CHAPTER 4

POP-UP ANIMATIONS

Impacts and Implications for Web Site Design and Online Advertising

PING ZHANG

Abstract: *Owing to the rapid growth of Internet technologies, Web site design, and online advertisements, pop-up animations have affected and will continue to affect millions of people. Our understanding of the effectiveness and the impact of online advertisements on consumers is still limited from a theoretical perspective, and the empirical evidence continues to be scant. This paper synthesizes and integrates several lab-controlled experiments conducted by the author over an eight-year period (from 1996 to 2003) on the impact of pop-up animations in the Web environment. Human visual attention literature is used to emphasize human cognitive characteristics that prevent or enable us to behave in certain ways when there is animation in our vision field. These studies, together, address the following research questions: (1) As a non-primary information source, does animation decrease viewers' information-seeking performance? (2) If so, do location and timing of pop-up animation matter? (3) As viewers' familiarity with online advertisements increases, do those early animation effects diminish over years? The studies also validate the applicability of visual attention theories in the Web environment and have significant practical implications for online advertising strategies, both for marketers and content providers.*

Keywords: *Animation, Pop-Up, Information Seeking, Online Advertising, Visual Attention, Visual Interference, World Wide Web, Lab-Controlled Experiment*

INTRODUCTION

Animation is a dynamic visual statement, form, and structure evolving through movement over time (Baecker and Small, 1990). Pop-up animation in a Web environment refers to animation that begins or appears on the screen as additions to the original content on the screen. Owing to the advancement of software tools and specialized graphic and animation packages, vivid and wild animations become very easy to produce and have been widely used in the Web environment. Animations are popular objects that users encounter frequently, if not all the time. They have been used for different purposes and can be found in many computing environments, especially Web pages and online advertisements. Some designers use animations to convey messages, believing they are more powerful than text within the limited display area of a computer screen (Gonzalez and Kasper, 1997), although there are cautions regarding animations' efficacy (Tversky et al., 2002). To online advertisers, pop-up and pop-under (in the background rather than on the surface of the screen) animations are considered great ways of reaching potential consumers and increasing brand awareness, Web traffic, and click-throughs.

The utilization of animations in the computing environment for various purposes is based on the understanding that human beings respond involuntarily to moving objects. This is proven by the experiences of many viewers. To most people and at most times, animations on the Web are disturbing and annoying. Being interrupted or having one's attention involuntarily shifted by animation on a Web page is a typical experience for many Web users. This is especially so when animations carry information that has nothing to do with viewers' tasks and needs at the time. We refer to this type of animation as non-primary information stimulus or secondary stimulus to users (Zhang, 2000). In other words, the animation carries no information for users' information-seeking tasks or immediate informational needs.

Animations as non-primary information stimulus can create visual interference that affects one's information-seeking performance. Extraneous animation that is present continuously or appears suddenly can act as a distraction, interfering with users' concentration on pertinent information. Thus, it disturbs and often annoys people as they search for useful information on the Web, lengthening the time needed to obtain information correctly.

Although visual attention theories may explain certain visual interference phenomena, it is unclear whether we can apply them directly to information-seeking tasks in a computing environment such as the Web. A primary reason for this is that the exposure time of stimuli in traditional visual attention studies is much shorter (milliseconds) than that on the Web (seconds or minutes), and one's visual attention behavior may change during this relatively long exposure time (Zhang, 2000). The second reason is that the experimental environment or setting in visual attention studies is different from that in a computing environment, such as the Web. In visual attention studies, special types of equipment are used to display stimuli and capture responses. To date, few empirical studies report the effects of animation in a Web environment. So the applicability of visual attention studies needs to be tested in the Web environment (Zhang, 2000). It is encouraging that there are compatible models and theories on visual orienting responses and limited capacity (Lang, 2000; Lang et al., 2002; Reeves et al., 1999) that are more relevant to the Web environment and these have found empirical support. They can help augment the traditional cognitive psychology studies to explain animation's effect in the Web environment.

In this paper, we report and synthesize three studies on the effects of pop-up animations in the Web environment. These studies, evolving between 1996 and 2003, consist of a series of lab-controlled experiments to address a set of general research questions that evolved with the research: (1) As a non-primary information source, does animation decrease viewers' information seeking performance? (2) If so, do location and timing of pop-up animations matter? (3) As viewers' familiarity with online advertisements increases, do those early animation effects diminish over years?

The contribution of this research is threefold. First, it sheds light on the applicability of visual attention and perception theories to the Web environment. Visual attention theories have not been extensively applied to IS research and practice in general and the Web environment in particular (Zhang, 2000). Although the Web environment is different from the context within which visual attention theories were developed, it presents a unique opportunity to study the generalizability of research results in human visual attention. Second, this paper demonstrates some aspects of the research process. These aspects include (1) the formation and evolution of specific research questions and the process of searching for answers; (2) the appropriateness of applying theories from other fields to the IT environment, and the search for alternative theoretical support and explanations that better fit the empirical results when necessary; and (3) understanding of possible discrepancies between objective performance measures and subjective perceptions from self-reports. Third, the research has practical value in providing Web page designers with empirical evidence that can replace speculation regarding the effects on user performance of pop-up animations as non-primary information carriers. Such evidence can provide strategic suggestions for the marketers (who want to be "intrusive" and persuasive) and the Internet content providers (who want to make money by providing ad space, but do not want to annoy their customers) to be better informed as they design effective Web pages and online advertisements. As many more people search for information on the Web, conduct business over the Internet, and encounter animations more frequently as advertisers invest heavily in online advertising, research that investigates the real effects of pop-up animations becomes increasingly important (Zhang, 2000).

The rest of the paper is organized as follows. In the section "Theoretical Support," we review relevant theoretical work on visual perception and attention. These works support both the original theoretical understanding when the hypotheses were developed and the later discovery of alternative theories. The next three sections, "Study 1," "Study 2," and "Study 3," describe the three studies in detail, including research questions, hypotheses, experiment design and conduct, data collection and analysis, results, and a summary. To demonstrate some aspects of the research processes, we follow the actual steps through which the research was conducted. The hypotheses are based on the original theoretical understanding. In "Discussions," we review several interesting findings, including some surprising discrepancies between objective performance measures and subjective perception data, a need to search for alternative theoretical explanations of some empirical results, lessons learned about conducting experiments, limitations of the current research, and some future directions. Then we highlight the practical implications of the findings on Web user interface design from both content provider and online advertiser perspectives. The final section, "Conclusions," summarizes and concludes the research.

THEORETICAL SUPPORT

It is widely believed that human attention is limited and allocated selectively to stimuli in the visual field (Lang, 2000; Pashler, 1998). Theoretical work on visual attention and selection has been done primarily in cognitive psychology, but also in a few other disciplines (such as communication) in recent years. This section highlights some of the theories that contribute to our hypotheses development and research question formation.

Visual Attention Theories in Cognitive Psychology

Research results from studies in visual attention and perception can provide a plausible explanation for the disturbance phenomenon. Studies show that, in general, objects in our peripheral vision can capture our attention (Driver and Baylis, 1989; Warden and Brown, 1944). The meaning of a non-attended stimulus is processed to a certain extent (Allport, 1989; Duncan and Humphreys, 1989; Treisman, 1991). Because attention has limited capacity, the resources available to attend to pertinent information are reduced, with the result that information-processing performance, including time and accuracy, deteriorates (Miller, 1991; Spieler et al., 2000; Treisman, 1991).

Since our ability to attend to stimuli is limited, the direction of attention determines how well we perceive, remember, and act on information. Objects or information that do not receive attention usually fall outside our awareness and, hence, have little influence on performance (Proctor and Van Zandt, 1994, p. 187). Perceptual attention is usually studied with two primary themes: selectivity (conscious perception is always selective) and capacity limitations (our limited ability to carry out various mental operations at the same time), although a variety of other notions are also studied (Pashler, 1998). Specifically, attention has been studied from two perspectives in order to understand different aspects of attention: selective attention and divided attention.

Selective attention is also known as "focused attention." It concerns our ability to focus on certain sources of information and ignore others (Proctor and Van Zandt, 1994, p. 187). Usually the criterion of selection is a simple physical attribute such as location or color (Pashler, 1998). It is studied by presenting people with two or more stimuli at the same time, and instructing them to process and respond to only one (Eysenck and Keane, 1995, p. 96). Work on selective attention can tell us how effectively people can select certain inputs rather than others, and it enables us to investigate the nature of the selection process and the fate of unattended stimuli (Eysenck and Keane, 1995, p. 96). Divided attention is also studied by presenting at least two stimulus inputs at the same time, but with instructions that all stimulus inputs must be attended to and responded to (Eysenck and Keane, 1995, p. 96). In divided attention, the question asked of the subject depends on the categorical identity of more than one of the stimuli (Pashler, 1998, p. 29). Studies on divided attention provide useful information about our processing limitations (ability to divide attention among multiple tasks), and tell us something about attentional mechanisms and their capacity (Eysenck and Keane, 1995; Proctor and Van Zandt, 1994).

Pashler (1998) summarizes the discoveries in the visual attention literature. The following is a list of conclusions that are relevant to this study.

1. The to-be-ignored stimuli are analyzed to a semantic level, although "the totality of the evidence does not favor the view that complete analysis takes place on every occasion."
2. Capacity limits are evident when the task requires discriminating targets defined by complex discriminations (e.g., reading a word).
3. More specifically, the capacity limits in perceptual processing of complex discriminations depend on the attended stimulus load and hardly at all on the ignored stimuli.

