Moving towards a new methodology:
The retirement of the “Web site heuristic evaluation”

By Susan Kaltenbach

June 10, 2002
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

Introduction

The Internet has evolved since the early days of node-based information sharing between the military, academia, and computer enthusiasts, and is now a ubiquitous tool used by the world for finding information, social interaction, purchasing items, and entertainment.

In the early days of Web site design, sites were created by HTML hackers – individuals who used HTML code to construct simple sites to serve as a nexus of information and hyperlinks. In the following years, Web sites became increasingly sophisticated and complex, and became increasingly important to businesses as another form of “real estate.” Thus Web site design became less and less a domain of technical programmers, and more a domain of business professionals and designers. Web site design was growing from an area of personal expression and creativity to an instrument of communications between people and organizations. The transition between the two types of design ethics was difficult, especially since there was no “right way” to design a web site.

In the CHI 96 workshop *HCI and the Web*, Jakob Nielsen presented a short position paper that included his history of Web design up to 1996. Below are some excerpts:

1993: Just having a server on the Web was enough to show that you were a pioneer!

1994: This was the year of home pages that were no more than glorified hotlists with long bulleted lists of links.

1995: **Less is more** definitely became a key design strategy.

1996: I think that web-surfing is dead…most users will probably spend the majority of their time with a small number of websites that meet their requirements...

It is in this presentation that Nielsen begins to express his ideas for Web site heuristic evaluation.

---

1 (Acronym for Human-Computer Interaction)
Recently, a real HCI contribution has become essential for Web design, with the need to draw upon our existing knowledge of icon design, knowledge elicitation (discovering appropriate information space structures), and usability testing…

At the end of his discussion, he concludes that “Heuristic evaluation would seem to be a perfect match for this need.”

The conference was in April 1996, and his “Top Ten Mistakes in Web Design”\(^2\) were made public in May of that year. It is obvious that Nielsen had been preparing to introduce his idea of Web site heuristic evaluation at that event, and that he also formulated a set of 10 design rules that would serve as the heuristics upon which Web sites would be evaluated.

Since that time, the 10 “Nielsen heuristics” have become a dominant standard for Web site usability studies, and it is impossible to find any discussions about the topic that do not mention them.

Of course, the Internet has now become a very complex organism and is used by a wide range of practitioners. Technically unskilled individuals now create and publish their own web sites using Web authoring software such as Microsoft® FrontPage© and Dreamweaver©. E-commerce sites have developed sophisticated systems for tracking your visit, and suggest items to purchase that might be of interest to you.

This paper will examine the veracity of existing Web site heuristics. I will discuss the evolution of these current heuristics, how they are currently used in practice, and their possible shortfalls. I will then propose the establishment of a “new heuristics,” and will discuss exciting research that informs a new philosophy of standards.

I hope that this paper makes a convincing argument for throwing out the non-empirical set of Web site heuristics currently in practice, and replacing it with a new, non-prescriptive set of guidelines. These new guidelines would better serve the Web design population to meet new challenges as the Internet continues to evolve.

\(^2\) Referred to hereafter as the “Nielsen heuristics.”
A brief discussion about Web heuristics

Definition of “heuristic evaluation”

Heuristics is a method of exploration that uses existing known elements as a beginning point. Rather than a “set of rules and standards,” heuristics implies a set of “best practices” that may be improved upon with future investigations. Thus, we refer to “adding to the body of heuristics.” The definition of heuristics as being a work in progress is especially relevant to the topic of Web site standards. As De Jong and van der Geest explain,

The term Web heuristics is used to refer to all the sets of process guides, principles, criteria, tips and tricks, and guidelines that are available to support Web designers. A heuristic is a discovery aid.  

