 1981; Zangemeister, Sherman, & Stark, 1995) gave additional credibility to the model. Support for the scanpath theory also comes from the finding that scanpaths occurred when persons only imagined a previously seen stimulus (Brandt & Stark, 1997; Gbadamosi & Zangemeister, 2001; Laeng & Teodorescu, 2002).

The scanpath theory has also received criticism. An early line addressed the dominant role of top-down, memory-guided control of saccadic eye movement it advocates (e.g., Walker-Smith et al., 1977). The contrasting bottom-up view holds that properties of the stimulus (e.g. the saliency of elements) control the progression of gaze which can also explain scanpath stability. However, bottom-up saliency models and algorithms (e.g., Itti & Koch, 2000; Privitera & Stark, 1998; or in the domain of web pages, Faraday, 2000) have often failed to reliably predict gaze sequences (Foulsham & Underwood, 2008; Privitera & Stark, 2000; Grier, Kortum, & Miller, 2007). Further, the bottom-up view does not account for individually different scanpaths. Groner, Walder, and Groner (1984) proposed a differentiation between larger-scale, top-down controlled global scanpaths and smaller-scale, stimulus-controlled local scanpaths. Integrative models that account for simultaneous top-down and bottom-up control have been suggested and adopted by many (e.g., Açık et al., 2009; Cave & Wolfe, 1990; Gray, Sims, Fu, & Schoelles, 2006; Rybak et al., 1998; Schill et al., 2001). More recently, Stark's successors have also embraced a less extreme notion, acknowledging some bottom-up influence (Privitera, 2006; Fujita, Privitera, & Stark, 2007).

It has further been stressed that there are additional factors influencing eye movement. In particular, the task in which a person is immersed has shown to influence scanning (Yarbus, 1967; Groner & Menz, 1985; Grier et al., 2007). A different implicit or explicit task on re-exposure can result in a different gaze sequence. Other factors are demographics, stimulus familiarity, and individual differences (Wedel & Pieters, 2006).

Another line of criticism questions the obligatory link between scanpaths and the recognition process. Locher and Nodine (1973) found that some participants recognized well but had low scanpath occurrence while others did not recognize stimuli but exhibited scanpaths. A subsequent old/new recognition test showed that the presence of scanpaths did not increase recognition perfor-

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

mance (Locher & Nodine, 1974). Also, a certain degree of recognition can be reached without any eye movements, for example under tachistoscopic stimulus presentation (Biederman, Mezzanotte, & Rabinowitz, 1982; Loftus & Mackworth, 1978; see Henderson, 2003; see Rayner, 1998).

Finally, some studies failed to replicate the empirical results of Noton and Stark (1971a, 1971b). Low scanpath occurrence was reported by Mannan, Ruddock, and Wooding (1997). Their analysis technique did not involve the definition of ROIs but a least squares method. While being objective, this method may have inadequately assessed fixation locations from two sequences as similar or dissimilar as it disregards stimulus features.

Applicability to Real-World Stimuli

The scanpath theory was most often tested with simple stimuli such as line drawings (Noton & Stark, 1971a, 1971b; Stark & Ellis, 1981), letter grids (Stark et al., 2001) or irregular checkerboards (Brandt & Stark, 1997). Whether the scanpath theory is applicable to more complex and practically relevant material was also studied. Pieters et al. (1999) repeatedly presented print advertisements to participants. While the duration of attention paid to ad elements decreased from presentation to presentation, the order of scanned elements remained largely stable. Foulsham and Underwood (2008) found evidence for scanpaths when repeatedly viewing photographs of natural scenes.

Josephson and Holmes (2002) tested the scanpath theory using web pages as stimuli. Characteristic about web pages as stimuli is a typically high information density and a high number of semantically and functionally separate regions. Also, the information on a web page may change from visit to visit. As opposed to static stimuli such as pictures, interaction takes place with typical activities being browsing or searching for information (Rada & Murphy, 1992). Josephson and Holmes (2002) presented a news page, an advertising page, and a portal page for 15 seconds each to four female and four male participants on three different days. They used string-editing (Wagner & Fischer, 1974) to measure gaze sequence similarity and visualized sequence distances using multidimensional scaling and cluster analysis (Borgatti, Everett, & Freeman, 1992). Josephson and Holmes found that clusters tended to include pairs of sequences from the same participant but also often the most similar sequences came from different participants. A result in clear support of the scanpath theory's assertion of individually different repetitive scanning should have shown a clearer instance of same-participant clusters. Therefore, the authors argue that other factors such as the stimulus or changing scan behavior over time may also have been influential. However, a definite conclusion on individuality and top-down versus bottom-up influences was refrained from. Such a conclusion would have been to some degree subjective and arbitrary because visualized sequence distances have to be judged for a tendency of clusters to contain sequences from the same versus from different participants.

Identification of Similarity in Eye Movement Sequences

For the analysis of eye movement sequences, two techniques have mainly been applied: Markov analysis (e.g., Groner & Groner, 1982; Groner & Groner, 1983; Ellis & Smith, 1985; Gbadamosi & Zangemeister, 2001; Pieters et al., 1999; Stark & Ellis, 1981) and string-editing (e.g., Brandt & Stark, 1997; Foulsham & Underwood, 2008; Josephson & Holmes, 2002; Pan et al., 2004). This section describes the two established techniques and contrasts them with T-pattern detection. All three techniques require to group fixation positions prior to analysis. This is typically achieved by defining ROIs in the stimulus.

Markov analysis

A Markov matrix describes the probabilities of states to transition into subsequent states. A stochastic process is a first-order Markov process if its future states depend on the present state only. In a second-order Markov process future states depend on the present and one previous state (2-state memory). In eye movement analysis, a state can be represented by an ROI and transitions are made to other ROIs. A Markov matrix is based on a theoretical assumption. For example, the size of ROIs or the duration of fixation on ROIs (Stark & Ellis, 1981) can be used as predictors. Figure 1 shows a first-order Markov matrix which assumes equal probabilities for the eye to travel from one stimulus region to another, i.e. when the eye rests in region a, the probability to look at region b, c, d, … k afterwards is 1/10 each. The goodness of fit of the

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

Markov matrix with an observed matrix of transition frequencies as shown in Figure 2 can then be tested.

	a	b	c	d	e	f	g	h	i	j	k
a	-	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
b	0.1	-	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
c	0.1	0.1	-	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
d	0.1	0.1	0.1	-	0.1	0.1	0.1	0.1	0.1	0.1	0.1
e	0.1	0.1	0.1	0.1	-	0.1	0.1	0.1	0.1	0.1	0.1
f	0.1	0.1	0.1	0.1	0.1	-	0.1	0.1	0.1	0.1	0.1
g	0.1	0.1	0.1	0.1	0.1	0.1	-	0.1	0.1	0.1	0.1
h	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-	0.1	0.1	0.1
i	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-	0.1	0.1
j	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-	0.1
k	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-

Figure 1. First-order Markov matrix showing probabilities for the eye to transition between stimulus regions a to k based on the assumption that the probabilities are equal

	a	b	c	d	e	f	g	h	i	j	k
a	-	1	0	0	0	0	0	0	0	0	0
b	0,25	-	0,25	0	0	0	0,25	0,25	0	0	0
c	0	0	-	0,75	0	0	0	0	0,25	0	0
d	0	0	0,25	-	0,25	0	0,25	0	0	0,25	0
e	0	0	0	0	-	0,67	0	0,33	0	0	0
f	0,33	0,33	0	0	0	-	0	0	0,33	0	0
g	0	0	0	0,25	0	0	-	0,5	0,25	0	0
h	0,17	0	0,17	0	0	0	0,33	-	0,17	0	0,17
i	0,33	0	0	0	0	0,33	0	0,33	-	0	0
j	0	0	0,5	0	0,5	0	0	0	0	-	0
k	0	0	0	0	0	0	0	0	0	0	-

Figure 2. Transition matrix derived from the three observed 10-second stimulus viewings shown in Figure 5

A critical assumption behind the Markov approach is that one or two preceding fixations will provide all information necessary to predict a future fixation. However, the process that generates eye movement sequences may well be more complex. Using longer-memory Markov processes however is usually not practicable because there are not enough data points available to fill all matrix cells. Another deficit of Markov analysis is the necessity for a stochastic assumption. The technique is not suited for direct comparison of observed sequences. Privitera and Stark (1998) describe a related technique for the comparison of observed transition matrices that results in a matrix distance metric. However, this remains a comparison of transitions, not of sequences.

String-editing

String-editing (Wagner & Fischer, 1974; Sankoff & Kruskal, 1983) measures the similarity of two sequences (strings). Editing operations (deletion, insertion, and substitution of elements) are performed to transform one sequence into the other. The more editing operations required, the more the sequences differ. There are several algorithmic variants for editing strings such as the Levenshtein distance (Levenshtein, 1966) which ensures that the smallest possible number of editing operations is applied for transforming one string into the other or the Needleman-Wunsch (1970) algorithm which incorporates gap penalties and scoring parameters that specify differences between string components. The outcome is a sequence distance index. Figure 3 shows an example.