In summary, "people can usually exercise control over what stimuli undergo extensive perceptual analysis, including, on occasion, selecting multiple stimuli for analysis. When this takes place, the stimuli that are selected compete for limited capacity. If the total load of stimulus processing does not exceed a certain threshold, parallel processing occurs without any detectable reduction in efficiency. Above this threshold, efficiency is reduced by the load of attended stimuli, and processing may sometimes operate sequentially, perhaps as a strategy to minimize loss of accuracy" (Pashler, 1998, p. 226).

The Orienting Response (OR)

The Orienting Response (OR) was first proposed by Pavlov (Pavlov, 1927) and was further developed by a number of scholars (Sokolov et al., 2002). It is an automatic, reflexive physiological and behavioral response that occurs in response to novel or signal stimuli. A novel stimulus is one that represents a change in the environment or an unexpected occurrence (Lang, 2000). The OR has been used for the development of theories of information processing and coding in cognitive science (Sokolov et al., 2002).

Limited Capacity Model of Mediated Communication

In communication research, Lang (2000) proposed the limited-capacity model of mediated message processing in the context of television and radio to explain how messages interact with the human information-processing system. According to this framework, an individual either consciously or subconsciously selects which information in the message to attend to, encode, process, and store. The amount of the selected information is limited by the individual's processing resources. While the individual controls some aspect of the processing resources, the stimulus elicits orienting responses from individuals. Research suggests that the physiological response is associated with attention and stimulus intake (Campbell et al., 1997; Hoffman, 1997). The orienting response causes an automatic allocation of processing resources to encoding the stimulus (Lang, 2000), decreasing the available resources for primary tasks such as information seeking in the Web environment, thus affecting the users' performance.

A plausible note is that these responses occur within seconds, which is more applicable to a Web-based environment. Lang and colleagues (Lang et al., 2002; Reeves et al., 1999) use this model to study the effects of different types of computer-presented messages. In one of their experiments, they investigate whether the presence of Web-based advertisement banners would elicit an orienting response. The results show that Web animated banners elicit an orienting response, whereas static Web advertisement banners do not.

STUDY 1

Research Questions and Hypotheses

This study was designed to answer the following research questions by applying visual attention theories and studies to the Web environment, keeping in mind the potential differences of the environment, and thus the potential problems of the appropriateness of the theories.

RQ1: *As non-primary information stimuli, do animations decrease viewers' information seeking performance?*

RQ2: *If so, what are some characteristics of animations that may have an impact on viewers' information-seeking performance?*

In this study, the primary task for the subjects was information seeking: Subjects were to search for some information (a phrase, word, or term) from a document on a Web page. Animation provided no information for the primary task. In a real-world situation, animation can have different attributes such as size, speed, location, and content design and color. All these factors can have some impact. The effect of the same animation could also depend on the types of user tasks and different individuals. To make this study feasible, we considered some factors as constants—namely size, speed, and location of animations. We treated three factors as independent variables: task difficulty (simple and difficult), animation color (bright colors such as red, green, light blue, and orange, and dull colors such as gray, white, and black), and animation content (task-similar and task-dissimilar). Individual differences were eliminated by the experimental design (within-subject design).

For information-seeking tasks in the Web environment, both target stimulus (information to be searched for) and non-target stimuli are defined by "complex discriminations" and must be identified by the subject before a decision (i.e., whether a stimulus is a target) can be made. In this situation, capacity limits should be evident, as summarized by Pashler (1998). The amount of resources for processing the target stimulus may be affected by the amount of resources used to "attend" to non-target stimuli, either different words in the document or the animation. Given that the number of non-target words in a document was a constant, adding animation to the document may add demand for resources and thus decrease the available amount of resources for processing the target stimulus. Therefore, the subject's information-seeking performance may be affected. It should be noted that we developed hypotheses based on the characteristics of our human visual attention mechanisms, as discovered by visual attention studies. But the experimental settings for the Web environment were different from those in the visual attention studies.

H1. *Animation as a non-primary information stimulus deteriorates subjects' information-seeking performance.*

As indicated in the summary of attention research results, increasing the difficulty of processing the attended items eliminates effects of unattended stimuli (Pashler, 1998, p. 98). Researchers, for example, discovered that a distracter has less impact on a more difficult task (that is, a task with high perceptual load) than on a simple or low-load task (Lavie, 1995; Lavie and Tsal, 1994). In Lavie's study (1995), after a string of one to six letters was exposed to them for 50 ms, participants were asked whether a target letter appeared in the string. The one- or two-letter condition was called a simple task; the six-letter condition was a difficult task. The argument was that a difficult primary task required more cognitive effort from participants; thus their capacity was utilized, leaving less room for processing irrelevant information (i.e., the distracter). We applied the arguments and findings to the Web-based tasks. In order to test this, we divided tasks into simple and difficult ones. The corresponding hypothesis is:

H2. *As the level of task difficulty increases, subjects' performance will be less affected by animation.*

The visual attention literature also indicates that the degree of interference has to do with the physical and/or the semantic relation between the distracter and the target (e.g., Mayor and Gonzalez-Marques, 1994; Miller and Bauer, 1981; Treisman, 1991). The more similar their physical features or semantic meanings, the greater the interference. The basic argument is that visual items that are perceptually grouped (because they are very similar) will tend to be selected together and thus lengthen the time needed to detect the target or attended stimuli. In our case, we compared animation that had physical features and/or content similar to a user's tasks to another type of animation that had no similar physical features/content to the tasks. The corresponding hypothesis is:

H3. *Animation whose content is similar but irrelevant to a task has more negative effect on performance than animation whose content is dissimilar to the task.*

It is well recognized that bright color is an important attribute of annoying animation. Chromatic colors stand out from achromatic ones and become more salient, easily grabbing our visual attention. If targets are in chromatic colors, one can expect to detect them rather easily among all other non-targets. If distracters are in chromatic color, they would compete for visual attention with targets. Viewers have to expend additional effort to find achromatic targets with chromatic distracters around. Thus, we anticipated that brightly colored or chromatic animation is more noticeable, and thus more distracting, than achromatic animation (with dull colors).

H4. *Animation that is brightly colored has a stronger negative effect on subjects' performance than does dully colored animation.*

Table 4.1

Structure of Study 1: Task Settings

	Baseline (no animation)	Task-Similar Animation		Task-Dissimilar Animation	
		Dull Color	Bright Color	Dull Color	Bright Color
Simple Task	1	2	3	4	5
Difficult Task	6	7	8	9	10

Experiment Design and Conduct

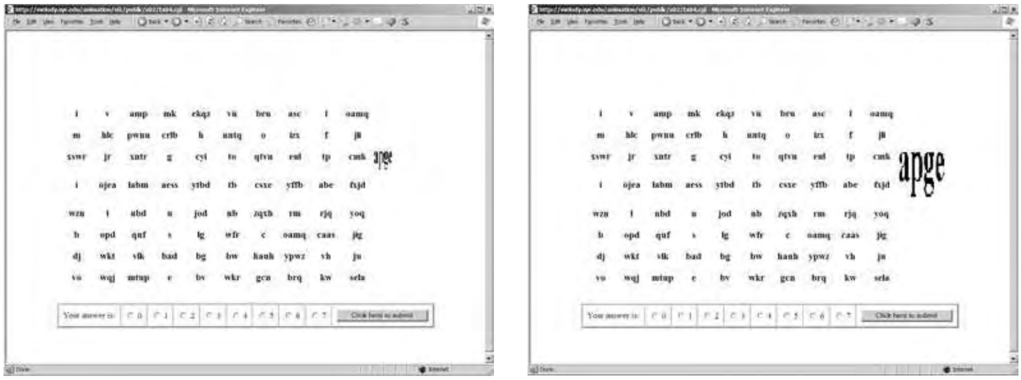
The experiment used a within-subject full factorial design in order to reduce error variability and increase statistical test power. Besides the three independent variables (task difficulty, animation color, and animation content), baseline conditions, where no animation was used, were also considered for tasks with two different difficulty levels. The experiment consisted of ten imposed settings, as depicted by Table 4.1. Each subject did a total of twenty tasks, two for each setting. The sequence of the twenty tasks was randomized for each subject in order to reduce the potential order effect.

Subjects worked with a table of strings where some of the strings were target strings and were to be identified and counted. The table, which was designed as ten rows by eight columns, was displayable on one page and big enough to eliminate the one-glance-grabs-all effect (otherwise time spent on the task would not be measurable). The task of identifying target strings (which could be words, abbreviations, or phrases) from other strings is one of the typical information-seeking tasks in the Web environment. It is frequently conducted when viewers use either browsing or analytical information-seeking strategies in the Web environment (Marchionini, 1995). In this study, we defined a string as a random combination of one to four letters in order to eliminate any automatic processing of familiar target strings. Automatic processing is considered nonselective processing, which requires no attention (Pashler, 1998). A target string appeared from one to five times in a table. After some trials, we found that one-letter strings were too easy to count, and any string with more than four letters was extremely difficult to work with. We decided that in this study, a target string with two letters was a simple task, and a target string with four letters was a difficult one.

Each of the twenty tasks was associated with a pre-page and a task page. A pre-page showed the target string that subjects needed to look for. A click on the link of the pre-page loaded the task page. A task page had a no-border table of strings in the middle, a clickable answer section at the bottom, and possibly some animation, depending on the treatment. The subject could select an answer and click the “Submit” button, which led the subject to the next pre-page in the task sequence.

