



Designing Web Sites for Older Adults: Expert Review of Usability for Older Adults at 50 Web Sites

February 1, 2005

Written by:

*Dana Chisnell
UsabilityWorks
San Francisco, California
www.usabilityworks.net*

And

*Janice (Ginny) Redish
Redish & Associates, Inc.
Bethesda, Maryland
www.redish.net*

Commissioned by and delivered to:

*Amy Lee, Director of Customer Service for Web
AARP
601 E ST, NW
Washington, DC 20049
www.aarp.org*

AARP is a nonprofit, nonpartisan membership organization that helps people 50+ have independence, choice and control in ways that are beneficial and affordable to them and society as a whole. We produce AARP The Magazine, published bimonthly; AARP Bulletin, our monthly newspaper; AARP Segunda Juventud, our bimonthly magazine in Spanish and English; NRTA Live & Learn, our quarterly newsletter for 50+ educators; and our website, www.aarp.org. AARP Foundation is our affiliated charity that provides security, protection, and empowerment to older persons in need with support from thousands of volunteers, donors, and sponsors. We have staffed offices in all 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands.

Table of Contents

- Foreword..... 1
- About This Project 2
 - Why did we create a new set of heuristics?.....2
 - How do these heuristics differ from others?2
 - What are heuristics, anyway?4
- Goals of the Study..... 4
- Methodology: How did we use the heuristics? 5
- What’s Different About Our Methodology?..... 6
- The “Participants”: Personas Matthew and Edith..... 7
 - How we “screened” the personas8
- Understanding the Findings 12
 - What the findings mean 13
 - What the findings don’t mean 14
- The Findings..... 15
 - General findings 15
 - Ease-of-use Ranking for Edith by Site Type 16
 - Ease-of-use Ranking for Matthew by Site Type 17
 - Findings by type of site 18
 - How well do Web sites support older adults?..... 18
 - Health Information 19
 - Health insurance and prescription drugs 22
 - Travel 27
 - Financial services and planning 29
 - Shopping 33
 - Portals and search engines..... 37
 - News..... 39
 - Hobbies and interests..... 43
 - What have we learned?..... 45
 - Key findings 45
 - Findings by heuristic 47
- Appendix 50
 - Heuristics 50
 - Interaction Design: Designing the way users work with the site..... 50
 - Information Architecture: Organizing the content..... 52
 - Visual Design: Designing the pages 53
 - Information Design: Writing and formatting the content 54
 - Tasks 55
 - Sites 55

Example evaluation worksheet..... 59

How to analyze the data1

Afterword 2

 Published, working precedents3

 Depth and breadth of review.....3

 Channeling personas4

 Coverage of the heuristics4

 Different “people,” different issues4

 Interpreting the data5

 A challenging method.....5

About the Authors 5

Other AARP Publications Related To This Report 6

Tables

Table 1. Which evaluators performed which tasks with which personas 12

Table 2. Sample data table 13

Table 3. Edith’s ranking of ease of use by type of Web site: Positive versus negative observations 16

Table 4. Edith’s failure rate by type of Web site: Site types listed from least number of failed tasks to most 16

Table 5. Matthew’s ranking of ease of use by type of Web site: Positive versus negative observations 17

Table 6. Matthew’s failure rate by type of Web site: Site types listed from least number of failed tasks to most 17

Table 7. Health information sites observations 19

Table 8. Health insurance sites observations 22

Table 9. Prescription drugs sites observations 23

Table 10. Health insurance and prescription drugs, total observations 23

Table 11. Travel sites observations 28

Table 12. Financial services and planning sites observations 30

Table 13. Shopping sites observations 34

Table 14. Portals and search engines observations 37

Table 15. News sites observations 40

Table 16. Hobbies and interests sites observations 44

Table 17. Task failures and problems totaled and then averaged per site visited 46

Table 18. Number of sites without task failures 46

Table 19. Total observations across all sites 47

Table 20. Heuristics that were most frequently satisfied 48

Table 21. Heuristics on which the personas most frequently encountered task failures49

Table 22. Who did what. One evaluator performed three of the high-level tasks (18 sites), the other did seven (46 sites). 5

Foreword

I buy a lot of heuristics reports. They've always been excellent, but they've missed a dimension that, until last year, I couldn't quite put my finger on. My focus is on "older adults", people over 50 years old. The existing heuristics seemed to me to be focused on people's disabilities rather than on people's abilities. Not everyone over 50 has eyesight poor enough to require maximizing the size or contrast of text of a web page. Not every person over 50 has problems with motor control or significant short term memory loss.

The diversity of this demographic group is stunning. Not everyone over 50 is new to the Web or afraid of their computer. Why are we trying to lump them all together like that?

Ginny Redish and Dana Chisnell took these questions to heart and came back with a new twist on using personas and heuristics that takes into account not only age, but the abilities, aptitude and attitude of the over-50 web site user. A web site may be perfectly functional for an 80-year-old who has been an avid web user for 10 years and dives into a new interface challenge with interest and excitement. The same site may be completely inaccessible to a 62-year-old who only turns on his or her computer to check email once a week. This new AARP Model takes those dimensions of the older user into account. The next step was to apply the model. We chose 50 web sites, from news sites to financial sites to health care and hobbies: Sites that are likely to be used by an older adult in the normal course of their day. We found that the World Wide Web is not an inherently hostile place for older adults, though there are some ways that sites could do better to increase their usability for everyone.

AARP represents "the power to make it better" for older adults anywhere in the world. This commitment doesn't just end with "the real world" concerns of Medicare, Social Security, aging in place or the myriad of other issues being addressed at any one time. It also extends into the click of a mouse.

Amy Lee
Director of Customer Experience for Web
AARP Services, Inc.

About This Project

In this document, we report our findings from an expert review of 50 Web sites. In a variation on the traditional heuristic evaluation methodology, we used heuristics derived from research about older adults and Web design to perform persona-based, task-based reviews of Web sites that older adult users are likely to go to. This is the final part of a multi-part project commissioned by AARP. The heuristics are extracted from our findings in *Designing Web Sites for Older Adults: A Review of Recent, Relevant Research*.

Why did we create a new set of heuristics?

- Available heuristics and guidelines didn't take into account factors important to improving the performance and experience of older adults as they use the Web.
- Some of the heuristics that practitioners have been using to evaluate Web sites were created for software. For example, there are software heuristics for implementing accelerators (ways to speed up interaction for advanced users) and for including help and documentation. Neither of these applies to most public Web sites.
- Most of the heuristics available don't really take into account people with different abilities. Since older adults have a wide range of physical and cognitive abilities that may or may not limit their use of, enjoyment of, and success with Web sites, it is important to consider those issues as well as their expertise with computers and the Web.
- Other guidelines we found for evaluating Web sites for older adults concentrated on Web sites for which older adults are the target audience. We focused instead on sites for general audiences that include older adults. There are, of course, many more of these sites than those that focus exclusively on older adults. If they are not easy for older adults to use, those potential users are missing out - and the organizations are missing a huge potential market. We wanted heuristics that we could apply to all sites where older adults might be relevant users.

As we discussed in our report on the research, the older adult audience is highly diverse in terms of age, ability, aptitude, and attitude. We wanted to create heuristics that practitioners could use to evaluate Web sites with these four factors in mind - on any Web site.

How do these heuristics differ from others?

Our heuristics are focused on usability and performance issues that older adults often have when using Web sites. We looked at several other sets of heuristics to get ideas about workable models:

- Jakob Nielsen's Ten Usability Heuristics, available at http://www.useit.com/papers/heuristic/heuristic_list.html

- Tec-Ed’s heuristics for evaluating user interfaces (unpublished)¹
- Ginny Redish’s heuristics for information-rich Web sites (unpublished)
- The National Institute on Aging and the National Library of Medicine’s checklist, “Making Your Web Site Senior Friendly” available as a PDF file at usability.gov/checklist.pdf or in HTML at <http://www.nlm.nih.gov/pubs/staffpubs/od/ocpl/agingchecklist.html>
- Czaja and Lee’s summary of interface design guidelines for older adults²

We borrowed elements of these ways of presenting heuristics to develop the form of our heuristics. Most heuristics and guidelines are too broad or too general. To actually use them, we must know more about what to look for. Therefore, in our heuristics set, each guidelines statement (for example, “Use adequate white space”) is followed by operationalizing questions (for example, “Is there line space of at least 2 pixels between clickable items?” and “Is body text broken up with appropriate and obvious headings?”).

We found this approach to be especially important and useful in this case because we were developing and using heuristics that apply to an audience with diverse skills and abilities. The questions point to the specific issues that the research shows or predicts older adults will have.

Our heuristics were derived from findings from our review of the research on Web site design and older adults. In reviewing the research, we focused on the disciplines of interaction and navigation, information architecture, presentation or visual design, and information design; and our heuristics also focus on those four areas.

We have three caveats for you to keep in mind about our heuristics:

1. Our focus is not on Web sites designed specifically for and about seniors. It is on Web sites of all types, from search engines to e-commerce sites, that older adults are likely to visit but that were not specifically developed only for older adults.
2. Although we are focusing on older adults, many of our heuristics may be important for all users. (As we drew the heuristics from research about older adults, we noticed that many of the issues raised were important to all users and show up often in lists of general heuristics for good design.) Thus, our findings may illuminate areas where good design for older adults is good design for everyone. However, that would remain to be tested with real users of all age groups.
3. Our heuristics are not a complete set of requirements for good design. We concentrated on issues that the research indicates are of particular importance to older adults. We did not include heuristics relating to all issues that might be important to all users. Thus, in evaluating or developing any specific site, you

¹ Tec-Ed is a usability consulting firm based in Ann Arbor, Michigan. We used their heuristics as input to this project with their permission.

² Czaja, Sara J. and Chin Chin Lee. “Designing Computer Systems for Older Adults.” Chapter in *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. pp 414-427. Julie A. Jacko and Andrew Sears, eds. Lawrence Earlbaum Associates, 2003.

should use the heuristics in this document in conjunction with other heuristics and guidelines, especially ones that are specialized for the type of site you are working on (Web application/software; information-rich Web sites; e-commerce; and so on).

What are heuristics, anyway?

Jakob Nielsen and Rolf Molich originated the methodology for heuristic evaluation in 1990. According to Nielsen's www.useit.com, "Heuristic evaluation is a usability engineering method for finding the usability problems in a user interface design so that they can be attended to as part of an iterative design process. Heuristic evaluation involves having a small set of evaluators examine the interface and judge its compliance with recognized usability principles (the 'heuristics')." ³

Wikipedia provides an excellent summary of the concept:

Heuristic is the art and science of discovery and invention. The word comes from the same Greek root (εὕρισκω) as "[eureka](#)," meaning "to find". A heuristic for a given problem is a way of directing your attention fruitfully to a solution. It is different from an [algorithm](#) in that it merely serves as a rule of thumb or guideline, as opposed to an invariant [procedure](#). Heuristics may not always achieve the desired outcome, but can be extremely valuable to problem-solving processes. Good heuristics can dramatically reduce the time required to solve a problem by eliminating the need to consider unlikely possibilities or irrelevant states...⁴

In psychology, heuristics are simple, efficient rules of thumb that people use to make decisions, typically when facing complex problems or incomplete information.