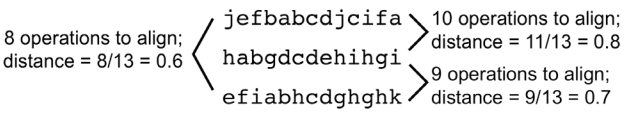


Figure 3. Three strings representing the three 10-second regionalized fixation sequences shown in Figure 5 compared for similarity using Levenshtein distance; the minimum number of editing operations required to align one string with another is divided by the length of the strings; an outcome of 1 would indicate maximally different strings and 0 identical strings

String-editing has some shortcomings too. It compares the complete sequence of ROIs whereas the scanpath may be just a smaller fraction of that sequence. Also, beginning and end of a scanpath cannot be detected with string-editing. Both, Markov analysis and string-editing require ordinal, sequence-only data and disregard the temporal characteristics of scanning. Furthermore, they are not capable of exposing repetitive subsequences within a sequence or across sequences.

T-Pattern detection

T-pattern detection was developed by Magnus S. Magnusson (2000, 2005, 2006) for finding temporal and sequential structure in behavior. The term T-pattern stands for temporal pattern. The algorithm is implemented in the software Theme distributed by Noldus Information Technology (Magnusson, 2004). It detects repeated patterns of intra- or inter-individual behavior coded as events on one-dimensional discrete scales. Sequences of time-coded ROIs can be used as input data (Unz, Werner, Mangold, & Burmester, 2005).

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

A minimal T-pattern consists of two event types. An event type is a category of observable behavior (e.g. to look at ROI 3) whereas an event is an instance of behavior occurring at a particular time unit without a duration (e.g. the event of looking at ROI 3 at time 134 ms). Two event types are considered a T-pattern if they both occur at least twice in the behavior record in the same order and their occurrence times are invariantly distributed over time, i.e. their time distances are unlikely random. Invariance of distribution is specified by two possible types of so-called critical interval (CI) relationships. A fast critical interval (the default type) is present if event type a is followed relatively quickly by event type b. A free critical interval is present if the time distances between occurrences of a and b are relatively similar. Figure 4 illustrates the two types of critical intervals.

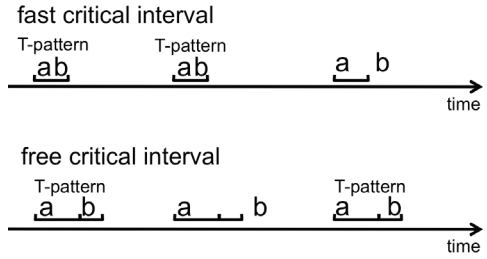


Figure 4. Event type temporal distributions within and outside fast and free critical intervals

Instead of an event type, a component of a T-pattern can also be another T-pattern. A T-pattern with m components $X_i..X_m$ can be formally expressed as a recurring ordered set

$$X_1\ [d_1, d_2]_1\ X_2\ [d_1, d_2]_2\ ..\ X_i\ [d_1, d_2]_i\ X_i{+}1\ ..\ X_{m-1}\ [d_1, d_2]_{m-1}\ X_m$$

where $[d_1, d_2]$ is the critical interval within which one component is followed by the next (d_1 is the temporal distance from the component after which the interval starts and d_2 the end of the interval). When searching for fast critical intervals, d_1 is set to zero (Magnusson, 2005).

The relationship between two pattern components can thus be described as $X_{left}\ [d_1, d_2]\ X_{right}$, i.e. the left component X_{left} is followed within the critical interval $[d_1, d_2]$ by the right component X_{right} (Magnusson, 2005). The critical interval algorithm uses the following procedure:

> The critical interval algorithm measures the time from each occurrence (end) of X_{left} to the first following or concurrent occurrence (beginning) of X_{right}. Using this distribution and some preset significance level, it searches for the longest possible interval $[d_1, d_2]$ such that (X_{left}) (ending at t) is, significantly more often than expected by h0, followed within $[t+d_1, t+d_2]$ by the beginning of another component (X_{right}); where h0 is that (X_{right}) is independently randomly distributed over the observation period $[t_1, t_2]$ with a constant probability per time unit = $N(X_{right}) / (t_2-t_1+1)$; where $N(X_{right})$ is the number occurrences of X_{right}. (Magnusson, 2005, p. 11)

The threshold for how far from random distribution temporal distances of pattern components have to be in the data is thus specified by the significance level (α). Few and shorter patterns are detected with a strict significance level while more and longer patterns are detected with a loose significance level. In an iterative search process, each detected T-pattern can be complemented by another event type or another T-pattern to form a higher-order T-pattern in the next analysis run. Figure 5 shows a T-pattern occurring in three 10-second stimulus viewings. Note the difficulty for a human observer to identify the pattern (abcd) although it is only interrupted by a few occurrences of other event types. The T-pattern consists of the sub-patterns (ab) and (cd) which are connected as ((ab)(cd)) one level higher. Only the longest, most complex patterns survive in the search process while shorter sub-patterns are discarded. The detection process stops at the point where no more critical interval relationships can be found.

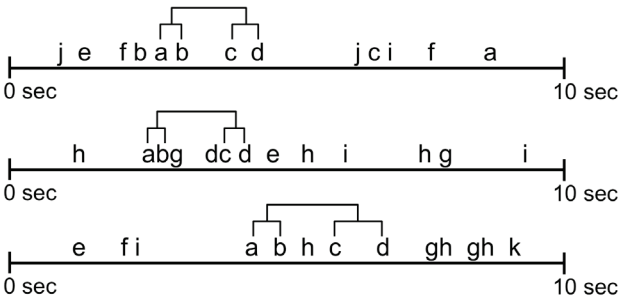


Figure 5. Three different 10-second stimulus viewings (letters represent fixations on regions a to k with successive fixations on the same region omitted) with T-pattern (abcd) occurring once in each viewing

Post-detection tools can be applied for filtering and analyzing patterns, for example based on their occurrence

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

frequency (e.g., show only patterns occurring five or more times), length (e.g., show only patterns consisting of at least 3 ROIs), or behavioral content (e.g., containing ROI 5).

Advantages of T-pattern detection over Markov analysis and string-editing

Compared to Markov analysis and string-editing, T-pattern detection has the following advantages for gaze sequence analysis:

(1) Recurrent patterns are extracted and become observable. For example, the beginning and end of a repetitive scanpath can be identified as the beginning and end of the associated T-pattern and the ROIs contained in the pattern can be studied. A T-pattern does not necessarily represent a continuous sequence in the ROI fixation data on each occurrence, i.e. on occurrences of the T-pattern, other ROIs may have been looked at between pattern components (see Figure 5). Thus, some variation in the eye movement sequences is accounted for.

(2) T-pattern detection does not require a reduction of eye movement data to ordinal sequence data or transition frequency data. The real-time temporal dimension is preserved and more accuracy retained regarding the identification of relevant similarities.

(3) The number of sequences to be compared in one analysis run is unlimited.

(4) The technique is robust to a high number of noise elements in the data. T-patterns need not represent continuous sequences of ROIs and a T-pattern may span just part of a sequence. For example, consider the two gaze sequences ‘gabcdegweiuhviuuwedeertiugearg’ and ‘tabcderqnmqxtcsxlfkfzcoqfoosxq’, each letter representing an ROI. They would show low similarity with string-editing and the recurrent pattern (abcde) at the beginning could go unnoticed, even if repeated across many stimulus exposures.

(5) The metrics provided by T-pattern detection (e.g., T-pattern quantity, length, and occurrence frequency) measure repetition not similarity. If many and long T-patterns are detected and they occur frequently this means that, possibly intermittent, subsequence repetition is high. String-editing in contrast quantifies the similarity of complete sequences. For the repetitive scanpath paradigm, the repetition perspective may be the more relevant criterion.

Hypotheses

The objective of the present study was to test if repeated viewing of web pages results in repetitive and individually distinct patterns of eye movement as would be expected if the scanpath theory applied to this type of stimulus and behavioral activity. In addition to the results obtained with string-editing (Josephson & Holmes, 2002), the T-pattern method allows to examine further aspects: to expose repetitive scanpath patterns so their characteristics can be studied, to assess the role of the individual person in scanpath occurrence, and to determine the stimulus exposure during which scanpaths initially occur. We assume that a detected T-pattern represents a scanpath if it meets certain minimum requirements for length, occurrence frequency, etc. Three hypotheses and one exploratory question were studied:

Universality hypothesis (H1). The scanpath theory suggests that the formation of repetitive scan patterns is an integral part of the visual recognition process (Noton & Stark, 1971a, 1971b). This implies that persons should normally show scanpaths on stimulus re-exposure. However, assuming a 100% criterion would not be reasonable. It is conceivable that other processes, for example related to goal-directed behavior may temporarily guide eye movement and thus no scanpath occurs. The scanpath theory does not state a concrete expectancy for the proportion of persons that should show scanpaths. Therefore, even though the T-pattern technique would allow, inferential statistics cannot be applied to this hypothesis.

> H1: Persons will consistently exhibit T-patterns meeting the requirements for a scanpath.