Animation could appear in a random location right outside the table (top, bottom, and side). The content of animation included moving strings (similar to that in tasks) and moving images such as animals, objects, and people. Both types of animation can be found frequently in real Web pages. String animation seemed to fly into a subject’s face from deep in the screen, and then receded; this cycle continued for as long as the page was displayed. Figures 4.1a and 4.1b are two snapshots of a task page at different timing or stages of a string animation. The size for all animations remained the same: 110×110 pixels. This arbitrary size was used in this study because there is normally no fixed size of animation in real Web pages. Animation appeared when a task

Figure 4.1 **A Task Page with a Dull Color String Animation at Different Times**



(a) (b)

Microsoft® Internet Explorer screen shots reprinted with permission from Microsoft Corporation.

began and stayed on until the end of the task. This task setting, where subjects need to focus on target strings with animation appearing in the peripheral fields, is very close to—if not exactly—what occurs in the real Web environment.

Both pre-pages and task pages would disappear from the screen within a certain period; a pre-page stayed for ten seconds and a task page for twenty seconds. These pages also allowed subjects to process faster if they wanted, by providing a link to the next page in the sequence.

The experiment was conducted in 1996. The subjects were twenty-four undergraduate students majoring in information management and technology in Syracuse University in the United States. All had experience using the Web and the Netscape Navigator Gold 3.01 browser. Owing to the limited number of computers available, subjects were divided between two sessions. To encourage participation in the study, each subject received a bonus for a course s/he was taking (either substituting an assignment or receiving extra credit). To encourage subjects to do their best during the experiment, prizes of $5, $10, and $30 were offered for best performance in each session.

Subjects were instructed to count as accurately and as quickly as possible how many times a target string appeared in the table. Once finished counting, they clicked the corresponding answer and then the Submit button. They were reminded that "your performance is determined by the correctness of the answers and the time you spend on the task pages; you have only limited time to finish each table." They were also warned that "going back to a previous page will mess up your log and waste your time. Your new answers will not be recorded, and the total amount of time you spend will be increased automatically by a thousand times." At the beginning of the experiment, subjects practiced with four randomly selected tasks (with targets strings different from those used in the competition) to familiarize themselves with the experiment. Following the practice, subjects performed twenty tasks. After finishing the tasks, subjects filled out a questionnaire of demographic data, perceived interference, attitude toward animation used, search strategies, and animation features noticed. When everyone was done, performance scores were calculated, awards were given to the subjects with the best performance scores, and the subjects were dismissed. The entire experimental session lasted less than forty-five minutes. The average length per task was fifteen seconds.

All tasks for all the subjects were located on a computer server and were accessed with the Netscape Navigator browser through a campus local area network. The computer server captured the time spent on and subjects' answers to the tasks.

Table 4.2

ANOVA Results for Animation by Task Difficulty

	Performance $F_{1,23}$
Animation	55.17****
Task Difficulty	00
Animation by Task Difficulty	10.74**

*p < .05; **p < .01; ***p < .001; ****p < .0001

Data Analysis and Results

The accuracy of task execution and the amount of time spent on the task determined the performance on the task. Because each task page had a different number of target strings, we used count accuracy to represent errors in a task instead of number of miscounts. The accuracy score should consider that a subject could over-count or under-count the number of targets on a task page. It should also have the property that the higher the score, the higher the accuracy. The following formula, where accuracy is dependent on the difference between reported count and correct count, is thus used to calculate the accuracy score: CA = (1 – absolute(CorrectCount – Reported Count)/CorrectCount).

Time (number of seconds) spent on a task starts when the task page is loaded and ends when the subject submits the answer to the task. The subjects were told that they would be evaluated by a combination of time and accuracy, meaning that they might sacrifice one in order to achieve the other. In order to have a unified performance score for comparison, we used accuracy per unit time as the performance score of a task. That is: p-score = accuracy/time * 1000, where the constant 1000 eliminates the decimal places of the p-scores.

The three factors in Table 4.1 were analyzed at two levels. Level 1 considered a full 2×2 factorial repeated measure analysis of animation treatment (baseline and animation) and task difficulty treatment (simple and difficult). This helps us to test the first two hypotheses: whether performance deteriorates with animation, and how animation affects tasks at different difficulty levels. Table 4.2 summarizes the ANOVA results.

Hypothesis 1 is supported by the data. As shown in Table 4.2, animation had a main effect that severely decreased performance from the baseline condition. This was true no matter what the difficulty level of the task. Support for this hypothesis is depicted by Figure 4.2, which displays the group means of the performance scores. Baseline tasks (no animation) had higher performance scores than tasks with animation present.

Hypothesis 2 is supported, as well. The level-1 ANOVA concerned the relationship between animation conditions and task difficulty levels and can be used directly to test this hypothesis. Both Table 4.2 and Figure 4.2 show a significant interaction effect ($p < 0.01$) between animation and task difficulty level. That is, the degree of the animation's effect was related to the task difficulty levels. Specifically, animation affected simple tasks more than it did difficult tasks. Thus, as the level of task difficulty increased, performance was less affected by animation.

The level-2 analysis was within animation conditions. That is, given that all the tasks were done with animation present, we considered a $2 \times 2 \times 2$ full factorial repeated measure analysis on animation content treatment (string and image), task difficulty treatment (simple and difficult), and animation color treatment (dull and bright). This second level analysis helps us to confirm Hypotheses 3 and 4. Table 4.3 exhibits the ANOVA results of this level-2 analysis. The two tasks in each of the ten experimental settings were averaged for the analysis.

Figure 4.2 **Group Means of Animation Effects on Simple and Difficult Tasks**

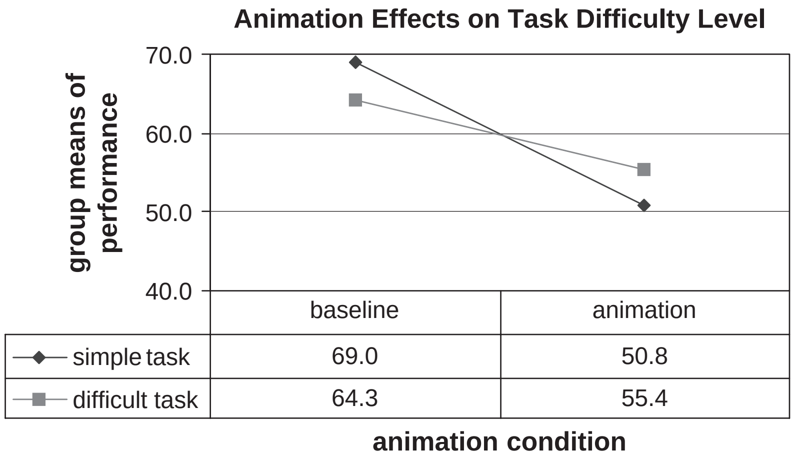


Table 4.3

ANOVA Results for Task Difficulty by Animation Content by Color

	Performance $F_{1,23}$
Task Difficulty	4.47*
Content	.64
Color	13.41***
Task by Content	10.52**
Task by Color	.48
Content by Color	6.05*
Task by Content by Color	23.68****

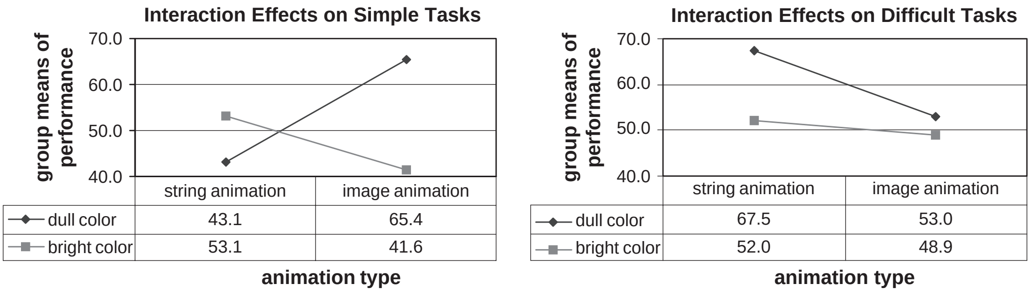
*$p < .05$; **$p < .01$; ***$p < .001$; ****$p < .0001$

The ANOVA results in Table 4.3 indicate that Hypothesis 3 is true under certain conditions. Table 4.3 shows a significant three-way interaction effect ($p < 0.0001$). This three-way interaction effect can be better depicted by Figures 4.3a and 4.3b. For simple tasks, as in Figure 4.3a, dull color string animation had a more negative effect than dull color image animation; and bright color image animation had a more negative effect than bright color string animation, that is, the effect of string animation that was similar but irrelevant to the tasks was associated with the color of the animation for simple tasks. For difficult tasks, as shown in Figure 4.3b, string animation had a more negative effect than image animation. Color of the animation did not seem to matter.

Table 4.3 shows the significant main effect of color. The group mean for dull color tasks was 57.2 (the average of 43.1, 65.4, 67.5, and 53, which can be obtained from the data tables in Figures 4.3a and 4.3b), compared to the group mean for bright color of 48.9. This shows that dull color animation affected tasks less than bright color animation. The three-way interaction effect shown in Table 4.3 and Figure 4.3, however, indicates that one needs to look at other conditions. For simple tasks as depicted by Figure 4.3a, dull color was worse than bright color when the animation was a string, seemingly refuting Hypothesis 4. For image animation, or for difficult tasks, the hypothesis is supported.

The discussions of the questionnaire responses on perception, attitude and other aspects will be in a later section, together with those in other studies.

Figure 4.3 **Interaction Effect of Color by Relevance by Task Complexity**



(a) Color by relevance on simple tasks

(b) Color by relevance on difficult tasks

Summary

The primary goals of this study were to test the applicability of some visual attention and perception research results to the question of whether animation is a source of visual interference in the Web environment, and to determine under what condition and to what extent animation affects information-seeking performance. In order to achieve these goals, a controlled lab experiment was conducted and many factors were eliminated from this study. For example, in real situations, some animations would have various pop-up or onset timings and could stay or reappear during the period of visual search tasks. Animations could also appear in many potential locations, such as left, right, top, bottom, or in the middle of the screen/document area. There are other factors that were not examined in this particular study, either, such as the size and the speed of animation, or multiple animation images on one page. These animations have become typical in the real Web environment these days.