Heuristic evaluation of the sort we were doing differs from an "expert review," which is a well-established and much older method in which experts review from the heuristics in their heads. Here, we make the heuristics and underlying questions overt. In doing this review, we limited ourselves to the issues raised in our heuristics. We might have made many other comments about the sites if we were in "expert review" mode.

Goals of the Study

We had these goals for the project:

- See how typical Web sites are working for older users.
- Identify common usability and design issues specific to older users.
- Show examples of designs or design elements that work well and that do not work well.

³ http://www.useit.com/papers/heuristic/heuristic_evaluation.html

⁴ <http://en.wikipedia.org/wiki/Heuristic>

- Provide guidance to designers and developers of Web sites who have older adults in their audiences.
- Add information about e-commerce Web sites and Web transactions to AARP's Older Wiser Wired (OWW) repertoire.
- Provide AARP internal and external partners with knowledge and tools for better design of Web sites for their AARP audiences.

Methodology: How did we use the heuristics?

Typically, heuristic evaluations are conducted using the heuristics as a checklist against which products are assessed. We took a different approach. We reviewed the sites through personas that represent major parts of the older adult audience and performed tasks that they would be likely to do.

Here is our process:

1. Create (or in our case, expand) personas or detailed user characteristics. The personas we used are described on page 7.
2. Using the personas and their characteristics, assess where they fall on our audience segmentation model. The model is described on page 50. Doing this helps evaluators understand how much complexity a persona can tolerate and how much training and support older adults might need on a particular Web site.
3. Define high-level tasks that the personas will perform on different Web sites. We had five high-level tasks: Start the day; do some research about options my doctor has given me; plan a trip; do the monthly household bookkeeping; plan and do a project, including shopping online and following up on leisure activities or hobbies. (More about this on page 55.)
4. Select Web sites to evaluate that are appropriate for the personas to use to carry out the high-level tasks. The list of sites we evaluated starts on page 55.
5. Taking into account the characteristics of the personas and where their attributes fall in the audience segmentation model
 - perform tasks that are realistic for that persona on sites from our list with the heuristics in mind. The heuristics we used start on page 50. As an example of tasks that the personas did, consider how the persona "Edith" did Start the day. She went to AOL, checked her email, then went to CNN.com and checked the news headlines; she also checked out BBC.com because she wanted to know the British perspective on the U S election, and so on. We include the list of sites visited with our summary of each type of site. You will also find a complete list of sites and which persona went to which on page 55.
 - record observations while performing the tasks on the Web sites.

6. Rate observations against the heuristics and questions using a worksheet like the one on page 59. Use the questions within each heuristic to specify what the issues and the successes are.

What's Different About Our Methodology?

In many heuristic evaluations, evaluators use the heuristics as a checklist, marking violations or successes for each heuristic. In many other heuristic evaluations, experts review a site with the heuristics in mind, going through the site methodically screen by screen.

We (Ginny and Dana) have both typically done heuristic evaluations more like usability testing. We ask our clients for descriptions of users and typical tasks and then exercise the site we are reviewing with those users and tasks in mind. Of course, because we are not (usually) the real users, we are still only predicting what will happen when users work with the sites.

In this project, we carried our version of heuristic evaluation to an even greater user-oriented, task-oriented extent.

The personas (Matthew and Edith) were much richer than the typical user profile we have worked with in the past. Calling on our many years of professional experience, we really “channeled” the personas. Matthew and Edith became real personalities with relationships, habits, and emotions beyond what was scripted in their original persona descriptions.

Instead of very specific tasks as in a typical usability test or the types of tasks we both usually agree on with the client in a heuristic evaluation of one specific site, we let the personas select the tasks within broad categories and within our list of 50 sites. Thus, our technique had elements of the broad, open-ended usability testing that Jared Spool and others use in some of their usability research.

We used this more open-ended task-based technique because we were reviewing 50 sites, not just one, and we had to do it within a limited time and budget. It also has the advantage of being realistic. Our personas acted as real users probably would, moving from task to task and site to site based on individual goals, individual needs, and elements of sites that attracted their interest.

In the guise of our personas, we made observations that were much like the “think aloud” verbal commentary that we typically get from participants in a usability test.

The heuristics came into play as we went back through the observations, noting which heuristics were relevant to the observation and how the persona would score the observation for those heuristics.

The “Participants”: Personas Matthew and Edith

Meet our “participants” - the personas we “channeled” as we reviewed Web sites:

Matthew



Matthew is a 54-year old attorney from New York City. He’s married, and he and his wife work full

time. Their income averages six figures, and it ought to for the hours they each work.

They own a one-bedroom condo in Park Slope in Brooklyn and a small cottage in Niantic, Connecticut where they go for weekends in the summer. Matthew commutes to work from Brooklyn on the subway. He and his wife don’t have kids, but their nieces and nephews come to visit for a week or two at school breaks.

Matthew doesn’t have much time for the Web. He uses email at work, and sometimes makes vacation plans or reads the newspaper online. Mostly, it’s a tool to get things done. Fast. When something doesn’t work right away, Matthew moves on. He doesn’t have the time or patience to figure it out.

Matthew doesn’t use the AARP Web site, but he expects that when it’s time to renew his membership, he’ll try to do it online this year and save himself the paperwork.

Matthew is still feeling fine although his doctor says he needs to exercise more. He wears contacts; his eyes aren’t what they were when he was younger.

Edith



Edith is 73 years old. She lives in Miami, Florida, with her husband of 49 years, Doug. They worked hard in the restaurant

business, and she’s glad that they have retired. Their income is about \$40,000 a year, from Social Security and what they got when they sold the restaurant.

They used much of the money from selling the restaurant and their house up north to buy a small retirement house in Florida.

They put down a lot of cash for the Florida house to keep their house payments low.

Edith and Doug like to joke that they can’t count how many grandkids they have anymore. It’s been too long since they’ve seen each other. Sometimes they get pictures through email (how do their kids do that?), and that’s nice. They can print them out.

Edith really doesn’t use the internet much - and neither does Doug. She’s never been to the AARP Web site before. It had not occurred to her that there was such a thing until she saw something about it in the AARP magazine last month. The magazine did a special on Boulder, Colorado, as a good place to visit. Edith thinks she’d like to go there with Doug for their anniversary. She wants to find about what AARP has to say about Boulder.

Edith’s hearing isn’t what it used to be. She likes that her hair style covers her hearing aids. She took off her glasses for the picture, but she needs them to read or look at the computer. She has slight arthritis in her hands so sometimes using the mouse is a problem.

Figure 1. Two personas for our persona-oriented, task-based review of Web sites

How we “screened” the personas

A key element in a successful usability test is selecting appropriate study participants. To conduct this heuristic review, we needed to identify user characteristics to represent types of users in the older adult audience. AARP has developed eight personas to use in designing AARP.org. Because we didn't have enough resources to use all of the personas in the evaluation, we selected two, Matthew and Edith.

Based on readings for our review of the recent research and our own experiences, segmenting the older adult population simply by age is not enough. A more nuanced approach is needed in order to

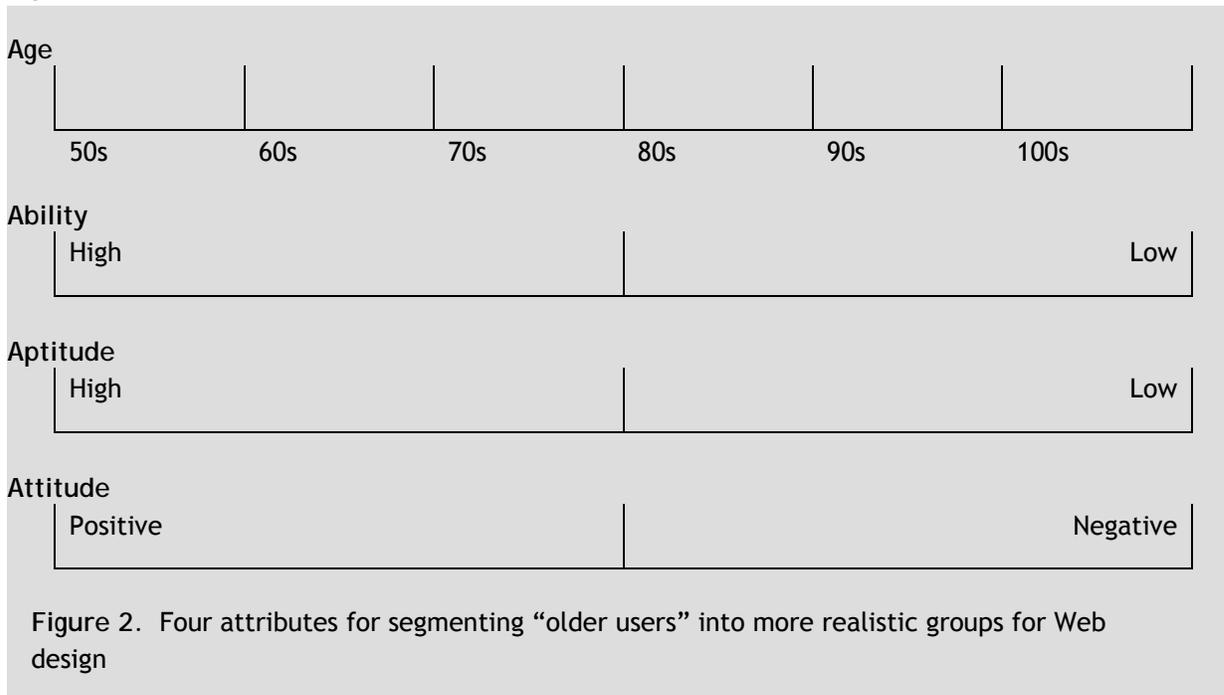
- understand the common attributes among people in a large and diverse audience
- gain insight into how much support users might need to use a given Web site
- examine how the levels of complexity designed into Web sites match up with what older adults can use effectively

We propose a new approach for quantifying characteristics in combination to identify more specific audiences within the population of older adults. The tool we propose could be used by Web design teams to help them make decisions about where their users fall along these dimensions and thus how best to serve their audiences. We developed this model in response to our review of research about older adults. However, it may work equally well for younger audiences, or audiences with physical or cognitive limitations. Our model and variations on it are untested.

Our approach looks at the four factors we have been discussing:

- age: including chronological age, but taking into account life experiences
- ability: cognitive and physical
- aptitude: expertise with the technology
- attitude: confidence levels and emotional state of mind

Figure 2 shows the four factors.



Finding 1. What goes into these four attributes?

Age - both chronological and experiential, along with maturity level; life events and experiences (i.e., various jobs, not just the most recent; military service; marriage, divorce, children, places lived); education level (including when it was achieved)

Ability - degrees of physical and cognitive limitations or restrictions⁵ requiring little remediation up to needing assisted living⁶

Aptitude - expertise with computers and the Web (being more relevant than straight measures of experience)⁷

Attitude - positive (forward looking, risk-taking, and experimental) or negative (fearful or diffident), confidence levels, and emotional need for support from another human being⁸

Finding 2. What are the implications of these four attributes?