Individuality hypothesis (H2). According to the scanpath theory, scanpath occurrence is linked to the individual person. The theory predicts that for the same stimulus, several gaze sequences of a given person will be similar but gaze sequences will be dissimilar between persons. With T-pattern detection, a high degree of sequence repetition results in more T-patterns detected and in longer T-patterns (consisting of more ROIs). The claim can therefore be tested by comparing the quantity and length of T-patterns detected in individual persons across visits (within-person patterns) with the quantity and length of T-patterns detected between persons across visits (between-person patterns). More and longer within-person than between-person T-patterns would indicate that the individual person plays an important role in pattern for-

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

mation. Identical quantity and length of within- and between-person T-patterns would indicate that the individual has little influence and other factors such as the stimulus may be more influential.

H2: The individual person is a decisive factor for the occurrence of scanpaths on web pages.

This is operationalized as:

H2a: There will be more within-person T-patterns than between-person T-patterns.

H2b: Within-person T-patterns will be longer than between-person T-patterns.

Origin hypothesis (H3). A central prediction of the scanpath theory is that scanpaths are formed on first exposure to a stimulus (Noton & Stark, 1971a, 1971b). With T-pattern detection, the formation rate of new patterns in each stimulus presentation can be measured and compared.

H3: T-patterns will initially occur during the first visit to a web page.

Again, the scanpath theory does not give a definite criterion for the expected proportion of scanpaths that should originate from the first presentation and it seems inappropriate to assume 100%. Therefore, no statistical test can be performed. Nevertheless, descriptive results can provide insight into how this condition is met.

Data exploration. As T-pattern detection makes scanpaths observable as T-patterns it is worthwhile to examine them qualitatively for included ROIs and temporal occurrence characteristics.

Method

Participants

32 participants were invited to the study but 7 were not admitted because of calibration issues and another 5 were disregarded in the analysis because of sporadic capturing issues. The remaining 20 participants with valid data were all university students (10 female and 10 male) aged between 20 and 29 (mean age 24.5). They had extensive Internet experience with at least two hours of daily usage. Participants had normal or corrected-to-normal vision.

Stimuli and Procedure

Participants were presented with web pages from two web sites corresponding to two common scenarios of recurring visits. The stimuli were the home page of the German broadcaster ARD (www.ard.de) captured on 10 different days and containing mainly general public news as well as 10 entry pages to product categories from the web shop of furniture retailer IKEA (www.ikea.de). Exposing participants 10 times to each web site allowed for a detailed analysis of the temporal origin of scanpaths in stimulus presentations. Every web page contained fixed elements as well as varying ones (see Figures 7 and 9). The varying ones varied on each presentation and were never shown twice. The layout of the pages was slightly manipulated so that regions were exactly identical in size for all pages of a site. In addition to the 10 ARD and 10 IKEA pages, 10 distractor web pages were inserted in the stimulus sequences to loosen up the viewing experience. Before each web page a white screen with a black dot in the middle was shown to unify initial fixation positions on the pages. In each presentation sequence, the order altered between the two web sites with a distractor page in between. The internal order of the 10 pages of each type (IKEA, ARD, distractor) was randomized for each participant. Presentation sequences were counterbalanced for initially shown site so that half of the participants saw an ARD page first and the other half saw an IKEA page first. The web pages were shown as screenshots for 10 seconds each. The dot screen was shown for two seconds. Participants had little to moderate experience with the two web sites. For ARD 11 and for IKEA 10 participants had not visited the respective web site during the past six months. The other participants had visited the sites less than five times during that period.

Task

Typical web user activities are searching and browsing (Rada & Murphy, 1992). When searching, users are in a goal mode (Hassenzahl, Kekez, & Burmester, 2002), looking for a particular and clearly defined piece of information, e.g., the director of a specific movie. Searching is characterized by extrinsic motivation and utilitarian benefits (Hoffman & Novak, 1996). When browsing, users are in an activity mode (Hassenzahl et al., 2002) with less precise intentions (e.g., to discover new movies), more easily inspired and distracted by the viewed material. Browsing is characterized by intrinsic motivation and hedonic benefits (Hoffman & Novak, 1996).

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

The instruction given was to "freely explore whatever is of interest". This browsing task was chosen for the following reasons: (1) The occurrence of repetitive, idiosyncratic scanpaths has mainly been reported under free-view paradigms when participants were not given any search targets (e.g., Josephson & Holmes, 2002; Noton & Stark, 1971a, 197b; Pieters et al., 1999). (2) A precise task may have directed participants to specific elements in the web pages which may have stimulated the formation of patterns. Instead, at this stage it was of interest if participants form patterns freely. (3) Browsing and exploring was considered a common activity on the web pages used as stimuli.

Apparatus

Eye movements were recorded with a Tobii 1750 remote infrared eye tracker integrated with a 17" TFT display running at its native resolution of 1280 by 1024 pixels. The eye tracker determines the point of regard using binocular tracking and the cornea reflection method combined with video-based pupil detection. Its frame rate is 50 Hz, the spatial resolution 0.25°, average accuracy 0.5°, time to tracking recovery below 100 ms, calibration drift over longer periods below 1°, and head movement compensation error below 1° for movement in three dimensions and head rotations (Tobii, 2006). Participants were calibrated using nine calibration points and correct calibration was verified with a real-time gaze display. Participants with unreliable calibration results were not admitted to the study.

Data Preparation

To determine fixations, the fairly established temporal threshold of 100 ms (Rayner, 1998) and a spatial threshold radius of 30 pixels were used. ROIs were first drawn based on the location of page elements (e.g., picture or headline). The defined ROIs were then compared with heatmap fixation data and corrected slightly so that ROI borders would not cross areas of high fixation concentration. The resulting ROI fixation data was transformed to event data for T-pattern analysis. Carryover fixations that started on the dot stimulus were discarded. Successive fixations within an ROI were aggregated into a single event (rather than one event per fixation or enter/leave ROI events), occurring at the time of first fixation in the ROI. This data reduction made detected T-patterns straightforward to interpret without losing meaningful detail. Distractor and dot pages were discarded. 20 participant datasets for ARD and 20 for IKEA were compiled, each containing the 10 exposures of the participant to the web site in the order of presentation. These datasets formed the basis for T-pattern detection.

Scanpath Definition for T-Pattern Detection

The transformed gaze data was analyzed with Theme version 5.0 (Magnusson, 2004). T-pattern detection parameters were specified so that each detected T-pattern would meet the minimum requirement for a scanpath still in conformance with the scanpath theory. To be considered a scanpath, T-patterns had to meet the following requirements:

Significance and pattern interruption. The significance level for T-pattern detection was set to 0.005, meaning that a probability of at most 0.5% was accepted for any critical interval relationship to occur by chance. Only fast critical intervals were permitted. These parameters corresponded to typically 0 to 2 ROIs occurring between the ROIs of a pattern. With regard to the scanpath theory this was considered appropriate because small deviations between the occurrences of a scanpath are expected (Noton & Stark, 1971a, 1971b).

Length. Every T-pattern had to have a minimum length of 3 ROIs. Patterns consisting of only 2 ROIs were not considered sufficiently meaningful.

Number of occurrences. Every T-pattern had to occur in at least 50% of the stimulus presentations. Noton and Stark (1971a, 1971b) presented each stimulus four times. Participants repeated the scanpath from presentation 1 in 65% of the subsequent presentations on average. This is equivalent to the scanpath occurring in 73.8% of all presentations. The criterion was 50% in the present study because the parameter represents the minimum requirement. Also, the stimuli were more complex with many more possible fixation targets. 31 ROIs were defined in the ARD pages and 36 in the IKEA pages. The distinctive features in Noton and Stark's stimuli were in the range of 10 to at most 20, depending on the stimulus and stringency of judgment.

First occurrence. T-patterns were required to initially occur during the first stimulus exposure as predicted by the scanpath theory.

These were the base parameters used when determining the number of participants showing scanpaths (H1).

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

Modifications for Testing Hypotheses 2 and 3

To test H2, 20 additional datasets were compiled for each web site. In distinction to the original within-person datasets, these were termed between-person datasets. In between-person datasets the participant changed on each serial position but viewing order remained intact. Participants and gender were balanced and no page viewing occurred twice throughout the datasets. T-pattern searches were run on both within- and between-person datasets with the same parameters. The requirement for patterns to initially occur during the first stimulus exposure was dropped because individuality (H2) was examined independent of the origin question (H3). T-patterns found in within-person and between-person datasets were compared for quantity and length to obtain information on the effect of a constant person on scanpath occurrence. For the length comparison, patterns with any length were permitted into the analysis (i.e., the length requirement was lowered to a minimum length of 2).

To test H3, the number of initially occurring patterns had to be compared between stimulus presentations (this requires to give up the first occurrence requirement). With a 50% occurrence criterion and 10 stimulus presentations, pattern counts can be compared at most between presentations 1 to 6 (the condition ≥ 5 occurrences in 10 presentations can not be met from presentation 7 on). However, there is a stochastic bias in this comparison: The probability decreases over time for a new pattern to still be repeated at least another 4 times. For example, a pattern A first occurring in presentation 1 has another 9 presentations (10 in total) in which it needs to occur at least another 4 times to meet the condition. A pattern B that first occurred in presentation 6 only has another 4 presentations in all of which it needs to occur to meet the condition. Thus, the tendency to exhibit new patterns that will meet the 50% condition decreases from presentation 1 to presentation 6. A simple comparison of first occurrence counts between presentations is therefore not valid.