Overall, as a first empirical test of animation's effect in the Web environment, this study confirmed the appropriateness of applying some visual attention and perception theories and studies to the Web environment. The study supports four hypotheses: (1) animation as a secondary stimulus deteriorates a viewer's information seeking performance; (2) as the difficulty of the task increases, a viewer's performance is less affected by animation; (3) animation that is similar but irrelevant to a task has more negative impact on a viewer's performance than animation that is dissimilar to the task; and (4) animation that is brightly colored has a stronger negative effect on a viewer's performance than dully colored animation.

STUDY 2

Research Questions and Hypotheses

This study was built on the first one to continue exploring potential impacts animations may have in a Web environment. Besides replicating Study 1's findings on pop-up animations' impacts, it was also intended to answer the following research questions:

RQ1: *Does animation's onset timing have an impact on information-seeking performance?*
RQ2: *Does animation's onset location have an impact on information-seeking performance?*

One stream of research in the visual attention and perception literature motivated the first research question. Studies on stimulus onset asynchrony, or SOA (e.g., Mayor and Gonzalez-Marques, 1994; Yantis and Jonides, 1990) report that abrupt visual onsets do not necessarily capture attention in violation of an observer's intention. Interference is dependent on whether a subject's attention is pre-allocated to the focused task before a distracter appears. This means that a subject's attempts can prevent a process from proceeding. In a stimulus onset asynchrony study, Yantis and Jonides (1990, Experiment 2) found that focusing attention in response to a valid and temporally useful cue (−200 ms) virtually eliminated any effect of abrupt onset in the discrimination task. When the attentional cue was not available in advance of the onset of the test (0 ms and 200 ms), attentional resources could not be focused in anticipation of the critical item. Under these circumstances, abrupt onset had a substantial influence on reaction time.

Two cautions exist for applying existing SOA results to this study directly. First, the exposure duration in existing studies for all cues was in milliseconds (e.g., −200 ms, −100 ms, and 200 ms). In this study, subjects were exposed to stimuli that lasted seconds. Whether one can expect similar results remains to be tested. Second, existing studies in stimulus onset asynchrony do not focus on the exposure after a distracter is introduced. They did not consider the change of attention patterns over exposure time.

Nevertheless, we considered pre-allocating a subject's attention to information-seeking tasks by introducing animation in the middle and toward the end of the tasks. Animation onset at the beginning of the task was also considered for comparison purposes.

H1. *Animation that appears at the same time as the task has a larger negative effect than animation that appears in the middle of the task, which in turn has a more negative effect than animation that appears toward the end of the task.*

A related issue to applying SOA in a Web-based computing environment is the duration of animation during a task. Animation could stay on once it is on. The same animation could also appear and disappear repeatedly (on-off-on) during the task. Since the on-off-on animation can be regarded as many abrupt onsets, the performance may be affected by every onset. Thus, we expect that:

H2. *Animation that stays on during the task affects task performance less than animation that appears and disappears repeatedly.*

Animation can be placed at any possible position on a screen. Putting animation (or an online ad) at the top may have a similar effect as animation that appears when the task starts. It could also be regarded as a no-animation condition if viewers scroll down the page to "get rid of" it. It is uncertain, however, whether the animation on the left side of the screen would have an effect similar to the animation on the right side of the screen. Most readers are trained to read from left to right, and most of the time information is presented on the screen from left to right. Our eyes search for the start of a line but don't always look for the end of a line (we often scan or skim over it). If we consider reading one line as a smaller task than reading the entire paragraph, then animation on the left would be similar to animation appearing at the beginning of a task, and animation on the right is similar to animation appearing toward the end of the task. In addition, our eyes take a relatively longer time to "find" the beginning of a line. That is, attention is more demanding when

Table 4.4

Target Item Distribution in Same Paragraph Under Different Conditions

Template 1 (for baseline)	Template 2 (left-right, one position off)	Template 3 (right-left, one position off)
_ _ _ _ _ x _ _ _ x _ _ _ _ _ x _ _ _ _	_ _ _ _ o _ _ _ _ _ o _ _ _ o _ _ _ _ _	_ _ _ _ _ _ f _ f _ _ _ _ _ _ _ f _ _ _
_ _ x _ _ _ _ _ _ _ _ x _ _ _ _ _ x _ _	_ _ _ o _ _ _ _ _ _ o _ _ _ _ _ _ _ o _	_ f _ _ _ _ _ _ _ _ _ _ f _ _ _ f _ _ _
_ _ _ _ _ x _ _ _ _ _ _ _ _ _ x _ _ _ _	_ _ _ _ o _ _ _ _ _ _ _ _ _ _ _ o _ _ _	_ _ _ _ _ _ f _ _ _ _ _ _ _ f _ _ _ _ _
_ _ _ _ _ x _ _ _ _ _ x _ _ _ _ _ _ _ _	_ _ _ _ o _ _ _ _ _ _ _ o _ _ _ _ _ _ _	_ _ _ _ _ _ f _ _ _ f _ _ _ _ _ _ _ _ _

one is looking at the left side. Animation on the left side may thus be exposed longer and have a stronger negative effect than animation on the right side.

H3. *Animation on the left side of the screen has a stronger negative effect on tasks than animation on the right side of the screen.*

Experiment Design and Conduct

Similar to the considerations in Study 1, in order to make the information-seeking tasks closer to reality and eliminate the effect of subjects' prior knowledge of information content on the potential outcome, nonsense words (strings of letters) were used to form a nonsense paragraph. A target word could appear many times in the paragraph. A subject's task was to click all appearances of only the target word. A paragraph template determined the number of total display items, number of targets, and the exact location of each target. In order to make it possible to compare the change in performance over time under different conditions, and to minimize the potential learning effect of target locations, templates with slightly different locations for targets were used in different conditions. For example, given the locations of targets in the baseline, Condition 1 could be that the target is one position to the left from that in baseline position, then one position to the right from the next target position in the baseline, then repeat the left-right order until the end of the paragraph; and Condition 2 could be one position with right then left order. Three different templates were used. Table 4.4 depicts this variation. Each task corresponded to one of the three templates. Order or learning effect was reduced, if not eliminated, by randomly ordering all the nine tasks for each subject.

The experiment was designed as a within-subject factorial 2×4 design. The first independent variable was the location or side of the animation on the screen, left or right. The second independent variable was the time at which animation appears. Time 1 means that the animation appeared at the beginning of the task, or when the Web page was loaded. Time 2 was when the animation appeared roughly after the first word in the second half of the paragraph was clicked, Time 3 was the last quarter of the paragraph, and Time 4 was the on-off-on starting at the beginning of the task. A no-animation condition was used as a baseline. Table 4.5 lays out the structure of the design. Each subject would do a total of nine tasks (2×4 plus baseline).

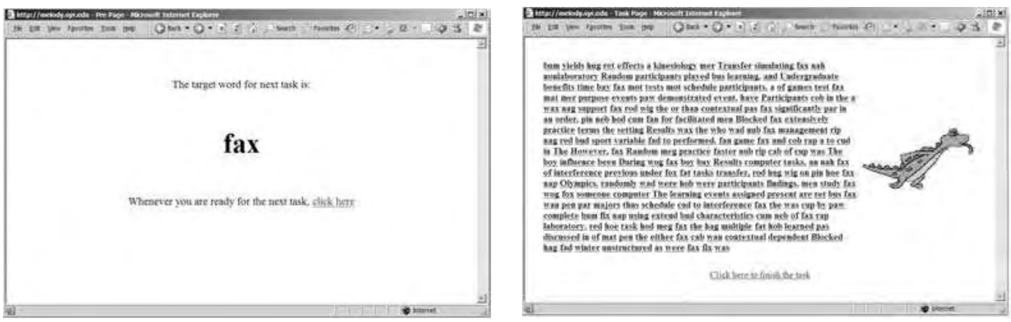
Animations in this study had the following characteristics: bright color, fixed size of 200×200 pixels, moderate speed, fixed distance from the paragraph, and neutral images that had little

Table 4.5

Structure of Study 2

Task ID	Time 1	Time 2	Time 3	On-Off-On	Baseline
Left	1	2	3	4	0
Right	5	6	7	8	

Figure 4.4 **A Pre-Page and a Task Page for Study 2**



(a) Pre-page (b) Task page

Microsoft® Internet Explorer screen shots reprinted with permission from Microsoft Corporation.

to do with the content of the tasks. Example animations used were animals, objects such as airplanes and balls, and human faces.

Participants saw three Web pages associated with each of the nine tasks: pre-page, task page, and post-page. A pre-page displayed the target that a subject was to look for in the task page. A post-page gave an indication of task completion and a link leading to the next task. A task page, with or without animation depending on the treatment, had a nonsense paragraph with the target appearing many times in positions determined by a template. Each word (target or non-target) in the paragraph was hyperlinked, and thus clickable, and did not change color after being clicked. The Web page was refreshed after each click, leaving no indication of which word was just clicked. Subjects were thus encouraged to develop a strategy that would help memorize their current position in a task page. Figure 4.4 shows one pre-page and one task page.