We can use these four attributes to judge the need for support and training and the level of complexity of features and functions that different users can be expected to handle. As we move towards the right on each of these dimensions, users are likely to need more support and training

⁵ Theofanos and Redish, 2005

⁶ Jacko, et al., 2002 for health and accessibility testing

⁷ See all of the Chadwick-Dias, et al. studies published in 2004 listed in our review of the research

⁸ Gregor, Newell and Zajicek, 2002; Hawthorn, 2003; Kantner and Rosenbaum, 2003

and less complexity in Web sites. Users far on the left of each of these dimensions are more likely to need less support and training and be able to handle more complexity.

Of course, the interesting and important issues are the trade-offs that are likely to be necessary for people who are in different places on different dimensions. That is, increased age is likely to require less complexity, but increased aptitude allows for more complexity. Higher ability (that is, physical and mental fitness) allows for more complexity, and higher ability is likely to also correlate with lower age.

We selected these two personas because they represent a fairly broad range of aptitude and ability (without reaching into major physical and cognitive accessibility issues). We also selected them because, although they both are positive about being online, their experiences so far with the Web have been rather different.

We can place each of our personas on the four dimensions to give us a sense of the differences in how they approach Web sites and what parts of the older adult audience they probably do and do not represent.

Figure 3 shows two personas aligned on the four factors.

M is Matthew and E is Edith.



Figure 3. Two personas along each of the four dimensions

Ideally, perhaps in another project, we would create specific assessments for each of the factors to add rigor to the model. For now, we have placed our personas intuitively, based on the information in the personas.

Matthew is more able than Edith, so sites that are complex may be easier for Matthew to use than they are for Edith. Some sites may be much more difficult for Edith if they don't follow conventional interaction practices or if they use computer and Web jargon. Edith's generally positive attitude may help her cope, however - she may be willing to explore a little bit to find what she needs. Matthew is confident enough to experiment with trying different links, but he isn't very patient.

Understanding the Findings

Because the methodology we have used is different from the classic checklist methodology for performing heuristic evaluations, the findings are different, too.

In the checklist methodology, on every page of a Web site, evaluators look for violations of each of the heuristics. The evaluators generally don't take into account the abilities and expertise of the users or of different types of users. Often, evaluators don't take typical tasks into account. A heuristic evaluation can be very system-oriented rather than task- and user-oriented.

In our methodology, we made observations "through" our two personas while we performed tasks as the personas. Before we took the personas to the Web, we specified the high-level tasks (for example: Start the day; Do monthly bookkeeping) and we developed a list of Web sites for them to choose from. Within those high-level tasks and within the list of Web sites, the personas formed their own detailed tasks and selected Web sites to use based on each persona's characteristics and attributes.

We viewed the tasks and sites as "natural" extensions of each persona. For example, Matthew visited Weather.com because he had plans at his country house on the weekend and wanted to know whether he would be able to be outdoors. Edith lives in Florida where the weather is the same most of the time, so she was not concerned about weather-dependent activities. Edith went to genealogy sites because that's an interest she has. Matthew's project revolved around his interest in birdwatching.

We used the personas to make predictions and generalizations about tasks, behaviors, and performance on types of Web sites rather than gathering the quantitative quality checks that you get through using heuristics as checklists.

Between them, our two personas used 50 sites, but they did not both visit all those sites. Edith went to 32 sites; Matthew to 30 sites. There were 11 sites that they both used. Even on those 11 sites, however, they did not perform the same detailed tasks using the same pages and the same steps. Each of us took a persona through an entire high-level task.

Task	Persona/Evaluator	
	Edith	Matthew
1. Start the day	Dana	Ginny
2. Research about medical / health options	Dana	Dana
3. Plan a trip	Dana	Dana
4. Do monthly bookkeeping	Ginny	Dana
5. Leisure activities or hobbies, shopping	Ginny	Dana

Table 1. Which evaluators performed which tasks with which personas

After we had each done the first high-level task, we reviewed the observations and ratings from that task to gauge how similar our observations had been. We found that we had both

approached the evaluation the same way and we had made similar observations. Our ratings were sometimes different but that was realistic because the personas were different. As with real users, the severity of the problems was different for the different types of users. For example, Matthew, who is younger and more computer savvy, might have rated an observation as a 2 because he could get past an obstacle or recover more easily from an error than Edith. Edith, who is older and much less computer savvy, might have given the same issue a 1 because she didn't know what to do to move on.

What the findings mean

The Findings starting on page 15 show that for the observations made during the tasks done on the sites visited, we rated the observations as

4 = No problem - satisfies the heuristic

3 = Minor hindrance - possible issue, but probably will not hinder this persona/user

2 = Serious problem - may hinder this persona/user

1 = Task failure - prevents this persona/user going further

Realistically, people in usability studies notice different things and have different types of problems. We didn't run a usability test, but the personas we used did "notice" different things sometimes, and they had similar problems but to different degrees. So, for example, a table such as the one below for news Web sites shows that we made a total of 54 observations as Edith and 36 observations as Matthew. ("Observations" comes from a running commentary in which the reviewers-as-personas made both positive and negative comments.) Within that number of observations, 20 of the observations that Edith made rated No Problem or Minor Hindrance; 14 of the observations that Matthew made rated similarly.

Scores on heuristics for Edith and Matthew for news sites:

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	8			12		
Minor hindrance	3	12			2		
Total for 4 & 3			20	37.0		14	38.9
Serious problem	2	20			19		
Task failure	1	14			3		
Total for 2 & 1			34	63.0		22	61.1

Table 2. Sample data table

The raw numbers of observations are not as important as the percentages. The percentages give us information about the proportion or ratio of positive to negative observations. Proportionately - based on the percentages - Edith and Matthew were similarly successful (and unsuccessful) using news sites. But the bottom line - the lowest rating - shows that there were only a few

issues that Matthew couldn't get by, whereas Edith was stopped or gave up much more often. This finding reflects both the abilities and expertise of the personas and the ease or difficulty for those personas on this type of Web site.

What the findings don't mean

The data in tables like the one above do not mean that Edith was successful 20 out of 54 times and that Matthew was successful 14 out of 36 times, and so on.

The findings do not mean that we rated every heuristic on every observation on every task on every site.

Many of the observations were single issue observations. For example:

Type is pretty small; makes me realize how bad my eyes are getting.

This observation only describes heuristic 14, Make elements on the page easy to read.

Other observations touched on multiple heuristics. For example:

Web page is really busy, but it is well organized. It's by topic. There's a lot here.

This observation describes heuristics 9, 11, 13, and 15:

- Make the structure of the Web site as visible as possible.
- Implement the shallowest possible information hierarchy.
- Make pages easy to skim or scan.
- Visually group related topics.

The same heuristic may have come up in more than one observation on a site. If it was on a different page in the site or for a different task within the site, we scored it again. In that way, sites that helped our persona in one place but that caused problems in another place may have gotten both a positive and a negative score for the same heuristic. Medicare.gov, for example, generally did a great job of writing plain language instructions for finding a discount drug card, but used jargon in a link for changing the size of type for the site. If the site really impressed the persona with a particular aspect in different places or for different tasks, it might have gotten two positive scores or two negative scores.

At the end of the observations for a particular site, some heuristics may have received no observations and, therefore, no scores for that site. The absence of data for a particular heuristic for a particular site does not mean that the site satisfied or did not satisfy that heuristic. It most often means that our persona did not go to pages where that heuristic was relevant. If no confirmation or error messages happened to come up during our visit to a site, heuristic 7 had no observations and no score for that site. It may also mean that our persona did

not notice that aspect of the site. Edith did not even look for a site map on many sites and so did not score heuristic 12 on those sites. We were following realistic paths for users like our personas and acting in ways that we believe strongly are very much what real users would do - and that meant that not all heuristics were touched on all sites.

The Findings

There are 20 heuristics that house about 100 questions. We report on the 20 heuristics, using the questions (or rather, the answers to them) to give depth to our findings. We do not report on all 100 questions.

Edith worked at a screen resolution of 800 x 600, sometimes on a 15-inch flat panel monitor and sometimes on a 19-inch flat screen monitor. Matthew worked at a screen resolution of 1024 x 768 on both monitors. Edith used AOL and its browser on the Start the day task; Edith and Matthew used Internet Explorer for the remaining tasks.

General findings

Clearly, the Web is a friendlier place to Matthew than it is to Edith. (See the personas and their segmentation on page 7.) We had expected that the same types of sites might be more difficult for Edith than for Matthew because of her lack of expertise on computers and the Web, and because of some age-related ability problems due mostly to having arthritis in her hands. This was indeed the case.

By the end of the first task, Start the day, it was clear that Matthew could perform tasks such as checking email and reviewing the day's news headlines on a variety of sites with relative ease. Edith had much more difficulty. For example, Edith had a tough time determining what was clickable and what was not and found scrolling to be a problem (perhaps because she set her screen at a lower resolution).

There was one exception: Health insurance Web sites. Edith is 73, giving her the privilege of being enrolled in Medicare. Her experience using Medicare.gov was perhaps the best of the 32 sites she used. It was the only situation in which she encountered no task failures and had more positive observations by proportion than Matthew did on his equivalent site.

The tables below show rankings for each of the personas by different measures. On the left side of each table the types of sites are ranked by the percentage of positive versus negative observations in those types of sites. The right side of each table shows how the sites rank in terms of failed tasks, with types numbering the fewest failed tasks at the top and the most failed tasks at the bottom.

It is important to remember that while each of the personas evaluated about the same number of Web sites, they did not evaluate all of the same sites; neither did they do the same tasks on the same types of sites. They did the tasks they "wanted" to do, so the findings are not strictly comparable. The outcomes of our methodology shouldn't be used as a "report card." Instead, they should be used (like a usability test) to provide insight into how a site is used and to predict where different types of users are likely to succeed and where they are likely to have problems.

Ease-of-use Ranking for Edith by Site Type

For Edith, by either measure, the health-related sites that she used were the most likely to satisfy the most heuristics. Travel sites also rated highly, but news and financial sites did not.

	Type ranking	Percent positive	Percent negative
Good	1. Travel	66.7	33.3
	2. Prescription drugs	66.7	33.3
	3. Health information	56.0	44.0
	4. Health insurance	55.6	44.4
	5. Shopping	55.4	44.6
	6. Portals / Search	50.8	49.2
Poor	7. Financial services and planning	47.1	52.9
	8. Hobbies and interests	40.5	59.5
	9. News	33.9	66.1

Table 3. Edith’s ranking of ease of use by type of Web site: Positive versus negative observations

	Type ranking	Failed tasks	Serious problems
Good	1. Health insurance	1	19
	2. Prescription drugs	3	11
	3. Health information	3	30
	4. Travel	3	35
	5. Hobbies and interests	8	36
	6. Portals / Search	12	19
	7. Shopping	15	14
Poor	8. News	16	23
	9. Financial services and planning	52	56

Table 4. Edith’s failure rate by type of Web site: Site types listed from least number of failed tasks to most

Ease-of-use Ranking for Matthew by Site Type

Ratings for Matthew are more of a mixed bag. Hobbyist and shopping sites were easier than they were for Edith, but health information also rated quite high for Matthew.