The approach chosen to test H3 while avoiding stochastic bias was to run a separate T-pattern detection where the window size for pattern occurrences was kept identical in all presentations. This was achieved by replacing the condition ≥ 5 occurrences in 10 presentations with ≥ 3 occurrences in 6 presentations, thereby maintaining the 50% criterion. All other T-pattern detection parameters remained identical to the aforementioned. When determining the pattern formation tendency of a presentation position, a window of only 6 presentations (i.e., the current one plus the subsequent 5) was considered. For example, when determining the pattern formation tendency of presentation 3, only patterns occurring at least 3 times during presentations 3 to 8 (current presentation + 5) with an initial occurrence in presentation 3 were considered. Patterns first occurring before presentation 3 did not count for presentation 3. Patterns occurring in presentations 9 and 10 likewise were disregarded. The patterns found with a 3/6 condition are not necessarily the same ones as found with a 5/10 condition but this was regarded an acceptable tradeoff. The measure taken allowed to validly compare newly formed patterns between the first five stimulus presentations because with a total of 10 presentations, a window of 6 presentations can at most be shifted to include presentations 5 to 10.

Validation of T-Patterns Against Chance Occurrence

Each T-Pattern search result was compared with the average result of 100 searches in datasets of equal length with the same events but occurring at random moments (random paths across the same ROI events). Those searches resulted in no or very few and short T-patterns being found (e.g., means of 0.0 T-patterns per person and 2.0 ROIs maximum length for H2 within-person datasets search validation), giving confidence in the validity of results.

Results

Universality hypothesis (H1). The universality hypothesis stated that repetitive scanpath eye movement should occur for nearly all persons consistently on both web sites. Table 1 shows how this prediction was met. 5 participants (25%) showed scanpaths on both web sites, 11 participants (55%) showed scanpaths on one of the two web sites, and 4 participants (20%) did not show any scanpaths on either web site. For ARD, 8 participants (40%) and for IKEA, 13 participants (65%) showed one or more T-patterns meeting the criteria for a scanpath, amounting to a mean of 53% between the two web sites. These data do not suggest a universal phenomenon and are not considered as supporting the hypothesis.

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

Table 1. Participants with occurrence of one or more T-patterns meeting the minimum scanpath criteria

Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
ARD	-	X	-	-	-	-	X	X	-	-	X	X	-	-	X	-	-	X	X	-
IKEA	-	X	X	X	X	X	-	X	-	-	X	X	X	X	X	X	-	-	-	X

Individuality hypothesis, quantity aspect (H2a). The first part of the individuality hypothesis stated that more T-patterns should occur in the stimulus viewings of an individual person (within-person) than in datasets with varying persons (between-person). The mean number of T-patterns detected within persons when viewing the ARD web pages was 6.1 T-patterns per person (SD = 9.8, min = 0, max = 29) and 19.8 for the IKEA pages (SD = 63.1, min = 0, max = 277), resulting in an overall mean of 13.0 detected T-patterns per person. Between-person datasets showed a mean of 0.3 T-patterns for ARD (SD = 0.7, min = 0, max = 3) and 0.2 for IKEA (SD = 0.5, min = 0, max = 2), resulting in an overall mean of 0.2 T-patterns. In the IKEA within-person datasets, two extreme outliers showed very high numbers of T-patterns (participant 2: 81 patterns; participant 15: 277 patterns). An inspection of their patterns revealed high pattern similarity and redundancy. This can be attributed to a lack of adequate redundancy reduction of the detection algorithm.

A 2 × 2 mixed-model analysis of variance (ANOVA) was performed with person stability (person stable, person varying) treated as between-subject factor, web site (ARD, IKEA) as within-subject factor, and pattern quantity as the dependent variable. When including the outliers, there was no main effect of person stability statistically significant at the 0.05 level ($F(1, 38) = 2.98$, $p = 0.09$, partial eta squared = 0.07). Because redundant detection was identified as the background, exclusion of the outliers seemed reasonable. When excluding the outliers, the effect of person stability on the number of patterns detected was significant ($F(1, 36) = 7.37$, $p = 0.01$, partial eta squared = 0.17). No significant main effect of web site ($p = 0.33$) and no significant interaction effect ($p = 0.32$) were found.

It can be concluded that substantially more patterns occurred when a stable person was present than between persons (13.0 vs. 0.2 patterns on average), supporting the hypothesis of a strong individual component involved in scanpath occurrence.

Individuality hypothesis, length aspect (H2b). The second part of the individuality hypothesis stated that within-person T-patterns should be longer (i.e., encompass more looked at ROIs) than between-person T-patterns. All datasets with at least one detected pattern of length ≥ 2 ROIs were submitted to this analysis (18 within-person and 19 between-person datasets).

The mean length of the longest pattern detected in individual persons was 3.7 ROIs for ARD (SD = 1.9, max = 7) and 3.5 ROIs for IKEA (SD = 1.8, max = 8), resulting in an overall mean of 3.6 ROIs. The mean length of the longest pattern in between-person datasets was 2.3 ROIs for ARD (SD = 0.5, max = 4) and 2.1 ROIs for IKEA (SD = 0.3, max = 3), resulting in an overall mean of 2.2 ROIs. A person stability (2) × web site (2) mixed ANOVA showed a significant main effect of person stability on the dependent variable pattern length ($F(1, 35) = 14.20$, $p = 0.001$, partial eta squared = 0.29). No significant main effect of web site ($p = 0.41$) and no significant interaction effect ($p = 0.89$) were found.

Scan patterns of individual persons were substantially longer than between-person patterns (3.6 ROIs vs. 2.2 ROIs maximum length). This result is in support of the hypothesis that the individual person plays an important role in scanpath occurrence.

Origin hypothesis (H3). The origin hypothesis stated that patterns should primarily emerge from the first visit to a web page. By applying the previously described procedure, pattern formation tendency was compared between stimulus presentations 1 to 5. Presentation percentages were calculated per participant first and total percentages formed from those values. Figure 6 shows the results.

While there seems to be a tendency for pattern formation during earlier stimulus exposures, patterns were not exclusively formed during the first viewing. 68.4% of the patterns do not emerge from presentation 1. The decrease of newly formed patterns from presentation 1 to presentation 4 and then near-stagnation in presentation 5 may indicate a process of pattern stabilization as the viewer becomes more accustomed to the web page. On the other hand, even at presentation 5, 2.4 new patterns were still formed on average per participant (excluding two outliers; average is 5.9 when including them). A qualitative inspection showed that many patterns indeed underwent a process of refinement and stabilization.

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

However, qualitatively different patterns to any previous ones also still emerged in later presentations. There seemed to be a tendency for certain participants to stabilize patterns and for other participants to keep forming new patterns. However such potential individual differences would have to be clarified by further studies.

There was rather high variation between participants in where patterns emerged from, resulting in large confidence intervals as shown in Figure 6. Still, it is obvious that a substantial proportion of the patterns were formed after presentation 1. The results therefore give little support to the hypothesis.

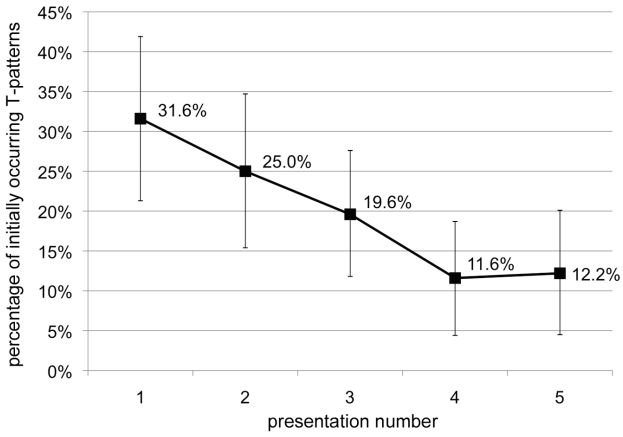


Figure 6. Emergence rates of new T-patterns in stimulus presentations 1 to 5 (means and 95% confidence intervals)

Data exploration. Detected patterns were examined to find out what web page elements were scanned repeatedly by participants. Of increased interest were long, often repeated, and particularly unique patterns exhibited only by certain individuals. Not all patterns presented subsequently initially occurred on the first viewing. The results of hypothesis 3 suggest that new patterns are formed continuously. Therefore patterns not in strict conformance with the scanpath theory were also considered. When interpreting the patterns, it is important to be aware that they usually do not represent continuous gaze sequences. On any occurrence of a pattern, a varying number of intermediate ROIs (typically between 0 and 2) may have been looked at between the ROIs of the pattern. Figures 7 and 9 employ a different type of illustration to show which page elements varied from presentation to presentation. Varying elements are blurry due to overlaid content. The other illustrations employ examples of pages as presented to participants. ROIs are overlaid.