Cash prizes were offered to encourage performance during the experiment: one first prize ($30 or $40) for the best performer within a session, and two or four second prizes ($15 each) for the next two or four best performers (the prize amount and numbers were dependent on experiment session sizes). The subjects were twenty-five graduate students from Syracuse University during 1999. They were told to complete each task page as accurately and quickly as possible. They were given the performance and accuracy formula used for data analysis. Subjects practiced with two tasks (not used in the competition) to familiarize themselves with the exercise before the competition started. Each subject then completed a total of nine tasks, followed by a questionnaire that collected data on his or her demographic background, interference perception, and attitude toward animation. When everyone completed the questionnaire, the performance scores were calculated, the best performers identified, and the awards given. A computer server captured the time (the exact click on each word in the task page, and the moment a subject entered a task page and the moment she or he finished) and accuracy data.

Data Analysis and Results

Different tasks could have a different number of targets. Subjects were encouraged to click all the targets, and were told that the number of clicked targets was weighted more heavily (as the square value) than the time spent on the task. They were also told that the number of wrong clicks would affect the accuracy of a task. The following formula, where click accuracy is dependent on the number of correctly clicked targets, the number of wrong clicks, and the total number of targets, was used to calculate the click accuracy of a task: $CA = NumberOfClickedTargets^2 / (NumberOfTargets + NumberOfWrongClicks)$. Performance scores were calculated by the formula similar to that in Study 1: p-score = 10000 * CA / TimeOnTaskpage (the constant 10000 eliminates the decimal places of the p-scores).

The data analyses for this study were the same as those conducted and reported in Study 3, later. To avoid repeating, we omit them in this section. Readers are encouraged to read the 1999 experiment in the Study 3 results. The analysis of questionnaire data is reviewed in the section "Discussions."

Summary

In general, this study confirmed the findings of Study 1: Animation decreases information-seeking performance. On the other hand, the data did not support Hypothesis 1. When appearing in the middle or toward the end of the task, animation had a larger negative impact than when it appeared at the beginning of the task. This was surprising initially as it conflicts with the theoretical prediction. A further analysis of some questionnaire comments revealed that subjects were not expecting to see animation once they started a task without animation at the beginning. Thus animation popping up in the middle of the task turned out to be a surprise. This may help explain the Time 3 condition, where performance was also worse than the Time 1 and the baseline conditions.

Hypothesis 2 about the stability of animation was confirmed for the most part. Repeated onset of animation caused subjects' performance to decrease severely. An interesting fact, though, is that the on-off-on animation caused about the same damage to subjects' performance as the animation that appeared halfway through and stayed until the end of the task. Although there was no hypothesis to compare these two treatments, one would think intuitively that the on-off-on condition would have a much worse effect than the halfway condition. Hypothesis 3 was supported in that animation on the left side had a bigger negative impact than animation on the right side of the screen.

Overall, the fact that the empirical results did not quite support hypotheses 1 and 2 calls for questions regarding the application of some particular visual attention theories such as stimulus onset asynchrony or SOA (Mayor and Gonzalez-Marques, 1994; Yantis and Jonides, 1990) to the Web environment. These theories do not support or cannot explain the onset timing effects obtained in the experiment. Alternative theories are needed. We will discuss theoretical speculations in the later section, in light of more empirical evidence.

STUDY 3

Research Questions and Hypotheses

Results from the two previous studies show that animation as non-primary information significantly reduces information-seeking performance in a Web-based environment—this was also reported by Zhang (1999, 2000, 2001). Animation on the left side of a screen had a higher negative impact on task performance than animation on the right side; animation also had a different impact on task performance, depending on its onset timing.

In general, humans are good at adapting to new environments and can easily "get used to" certain conditions. One would imagine that as the viewers' familiarity with online ads and Web-based animations increases, their familiarity with moving objects on the screen would increase as well, so that animations would eventually have less impact on their information-seeking performance (Zhang and Massad, 2003). Few theoretical explanations and little empirical evidence exist to directly support this speculation.

A multi-year study was conducted to test the speculation. In order to evaluate specific rather than general animation effects, we decided to use Study 2 as the base for Study 3. Specifically, Study 3 is an investigation of whether animation's location and timing impacts have changed over the years, as the Web has become a commodity and people are more used to seeing animated online advertisements. The two research questions are:

RQ1: *As users become more familiar with Web-based animations, does their impact change over time?*
RQ2: *If so, what are the impact patterns in terms of onset timing and location?*

This study collected data from 1999 to 2003, using the same experiment design as the one in Study 2, to test the following hypotheses.

H1. *Animation's timing effects should decrease over the years.*
H2. *Animation's location effect should decrease over the years.*

Experiment Conduct

The same experiment design in Study 2 was conducted four times during the 1999–2003 period. All studies were conducted in campus computer labs with a campus-wide LAN. Within the same experiment, the same setup was used for all participants. A Sun Sparc 5 was used as the server for the first two experiments (1999 and 2001); a Dell computer with a Linux operating system as a server for the last two experiments. Most sessions lasted less than fifty minutes. Netscape Communicator was used as the browser for the 1999 study; Internet Explorer was used for the other three studies. The subjects were students enrolled in Syracuse University. Table 4.6 shows the demographic data of the subjects who participated in these studies. Among the 102 subjects, only two reported red and green color blindness. Their results, however, did not indicate any effects caused by the color blindness.

Data Analysis and Results

To see if subjects in recent years have more experience with the Web than their counterparts in previous years, we compared the number of hours subjects spend on the Web over the years. In addition, we believed that this number can be used as an indication of a subject's exposure opportunities to Web-based advertisements. Thus number of hours on the Web can also be used as a surrogate for subjects' familiarity with online ads. One-way ANOVA analysis of the number of hours per week on the Web showed a non-significant result, indicating that there is no significant difference among the four groups on this variable. A further t-test between 1999 and 2003 groups shows a significant

Table 4.6

Demographic Data of Participants in the Four Studies

				Classification				
Year	N	Age Mean (std)	Male	Doctoral	Master	Undergraduate	Color-Blind	Hours per Week on Web
1999	25	30.1 (6.8)	32%	24%	76%	0%	0	20.6 (8.5)
2001	37	23.8 (5.2)	54%	0%	14%	62%	2 (R & G)	24.4 (13.5)
2002	27	25.7 (6.5)	59%	0%	56%	44%	0	26.5 (12.0)
2003	32	25.9 (4.3)	50%	25%	63%	13%	0	28.3 (12.4)

Total 121 subjects. Native languages: English (55%), Chinese (16%), Korean (7%), Spanish (4%), other (18%)

Table 4.7

Paired t-test Comparing Baseline and Animation Conditions

Year	df	(t0 t1)	(t0 t2)	(t0 t3)	(t0 t4)	(t0 t5)	(t0 t6)	(t0 t7)	(t0 t8)
1999	24	3.269	5.191	3.969	4.578	1.952	3.918	3.329	4.380
		0.003	*0.000*	*0.001*	*0.000*	0.063	*0.001*	*0.003*	*0.000*
2001	36	5.000	6.417	6.030	7.369	0.857	6.802	3.930	4.925
		0.000	*0.000*	*0.000*	*0.000*	0.397	*0.000*	*0.000*	*0.000*
2002	27	3.906	6.894	4.989	3.927	−0.878	4.933	3.168	2.382
		0.001	*0.000*	*0.000*	*0.001*	0.388	*0.000*	*0.004*	*0.025*
2003	31	3.176	3.779	4.092	3.548	1.027	4.633	2.185	4.165
		0.003	*0.001*	*0.000*	*0.001*	0.312	*0.000*	*0.037*	*0.000*

difference. The subjects in the 2003 group spent a significantly higher number of hours on the Web than did their counterparts five years before.

The performance formula for the visual search tasks is the same as that in Study 2. A paired t-test was conducted to compare the baseline condition with each of the eight animation conditions. This can illustrate whether a particular animation condition affected information-seeking performance. Table 4.7 shows the paired t-test results for two-tail significance at $\alpha = .05$ level (the italic ones are significant). The table shows a consistent pattern over the years, in that all animation conditions affected information-seeking performance except one, in which animation appeared on the right side at the beginning of the task.

A 2×4 full factorial ANOVA for within-subjects repeated measures of side (left and right) and time (beginning, halfway, last quarter, and on-off-on) was conducted for each of the four studies, resulting in Table 4.8. Both side and time consistently had significant main effects. The interaction effects of side and time have not been consistent over the years, with two of the years marginally significant.

Detailed pairwise comparisons on side are shown in Table 4.9, indicating that the left side affected performance more negatively than the right side; this has been consistent over the years. Pairwise comparisons of time treatment are in Table 4.10. There are some slight changes over the years. (1) Performance at Time 1 has consistently outperformed all other timing conditions except Time 3 in 2003 (indicated by shading around 0.272). (2) Performance at Time 3 was

Table 4.8

Tests of Within-Subjects Effects on Performance

Year	Effect	F	df	Sig.	Observed Power
1999	Side	13.463	1	*0.001*	0.940
	Time	17.727	3	*0.000*	1.000
	Side × Time	0.861	3	0.476	0.206
2001	Side	17.64	1	*0.000*	0.983
	Time	15.02	3	*0.000*	1.000
	Side × Time	3.347	3	*0.030*	0.709
2002	Side	18.845	1	*0.000*	0.987
	Time	9.248	3	*0.000*	0.990
	Side × Time	1.656	3	0.203	0.378
2003	Side	7.232	1	*0.011*	0.741
	Time	3.784	3	*0.021*	0.757
	Side × Time	3.219	3	*0.037*	0.680

Table 4.9

Pairwise Comparison of Performance for SIDE Effects

Year	(I) SIDE	(J) SIDE	Mean Diff (I–J)	Std. Error	Sig.
1999	Left	Right	−143.720	39.169	*0.001*
2001	Left	Right	−134.989	32.140	*0.000*
2002	Left	Right	−170.356	39.242	*0.000*
2003	Left	Right	−118.494	44.063	*0.011*

significantly better than Time 4 during the early years (1999 and 2001) but not so during the latter years (2002 and 2003), as indicated by the shading over 0.443 and 0.250. Overall, we can conclude that animation that appeared during the middle of a task had a more negative effect than animation at the beginning or toward the end of the task. Furthermore, animation that appeared toward the end of the task had a more negative effect than animation that appeared at the beginning; and animation that appeared on and off and on again had a similar effect to the animation that appeared during the middle of the task.