	Type ranking	Percent positive	Percent negative
Good	1. Hobbies and interests	94.5	5.5
	2. Shopping	91.2	8.8
	3. Health information	80.8	19.2
	4. Portals / Search	73.8	26.2
	5. Financial services and planning	71.9	28.1
Poor	6. Prescription drugs	65.0	35.0
	7. Travel	59.1	40.9
	8. Health insurance	41.4	58.6
	9. News	38.9	61.1

Table 5. Matthew’s ranking of ease of use by type of Web site: Positive versus negative observations

	Type ranking	Failed tasks	Serious problems
Good	1. Hobbies and interests	0	3
	2. Shopping	0	13
	3. Health information	1	9
	4. Portals / Search	2	15
	5. Travel	2	34
Poor	6. Financial services and planning	2	43
	7. News	3	19
	8. Prescription drugs	3	25
	9. Health insurance	7	10

Table 6. Matthew’s failure rate by type of Web site: Site types listed from least number of failed tasks to most

It is important to note that health information sites were fairly easy for both Edith and Matthew. Also, although Matthew “experienced” relatively few failed tasks (scored as 1) - points at which

he was prevented from going further - he still encountered many, many serious problems (scored as 2) that might hinder real people and create obstacles to task success. These serious problems weighed heavily in the rankings above.

Findings by type of site

Through the personas, we reviewed and evaluated 50 Web sites in eight categories:

- Health information
- Health insurance and prescription drugs
- Travel
- Financial services and planning
- Shopping
- News
- Hobbies and interests
- Portals and search engines

How well do Web sites support older adults?

Some of the sites included in the study seem to be paying attention to the older adults in their audiences. Information Web sites related to health concerns are the most tuned in. We think that people like Edith and Matthew will find sites that are like the health information sites we reviewed to be useful and usable.

Many of the other types of sites were at best difficult, and at worst agonizing, for Edith, our older, less Web-expert persona. In her “best” experiences, she had serious problems or task failures only about a third of the time. But on average, on the rest of the sites, Edith met with serious problems or task failures half of the time. This level of difficulty diminishes the usefulness and desirability of doing things online like shopping, paying bills, reviewing retirement accounts, or looking deeper than news headlines.

Although some types of sites do well at supporting older adults, these sites are generally those that think of older adults as a large part of their audience. For example, health information and prescription drug sites probably expect that most of their visitors are age 50 or older because that is when many of the conditions covered begin to set in. Health information and sites about prescription drugs represent only 16% of the sites we reviewed (8 of 50).

Types of sites with audiences in a wider age range, such as travel (7 sites reviewed), shopping (10 sites), and financial services (8 sites) could do more to support older adults while still making their sites easier for younger people, too. For example, implementing consistent link treatments within sites, including Close buttons on pop-up windows, and taking a minimalist approach to layout and content would help older adults avoid confusion and use sites more efficiently. The same is probably true for younger users as well.

Younger users are more familiar with the language of the Web than older adults are. This becomes important as older adults encounter labels on navigation, buttons, fields, and links that use language that is unfamiliar to them and prevents them from inferring what the next content or step might be if that interaction widget is clicked. (For example, “Bookstore” as a navigation item when you are already at a bookstore’s Web site; “Add Drug” when it is unclear what the drug is being added to.) But the full age range of users for travel, shopping, and financial sites would be helped by clearer instructions and error and confirmation messages, minimal introductions, and task-oriented headings, rather than marketing- or organization-oriented headings. Focusing on users’ tasks and goals is the best kind of marketing.

Note: We have written the findings as if a real person performed each of the tasks. Keep in mind that Edith and Matthew are fictional representatives of composite user characteristics. It was simply easier to write the findings as if the evaluators had observed the personas performing the tasks.

Health Information

Americanheart.org
 Cancer.org
 Drkoop.com
 NHLBI.NIH.gov
 WebMD.com

Edith and Matthew used five health information Web sites between them. Edith used three; Matthew used two. Although the personas reviewed none of the same sites, both personas found the sites they used to be easy to skim and read. For the most part, the sites did a good job of writing to a general consumer audience. However, there were some spotty exceptions to this within sites (more about this below).

The sites satisfy many of the heuristics. Matthew had relatively few problems. Edith had a number of serious problems but fewer task failures than on other types of sites.

The table below shows the number of observations that reviewers rated as 4, 3, 2, or 1 for each of the personas on health information sites.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	20			23		
Minor hindrance	3	22			19		
Total for 4 & 3			42	56.0		42	80.8
Serious problem	2	30			9		
Task failure	1	3			1		
Total for 2 & 1			33	44.0		10	19.2

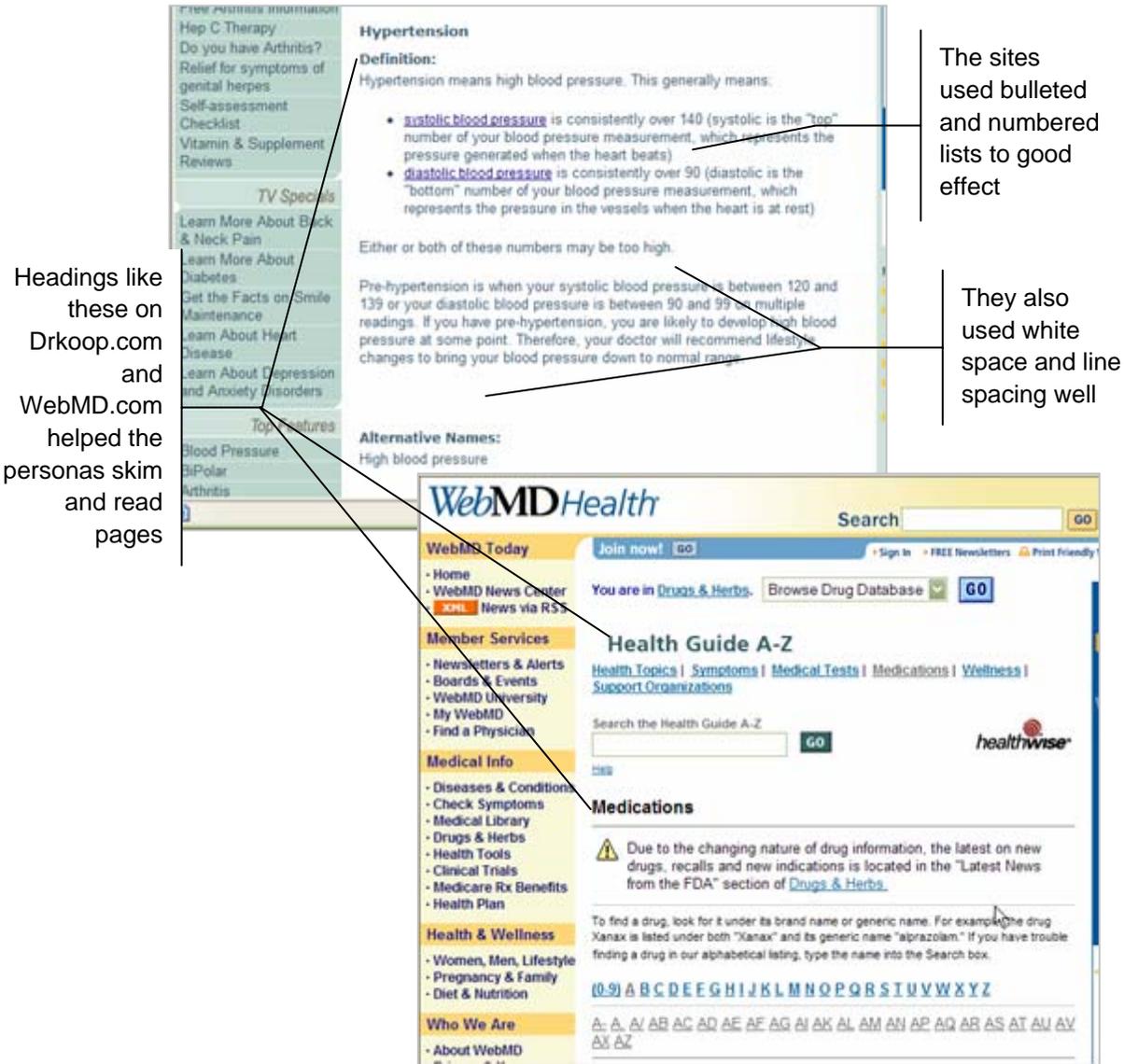
Table 7. Health information sites observations

Finding 3. What the numbers mean

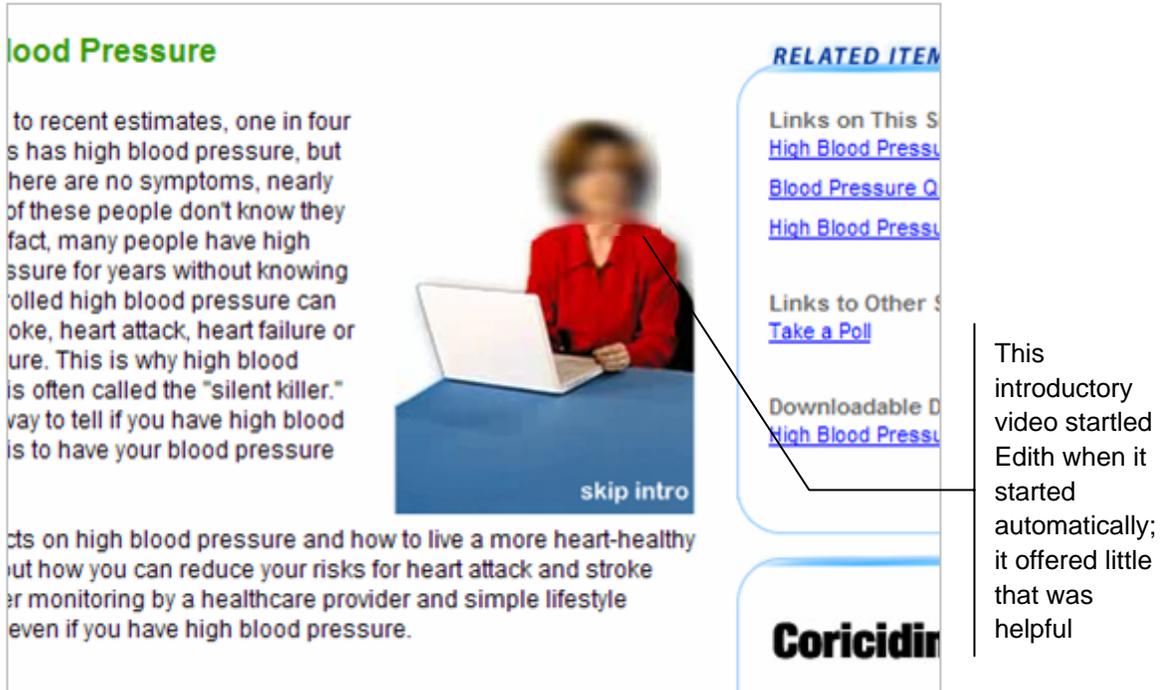
Health information sites caused Edith relatively few task failures. Clearly, Matthew’s experiences on health information sites were also very good, but there were still problems.