The actual phrase “heuristic evaluation” was promoted by Nielsen during his publishing of the usability techniques literature. Nielsen proposed heuristic evaluation as a component of usability studies. Using this technique, each usability expert would evaluate a product individually, and later would discuss their findings with the other experts. A consolidation of findings would be presented, along with suggestions on how to improve the product’s usability. Nielsen describes the utility of heuristic evaluations as follows:

Heuristic evaluation does not provide a systematic way to generate fixes to the usability problems or a way to assess the probable quality of any redesigns. However, because heuristic evaluation aims at explaining each observed usability problem with reference to established usability principles, it will often be fairly easy to generate a revised design according to the guidelines provided by the violated principle for good interactive systems. Also, many usability problems have fairly obvious fixes as soon as they have been identified.

When he presented his “10 heuristics,” Nielsen proposed that the practice of Web site design could also benefit from such an evaluation. The heuristic evaluation method would be identical to the one described above, only the criteria to be followed were the “Nielsen heuristics.” Nielsen was then positioned as the one true Web site usability expert, and received the

---

3 de Jong and van der Geest, p. 311.
4 Nielsen, “How to conduct a heuristic evaluation.”
appellation “Guru of Web Page Usability” by the New York Times, by which he is still referred today.

The use of the term “heuristics” itself is an interesting one. When Nielsen positioned himself with his group of heuristics as the “usability guru,” the field of Web design was populated with a diverse range of people, including many (such as programmers) who were not familiar with the more academic terms used by the usability profession.

I have always maintained that the use of the term “heuristic” distances Web designers from usability experts, by implying an authoritative, academic expert qualified to pass judgment on the designers’ work. This is a serious error. If the two fields don’t share a common language, then actually applying web usability principles to web site designs will be a difficult (and non-cooperative) task. (See below, Why is Jakob Nielsen’s web site so ugly? for a discussion of how this type of language has distanced designers from the usability community.

**Existing Web site heuristics**

**Heuristics in place today**

The development of current Web site heuristics was informed greatly by previously existing usability standards, but it would be impossible to trace the parentage of each element of current Web site heuristics. This section will not discuss the historical basis for the existing heuristics (see the section Have heuristics been properly developed? for a more in-depth discussion of the origins of existing Web site heuristics.) Instead, this section will discuss what heuristics are in place today, along with methods to distinguish them by their nature and applications.

In their study of existing Web site heuristics, De Jong and van der Geest distinguished existing heuristics on the basis of their foundations:

- **Standards-based heuristics** come from fixed standards such as the W3C accessibility guidelines, or the Federal government’s Section 508. As De Jong and van der Geest dryly comment “agreed-on rules do not necessarily have a strong relationship to the usability and effectiveness of Web sites.” Standards-based heuristics will not be discussed in this paper.
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

- **Theory-based heuristics** come from accepted theories and practices. In the case of Web site heuristics, the theories were adapted from other “relevant areas” as HCI, text comprehension, rhetoric, or visual design. An example of theory-based heuristics could be the majority of Jakob Nielsen’s contributions to the field, which represent guidelines based on observations and anecdotes, rather than empirical data.

  Most of the heuristics that we found through our Internet search seem to come from the same field as the Instone/Nielsen heuristics: the field of user interface design and usability studies...They often include a limited number of items, a fact that makes them arbitrary or simplistic in light of the many factors that can influence the usability of a Web site.

- **Research-based heuristics** depend on data from usability studies applied to the Web.

- **Practitioners’ heuristics** are expressed in the common language of Web site developers and are based on views and experiences that practitioners use to build a set of “best practices.” Practitioner heuristics vary wildly in quality. Since many early practitioners of Web design were technically adept but not competent in areas of design, these heuristics usually lack any basis in the simplest forms of usability. These heuristics will also not be discussed in this paper. (pp. 312-313)

This type of classification helps us understand the origin of the heuristics, which is an important component in determining their authenticity.

An important distinction to make in the study of Web heuristics is the difference between empirical and non-empirical guidelines. Empirical heuristics are informed by research, often in the form of controlled usability studies and results published in peer-reviewed journals. Non-empirical heuristics are usually based on a subjective vision of “best practices” and “common sense,” and are often generated by practitioners who are not required to defend their findings. It is the latter set of heuristics which I criticize in this paper.