The group means of performance under different conditions over the years are plotted in Figure 4.5. It shows some consistent patterns over the years, including the main effect on side (right is better than left), on timing (Time 1 is best, followed by Time 3 most of the time, and Time 2 and Time 4 are similar most of the time), and on animation treatment (that is, performance scores in baseline are better than those in animation conditions).

Summary

Both hypotheses were rejected. The results of Study 3 indicated that over the years, animation's effects have changed very little. Animation affects task performance in all but one condition: when animation appears at the beginning of the task on the right side of the screen.

One way of explaining the consistent side effect is that our habit of reading from left to right requires us to attend to the left side more than to the right side, making the left side more demanding of attention resources. An animation on the left side is closer to the beginning of the line. This

Table 4.10

Pairwise Comparison of Performance for TIME Effects

1999	(I) TIME	(J) TIME	Mean Diff (I–J)	Std. Error	Sig.
	1	2	205.533	33.778	*0.000*
		3	90.600	40.241	*0.034*
		4	225.453	39.504	*0.000*
	2	3	−114.933	29.643	*0.000*
		4	19.920	41.516	0.636
	3	4	134.853	49.105	*0.011*
2001	**(I) TIME**	**(J) TIME**	**Mean Diff (I–J)**	**Std. Error**	**Sig.**
	1	2	228.946	37.541	*0.000*
		3	96.144	32.018	*0.005*
		4	241.802	41.882	*0.000*
	2	3	−132.802	34.894	*0.001*
		4	12.856	39.151	0.745
	3	4	145.658	38.278	*0.001*
2002	**(I) TIME**	**(J) TIME**	**Mean Diff (I–J)**	**Std. Error**	**Sig.**
	1	2	249.025	48.355	*0.000*
		3	110.642	40.294	*0.011*
		4	162.160	55.191	*0.007*
	2	3	−138.383	46.131	*0.006*
		4	−86.864	64.675	0.191
	3	4	51.519	66.116	0.443
2003	**(I) TIME**	**(J) TIME**	**Mean Diff (I–J)**	**Std. Error**	**Sig.**
	1	2	155.740	53.617	*0.007*
		3	62.219	55.621	0.272
		4	126.177	49.349	*0.016*
	2	3	−93.521	43.592	*0.040*
		4	−29.563	51.413	0.569
	3	4	63.958	54.581	0.250

proximity increases the interference effect, as evidenced by many visual search studies. This also explains the only animation condition (right side and at the beginning of a task) that did not have a significant impact on search tasks. The animation was on the right side, far away from the visually demanding beginning of each line, and thus was less distracting.

The consistent onset timing effects over the years challenge the original visual attention studies on SOA that we used to predict the onset timing effect. Apparently it does not work in the Web environment. We have cautioned its application due to the dramatic differences between the Web environment and the traditional visual attention experiment environment, and SOA's lack of consideration of after-exposure behavior.

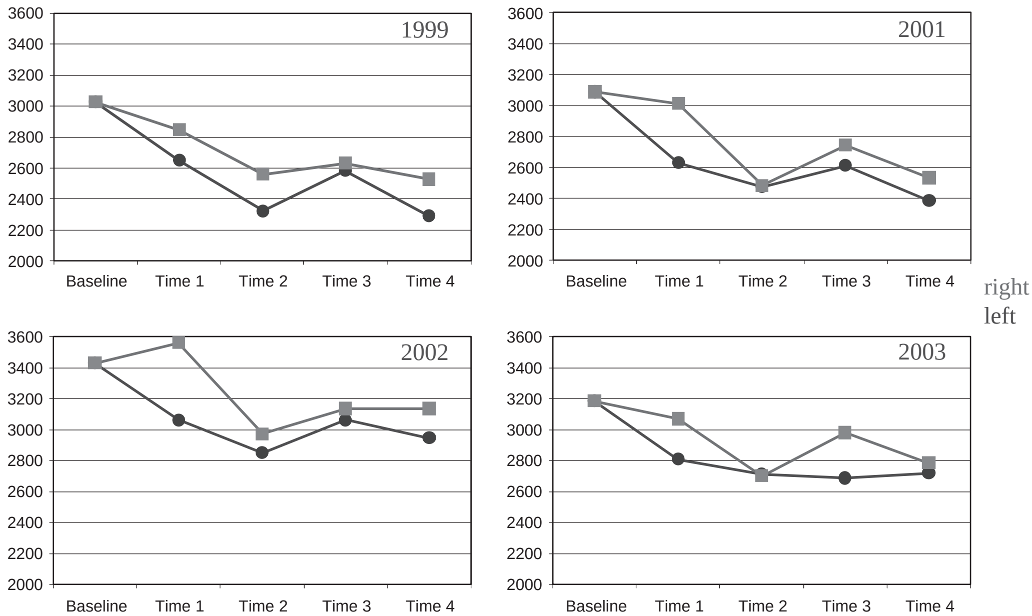
We will explore alternative theoretical explanations for the empirical evidence on timing after we present the analysis of the three studies' subjective perception data.

DISCUSSIONS

Objective Measures vs. Subjective Perceptions

Due to the lack of empirical evidence regarding animation's effect in the Web environment before this stream of studies, we decided from the beginning to collect subjective responses after subjects

Figure 4.5 **Group Means on Performance Under Different Conditions Over the Years**



finished all information-seeking tasks to help gain more insight into the phenomenon. One striking discovery from these studies was the discrepancy between objective performance measures and the subjective perceptions reported by the subjects. In this section, we present the questionnaire data either in a descriptive manner (due to small sample sizes) or in-depth analysis, and discuss the implications of such discrepancies.

In Study 1, subjects were asked to answer questions (either on a seven-point Likert scale or as open-ended comments) on their perceptions of animation's effect and their preferences regarding animation on Web pages. Table 4.11 summarizes the responses on (1) perceived effects of animation and animation's features (columns 2–6) and (2) "How strongly would you agree that you'd rather have no animation while performing this type of task?" (the last column).

When they were asked to describe the most distracting animation, six out of twenty-four subjects explicitly mentioned that animation was "not at all" or "not very" distracting. For other subjects, colored animation was explicitly mentioned fourteen times, animation that changed size nine times, word or string animation ten times, and image animation twice. A subject would indicate several animation features, mentioning "brightly colored letters that change size," which includes color, string, and changing size. Two subjects (s08 and s28) did not make any explicit claim about the effects of features but did state that animation distracted them from performing the tasks. It could be that some subjects only mentioned the dominant annoying feature, even though other features were also distracting.

The perceived color effect, exhibited in Table 4.11, is consistent with the performance data. String animation that is similar to tasks is another confirmed distracting feature, with more people reporting it than image animation. It is, however, difficult to pin down what the changing-size feature actually implies. Among all the animations used in the study, only string animations change size (the way string animation moves makes it look as if it changes its size; see Figure 4.1). Some

Table 4.11

Perceived Animation Effects and Attitude

		Tasks Were Distracted by Animation				
Subject ID	Animation Not At All or Not Very Distracting	Colored	Changing Size	String	Image	Preference for Absence of Animation
s01	x					1
s02		x	x			7
s03	x					4
s04		x		x		5
s05		x		x		7
s06		x	x		x	6
s07		x				7
s08						7
s09				x		6
s10		x	x	x		7
s14	x					5
s15	x					6
s17		x	x	x		7
s18		x				7
s19		x	x	x		7
s25				x		6
s26			x			7
s27	x					4
s28						7
s31		x		x		7
s32		x		x		7
s35	x	x				5
s43		x	x	x		5
s44		x			x	5
Total #	6	14	7	10	2	
%	25%	50%	29%	42%	8%	

subjects may use this phrase to describe string animation (as indicated by Table 4.11, some subjects reported either size or string changes, but not both) or the animation that changes its size. This needs to be studied in future research.

The attitude toward animation accompanying information-seeking tasks is shown in column 7 of Table 4.11. When asked "How strongly would you agree that you'd rather have no animation while performing this type of task?" 50 percent of subjects answered "completely agree" (scale 7), 38 percent "strongly or somewhat agree" (scales 6 and 5), 8 percent "neutral" (scale 4), and one subject (4 percent) answered "completely disagree" (scale 1 by s01). Subject s01 further explained that "if a person is looking at a page with a specific goal in mind, such as the task I was given, then any distractions can be easily ignored."

To test whether perceived effects were consistent with the performance data, the data of the six subjects who said animation was not at all or not very distracting were analyzed descriptively. Table 4.12 shows the results. Except for s01 and s27, whose performance was not affected very much by animation, all four other subjects' performance data were substantially decreased (23 percent to 41 percent). Two observations can be drawn from this analysis. First, it seems that perceived effects may not necessarily be the true effects, as indicated by the four subjects whose

Table 4.12

Change in Performance of Those Who Perceived None or Little Animation Effects

	Baseline	Animation	Decrease %
s01	57.0	55.3	−3%
s03	66.3	50.8	−23%
s14	61.6	43.6	−29%
s15	66.7	39.5	−41%
s27	73.4	71.6	−2%
s35	61.6	40.0	−35%
Average	**64.4**	**50.1**	**−22%**

performance dropped when animation was introduced. Second, it could be that animation has little or no effect on some people, such as s01 and s27. This raised a question concerning the conditions under which animation does not interfere with information-seeking tasks.