Health information Web sites placed third for both Edith and Matthew in terms of percentage of positive observations. They rated second for Edith and third for Matthew for fewest failed tasks.

What made these sites usable and useful to people like Edith and Matthew is how easy they were to skim and read pages: They presented clean, well organized pages with large, noticeable headings and made excellent use of bulleted and numbered lists. They used white space for grouping topics that are related. WebMD.com, Drkoop.com, and Americanheart.org showed especially well in these areas, but Cancer.org, and NHLBI.NIH.gov (the National Heart, Lung, and Blood Institute of the National Institutes of Health) also did very well.



The sites generally did a good job of addressing the content to the audience and meeting their purposes with appropriate language, but there were a few instances in which the sites were less successful in satisfying these heuristics.



Images on screen shot are not reproduced per copyright restrictions.

While almost all of the headings on Cancer.org were very good, small things on what could be important topics kept Matthew from finding content and understanding it once he got there. For example, a heading “How is Thyroid Cancer Staged” could have been recast to be simpler: “What are the stages of thyroid cancer?” Matthew also had difficulty interpreting information about survival rates.

Edith encountered similar problems but stopped using WebMD.com when she had to agree to a disclaimer written in legalese about the reliability of the information about prescription drugs. She wasn’t sure what the disclaimer was saying, and she was offended at the change in tone of content from being helpful to being defensive.

Both personas encountered unevenness in the level of information presented within sites. For example, some of the information on Americanheart.org was descriptive and useful, but there were needless introductions and overviews that made the structure of the site deeper than it should have been.

Edith was also very surprised on Americanheart.org when a video of a person started addressing her as soon as the page loaded. The person doesn’t introduce herself. At the end of the clip, the video person says, “Let us know if we can help you” but doesn’t say how to do that. Something that was intended to be friendly and reassuring actually offered nothing helpful; and that was a bit startling.

Edith also complained about scrolling. Especially on pages listing drug names or conditions, Edith needed a better way to manage the large numbers of items besides having to scroll the length of the pages.

Health insurance and prescription drugs

Aetna.com
 Drugs.com
 Medicare.gov
 Medicare.org
 MedlinePlus.gov
 Purplepill.com

This is the only category of Web site in which Matthew had more negative observations and more task failures than Edith did. Both personas used MedlinePlus.gov and Drugs.com to look up information about prescription drugs. However, their insurance needs were different - Edith’s overall positive experience is due mainly to the difference in usability between Medicare.gov and Aetna.com.

If we look only at the totals for this category (Table 10 on the next page), the experience seems quite even for Edith and Matthew - which is unusual in itself. However, if we look more deeply into the totals (comparing Table 8 to Table 9), we find that Edith had a better experience on her health insurance site than Matthew did on his. Edith went to Medicare.gov; Matthew went to Aetna.com. This shows how valuable satisfying usability principles for older adults can be. Even though Medicare.gov was not architected to support the task Edith did there, it was easier for her to use than the site Matthew went to was for him to use.

The tables below show the number of observations that reviewers rated as 4, 3, 2, or 1 for each of the personas on health insurance and prescription drug sites.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	20			4		
Minor hindrance	3	5			8		
Total for 4 & 3			25	55.6		12	41.4
Serious problem	2	19			10		
Task failure	1	1			7		
Total for 2 & 1			20	44.4		17	58.6

Table 8. Health insurance sites observations

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	16			30		
Minor hindrance	3	12			22		
Total for 4 & 3			28	66.7		52	65.0
Serious problem	2	11			25		
Task failure	1	3			3		
Total for 2 & 1			14	33.3		28	35.0

Table 9. Prescription drugs sites observations

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	36			34		
Minor hindrance	3	17			30		
Total for 4 & 3			53	60.9		64	58.7
Serious problem	2	30			35		
Task failure	1	4			10		
Total for 2 & 1			34	39.1		45	41.3

Table 10. Health insurance and prescription drugs, total observations

Finding 4. What the numbers mean

It appears that the evenness of experience on these Web sites between Edith and Matthew is due to the high usability of the two government-owned sites: MedlinePlus.gov and Medicare.gov.

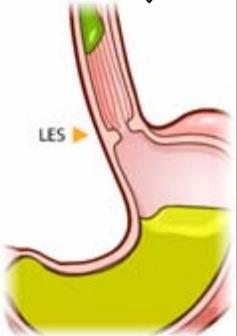
Most of the sites reviewed did very well at supplying content in plain language and meeting the personas' level of domain knowledge. While Matthew's being an attorney might give him an edge on understanding complex language, he had never researched prescription drugs before; Edith had more experience with this type of task. The sites did especially well at introducing and explaining technical or domain-specific terms.

These sites also used cross-reference links and redundant links well. They followed link conventions. And they made things easy to find on pages by implementing minimal text, clear headings with plenty of white space between sections, and bulleted lists. Pages at every level were easy to skim and easy to read. Purplepill.com got points from Matthew for its simple, short, and clear animated sequences of what the drug does and how it works.

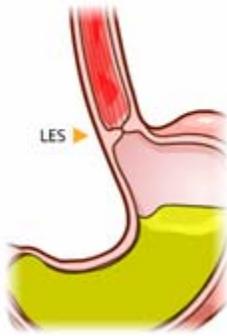
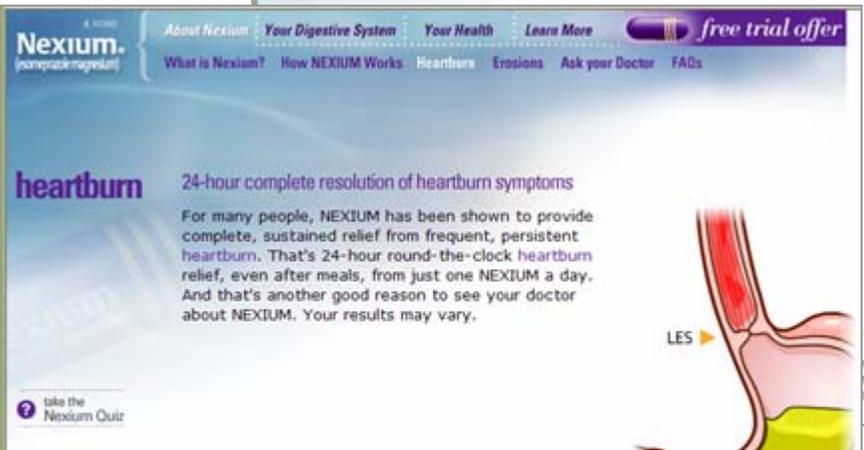


The animation highlights the affected area, and then zooms in on it

While the background detracts from the usability of the page, the navigation is large and user-oriented and the animated sequence has a white background



The Lower Esophageal Sphincter (LES) is a muscle at the bottom of your food pipe or esophagus.



This leaves the esophagus red & irritated

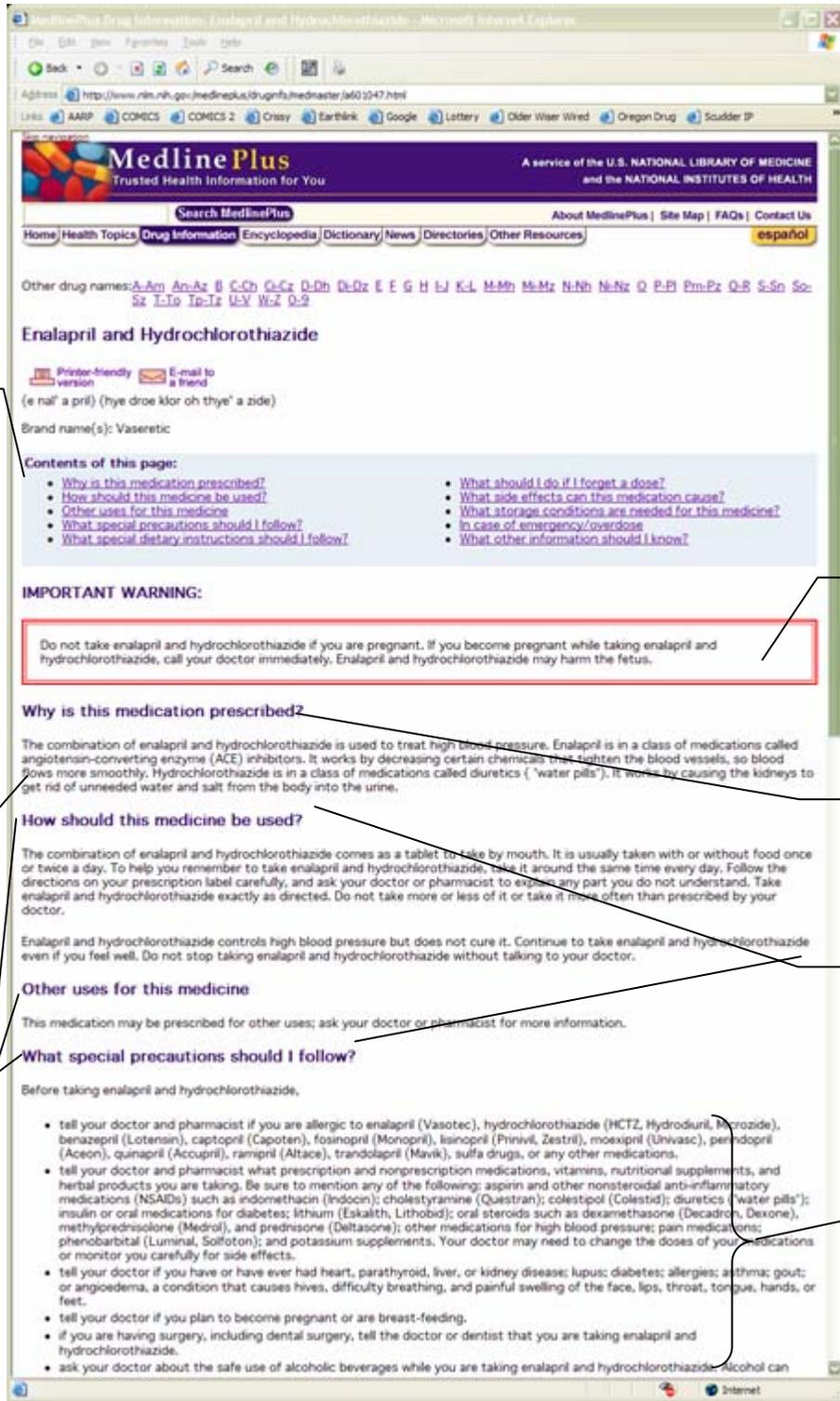
Erosions heal in 4 to 8 weeks with minimal pain. Symptom relief does not occur until NEXIUM has dissolved it with water.