The longest patterns found on the ARD pages involved 7 ROIs. Participants 8 and 12 showed two patterns each of that length. One of the patterns by participant 8 is visualized in Figure 7. It contains several pictures and captions in the right column in a top-to-bottom scan order. It initially occurred in its complete form in presentation 3. Various similar but shorter variations of this pattern initially occurred in presentation 2. This may indicate a process of pattern stabilization or refinement.

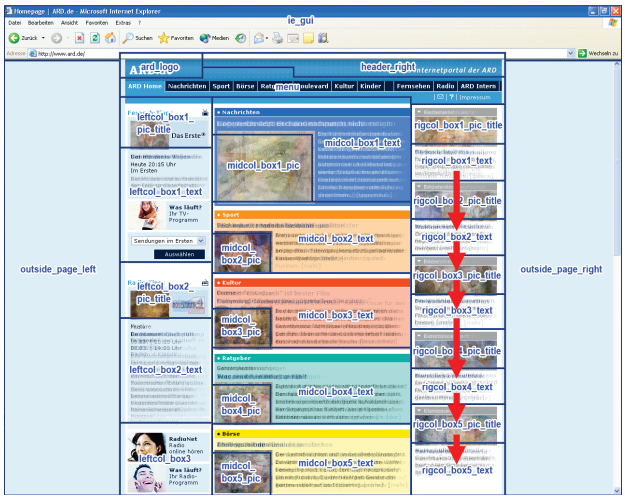


Figure 7. Long ARD pattern involving 7 ROIs, occurring in 5 out of 10 presentations, initially in presentation 3 (participant 8); blurred elements varied from presentation to presentation

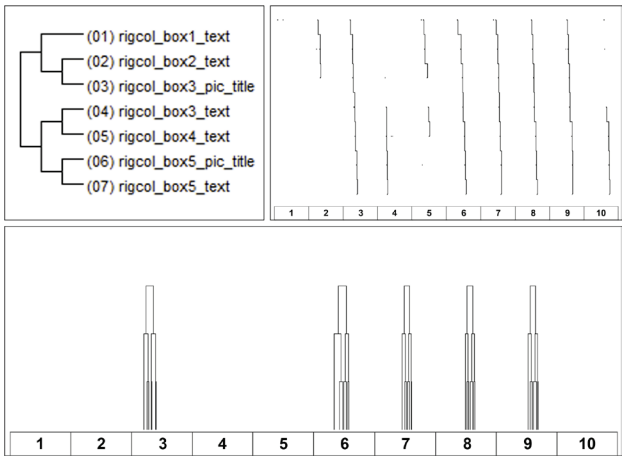


Figure 8. T-pattern diagram of pattern visualized in Figure 7; pattern structure in upper left, pattern occurrences in upper right and bottom, presentation positions at bottom of graphs

Figure 8 shows the associated T-pattern diagram. In the upper left the pattern tree graph shows the ROIs included and how they are connected as sub-patterns. The

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

graph in the upper right visualizes the real-time occurrence distribution of the pattern and smaller sub-patterns. ROI occurrences are indicated by dots and the lines connecting the dots represent (sub-) pattern occurrences. The numbers 1 to 10 at the bottom denote the 10 stimulus presentations of 10 seconds each. The graph at the bottom also shows the occurrence distribution of pattern instances but without sub-patterns and using the same hierarchical pattern visualization as in the upper left. From this graph the stimulus presentations with pattern occurrences and the moments and lengths of T-pattern occurrences can be conveniently observed. The pattern occurred 5 times in total, initially in presentation 3.

The longest patterns on the IKEA pages involved 8 ROIs. Only participants 2 (1 pattern) and 15 (6 patterns) exhibited patterns of this length. The pattern by participant 2 is shown in Figure 9. It first occurred in presentation 3. The order of scanning may follow a strategy of least effort, avoiding unnecessary and long saccades, thus overriding the typical left-to-right scanning order in the upper row of products. An inspection of the eye movement data supports this presumption as few and if any, short saccades were made between the pattern elements. A similar shorter version of this pattern (5 ROIs, 8 occurrences) initially occurred in presentation 2.

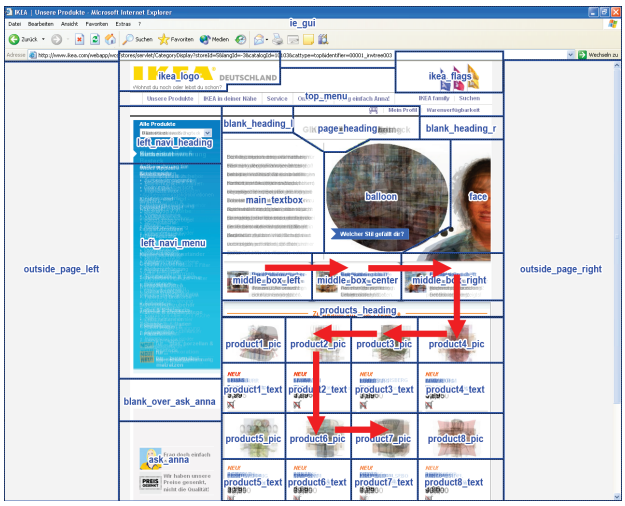


Figure 9. Long IKEA pattern involving 8 ROIs, occurring in 5 out of 10 presentations, initially in presentation 3 (participant 2); blurred elements varied from presentation to presentation

The most often-repeated scan patterns occurred in all 10 presentations. Only participant 15 showed patterns with occurrences in all presentations (one on each web site, involving 3 ROIs each). 6 participants still exhibited patterns with 8 or more occurrences. An often-repeated ARD pattern is shown in Figure 10. The participant also exhibited several longer but less often repeated versions of this pattern involving up to 8 ROIs in the middle column from top to bottom. Other participants showed similar patterns in the middle column.



Figure 10. Often-repeated ARD pattern involving 4 ROIs, occurring in 8 out of 10 presentations, initially in presentation 1 (participant 12)

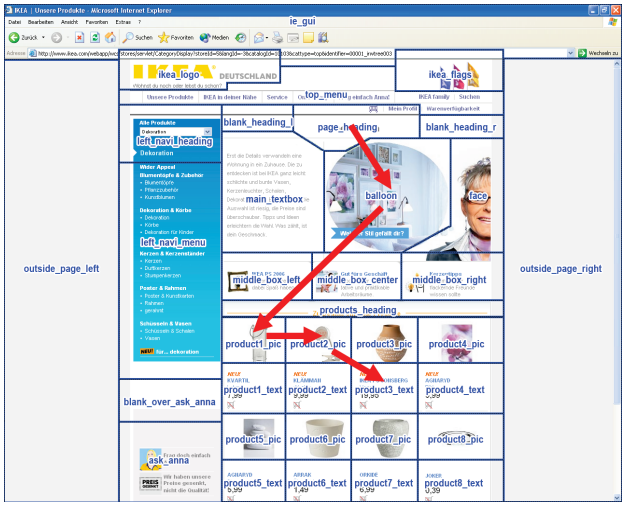


Figure 11. Often-repeated IKEA pattern involving 5 ROIs, occurring in 8 out of 10 presentations, initially in presentation 1 (participant 15)

Figure 11 shows an often-repeated pattern by participant 15 which has an interesting occurrence series. The pattern first occurred on the first presentation, did not occur during presentations 2 and 3 but then occurred in

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

each presentation from the fourth onwards. This participant apparently reverted to the old pattern and kept it after trying some alternative scanning. The page heading, face, and balloon ROIs were contained in the patterns of many other participants in varying order (eight participants showed at least one pattern containing two or more of these ROIs).

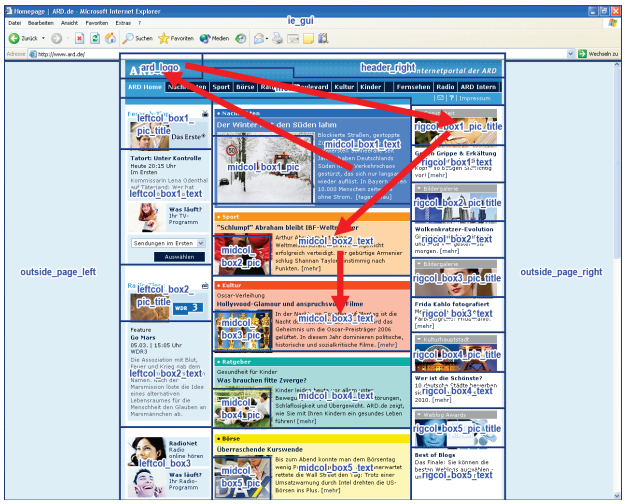


Figure 12. Unique ARD pattern involving a fixed element (ARD logo) and shifting between columns, occurring in 5 out of 10 presentations, initially in presentation 2 (participant 2)

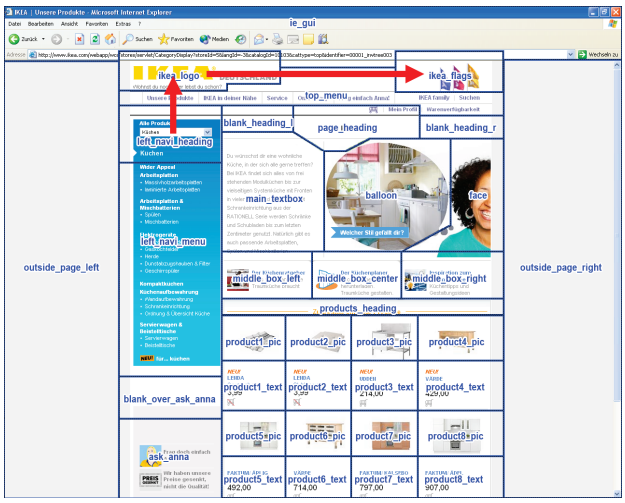


Figure 13. Unique IKEA pattern involving 1 varying and 2 fixed ROIs, occurring in 5 out of 10 presentations, initially in presentation 1 (participant 16)

Figure 12 shows a unique ARD T-pattern of participant 2. It includes a fixed element, the ARD logo in the upper left, which rarely appeared in other patterns. Also, the order of scanning is highly individual and rare, with a specific order of shifting between columns. The pattern first appeared in presentation 2 and five times altogether.