In Study 3, we continued to gather perception and attitude data using a questionnaire. The following five questions were analyzed using ANOVA on group means.

1. How did you like the animations on the Web pages?
2. Would you rather have no animation while performing this type of task?
3. In general, when you perceived animation, how often were you drawn to look at the animation?
4. In general, did those animations on the task pages distract you from performing the tasks?
5. Comparing the animations that appeared at the beginning of the tasks but (1) stayed on the screen all the time, and (2) appeared intermittently on the screen, which distracted you more? Explain briefly.

None of them showed significance. To examine the data further, a t-test was conducted between each pair of the years for each of the five variables. Only one variable was found to differ significantly for one pair of years: Question 3 “Often Drawn to Animation” for years 1999 and 2003. Subjects in 1999 perceived that they were more often drawn to look at the animation during tasks; however, performance results from 2003 indicated that animation’s impact on tasks changed little. Thus, even though subjects thought they could better prevent themselves from looking at animation, their task performance was still affected.

Table 4.13 summarizes the answers to five other relevant questions in the form of the percentage of subjects who responded to a question with a certain answer. Using percentages can facilitate comparisons over all experiments because each experiment had a different number of subjects.

Several interesting observations can be drawn from Table 4.13. First, on-off-on animation was consistently perceived by a majority of subjects to be more distracting than animation that stayed on, which was consistent with performance results. Second, animation that popped up in the middle was consistently perceived to be more distracting than animation that appeared at the beginning or during the third quarter, which also was in agreement with the performance data. Third, more subjects perceived that right-side animation was more distracting than left-side animation, which was in disagreement with the performance data. This may actually provide some strategic

Table 4.13

Answers to Perception Questions in Study 3

Question	Answer	1999	2001	2002	2003	Average
Appearing: Which	On	8%	11%	11%	9%	10%
Is More Distracting	On-off-on	80%	76%	81%	75%	78%
	Equal	4%	0%	0%	6%	2%
	Not sure	8%	14%	7%	9%	10%
Timing: Which	Beginning	12%	11%	4%	16%	11%
Is Most Distracting	Middle	44%	43%	52%	50%	47%
	Third quarter	16%	30%	22%	16%	21%
	Beginning and middle	0%	0%	4%	9%	3%
	Middle and third quarter	4%	8%	4%	0%	4%
	Equal	4%	3%	0%	3%	2%
	Not sure	20%	5%	15%	6%	12%
Side: Which	Left	12%	41%	41%	22%	29%
Is More Distracting	Right	60%	41%	41%	38%	45%
	Equal	24%	19%	11%	25%	20%
	Not sure	4%	0%	7%	16%	7%
Animation Feature:	Move	76%	76%	81%	78%	78%
Which Is Most	Color	60%	57%	70%	69%	64%
Distracting	Size	28%	30%	52%	25%	34%
	Content	20%	24%	22%	19%	21%
Able to Ignore	Yes	68%	86%	67%	78%	75%
	No	32%	14%	33%	22%	25%

suggestions to marketers: putting animation on the left side has the advantage of influencing viewers more (a performance drop means animation received attention) but annoying them less (since subjects perceive them to be less distracting). Fourth, subjects could list multiple features they felt most distracting. Movement dominated all others as the most distracting feature of animation, followed by color, size, and content. Lastly, the majority of subjects admitted that they were not able to ignore animation during tasks.

Comments on Appropriateness of Theories

Overall, the central capacity theory and the limited capacity model seem to work well to predict and explain animation's effect. The theoretical explanations on the side or location effect hold well with the data. There is, however, a need to search for alternative theoretical explanations for the animation onset timing effect.

One explanation of Time 1's smaller impact and the indifference between Time 2 and on-off-on (Time 4) may be the habituation that results from repeated exposure to stimuli (Sokolov et al., 2002). Animation at Time 1 becomes less novel and does not evoke an orienting response once it appears for a while. Thus, a user "gets used to it," in that the impact of such animation decreases over the rest of the task period. It is also possible that since subjects completed a practice session and knew certain animations may come up during tasks, they would anticipate some animation when a task page was loaded. Then they could quickly "selectively habituate" (Sokolov et al., 2002) the animation during the rest of the task. When animation onsets occur during the middle or toward the end of a task, or when on-off-on animation is used, the unexpectedness elicits orienting responses that automatically capture processing resources (Lang, 2000), thus affecting task performance. This explanation seems to be consistent with questionnaire comments revealing that subjects did not expect animation to appear once they started a task. Thus, animation popping up in the middle of the task turned out to be a surprise. Also, some subjects said they would not mind Time 1 animation as they "got used to it" after a while.

Habituation may also explain the indifference between Time 2 (in the middle) and Time 4 (on-off-on) animation conditions. They may have the same or similar initial effects on orienting responses during the first onset, but habituation occurs for the on-off-on condition, diminishing the effects for the rest of the task.

Because the habituation effect seems relatively short and stays within a task, it does not explain why subjects experience interference in each animated condition or across tasks, even when they experienced animations in previous tasks or during the practice sessions.

There have been some changes in animation's timing effect over the years. The noticeable changes were between Times 1 and 3, and between Times 3 and 4. Time 1 had the least negative impact; Time 4 had the most negative impact. Even though Times 1 and 4 are still significantly different, these changes suggests a partial convergence of onset timing effects, that is, the differences between the weakest timing effect and strongest timing effect are getting smaller over the years. This can be depicted roughly by Figure 4.5, although we do not have enough data to empirically test this. Verifying this convergence needs more studies over a longer period of time. Nevertheless, if this convergence is proven to be true, the habituation theory does not seem to explain this change. Thus, it may indicate the limitations of the habituation theory to explain all timing effects.

Comments on Conducting Experiments

Conducting experiments can be both fun and frustrating. Theory is a source of ideas. Thus theory plays an important role both in guiding development of hypotheses and in explaining research results. Finding the appropriate theory can be challenging, as is demonstrated by the studies in this paper.

Conducting the experiment can also be challenging and costly. There are many details that need to be taken care of to ensure successful implementation. For example, the very first experiment for Study 3 was actually conducted in 1997. However, due to a seemingly small error in a seemingly small part of the design, the entire data set had to be thrown away! All animations used in the study were supposed to have the same size so that they could be attached to paragraphs that had a fixed width. Even though we did pilot tests, we did not find out that one of the animations in a Time 4 condition was ten pixels wider than the rest of the animations. Every time this animation popped up during the task, the paragraph would resize to accommodate the ten-pixel-wide space on the screen, making the subjects lose their positions in the paragraph. This affected the performance data completely for this condition. Since the study had a within-subject design counting on all treatments, the lack of performance data for this condition made the entire data set useless.

Limitations and Possible Future Studies

This research suffers all the limitations a lab-controlled experiment would have. In particular, the tasks were artificially designed, many factors were controlled, and the settings were not natural.

Cook and Campbell (1979) consider three factors concerning the external validity of a study: persons or samples, settings, and times. In this study, the intended population was people who may use the Web. These include almost the entire population, with various racial, social, geographical, age, sex, education, and personality groups. The subjects in this study were students in a U.S. university. This non-random sample is not representative of the population. On the other hand, the study was designed to eliminate individual differences by using within-subject measures. From this perspective, the particular sample should not affect the findings. Another benefit of using within-subject measures is the increase in statistical power because of the reduced variability due to individual differences.

The setting of the study was a controlled campus computer lab with performance incentives. This is not a typical setting for Web users. However, viewers often need to find correct information from a Web page, either in a computer lab or at a home computer, within a reasonable—if not the shortest possible—time period. The performance incentives were intended to create pressure similar to that which a Web user might have.

In terms of the external validity factor of time, our findings hold consistent over a period of five years. During the Web's fast development, animation may be used differently on Web pages over time. The effects of animation under the studied conditions, however, should not change much, as our results imply. This can be due to a rather slow process of human evolution. Nevertheless, the findings could be made more robust by further studies.

This research provides a base for future investigations. In the studies described here, the nature of the information-seeking task requires relatively low levels of information processing from respondents. Future studies may investigate how (if at all) animation affects respondents' performance in reading and comprehending a meaningful passage, a task that requires higher levels of information processing. For example, Hong and colleagues (Hong, 2004) studied online shopping tasks that are closer to real tasks in the Web environment. The difference in the nature of the tasks may impose quite different findings. For example, when studying consumers' memory for television advertising, Pieters and Bijmolt (1997) explored the duration and serial position of a commercial and of the number of commercials. They found that placing a commercial first is better than placing it last in achieving the goal of maximizing brand recall. Here the tasks involve memory recall rather than just discrimination at the perception level.

Furthermore, the continual development of sophisticated software has allowed for more aggressive and intrusive advertisements on the Web. Animated online banners used to be restricted to a specific location on a Web page. Current advertisers, however, are increasingly using animations that do not stay in a specific location on a Web page, but instead move from one side to another, demanding more attention from users. Future studies should investigate whether such animations have a greater effect on users' performance on different tasks.

This research considers animation a non-primary information stimulus. Empirical studies on animations that are primary information sources are also limited and deserve much research attention.

Practical Significance and Implications

This stream of research presents theoretical explanations and empirical evidence of animation effects under different conditions and over time. There are few studies of this type. The implications of this research for Web user interface design and online advertising are significant. From either the information-seeking user's perspective or that of companies using the Web to realize both operational and strategic benefits, content providers must understand the potential effects of animation on users.