The captions are concise and direct, in large type, and are presented for several seconds - long enough to read

The lack of contrast between the text and background may actually help focus attention on the useful information in the animation

There were a few spots on Drugs.com and MedlinePlus.com where the language could have been clearer and plainer.

The Contents box presents an advanced organizer of the page and acts as a set of cross-reference links



Important information is close to the top, called out clearly

The type is large enough to read

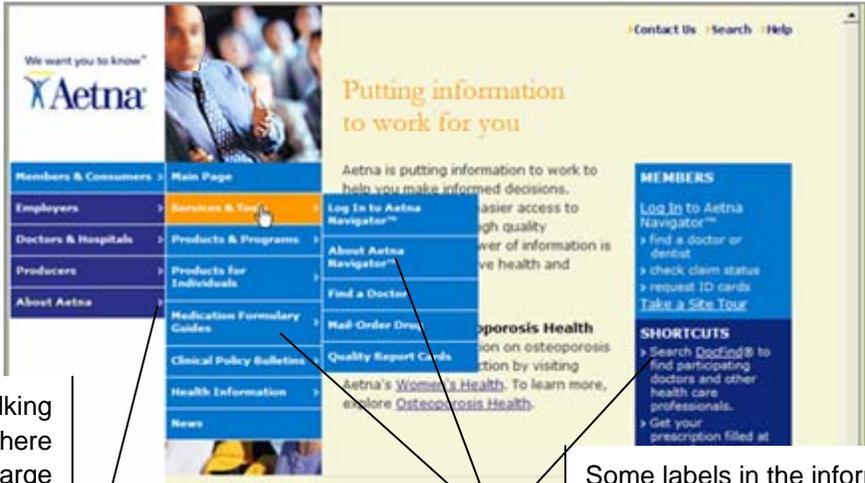
Headings are cast in patient-friendly questions

Headings are noticeably larger than body type

There is ample white space between sections

Although the site makes ample use of bulleted lists, this example could have been implemented better to help patients skim information more efficiently

But the site most likely to violate heuristics related to plain language and understanding the user's point of view was Aetna.com. Through its content and information architecture, Aetna's recently redesigned site demonstrates that it isn't quite as customer-oriented as it could be by exposing its company-centric and product-centric organization.



Images on screen shot are not reproduced per copyright restrictions.

The walking menus here provided large targets and the type was large and easy to read

Some labels in the information architecture exposed a marketing orientation rather than customer centeredness

If you're looking for a health insurance plan, how can you tell from this what the plans are and what they offer?

This list only works if you're looking up a plan that you already have and you know the name of the plan



Matthew's serious problems on the drug Web sites were spread fairly evenly among the three he used: Drugs.com (10), MedlinePlus.gov (7), and Purplepill.com (8). His issues on Drugs.com had mostly to do with interaction elements. Matthew found that the things that were clickable were not always obvious or easy to click on because of their small size. MedlinePlus.gov and Drugs.com were much larger sites, and Matthew had difficulties with navigating through the deep information architectures on both sites. Matthew also found Drugs.com to be very difficult to skim, the type in many places was too small to read easily, and he had problems finding things on pages.

Edith's task failures with drug sites happened on Drugs.com. (She had one task failure on Medicare.ORG - not to be confused with Medicare.GOV.) Matthew did have quite a few serious problems on Drugs.com, but nothing that stopped him. While Edith had no task failures on MedlinePlus.com, Matthew got stuck three times there. Why the difference?

It was not the case that a site designed to be simple and usable for people like Edith was too limited for people like Matthew. Matthew used MedlinePlus.gov to get information about Nexium, which is a brand name, but MedlinePlus indexes prescription drugs by their active ingredients. When he searched on the brand name, the site gave him 24 links to active ingredients, supposedly related to Nexium, but he saw no way to tell.

Both personas had difficulty going backward in MedlinePlus.gov - Edith within a content page with anchored links, Matthew wanting to back out of Search results to try a different navigation method. The Back button worked for Edith; Matthew never found what he wanted.

Although Medicare.gov did the best job of satisfying the most heuristics the most consistently of the sites, there were a few odd things that could be confusing to people like Edith:

- There's a disclaimer about using the drug information on the site for purposes other than personal information that seemed perhaps too threatening.
- While there is a link for changing the type size on pages (which appears only in Internet Explorer but not in other browsers such as Netscape and Mozilla), the label for the link toggles between "Use Larger Font" and "Use Default Font." Using the word "type" rather than "font" might be better, but the word "default" is definitely not something that most of the Medicare audience would be familiar with.
- In addition, although the "Quick Search" for finding discount drug cards does a decent job of shepherding users through the process, some of the button labels could be clearer, and there's no indication of how many steps there are in the process and how many remain.

Travel

AAA.com	Both personas reviewed one site - Expedia.com - but also reviewed another six sites between them. Both had fairly positive experiences with the travel sites in that they were able to complete most of the tasks they attempted. The travel booking sites satisfied many of the heuristics. The mapping sites - AAA and Mapquest.com - did well on visual and information design, but AAA has some issues to work on in information architecture. The Web sites by publishers of tour book presented two very different
Expedia.com	
Elderhostel.org	
Fodors.com	
Frommers.com	
Mapquest.com	

Travelocity.com

user experiences, mainly assuming different levels of Web savviness on the part of users. Expedia satisfied the most heuristics most of the time; Frommers.com was crowded, with small type and difficult to navigate.

The table below shows the number of observations that reviewers rated as 4, 3, 2, or 1 for each of the personas on travel sites.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	54			27		
Minor hindrance	3	22			25		
Total for 4 & 3			76	66.7		52	59.1
Serious problem	2	35			34		
Task failure	1	3			2		
Total for 2 & 1			38	33.3		36	40.9

Table 11. Travel sites observations

Finding 5. What the numbers mean

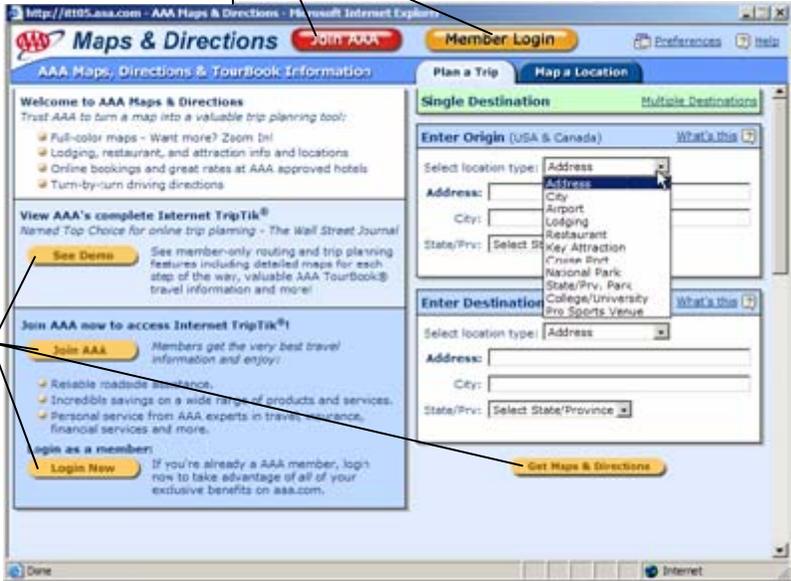
Edith’s experience with travel sites was more positive than Matthew’s by proportion. Even though she had negative observations about a third of the time, she only encountered three task failures in the four sites she visited - in the same range as the health information and prescription drug information sites she visited.

Edith found much more to like about the sites she used than Matthew did. Overall, these sites did a good job of visual and information design. They had obvious headings, clear instructions and messages that were called out in obvious, non-threatening ways. They made good use of white space and contrast.

A particular winner in these areas was the regional AAA Web site that Edith used. It had large buttons that were obviously clickable and had clear labeling. The same was true of Travelocity and Expedia (although these sites weren’t consistent from section to section). Matthew also gave high marks to Expedia for making its buttons and links obviously clickable. Fodors.com and Mapquest.com also did well in this area.

The most important buttons to AAA are large and clearly labeled

These buttons are obviously clickable because of their color, shape, labeling, and the slight shadow that gives them dimension



Both Matthew and Edith used Expedia. They both found it easy to skim - with its large headings and chunked pages - and fairly easy to read, although the size of type was not consistent and in some places was very small. Unfortunately, Expedia.com does not allow users to change the text size on pages. Travelocity does a much better job of this: the text in the main part of the screen all changes, although the selections for modifying searches in the left columns remain small.

As good as AAA was on the visual and information design, it was difficult to get around. The structure of the site was invisible to Edith; its information architecture was rather deep and narrow, and the labeling of navigation links was sometimes unhelpful. Matthew had similar complaints about Frommers.com. It was difficult to read, and it assumed that he knew more about Web terminology than he actually did know.

Financial services and planning

- Americanfunds.com
Tiaacref.org
Vanguard.com
Schwab.com
Chevy Chase Bank.com
Citibank.com
Wellsfargo.com
Smartmoney.com

Of eight financial Web sites, three were reviewed by both Edith and Matthew. Edith also used two sites that Matthew did not and Matthew used three sites that Edith did not. Matthew's experiences on the sites were far more positive than Edith's. Although the banking sites satisfied many of the heuristics, they were inconsistent from section to section within their sites, creating uneven experiences. The mutual fund Web sites did fairly well at satisfying heuristics in all of the areas - both Edith and Matthew worked especially well on Americanfunds.com. However, the more complex sites presented more opportunities for problems. While that correlation makes sense, it was surprising that sites like Citibank.com and Chevy Chase Bank.com backed by huge corporations could be so difficult to use in



some areas. Their issues pale in comparison to Smartmoney.com, however, with its unorthodox menu navigation, unclear labeling, and sheer clutter of content.

The table below shows the number of observations that reviewers rated as 4, 3, 2, or 1 for each of the personas on financial services and planning sites.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	77			83		
Minor hindrance	3	19			32		
Total for 4 & 3			96	47.1		115	71.9
Serious problem	2	56			43		
Task failure	1	52			2		
Total for 2 & 1			108	52.9		45	28.1

Table 12. Financial services and planning sites observations

Finding 6. What the numbers mean

Across the sites reviewed, Edith’s experience was fairly even between positives and negatives; Matthew’s experience was far more positive by proportion.

Several of the sites did well with labeling in the navigation and links, with Americanfunds.com doing the best in information architecture in general and Chevy Chase Bank.com following closely. The banking, brokerage, and mutual fund sites supported easy skimming and reading with good sized type, lots of contrast, and clean, uncluttered page layouts. They also did well in information design and content by providing clear explanations or other text content in plain language.

Some sites were inconsistent about following link conventions, such as Wells Fargo.com and Chevy Chase.com. However, all of the sites implemented links in recognizable ways, even if they weren’t completely conventional or consistent.