In “Challenges in Developing Research-Based Web Design Guidelines,” Mary Evans of NOAA outlines her efforts to formulate a set of design guidelines for future use in Web page construction. Evans first researched the literature to find existing guidelines, but soon found that many of these guidelines lacked relevant research.

In most cases, the authors described the basis of their guidelines as general design principles. Though a few compilations of guidelines were described by
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

their authors as research-based, only in a very few cases were specific studies cited to support particular guidelines. There seemed to be a gap to be filled by a set of web design guidelines based on relevant research. (p. 303)

Because of the uncertain origins of non-empirical heuristics, we should analyze whether these findings are, in fact, applicable to the Web site interaction experience.

The often-discussed “Nielsen heuristics” first appeared in the form of a list of “don’ts” titled “Top Ten Mistakes in Web Design” that appeared in Jakob Nielsen’s online bulletin Alertbox for May 1996. Nielsen’s top ten mistakes were:

1. Using Frames
2. Gratuitous Use of Bleeding-Edge Technology
3. Scrolling Text, Marquees, and Constantly Running Animations
4. Complex URLs
5. Orphan Pages
6. Long Scrolling Pages
7. Lack of Navigation Support
8. Non-Standard Link Colors
9. Outdated Information
10. Overly Long Download Times

The explanation following each caveat explains in detail Nielsen’s rationale for including it in the list. Amazingly, it is this simple list of “top ten mistakes” which formed the footing for a large culture of Web site usability studies and products.

How do we test whether the heuristics are correct or not? The current practice seems to be that you review the existing accepted literature, identify something you find lacking, formulate a theory, and test your theories in a usability lab. But is this sufficient evidence that a heuristic works? This is a common question that receives an increasing amount of research to resolve.

---

5 Although it is possible that these heuristics were formalized elsewhere, the Alertbox archive was the only source for the original heuristics that I could find. Of course, countless interpretations have evolved over the years to expand on the originals.
Moving towards a new methodology:  The retirement of the “Web site heuristic evaluation”

In “Empirically Validated Web Page Design Metrics,” Ivory et al investigate this same question. They begin by noting that Nielsen’s original 10 heuristics and subsequent follow-ups are based on anecdotal observational evidence and reasonable assumptions, rather than empirical data. They go on to mention that

Furthermore, there is no general agreement about which web design guidelines are correct. A recent survey of 21 web guidelines found little consistency among them. We suspect this might result from the fact that there is a lack of empirical validation for such guidelines.

De Jon and Van der Geest suggest that validation research could be used to back up the heuristics. But what constitutes real validation research? They suggest that “The most convincing proof of benefit would be seen when a study demonstrates that a site designed with a set of heuristics is better than one designed without heuristics.” They go on to discuss a study by Gerhardt-Powals, who designed an interface based on her principles, and then compared it to two other interfaces not based on those principles. Not surprisingly, Gerhardt-Powals concluded that the interface based on her principles was superior to the others.

Is it possible that the current methodology for expanding the heuristics is flawed? In traditional research, the practice of using a blind or double-blind test helps prevent experimenter bias. However, as Evans discovered, approximately half of the contributing research were obtained using this more formal experimental method. The remainder were based on results from less rigorous sources, such as usability tests or surveys.

It would be preferable if experimenters used some kind of blinding in their usability research; this would contribute a more objective element into the findings. Instead, we have practitioners testing their own practices. The development of heuristics is not exclusively based on empirical results, but a stricter adherence to the traditional methods of experimentation would help establish a validity currently lacking.

This environment could easily contain many types of bias. If a group of researchers created a set of heuristics and then conducted the usability tests themselves, they might be
naturally prepossessed to find success with their findings. This of course is not necessarily a deceitful act, but rather a situational positioning which might prevent truly unbiased results.