A unique IKEA pattern is shown in Figure 13. Untypical about the pattern is that it involves 2 fixed ROIs. The first ROI contained is the name of the current page's product category (varying element), the second the IKEA logo (fixed) and finally the IKEA flags in the upper right (fixed). All of these ROIs were rarely contained in any patterns of other participants. Possibly, this participant had a strong need for orientation on the web site. However, the participant did not show any patterns with fixed elements on the ARD pages.

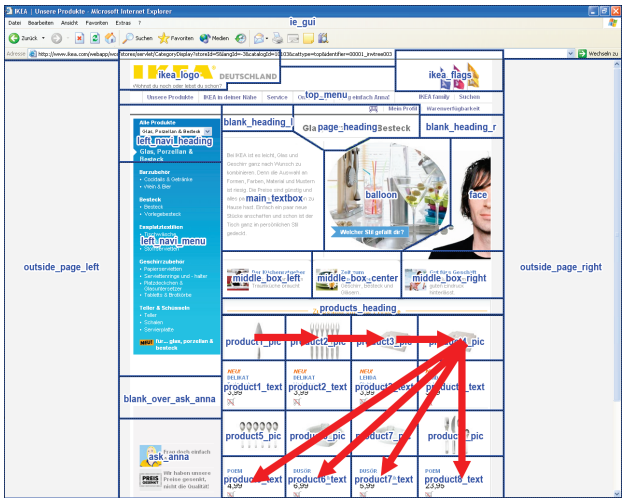


Figure 14. Four IKEA patterns with a common beginning, involving 5 ROIs each, occurring between 5 and 6 times each, 3 of the patterns initially in presentation 1, the other initially in presentation 6 (participant 8)

Participant 8 showed several scan patterns on the IKEA pages that share the same series of ROIs in the beginning but differ at the end. Four patterns are visualized concurrently in Figure 14. At the beginning of all four patterns product pictures in the upper row are contained in a left to right direction. Each pattern ends with a different product text ROI in the lower row. Inspection of the eye movement data revealed that the participant often scanned over the lower row of products but in a different order on each presentation, resulting in four different patterns rather than one longer pattern including more ROIs from the lower row. The occurrence series of the participant's patterns is interesting too. The participant did not seem to apply any consistent scan strategy between presentations 2 to 5 (no patterns). However, in

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

presentation 6, the participant reverted to the pattern from presentation 1 and stayed with it until the end.

In summary, the inspection of detected patterns revealed that patterns were heterogeneous between participants in the regions included and the order of scanning, giving further support to the result of hypothesis 2. Most often, patterns contained elements that changed from presentation to presentation such as pictures and text in content areas. Fixed elements such as the menu, logo, etc. were rarely contained. Not surprisingly, long patterns tended to be seldom-repeated and often-repeated patterns tended to be shorter. Some patterns seem straightforward to interpret, e.g., the pattern shown in Figure 7 representing a natural and efficient scanning strategy for a column. Other patterns are rather unexpected, e.g., the IKEA pattern in Figure 13 with its two fixed ROIs. Parts of some patterns followed a reverse-to-normal scanning order such as the pattern shown in Figure 9 going in part from right to left or other patterns that went from bottom to top in the columns on the ARD pages (not illustrated). This may be related to effort reduction strategies applied to minimize the use of long saccades (Ballard, Hayhoe, & Pelz, 1995). With regard to the scanpath theory it is important to note that 74% (on average between the two web sites) of the participants who showed patterns did not just show one but several patterns. In some cases, this was due to redundant patterns and the existence of shorter versions of similar more complete patterns. For other participants, indeed, multiple different patterns were found.

Discussion

Although repetitive scanning in part heavily occurred, not all predictions of the scanpath theory could hold. It was hypothesized (H1) that the majority of participants should exhibit repetitive visual scan patterns meeting the minimum requirements for a scanpath if indeed, scanpath eye movement is essential during repeated perception of a web page (Noton & Stark, 1971a, 1971b). Only 25% of the participants showed such scan patterns on both web sites. On the other hand, 80% of the participants showed scanpaths on at least one web site. A post-hoc lowering of the minimum repetition requirement from 50% to 30% of the stimulus presentations coupled with giving up the initial occurrence requirement resulted in all participants showing patterns on both web sites. This shows that some patterning always occurs but the extent was often below the minimum criteria for a scanpath. For those cases where scanpath-qualifying patterns occurred, there were also large differences in pattern length. Thus, repetitive scanning of the extent described by the scanpath theory (Noton & Stark, 1971a, 1971b) does occur on web page revisitation but not by far as reliably as hypothesized. There did not seem to be much consistency in pattern occurrence or pattern length for a given participant across web sites or for a given web site. Scanpath eye movement has shown to be consistently applied (Brandt & Stark, 1997) and necessary (Laeng & Teodorescu, 2002) for optimal recall during visual imagery. On web page revisitation, it is not consistently applied, at least under the here induced circumstances with a browsing task and some variable content.

In line with the scanpath theory are the results of hypothesis 2, suggesting an important role of the individual person in pattern formation. Repetitive scanning occurred to a much lesser extent between persons than within persons (0.2 vs. 13.0 patterns and 2.2 vs. 3.6 ROIs length on average). Therefore, the stimulus did not account for scanpaths. In the previous study by Josephson and Holmes (2002) it remained unclear how much influence the individual had on scanpath occurrence. The present results suggest strong top-down influence as scanpaths nearly did not occur between persons. This is consistent with the failure of bottom-up saliency models for predicting gaze sequences on web pages (Grier et al., 2007).

As the scanpath theory asserts that patterns are formed on first stimulus encounter, it was further hypothesized (H3) that visual scan patterns should initially occur on the first page viewing. While there was a tendency for patterns to be formed on earlier page viewings, only 31.6% of the patterns emerged from presentation 1 while 68.4% emerged from presentations 2 to 5. Hence, it seems that the “learning phase” (Noton & Stark, 1971a, 1971b) for web pages is not limited to the first stimulus exposure but new patterns are continuously formed or existing ones modified. One of the reasons for the continuous occurrence of new patterns was pattern refinement as revealed by the qualitative inspection. It is also conceivable that the high information density on a web page and the resulting higher processing effort contributed to longer periods for pattern establishment.

The qualitative inspection of detected patterns revealed page elements repeatedly included in eye move-

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

ment sequences. The individuality of patterns was confirmed from a pattern content perspective. Unexpected persistent inclusions of invariant page elements such as the company logo were discovered. Indication for saccade-minimizing strategies in patterns with partially reversed scanning orders was found. Further, pattern refinement and gradual modifications were observed. Myers and Gray (2007) also describe such processes and provide a theoretical foundation in soft constraints theory (Gray & Fu, 2004; Gray, Sims, Fu, & Schoelles, 2006). Most importantly though, the qualitative sighting showed that "the" scanpath of a person usually could not be identified. Participants typically showed a number of different patterns and sub-patterns repeated at seemingly irregular intervals. Part of the large pattern quantity can be explained with redundant detection and another part with refinement processes. But often, a given participant exhibited several patterns with entirely different content. This objects the scanpath theory's assumption of one consistent scanpath per person.

Two limitations of the present study should be noted. First, all web pages were presented as a single sequence on a single day. At least for the ARD pages, it would have been more realistic to invite participants on different days. Under given practical restraints, the chosen approach however allowed for a high number of stimulus repetitions and therefore an analysis of the temporal origin of patterns, which was considered more important. Second, the participants were presented with static mockups and could not navigate. Free navigation poses several challenges to data analysis but should be considered in future research for higher ecological validity.

Judging from the obtained data, direct applicability of the scanpath theory to the paradigm of web page revisitation does not seem plausible. While the presence of an individual person was a prerequisite for the occurrence of scanpath-qualifying patterns, pattern occurrence and the extent of patterning were too variable, often several different patterns per person were found, and patterns were not primarily formed on the first exposure. Still, the many long, often-repeated, and individually distinctive patterns found are noteworthy.