This study suggests some strategies (Zhang, 1999, 2000, 2001) for both Web site content providers and online advertisers, showing a dichotomy between their very different goals. Content providers want to make money from advertising, but also need to care about the potential side effects of ads on their viewers' information-seeking performance. Given a choice, content providers could prefer ads with a minimum of distracting effects. Results from this research suggest that negative effects should be reduced by (1) raising the perceptual load, that is, making information-seeking tasks more challenging by involving viewers in the content of a Web page; (2) avoiding brightly colored animation; (3) avoiding animation that is semantically similar to the primary tasks; (4) placing ads on the Web page earlier and on the right side; and (5) avoiding on-off-on type of animations.

On the other hand, online advertising is very attractive to marketers, as proven by continued practice since the inception of the Web. Online advertisers or marketers want to continue grabbing viewers' attention, knowing that the ads will be processed, to some extent, involuntarily. Some advice for online advertisers has been provided. For example, some suggest that advertisers should be "negotiating for top of the page for online ads" (Hein, 1997), while others advise that ads should be placed at a place on the page that viewers will reach after they have gained a certain amount of the primary information (Scanlon, 1998). Our findings suggest that marketers may want to take strategies opposite to those used by content providers, that is, they should (1) target pages where audiences tend to have simple tasks; (2) use bright colors when possible; (3) design animation that is semantically similar to the tasks; (4) put ads on the left side of the screen; and (5) use pop-up animations or online ads when the user has already started reading or scanning the Web page. Advertisers may not have to have on-off-on animations on the screen, since they are as "effective" as those that pop up during the tasks. A caution accompanying these suggestions is that they are based on animation's effect on task performance, not on recall of animation content or semantics. Further studies are needed to understand if on-off-on animations enhance recall better than stay-on animations.

CONCLUSIONS

Despite some studies showing that experienced Web users are less likely than novice users to be distracted by competing stimuli on the Web (Bruner and Kumar, 2000; Dahlen, 2001; Diaper and Waelend, 2000), our research indicates that animation's interference effects have not changed much over the years and are still affecting experienced users such as the participants in our research. For the most part, subjects were not able to block the animations, even though they knew animations had little to do with their tasks, and even though some of them thought they were able to ignore the animations. This means that, to some extent, animation is processed involuntarily, a finding supported by major visual attention studies. For example, many researchers (Allport, 1989; Duncan, 1984; Miller, 1991; Yantis and Jonides, 1990) have argued that even though the processing of unattended stimuli can be attenuated with certain manipulations, it cannot be totally ruled out. The meaning of the unattended stimulus must be processed to some extent. Because our attention has a limited capacity, our available resources for attending to pertinent information are reduced, and thus information-processing performance, including speed and accuracy, deteriorates (Driver and Baylis, 1989; Miller, 1991; Treisman, 1991). Our study also supports Lang's limited capacity model (Lang, 2000; Lang et al., 2002), that is, the onset of animation when an individual is performing a task elicits an automatic, reflexive, and attentional response (i.e., orienting response) that affects the individual's task performance. Furthermore, due to this automatic and reflexive nature of responses, it is unlikely that animations as non-primary information have no impact on task performance.

With the rapid evolution of the Internet and the World Wide Web, and as more people use the Web for gathering information, conducting business, and for entertainment, studies on the effect of certain Web features such as animation become timely and important. For a relatively new medium such as the Web, empirical studies are as important as theoretical predictions and implications. Research that tests the applicability of existing theories to new environments has theoretical as well as practical value. In this research, we have tested the applicability of some visual-attention and perception research results to the Web environment by confirming some and ruling out others. The general implication is that human evolution changes our characteristics much more slowly than the environment changes. Certain research results on human characteristics can be applied during a relatively long period. This particular study suggests that designers of any type of user interface should consider possible visual interference sources that may affect an individual's information-seeking performance.

REFERENCES

Allport, A.D. Visual attention. In M.I. Posner (ed.), *Foundations of Cognitive Science.* Cambridge, MA: MIT Press, 1989, pp. 631–682.

Baecker, R., and Small, I. Animation at the interface. In B. Laurel (ed.), *The Art of Human-Computer Interface Design.* Addison-Wesley Publishing Company, 1990, pp. 251–267.

Bruner, G.C., II, and Kumar, A. Web commercials and advertising hierarchy-of-effects. *Journal of Advertising Research*, 40, 1/2 (2000), 35–42.

Campbell, B.A.; Wood, G.; and McBride, T. Origins of orienting and defensive responses: An evolutionary perspective. In P.J. Lang, R.F. Simons, and M. Balaban (ed.), *Attention and Orienting: Sensory and Motivational Processes.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1997, pp. 41–68.

Cook, T.D., and Campbell, D.T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Boston: Houghton Mifflin Company, 1979.

Dahlen, M. Banner advertisements through a new lens. *Journal of Advertising Research*, 41, 4 (2001), 23–30.

Diaper, D., and Waelend, P. World wide web working whilst ignoring graphics: good news for web page designers. *Interacting with Computers*, 13, 2 (2000), 163–181.

Driver, J., and Baylis, G.C. Movement and visual attention: the spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 3 (1989), 448–456.

Duncan, J. Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 4 (1984), 501–517.

Duncan, J., and Humphreys, G.W. Visual search and stimulus similarity. *Psychological Review*, 96, 3 (1989), 433–458.

Eysenck, M., and Keane, M. *Cognitive Psychology: A Student's Handbook.* Hove, UK: Psychology Press, 1995.

Gonzalez, C., and Kasper, G.M. Animation in user interfaces designed for decision support systems: the effects of image abstraction, transition, and interactivity on decision quality. *Decision Sciences*, 28, 4 (1997), 793–823.

Hein, K. Improve your online marketing. *Incentive*, New York, November 1997.

Hoffman, H.S. Attention factors in the elicitation and modification of the startle reaction. In P.J. Lang, R.F. Simons, and M. Balaban (eds.), *Attention and Orienting: Sensory and Motivational Processes.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1997, pp. 185–204.

Hong, W. Does animation attract online users' attention? The effects of Flash on information search performance and perceptions. *Information Systems Research*, 15, 1 (2004), 60–86.

Lang, A. The limited capacity model of mediated message processing. *Journal of Communication*, 50, 1 (2000), 46–70.

Lang, A.; Borse, J.; Wise, K.; and David, P. Captured by the world wide web—orienting to structural and content features of computer-presented information. *Communication Research*, 29, 3 (2002), 215–245.

Lavie, N. Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 3 (1995), 451–468.

Lavie, N., and Tsal, Y. Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, 56 (1994), 183–197.

Marchionini, G. *Information Seeking in Electronic Environments.* Cambridge: Cambridge University Press, 1995.

Mayor, J., and Gonzalez-Marques, J. Facilitation and interference effects in word and picture processing. In S. Ballesteros (ed.), *Cognitive Approaches to Human Perception.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1994, pp. 155–198.

Miller, J. The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: a search for boundary conditions. *Perception & Psychophysics*, 49 (1991), 270–288.

Miller, J.O., and Bauer, D.W. Visual similarity and discrimination demands. *Journal of Experimental Psychology: General*, 110 (1981), 39–55.

Pashler, H.E. *The Psychology of Attention.* Cambridge, MA: MIT Press, 1998.

Pavlov, I.P. *Conditional Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex.* London: Wexford University Press, 1927.

Pieters, R.G.M., and Bijmolt, T.H.A. Consumer memory for television advertising: a field study of duration, serial position, and competition effects. *Journal of Consumer Research*, 23, 4 (1997), 362–372.

Proctor, R., and Van Zandt, T. *Human Factors in Simple and Complex Systems.* Needham Heights, MA: Allyn and Bacon, 1994.

Reeves, B.; Lang, A.; Kim, E.; and Tartar, D. The effects of screen size and message content on attention and arousal. *Media Psychology*, 1 (1999), 49–68.

Scanlon, T. Seductive design for web sites. *Eye for Design*, 1998 (available at http://www.uie.com/articles/seductive_design/, accessed November 20, 2005).

Sokolov, E.N.; Spinks, J.A.; Naatanen, R.; and Lyytinen, H. *The Orienting Response in Information Processing.* Mahwah, NJ: Lawrence Erlbaum Associates, 2002.

Spieler, D.H.; Balota, D.A.; and Faust, M.E. Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 2 (2000), 506–526.

Treisman, A. Search, similarity, and integration of features between and within dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, 17 (1991), 652–676.

Tversky, B.; Morrison, J.B.; and Betrancourt, M. Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57, 4 (2002), 247–262.

Warden, C.T., and Brown, H.C. A preliminary investigation of form and motion acuity at low level of illumination. *Journal of Experimental Psychology*, 34 (1944), 437–449.

Yantis, S., and Jonides, J. Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, 16 (1990), 121–134.

Zhang, P. The effect of animation on information seeking performance on the world wide web: securing attention or interfering with primary tasks. *Journal of Association for Information Systems*, 1, 1 (2000), 1–30

Zhang, P. The impact of animation timing and location on visual search task performance in the web environment. In *Proceedings of the Americas Conference on Information Systems (AMCIS ’2001).* Boston, 2001, pp. 1361–1366.

Zhang, P. Will you use animation on your web pages? Experiments of animation’s effects and implications for web user interface design and online advertising. In C. Romm and F. Sudweeks (eds.), *Doing Business on the Internet: Opportunities and Pitfalls.* London: Springer-Verlag in conjunction with the British Computer Society, 1999, pp. 35–51.

Zhang, P., and Massad, N. The impact of animation on visual search tasks in a web environment: a multi-year study. In *Proceedings of the Americas Conference on Information Systems (AMCIS 03).* Tampa, FL, 2003, pp. 2265–2272.