**Existing Web site heuristics**

**How heuristics are used today in the Web design process**

The existing set of heuristics can be used for several reasons. Methods include a planning-oriented and evaluation-oriented heuristics, a process-oriented and product-oriented, troubleshooting and verifying,

Heuristic evaluation is seen as an important component of Web site usability research. The operating theory is to take from two to five professionals and have them evaluate a web site using their expert knowledge. The heuristic evaluation can take place during many parts of the design process.

Heuristic evaluation of a Web site is considered to be an essential part of Web design, and it’s often seen as a standalone method to perform quick, easy and inexpensive usability evaluation of the interface.\(^6\) The expert evaluation method can be used as a discount method for quick, cheap, and easy evaluation of the user interface. From Kantner, “the heuristic evaluation makes a first pass at catching the most visible usability problems (‘the low-hanging fruit’), enabling laboratory testing to focus on deeper issues.” (p. 160)

How does the traditional heuristic evaluation operate? A description of the process is provided by the Online Computer Library Center, Inc. (OCLC) on their Web site [Heuristic evaluation](http://www.oclc.org).

1. Several HCI experts compare a prototype to a set of heuristics (‘rules of thumb’) for developing easy-to-use user interfaces. We use a set of 14 heuristics. A mixture of domain and design experts work best as evaluators.
2. Any usability problems found are evaluated for their severity and extent. It is important that the evaluators do not discuss their findings with other evaluators during this phase.
3. The reports of each evaluator are compiled and grouped based on severity.

\(^6\) It is often marketed to the public as a fast, cheap Web site usability evaluation.
Moving towards a new methodology: The retirement of the "Web site heuristic evaluation"

4. All the evaluators participate in a brainstorming session. During this session:
   (1) Reports are grouped into a common set of problems and severity/extent ratings, (2) This set of problems is further refined into several problem categories, and (3) Solutions to the problems are proposed.

5. A summary report of the brainstorming session is distributed to interested parties, and a course of action is determined.

Not surprisingly, OCLC uses a set of fourteen heuristics “based on Nielsen’s 10 Heuristics.”

Although the definitive heuristic evaluation uses several HCI experts to evaluate usability, I found several references in my research to evaluators whose qualifications came from their attending half-day training workshops on the subject, or from their physical proximity to the test area (i.e., “just grab anyone around here!”). This reflects the current practice of seeing heuristic evaluation as a fast, cheap usability testing method.

And so now we face a large population of Web site designers who can claim to have performed their own “heuristic evaluation.” The definition has so relaxed that the concept of a group of experts conducting analysis on a Web site has evaporated. Instead, we have a process where a person agrees to help their friend/coworker/etc. by sitting down and try to map any problems discussed in the Nielsen heuristics

**Have heuristics been properly developed?**

In her paper, Evans describes her efforts to formulate empirically-based heuristics by using findings (mostly) from published and/or peer-reviewed sources. Of these resources, one-third was from Web site, hypertext, or intranet system systems. These findings are significant. In essence, only one-third of Web heuristics come from Web studies. The remainder comes from other sources such as HCI, print usability, psychological studies, etc.⁷

If Web heuristics aren’t generated from direct Web studies, where did they come from? A great deal of Web heuristics originates in research performed to improve usability of print documentation. These studies, in turn, were often informed by psychological, psychophysical,

---

⁷ Although Evans didn’t report on the age of her references, my personal estimate has been that 80% of these findings were published before 1998.
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

and visual perception findings in areas including attention span and visual discrimination. I will characterize these heuristics as “empirical heuristics” because, although they don’t come from direct Web studies, they rely mostly on research observations.

Many authors (including Evans, Williams et al) recognize the importance of basing heuristics in a firm empirical background, and have consolidated relevant research and findings from these fields to construct such standards. I consider these “empirical heuristics” and will not discuss them further.