These are links, but they are red without underlines

There are many links on Wells Fargo's home page that are underlined, but they are red

[Wells Fargo](#) | [Employment](#) | [PRIVACY, Security & Legal](#) | [Report Email Fraud](#) | [Home](#)
[Line Access Agreement \(48/04\)](#) | [Important Notice on Trading in Fast Markets](#)
 © 1999 - 2004 Wells Fargo. All rights reserved. Member FDIC.

These are links, but they are blue without underlines



Inside the site, most of the links are blue and underlined

Note: Since we performed the evaluation in November 2004, Wells Fargo has changed the color scheme on its Web site so that almost all of the links are blue and most are underlined.

Some sites were inconsistent in approaches to labeling, such as Tiaacreff.org and Chey Chase bank.com. Edith thought Citibank.com was difficult to skim, but Matthew found it easy to skim - they did different tasks, however.

These issues of consistency suggest that sections may be managed by different departments. This exposes to customers the complications of how products are managed within companies versus a unified company view of supporting customers. It also can confuse users like Edith.

Most of the points at which Edith got completely stuck were on Chevychasebank.com, with 20 observations rated as “task failures”.



While the clickable area is very large in the navigation blocks, Edith expected to click on the labels, so she was surprised when the menu appeared

When trying to click an item in the menu above, Edith had trouble selecting because her mouse hovered close enough to the choices below to open that menu, obscuring the item she wanted to click

Images on screen shot are not reproduced per copyright restrictions.

Edith had problems with managing the floating, moving navigation menus on the home page of Chevychasebank.com, as well as with the unconventional and inconsistent link treatments and the sometimes low contrast between background and text. But she also encountered great difficulty on Citibank.com, Smartmoney.com (with 11 observations rated as task failures on each

of those sites), and Tiaacref.org (with 10 task failures). The labeling in all three sites was not always supportive of what Edith tried to do. Often, labels and links used jargon rather than plain language or the language was not appropriate for the level of domain expertise someone like Edith might have. Citibank.com was difficult for Edith to skim and read. Edith felt she had little control over when and how to log in and log out of Tiaacref.org. Matthew’s two task failures happened on Vanguard.com, a site that Edith did not visit. His issues there concerned several graphics that were presented to ostensibly explain and illustrate the mutual fund company’s investment philosophy. They were actually somewhat misleading and confusing, leaving Matthew with more questions.

Smartmoney.com figured the most often in serious problems and task failures for both personas, with problems of density and clutter, small type size, deep and difficult information architecture, and unconventional and difficult menu interaction. While Smartmoney.com perhaps should have been reviewed in the news category (it is the Web site for Smart Money magazine), it would have fared no better in comparison except to have its possible mate in ESPN.com, which was also difficult for both personas.

Shopping

- Amazon.com
- BarnesandNoble.com
- Buy.com
- Carmax.com
- eBay.com
- eToys.com
- Fisher-Price.com
- Landsend.com
- LLBean.com
- Travelsmith.com

Although we had originally planned to review fewer shopping sites, Edith and Matthew took their own paths - not unlike real people might - because of their different interests and priorities. Edith used five shopping Web sites, Matthew used six, all different except for both using eBay.com. There were two book/media sites; two toy sites; three clothing sites, one car site, one consumer electronics site; and eBay - a site in a class of its own.

There are a lot of good things to say about many of these sites. Their interaction designs and information architecture and navigation were generally clean, smooth, and easy for both users. Issues had mainly to do with the insider language that has developed in the eBay world and managing secondary windows that didn’t have buttons to close them.

The table below shows the number of observations that reviewers rated as 4, 3, 2, or 1 for each of the personas on shopping sites.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	30			98		
Minor hindrance	3	6			36		
Total for 4 & 3			36	55.4		134	91.2
Serious problem	2	14			13		
Task failure	1	15			0		
Total for 2 & 1			29	44.6		13	8.8

Table 13. Shopping sites observations

Finding 7. What the numbers mean

The shopping sites were much easier for Matthew to use than they were for Edith. Although the personas had a similar number of serious problems, Edith encountered many times in which she could not go further. Matthew met with no major obstacles.

All of the sites used link conventions, although Carmax.com, eBay.com, and Fisher-Price.com were inconsistent about this within their sites. All of the sites had excellent high-level information architectures, which clearly laid out their departments and sections. Although many of the sites were rather deeply layered, many did a good job of managing navigation through the shopping area and into the checkout process.

Most of the sites did very well on one very important element of the online shopping experience: clear feedback. They had clear instructions in the right place and good confirmation and error messages that helped the personas see where the problems were and how to correct them.

Matthew generally found that the size of the type and the layouts of the pages of most of the sites he visited facilitated skimming and reading, even when there was a lot of content on some of the home pages and product pages. Edith was a bit more discriminating, but found that Landsend.com and eToys.com made it easier for her pick out what was on the site than the other sites she used.

Edith got stuck in Carmax.com and Fisher-Price.com because of issues of user control. In one case, she had difficulty managing a pop-up window. In another case, the site insisted that she needed a Flash plug-in to see content (which she declined to download).

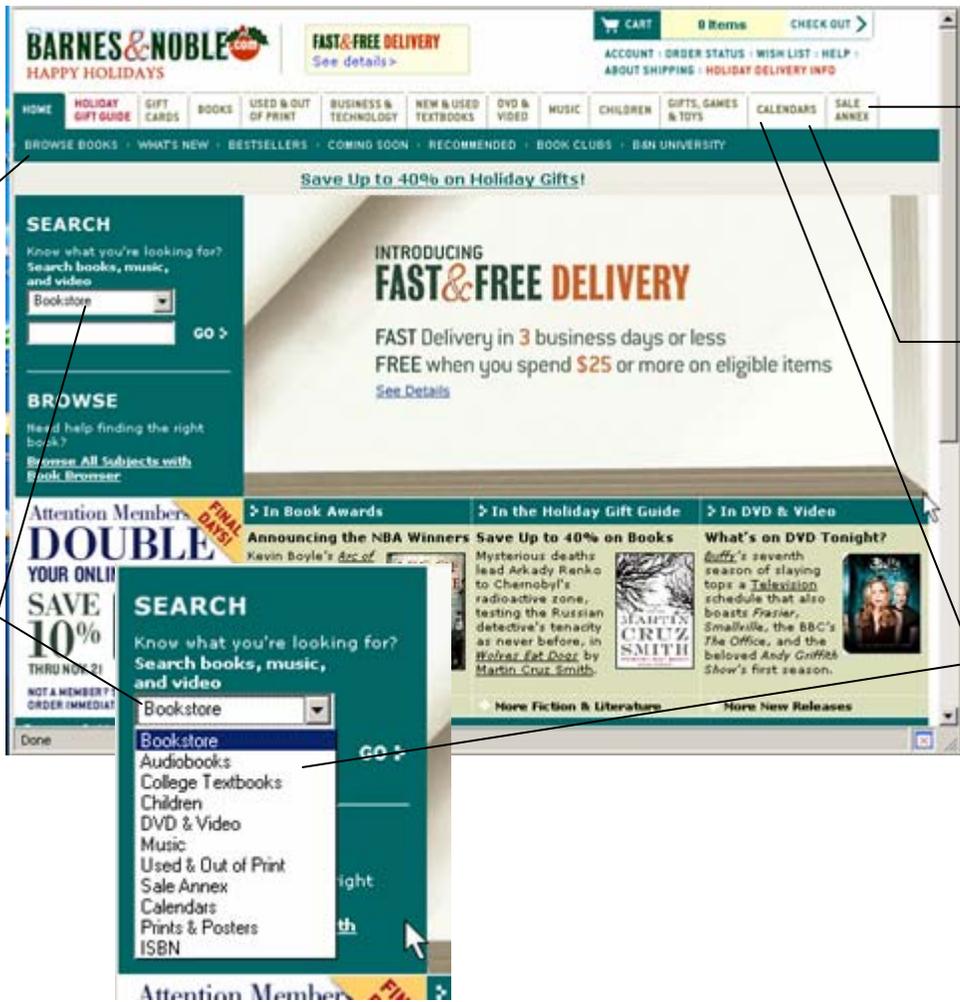
Although the secondary navigation looks like it should change with each choice in the top level, it does not

The type is large, but all caps is difficult to read

The information architecture is nice and broad

We wondered: Aren't you at a bookstore's Web site?

Some of the choices are the same in both places, some are not – identical redundancy would be better



Matthew would have found BarnesandNoble.com easier if it had been better organized and used clearer labeling in its information architecture. There was a choice in the Search called “Bookstore” along with choices that could also be found in the “Bookstore” that were more specific. This lack of parallelism in the list items in the navigation as well as the lack of consistency in the redundant navigation confused Matthew.

One of the most interesting problems was with labeling on eBay, a site for which a unique culture and community of users has developed over the years. Some of the issues had simply to do with inconsistent or apparently different redundant links. Other issues arose from eBay's use of terms specific to its domain and culture such as Anything Points and Completed Listings.

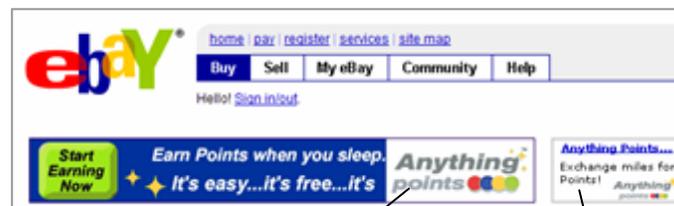
Edith and Matthew both had to spend some time getting oriented on results pages, but this was a small and momentary hindrance to Matthew; Edith did not recover from the busy inconsistency of the presentation of results. Mainly, the lack of content such as pictures on all of the items and the wide ranges of prices for the same items was rather confusing - and part of the sort of online flea market that eBay can be.

Although our personas used eBay.com as a shopping site, it could for other types of users slip into the "leisure / hobby / pastime" category and receive different ratings. People who are interested in auctioning their wares may be positively challenged by learning the intricacies of how to work the system, observing others bid on offerings, and then reaping the rewards of the transactions. They would approach the site with different tasks and attitudes that would generate a different type of experience from simply shopping on the site.

The clothing sites (LLBean.com, Landsend.com, and Travelsmith.com) as well as Amazon.com most consistently satisfied the most heuristics.

Fisher-Price.com didn't seem to take grandparents into account in their shopping audience. The colors used on the site may be attractive to kids, but reading white type on a yellow, lavender, or other light-colored background is difficult for older adults. In addition, images of products were much too small and no larger views were provided. Getting to a place online where one could actually purchase a toy was an agonizing experience for Edith. The path she followed through Toys by Age took her to what could easily have turned into a dead end: a link called Where to Buy goes to a page where users can locate retail locations. Persistent, Edith noticed a link to the Fisher-Price Online Store, where she had to start her shopping all over again.

Buy.com was manageable for Matthew, but Edith (who did not use this site) probably would have been overwhelmed at the amount of content to take in on every page and frustrated with the small type and interaction elements.