That scan patterns occurred even though content changed means that users must have the ability to develop a kind of semantic or categorical patterning that abstracts from concrete visual content. Steady content does not seem to be a prerequisite for the occurrence of scan patterns. There are two other studies known to us that investigated scanpath occurrence under changeable stimuli. Privitera (2006) describes an experiment where the image contrast of pictures was manipulated but no semantic changes were introduced. Gaze sequences remained fairly stable across exposures. Related research on the effect of contrast changes on the allocation of fixations (Açık et al., 2009; Nyström & Holmqvist, 2008) has shown that such manipulations may or may not influence fixation locations, depending on the type of image. Stark et al. (2001) report an experiment on scanpath occurrence where the letter "E" was distributed in a specific configuration on a grid. In subsequent stimulus presentations, it was exchanged with the letter "W" while maintaining the configuration. This semantic change resulted in gaze sequences to change. Stark et al. (2001) concluded that scanpath occurrence depends on semantic, or symbolic, binding. In view of this result it is remarkable that a high amount of patterning occurred in the present experiment, despite semantic changes. A possible explanation is that the changed letters on the otherwise blank grid to participants suggested the presence of a new stimulus. On the web pages, the spatial layout, visual appearance, and some elements (e.g., navigation or company logo) remained identical across visits, providing a stable global context. Further, the elements varying from page to page had a common theme: e.g., on the IKEA pages, the product pictures themselves varied but at their location only product pictures could be found. Contextual cues (Chun & Jiang, 1998) provided by invariant global properties combined with the categorical stability of the varying elements may have resulted in the stimulus still being considered the same and provided a sufficient anchor for patterning to occur.

To us the most important open question is the reason for the high variability in pattern development. When a user repeatedly visits a web page, what causes scans to be stable or variable? Noton and Stark (1971a, 1971b) assumed that stimulus recognition is the reason for scanpath stability and most subsequent authors who found scanpaths adopted this explanation (e.g., Josephson & Holmes, 2002; Pieters et al., 1999). While this assumption was not specifically tested in the present study, it appears unlikely that recognition underlies the scan patterns found on web pages. A mechanism as basic as recognition can be expected to operate similarly in all persons. The many cases where participants' pattern development was below the criteria for a scanpath are in

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

disagreement with this view. Also, the exposure durations during which web page scanpaths were found (10 sec in the present study and 15 sec in Josephson and Holmes) seem too long to be sufficiently explained with stimulus recognition.

It has been shown that the information acquired with a single fixation foveally and peripherally suffices to attain a fair degree of recognition (e.g., Biederman et al., 1982; Friedman, 1979; Loftus & Mackworth, 1978). In Locher and Nodine's (1974) experiment scanpaths occurred in about half of the cases but the presence of scanpaths did not increase recognition performance. In fact, Noton and Stark (1971a) originally assumed that the scanpath is not necessary for attaining recognition: "Under more normal viewing conditions, when recognition requires few or no eye movements, the eye movements are assumed to be replaced by shifts of an internal attention mechanism …" (p. 310). In the original experiment stimulus visibility was reduced to force participants to make active and observable eye movements instead of covert attention shifts. Subsequent studies (e.g., Josephson & Holmes, 2002; Pieters et al., 1999; Stark & Ellis, 1981) did not take such measures but still found scanpaths. They carried over the recognition explanation for scanpaths found under quite different, active looking conditions. While the definition of recognition is somewhat debatable (see Goldstein & Gigerenzer, 2002; Liter & Bülthoff, 1998; Mandler, 1980; Ullman, 1996), an old/new recognition test using a signal detection paradigm (Green & Swets, 1966) and mental chronometry (Donders, 1869; Posner, 1978) to determine the time and number of saccades necessary for successful recognition of previously viewed web pages could illuminate the recognition account. Likely, recognition is completed quite early and other cognitive mechanisms then take over which cause the observed patterning. It is also possible that recognition of a web page is attained without dedicated eye movements, parallel to following active strategies of information acquisition.

The observed patterns may well represent more conscious and goal-driven strategies or heuristics of information acquisition (Peterson, Beck, & Vomela, 2007; Pirolli et al., 2003). Perhaps through gradual refinement (Myers & Gray, 2007) one or several scan patterns are developed which satisfy certain criteria such as covering all interesting regions and minimizing saccadic effort (Ballard et al., 1995; Shah & Oppenheimer, 2008). These patterns are kept until the goals and associated information requirements change. However, this would have to be verified by further research.

Conclusion

The occurrence of scan patterns was linked to the individual person (i.e. patterns occurred to a much lesser extent when searching between persons) and the detected patterns were individually different. This can be regarded in line with the scanpath theory. However, occurrence and extent of patterning were variable: not all participants showed patterns meeting the criteria for a scanpath, participants who did on one web site often did not on the other and scanpaths were usually just repeated on a fraction of a person's page visits. Also, most participants showed several different patterns rather than one definite scanpath as described by scanpath theory. While there was a tendency for pattern formation during earlier stimulus exposures, patterns were not primarily formed on the first stimulus exposure as would have been expected according to scanpath theory.

We can conclude that applicability of the scanpath theory does not seem plausible in view of the obtained results. However, the substantial patterning occurring in many cases is remarkable. The detected repetitive patterns were validated against chance occurrence, they did not occur to a large extent between persons, and they were stable across content variations. Often, the same eye movement pattern occurred in 50% or 60% (up to 100%) of the page viewings, involving 4 or 5 (up to 8) regions of interest. This can be considered a rather strong degree of gaze sequence repetition. The question remains what causes the extensive patterning observed in some cases yet near-absence in other cases. We outlined that stimulus recognition as the cognitive background discussed in the literature on repetitive scanning seems unlikely.

In comparison to the previous study by Josephson and Holmes (2002), the present study showed that repetitive scanpath patterns can also occur when page contents change, as is commonplace for example on the home pages of large web sites. Occurrence of patterns was found to be linked to the individual. The temporal origin of scan patterns in web page exposures was determined and continuous new emergence with a trend for formation on earlier page viewings was found. The T-pattern technique allowed exposing and visualizing scanpaths. This

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

for example revealed ongoing refinement and the existence of several different patterns per person.

Future Work

To understand the reasons of pattern development, factors have to be identified that support it. Additional research is required on the determinants of pattern occurrence. It should address such questions as: (1) What distinguishes the persons who develop patterns from the ones who do not? (2) Are there fixed inclinations of certain persons for repetitive scanning or can the inclination change within a person as a function of other factors? (3) Are there explicit or implicit tasks, goals, heuristics, or search strategies that support pattern occurrence? And how does a common task affect the individual distinctiveness of patterns? (4) Are there web pages that support pattern formation? (5) At what points in time does scanning behavior change and what are the triggers?

Effort-reducing heuristics (Ballard et al., 1995; Shah & Oppenheimer, 2008) could be a particularly promising direction for further research. Eye tracking and T-pattern detection could be combined with retrospective think-aloud (Guan, Cuddihy, & Ramey, 2006) in order to elicit user heuristics and search strategies. The onset of new heuristics or search targets could then be checked for temporal correlation with the onset of new scan patterns. If certain determinants can be identified as decisive, it may be possible to infer behavioral rules behind web users' repetitive scanning. These rules could then be validated and contribute to improving web design to better match users' habits, preferences, and cognitive characteristics.

References

Açık, A., Onat, S., Schumann, F., Einhäuser, W., & König, P. (2009). Effects of luminance contrast and its modifications on fixation behavior during *free viewing of images from different categories. Vision Research, 49,* 1541-1553.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7,* 66-80.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143-177.

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience,* 9, 27-38.

Borgatti, E. F., Everett, M., & Freeman, L. C. (1992). UCINET IV (Version 1.0) [Computer software]. Columbia, SC: Analytic Technologies.

Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology, 22,* 225-271.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28-71.

Cockburn, A., & McKenzie, B. (2001). What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies, 54,* 903-922.

Donders, F. C. (1869). On the speed of mental processes. In W. G. Koster (Ed.), *Attention and Performance II. Acta Psychologica, 30,* 412-431. (Original work published in 1868)

Ellis, S. R., & Smith, J. D. (1985). Patterns of statistical dependency in visual scanning. In R. Groner, G. W. McConkie, & C. Menz (Eds.), *Eye movements and human information processing* (pp. 221-238). Amsterdam: Elsevier.

Faraday, P. (2000). Visually critiquing web pages. *Proc. 6th Conference on Human Factors and the Web,* Austin, Texas.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8,* 1-17.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108,* 316-355.

DOI 10.16910/jemr.3.4.5
This article is licensed under a
Creative Commons Attribution 4.0 International license.
ISSN 1995-8692

Fujita, T., Privitera, C. M., & Stark, L. W. (2007). Image-type dependent eigen-regions-of-interest define conspicuity operators for predicting human scanpath fixation. *Computers in Biology and Medicine, 37,* 965-974.

Gbadamosi, J., & Zangemeister, W. H. (2001). Visual imagery in hemianopic patients. *Journal of Cognitive Neuroscience, 13,* 855-866.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75-90.

Gray, W. D., & Fu, W.-T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science, 28,* 359-382.