Although empirical heuristics are available to many practitioners, it is the classical “Nielsen list” heuristics that are discussed by the public at large. These heuristics stem from Nielsen’s research into usability studies, from which he published books in the early 1990s. Nielsen introduced the phrase "heuristic evaluation" in usability studies. It has become common in the Web design vernacular. A recent Google search provided 731 matches on the phrase “Nielsen website heuristics.” Obviously the concept of Web site heuristic evaluation is very popular in the Web design community. Where do these heuristics originate from?

As mentioned above, many of the elements for Web heuristics originate in the field of usability research. Nielsen’s studies into usability research and hypertext media were later formulated into a series of heuristics, which Nielsen brought to the public in “Top Ten Mistakes in Web Design.” These heuristics were soon seized by the Web design community and became known as the Nielsen list of 10 design principles. Nielsen’s “rules” soon came to dominate studies of Web design usability and became a common referent in further usability research. As de Jong and van der Geest observe,

Most of the heuristics that we found through our Internet search seem to come from the same field as the Instone/Nielsen heuristics: the field of user interface design and usability studies…They often include a limited number of items, a fact that makes them

---

8 I’m not sure how “heuristic evaluations” are different from “expert evaluations,” since both use internalized information and concepts of standards to evaluate a product. I suspect the term “heuristic” was introduced to make the concept more memorable in the common language.

9 The list is what I will continue to call the “Nielsen heuristics.”
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

arbitrary or simplistic in light of the many factors that can influence the usability of a Web site.10

Nielsen’s behavior and statements sometimes alienated members of the Web design community. His dogmatic approach to web design is criticized by many due to its inflexibility in the face of the growing diversity of the internet and its population. He considers his heuristics to be the only valid standard for all Web designs, and appears unable to entertain the idea that his personal recommendations could be improved by further examination of current Web usability issues, including research done outside his organization.

We are now in an incongruent position. The Nielsen heuristics continue to motivate much of Web usability research and Web design standards, and yet a large portion of the population is uncomfortable with their authenticity and relevance.

Not surprisingly, many researchers [Sutcliffe] [Evans] [de Jong and van der Geest] have focused on evaluating the reliability of current web heuristics which, in turn, were informed by Nielsen. In de Jong and van der Geest’s study, they discovered that the Nielsen heuristics were not “straightforwardly applicable to Web design.” The heuristics were too specific and thus inflexible, and could only be used in specific contexts.

What follows is a humorous aside and is not intended as an ad hominem attack on Jakob Nielsen, but as an interlude before the second portion of the paper.

**Why is Jakob Nielsen’s site useit.com so ugly?**

This is a very popular Internet topic. Why is Jakob Nielsen’s useit.com site so ugly?

Nielsen explains his rationale for the current style of his site in *Why this web site has no graphics*:

Download times rule the web, and since most users have access speeds on the order of 28.8 kbps, Web pages can be no more than 3 KB if they are to download in one second which is the required response time for hypertext navigation. Users do not keep their attention on the page if downloading exceeds 10 seconds, corresponding to 30 KB at modem speed. Keeping below these size limits rules out most graphics.

10 De Jong and van der Geest, p. 321.
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

Nielsen supports these figures with references including Miller’s 1968 article *Response time in man-computer conversational transactions*. But why is the site text-only with very few graphics? He continues:

I am not a visual designer, so my graphics would look crummy anyway. Since this website is created by myself (and not by a multidisciplinary team as I always recommend for large sites) I didn’t want to spend money to hire an artist.

The useit.com style format uses a 12-point Verdana font without appropriate leading between lines, and it is difficult to read. It has been the subject of many heuristic evaluations and is notorious for its lack of usability.\(^\text{11}\)

Although one can view the useit.com formatting as an amusing eccentricity, some usability professionals are embarrassed by its lack of professionalism. During an interview with Digital Web Magazine, one usability professional stated that

I think in many ways Jakob did us all a great disservice when he alienated designers. He forgot his own principles: speak to your audience in their own language. His site is so ugly most designers just throw their hands in the air and run away. I can’t blame them.\(^\text{12}\)

Where should we go from here?

**What should we do with the current non-empirical Web heuristics?**

As we have seen, the current body of non-empirical Web site heuristics has largely been derived from a series of dogmatic instructions introduced by one usability specialist in 1996. Many (if not most) of the members of the Web design and usability community agree that these heuristics need to evolve to encompass the growing diversity of Internet vehicles and Web site

---

\(^\text{11}\)See, for example, Todd Lisonbee’s sophomore paper *An examination of Jakob Nielsen’s Website Useit.com*, CDES 215 - Project 1, California State University Chico. Mr. Lisonbee apologizes for the cannibalistic method of examining Nielsen’s site with his own heuristics, but he does conclude with several suggestions for Nielsen’s site.

Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

purposes. However, the papers I’ve seen have simply been reflections of the “Nielsen heuristics,” albeit expanding the rules to encompass newer features available to designers. Thus we maintain a system whose core is based on a culture of usability studies conducted by a small core of practitioners.

I suggest we throw the whole thing out.

The current heuristics motivate us to ask questions such as: of what use is it to include a biography of a Web page’s author? Why should we even care about whether a link is blue and underlined, when we now have a well-established mental model that a link is the active part of a web page? Why shouldn’t we allow a link to open a new browser window, so that the visitor can review the original site’s content while also viewing the new site? These instructions are all purely subjective, and have little relevance to real Web site design. Of what use are these “classic” heuristics when the Internet demonstrates that they are out of sync with the reality of Web site diversity?

My suggestion is to discard these historical heuristics. They do not meet the current requirements of the diverse Web designer population. They simply do not work. Insisting that we continue to meet these standards is like trying to convince computer artists to only render two-dimensional images. Sooner or later, the creator will evolve new uses for a technology, and this discovery process adds to the heuristics by showing what works in the new framework.

In 1996, the field of Web design was a wild frontier, and Nielsen’s heuristics served as a guiding standard to allow everyone effective, equal footing from which to launch their projects. Six years later, the Internet and Web technology has evolved far beyond the early days of the “HTML Hackers.” We are now able to explore new ways to communicate information and symbolize actions so that the Web site visitor can evolve with us. It is the user who drives the heuristics, not vice versa.

*The first thing we need to throw out is the term “heuristics.”* This term has little legitimacy outside the usability and academic professions. If we are truthfully motivated to help Web
designers, it is necessary to use a common language that does not use words to separate the two groups.

**What will the new, non-empirical Web heuristics be like?**

If we throw it all out and keep the (relevant)\(^{13}\) empirical heuristics, where do we go next? I believe that a new set of guidelines can be created to accompany the empirical heuristics. These guidelines would suggest rather than dictate, design concepts and rhetorical models.

These new guidelines would be instructive, rather than prescriptive.

Rather than dictating that “users should know where they are in a navigational structure,” the heuristic could suggest “include an image or map so that the user knows how to return along her original path to recapture those earlier steps.” Another example could be “make sure the user knows she is going to open a new browser window by following a hyperlink” rather than “don’t open new browser windows since it is a user-hostile message implied in taking over the user’s machine.”\(^{14}\)

In capturing ideas for the “new, non-empirical heuristics” for Web design, we should review recent literature that has not been outdated by the contemporaneous technology. For example, a discussion on total download time should be based current technologies and trends for internet connections. Discussions on “Web site size” should be informed by the statistics of what the current population uses. And discussions about “non-standard interfaces” will need to be ruled by the overlying concept of creative expression.

\(^{13}\) A topic which is rich with possibilities is the study of whether usability tests incorporating paper models are actually relevant to Web usability issues. The same might be explored with older psychological studies of attention spans and concentration. The visual perception of subjects facing a CRT or LCD monitor may be significantly different from paper models: the monitor not only does not simulate paper well, but it is presented normal to the plan of the viewer’s eye span, requiring a special position to interact with it. It could also be argued that the national psyche continues to evolve along with our technological referents, and the attention span of a person in 1965 may be different from that of a person in 2002.