Gray, W. D., Sims, C. R., Fu, W.-T., & Schoelles, M. J. (2006). The soft constraints hypothesis: A rational analysis approach to resource allocation for interactive behavior. *Psychological Review, 113,* 461-482.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Grier, R. A., Kortum, P., & Miller, J. T. (2007). How users view web pages: An exploration of cognititve and perceptual mechanisms. In P. Zaphiris & S. Kurniawan (Eds.), *Human Computer Interaction Research in Web Design and Evaluation* (pp. 22-41). Hershey: Idea Group Publishing.

Groner, R., & Groner, M. (1982). Towards a hypothetico-deductive theory of cognitive activity. In R. Groner, & P. Fraisse (Eds.), *Cognition and eye movements* (pp. 100-121). Amsterdam: Elsevier.

Groner, R., & Groner, M. (1983). A stochastic hypothesis testing model for multi-term series problems based on eye fixations. In R. Groner, C. Menz, D. Fisher, & R. A. Monty (Eds.), *Eye movements and psychological functions: International views* (pp. 257-274). Hillsdale, NJ: Lawrence Erlbaum.

Groner, R., & Menz, C. (1985). The effects of stimulus characteristics, task requirements and individual differences on scanning patterns. In R. Groner, G. W. McConkie, & C. Menz (Eds.), *Eye movements and human information processing* (pp. 239-251). Amsterdam: Elsevier.

Groner, R., Walder, F., & Groner, M. (1984). Looking at faces: Local and global aspects of scanpaths. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 523-533). Amsterdam: Elsevier.

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. *Proc. CHI.* ACM Press, 1253-1262.

Hassenzahl, M., Kekez, R., & Burmester, M. (2002). The importance of a software's pragmatic quality depends on usage modes. In H. Luczak, A. E. Çakir, & G. Çakir (Eds.), *Proceedings of the 6th International Conference on Work With Display Units (WWDU)* (pp. 275-276). Berlin: Ergonomic.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7,* 498-504.

Hoffman, D. L., & Novak, T. P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing, 60,* 50-68.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 1489-1506.

Jeannerod, M., Gerin, P., & Pernier, J. (1968). Déplacements et fixations du regard dans l'exploration libre d'une scène visuelle. *Vision Research, 8,* 81-97.

Josephson, S., & Holmes, M. E. (2002). Attention to repeated images on the World-Wide Web: Another look at scanpath theory. *Behavior Research Methods, Instruments, & Computers, 34,* 539-548.

Laeng, B., & Teodorescu, D. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26,* 207-231.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10,* 707-710.

Liter, J. C., & Bülthoff, H. H. (1998). An introduction to object recognition. *Zeitschrift für Naturforschung C,* 53, 610-621.

Locher, P. J., & Nodine, C. F. (1973). Influence of stimulus symmetry on visual scanning patterns. *Perception & Psychophysics, 13,* 408-412.

DOI 10.16910/jemr.3.4.5

This article is licensed under a
Creative Commons Attribution 4.0 International license.

ISSN 1995-8692

Locher, P. J., & Nodine, C. F. (1974). The role of scanpaths in the recognition of random shapes. *Perception & Psychophysics, 15,* 308-314.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 565-572.

Magnusson, M. S. (2000). Discovering hidden time patterns in behavior: T-patterns and their detection. *Behavior Research Methods, Instruments, & Computers, 31,* 93-110.

Magnusson, M. S. (2004). Theme (Version 5.0) [Computer software]. Wageningen, Netherlands: Noldus Information Technology.

Magnusson, M. S. (2005). Understanding social interaction: Discovering hidden structure with model and algorithms. In L. Anolli, S. Duncan Jr., M. S. Magnusson, & G. Riva (Eds.), *The hidden structure of interaction: From neurons to culture patterns* (pp. 3-22). Amsterdam: IOS Press.

Magnusson, M. S. (2006). Structure and communication in interactions. In G. Riva, M.T. Anguera, B.K. Wiederhold, & F. Mantovani (Eds.), *From communication to presence: Cognition, emotions and culture towards the ultimate communicative experience. Festschrift in honor of Luigi Anolli* (pp. 127-146). Amsterdam: IOS Press.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252-271.

Mannan, S. K., Ruddock, K. H., & Wooding, D. S. (1997). Fixation sequences made during visual examination of briefly presented 2D images. *Spatial Vision, 11,* 157-178.

Myers, C. W., & Gray, W. D. (2007). Are eye movements involved in cued target recall from repeating spatial contexts? In *D. McNamara, & G. Trafton (Eds.), Proceedings of the 29th Annual Conference of the Cognitive Science Society, Nashville, TN* (pp. 515-520). Hillsdale, NJ: Lawrence Erlbaum.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48,* 443-453.

Noton, D., & Stark, L. (1971a). Scanpaths in eye movements during pattern perception. *Science, 171,* 308-311.

Noton, D., & Stark, L. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research, 11,* 929-942.

Nyström, M., & Holmqvist, K. (2008). Semantic override of low-level features in image viewing – both initially and overall. *Journal of Eye Movement Research, 2,* 1-11.

Obendorf, H., Weinreich, H., Herder, E., & Mayer, M. (2007). Web page revisitation revisited: Implications of a long-term click-stream study of browser usage. *Proc. CHI.* ACM Press, 597-606.

Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004). The determinants of web page viewing behavior: An eye-tracking study. *Proc. ETRA.* ACM Press, 147-154.

Parker, R. E. (1978). Picture processing during recognition. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 284-293.

Peterson, M. S., Beck, M. R., & Vomela, M. (2007). Visual search is guided by prospective and retrospective memory. *Perception & Psychophysics, 69,* 123-135.

Pieters, R., Rosbergen, E., & Wedel, M. (1999). Visual attention to repeated print advertising: A test of scanpath theory. *Journal of Marketing Research, 36,* 424-438.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Privitera, C. M. (2006). The scanpath theory: its definition and later developments. In B. E. Rogowitz, T. N. Pappas, & S. J. Daly (Eds.), *Human Vision and Electronic Imaging XI. Proc. SPIE-IS&T Electronic Imaging, SPIE Vol. 6057* (pp. 87-91). The International Society for Optical Engineering.

Privitera, C. M., & Stark, L. W. (1998). Evaluating image processing algorithms that predict regions of interest. *Pattern Recognition Letters, 19,* 1037-1043.

Privitera, C. M., & Stark, L. W. (2000). Algorithms for defining visual regions-of-interest: Comparison with eye fixations. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22,* 970–982.

Rada, R., & Murphy, C. (1992). Searching versus browsing in hypertext. *Hypermedia, 4,* 1-30.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 3, 372-422.

DOI 10.16910/jemr.3.4.5

This article is licensed under a
Creative Commons Attribution 4.0 International license.

ISSN 1995-8692

Rybak, I. A., Gusakova, V. I., Golovan, A. V., Podladchikova, L. N., Shevtsova, N. A. (1998). A model of attention-guided visual perception and recognition. *Vision Research, 38,* 2387-2400.

Sankoff, D., & Kruskal, J. B. (Eds.). (1983). *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison.* Reading, Mass.: Addison-Wesley.

Schill, K., Umkehrer, E., Beinlich, S., Krieger, G., & Zetzsche, C. (2001). Scene analysis with saccadic eye movements: top-down and bottom-up modeling. *Journal of Electronic Imaging, 10,* 152-160.

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134,* 207-222.

Stark, L. W., & Ellis, S. R. (1981). Scanpaths revisited: Cognitive models direct active looking. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 193-226). New Jersey: Lawrence Erlbaum Associates.

Stark, L. W., Privitera, C. M., Yang, H., Azzariti, M., Ho, Y. F., Blackmon, T., & Chernyak, D. (2001). Representation of human vision in the brain: How does human perception recognize images? *Journal of Electronic Imaging, 10,* 123-151.

Tauscher, L., & Greenberg, S. (1997). How people revisit web pages: Empirical findings and implications for the design of history systems. *International Journal of Human-Computer Studies, 47,* 97-137.

Tobii (2006). *User manual Tobii eye tracker and ClearView analysis software.* Danderyd, Sweden: Tobii Technology AB.

Ullman, S. (1996). *High-level vision.* Cambridge, USA: MIT Press.

Unz, D. C., Werner, B., Mangold, R., & Burmester, M. (2005). Information seeking in the WWW: Detecting T-patterns in eye movements and navigational behaviour. *Proc. Measuring Behavior: Conference on Methods and Techniques in Behavioral Research, Wageningen, The Netherlands.*

Wagner, R., & Fischer, M. (1974). The string-to-string correction problem. *Journal of the Association for Computing Machinery, 21,* 168-173.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception, 6,* 313-326.

Wedel, M., & Pieters, R. (2006). Eye tracking for visual marketing. In J. Eliashberg, T. H. Ho, & M. F. Luce (Eds.), *Foundations and trends in marketing, Vol. 1, Issue 4.* Delft, NL: now Publishers.

Yarbus, A. L. (1967). *Eye movements and vision.* New York: Plenum Press.

Zangemeister, W. H., Sherman, K., & Stark, L. (1995). Evidence for a global scanpath strategy in viewing abstract compared to realistic images. *Neuropsychologia, 33,* 1009–1025.

DOI 10.16910/jemr.3.4.5

This article is licensed under a
Creative Commons Attribution 4.0 International license.

ISSN 1995-8692