\(^{14}\) Nielsen, “The Top Ten New Mistakes of Web Design.”
The following is a personal example for establishing a non-empirical heuristic. What size of Web site should we design for? The existing heuristic dictates that we accommodate all visitors, regardless of how big their monitor may be.

Under the world of “new practices,” I asked myself “how can we possibly design a Web site that will accommodate users ranging from the classic “640 x 480 pixels” user to someone like me, who has a desktop consisting of two LCD monitors at a resolution of 2560 X 1024? Is there any way we can serve both users?”

My answer would be that we cannot accommodate both users, and in many cases we shouldn’t even try to. The size of a Web site should be dictated by the estimated characteristics of the plurality of web site visitors. For example, if I was designing a web site for a luxury leather goods e-retailer, I would research any existing material available to learn what people at a certain income level report as their home computer system. If such research doesn’t exist, I would then make a logical assumption that (a) people who want to purchase luxury leather goods are wealthier than others, and (b) anyone visiting an e-retailer would be a computer user, then (c) the visitors to my site would be wealthy computer users who have a large monitor at home.

If I was designing for a government information source, or any other public service, I would make sure to design for even the smallest monitor, in order to accommodate every visitor that needs information.

Thus a new guideline is born: Size your web site to your intended audience’s monitor site.

What’s coming up ahead?

It has been suggested that the next evolutionary step of the Internet will be as a fully commercial enterprise, populated by large corporate presences. In fact, the trend in Internet transformation is towards information portals, entertainment sites, and retail presences. The days of the “new economy” of dot.coms providing intangible “products” and “services” have passed,
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

And the original user community (including hackers$^{15}$ and information junkies) can continue to use the Internet as a gathering place for information exchanges.

Continuing with that thesis, we can formulate ideas about what information is needed and how to gather it. What will make users enjoy web sites more? What will make them buy my products? What will improve their ability to get information from my company so that they won’t drive up my personnel costs due to help desk calls? What will entertain them so that they’re motivated to pursue the stickiness of my site in the future?

Sources for this type of non-empirical research could include findings from the *Journal of Interactive Marketing, Psychology and Marketing, and International Marketing Review*, among other social science journals. The efficacy of the heuristics would be evaluated on the basis of monetary rewards, rather than concurrence with current theories. Thus, Web sites evaluated on an economic basis face will face a set of heuristics that come from a reliable source – the marketplace. From Sutcliffe,

The next step is to adopt grounded theory techniques from social science to develop more reliable questionnaire inventories [20] for assessing subjective qualities in websites, and then see how these correlate with observed user behaviour. $^{16}$

**Part Two: What don’t I see in existing heuristics that I think should be included?**

We must also remember that any heuristics we develop need to be able to evolve along with the new technologies. Studies have already begun on the usability of PDAs, psychophysical differences in visual perception between a CRT and a LCD monitor, mobile computing appliances, and retina-scanning laser displays. These are just a few of the technologies evolving

$^{15}$ The traditional meaning of “hacker” was an individual who spent large amounts of time “hacking” in system-level computer languages in order to increase proficiency in computer systems. It has since become synonymous with “cracker,” a hacker with malicious intent.

$^{16}$ Sutcliffe, p. 10.
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation” at the time of this writing which could become a viable and popular technology in a few years’ time.

However, a refreshing trend has appeared in the gathering of the “new heuristics” – a trend to actually empirically test prior assumptions to make sure of their accuracy. For example, it is assumed by all practitioners that Web site animation is distracting to users, and most evidence was based on subjective responses solicited from usability study participants. But subjective reports can be deceiving; we may not always know what we like, or how we act, or what we remember, even though we think we do. Thus, an individual who says “the animation distracted and irritated me” is expressing a subjective account of what happened.

But when she was performing the action, did she act distracted? In “The Effects of Animation on Information Seeking Performance on the World Wide Web: Securing Attention or Interfering with Primary Tasks?” Zhang investigates whether such an assumption is scientifically reproducible. In his study, he asked users to perform a task in different animation scenarios. Not surprisingly, Zhang’s results showed that users performing tasks were distracted by animations. But when the results are examined in finer detail, there are some very interesting results:

1. Animation as a secondary stimulus deteriorates viewer information seeking performance.
2. As the difficulty of the task increases, viewer performance is less affected by animation.
3. Animation that is similar but irrelevant to a task has more negative impact on viewer performance than animation that is dissimilar to the task.
4. Animation that is brightly colored has a stronger negative effect on viewer performance than does dull colored animation.\(^\text{17}\)

Thus one can assume from this research that using dull-colored animation, dissimilar to a user’s task, and in conjunction with a highly complex task, causes the user to be less distracted than in a scenario involving animation with the opposite values. Zhang discusses possible further research, noting that “it would be interesting to know the extent to which viewers semantically process non-primary animation or banners. In other words, perhaps viewers’ intentionally ignoring

\(^{17}\) Zhang, p. 1.
animation/banners does not mean they are not aware of the semantic content of the ignore features.”\textsuperscript{18} He dryly adds that this information would be relevant to online advertisers.

Another interesting study by Sutcliffe adds the element of “attractiveness” into web site heuristics, defending his choice by arguing that when given a choice of web sites to visit, users will prefer the one which makes a better initial impression. He defines attractiveness as having three complex variables: Arousal, Motivation, and Perceived Utility Gain. These categories are discussed in fine detail with many references to the literature of Psychology.

Sutcliffe proposes a new model for the Web site visitor experience, shown in Figure 1 below. This model depicts the user Web site experience into four components: Web site discovery, initial attraction, exploration/navigation, and transaction. Sutcliffe’s research incorporated questionnaires as the measurement of attractiveness, but he also suggests that attractiveness can be measured by visitor dwell-time on Websites.\textsuperscript{19}

\textsuperscript{18} Zhang, p. 24.

\textsuperscript{19} In my opinion, the questionnaire in this measurement would be of low value; even though users might argue that a site is attractive, the truly attractive site is the one in which they linger for an extended period of time.
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

By formalizing the role it plays in the Web site user experience, Sutcliffe has made it easier to incorporate empirical results of attractiveness theories into the Web site design process. Rather than relying on the old, prescriptive instruction “avoid ugly or clashing colors,” designers now have an instructive heuristic helping them incorporate attractiveness into their Web sites.

**Conclusion**

In essence, it is time for us to release this inherited body of subjective, non-empirical Web site heuristics and replace them with a new methodology for finding best Web practices. The end of the term “Web site heuristics” would engender a new sense of openness and opportunity with the Web design community previously alienated by obscure terminology and autocratic attitudes.

---

**Figure 1. Web site visit model by Sutcliffe. In this model, initial attraction is the first component of the visitor’s experience with a Web site.**
Moving towards a new methodology: The retirement of the “Web site heuristic evaluation”

The evolving nature of the Internet dictates that there can be no “Web site heuristics.” Unlike other more firmly entrenched communications methods, the boundaries of Internet technology is not yet in sight, and it would be foolish for us to try to create static standards for Web design.

It is tempting to create new buzzwords or slogans to describe a proposed change in the current methodology. We have been introduced to such terms as “a new epistemology” and “heuristic evaluation.” It would be natural for me to suggest that we name this new approach of adding to Web site best practices “the new heuristics,” but that would deny the very nature of this new methodology – the sense of exploration and elasticity needed to evolve along with technology.
Moving towards a new methodology: The retirement of the "Web site heuristic evaluation"

Bibliography


Evans, Mary. “Challenges in Developing Research-Based Web Design.” IEEE Transactions on Professional Communication, Volume 43 No. 3 (September 2000) 302-312


“Top Ten Mistakes in Web Design” Useit.com Alertbox, May 1996

Position Paper, CHI 96 Workshop: HCI and the Web, April 1996

“How to conduct a heuristic evaluation” Undated.