There were a few terms used by eBay – like Anything points – that signify an exclusive culture on the site

These graphics showed up in an area in which the content rotated, so to learn about this concept you need to be lucky or persistent

Portals and search engines

AOL.com
 Google.com
 Iwon.com
 Yahoo.com

We defined “portal” as “home page” or where interacting with the Web starts. We included AOL because so many of the participants in AARP’s 2003 usability studies mentioned that they used it as their primary connection to the Internet. Many participants in these studies also mentioned starting out their Web interactions at a search engine like Google.com or Yahoo.com. Several said they had set up Iwon.com as their home page because they had fun taking part in the many contests and other activities there.

Three out of the four of these portal sites satisfied most of the heuristics most of the time. AOL, with its own unconventional models for interaction, visual design, and information architecture violated at least half of the heuristics. Edith, like the oldest participants in AARP’s 2003 usability studies, was the AOL user.

The tables below show the number of times that reviewers rated observations as 4, 3, 2, or 1 for each of the personas on Web portals.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	21			34		
Minor hindrance	3	11			14		
Total for 4 & 3			32	50.8		48	73.8
Serious problem	2	19			15		
Task failure	1	12			2		
Total for 2 & 1			31	49.2		17	26.2

Table 14. Portals and search engines observations

Finding 8. What the numbers mean

Edith’s experiences with the portal/search sites were somewhat less positive than Matthew’s. This is largely due to her issues with the AOL browser interface, which has a number of unconventional elements to it. For one, there’s no way to go Back. Rather than using conventional page layout techniques, AOL generates multiple objects and pop-up windows that are difficult to manage. For example, the navigation block can be moved around anywhere within the workspace. It is very cluttered and crowded, especially at low resolution, which also leads to difficulty skimming, reading, and finding items on the page. If Edith had used Yahoo.com as a home page, she may have faired better, and she would have learned more about conventional design elements that might make it easier for her to use many other Web sites.

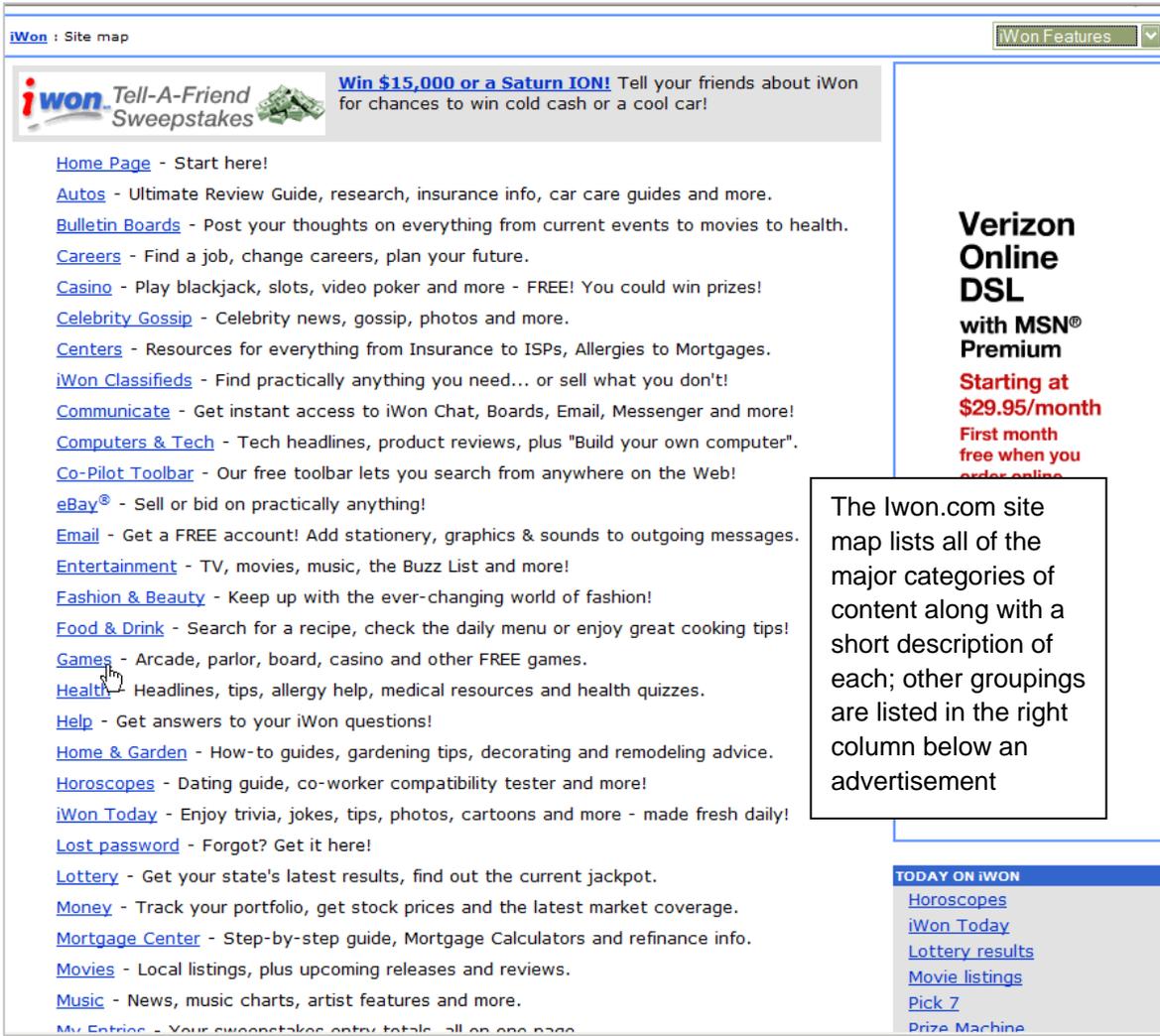
As a portal and search engine, Yahoo.com got high marks from Matthew. It is easy to find items within pages, in part because they are easy to skim, implementing obvious headings, lots of white space and clear grouping of the directory-style links.



Images on screen shot are not reproduced per copyright restrictions.

Both personas used Google.com with great success. Edith's one complaint was that she couldn't always tell what site she would end up on by clicking something - not necessarily something that Google can do anything about.

Another site we included, Iwon.com, is geared to people who are attracted to sweepstakes and winning prizes - a sort of Publisher's Clearinghouse for the computer. Edith had a difficult time figuring out what the site was all about because they don't talk about getting "points"; they call it "entries." Other unfortunate and unhelpful instructions included, "All Links = 5 entries" and "No entries for clicks on results." Iwon does have an excellent site map; Edith found it because she was really trying to understand how the site works and couldn't find a link on the home page that explained the site. She spent most of her time there figuring out what "entries" was all about.



News

- BBCnews.com
- CNN.com
- ESPN.com
- NYTimes.com
- Weather.com

Edith and Matthew did tasks on five news sites; each persona visited at least three. Interestingly, both personas met with hindrances about *two thirds* of the time. While both personas met with small or momentary obstacles that might prevent some real users from progressing, Edith stopped completely or gave up much more often than Matthew.

BBCnews.com and NYTimes.com were the most supportive of the needs of our older adult personas; ESPN.com was the least.

The table below shows the number of observations that reviewers rated as 4, 3, 2, or 1 for each of the personas on news sites.

		Edith			Matthew		
Ratings		observations	total	percent	observations	total	percent
No problem	4	8			12		
Minor hindrance	3	12			2		
Total for 4 & 3			20	33.9		14	38.9
Serious problem	2	23			19		
Task failure	1	16			3		
Total for 2 & 1			39	66.1		22	61.1

Table 15. News sites observations

Finding 9. What the numbers mean

Reviewers rated a similar number of observations as serious problems for both personas, but Edith met with many more points at which she could not go on than Matthew did.

The personas had their best experiences on the NYTimes.com and BBCnews.com sites. Both sites did very well at making it obvious what was clickable. BBCnews was the most readable of the sites, with large type, plenty of white space, and clear grouping of related topics. But Matthew also found NYTimes.com to be easy to skim in part because its structure was clear not only through its extensive navigation list, but through cross-reference links throughout the site.

Generally, news sites respected the conventions for link treatments, but ESPN had a mixture of link treatments which could be extremely confusing to older adults.



Although the New York Times site does not satisfy some of the heuristics, it does satisfy the heuristic for using conventional link treatments



ESPN.com uses at least four link treatments, which can be confusing to some older users

Images on screen shot are not reproduced per copyright restrictions.

Most difficulties with news Web sites have to do with visual design. While news pages were often very well organized and the related content was grouped together well, these pages were very dense with information. We don't consider this a failure, per se. It is a fact of media Web sites, and the density is similar to a printed newspaper. Because of some of the unique issues of delivering text online, however, some older adults may experience difficulties in grasping of the scope of the site and finding wanted links.

Mainly because the pages are so densely packed with information, they are difficult to read. None of the news sites we reviewed had obvious ways to increase the type size. Even if the user knew how to do it through the browser, most of the sites were not set up to increase the size of text and buttons together. This only worked on Yahoo!. On NYTimes.com, the text in the main column changed size, but the buttons and the navigation did not.

By the way, there is no way to change the size of text in AOL (see also the findings on portal Web sites on page 37). When viewing sites through it, the only way to increase the size of text on Web sites is to hold the CTRL key while turning the scroll wheel. We're assuming that few people know this trick - and it only works if the site is set up to allow changes in type size.

On most of the sites the type size starts out too small if the user has the screen resolution set fairly high - probably not 12-point by default - and the New York Times used serif type as the default, which can be difficult to read because it breaks up. People who view Web sites at low resolution will have particular difficulty with serif types breaking.

ESPN.com offered a version of its site called "lite site" that showed less content on every page and all of the content was presented larger. Unfortunately, the link for "lite site" is in 6-point gray type in the footer of the pages.



ESPN.com's footer at actual size at a resolution of 1024x768 and a useful link that is nearly impossible to find

Added to that, sites such as ESPN implement a color palette range that is mainly dark colors such as gray or blue on darker backgrounds. There is little contrast between links, text, and other interactions and the backgrounds they are on.

Most of the task failures in the news sites came from problems of interaction design. One factor that makes these sites particularly difficult to use is the large amount of advertising on most of them. (BBCnews.com has no advertising.) Often the advertising is animated as well, which can be disturbing for many older adults both in losing control of the page content and in distracting them from what they were on their way to doing. People who have low Web expertise and who experience age-related memory loss and loss of concentration may be particularly distressed by some of the large, animated sequences we encountered during the review. These large animations are very likely to prevent older adults like Edith from even entering a site. If people are able to bypass the first presentation of animated text, they may be so distracted that they will be unable to complete tasks on the site.

Related to user control are issues of helping users know what to expect and telling them what has happened. On NYTimes.com, Edith clicked a link to presidential poll results, but ended up in a PDF, which surprised her and took her away from the task she was trying to perform. A similar problem came up on the CNN Web site when Edith tried to view a video from the site, but had to select a media player. The pop-up for the selection dialog was not sized properly for use at a resolution of 800 x 600 - the OK button was hidden and she had no way to move the dialog.

The ESPN home page is very dense and cluttered



Grouping of related topics isn't obvious

The type size is too small and does not adjust

Images on screen shot are not reproduced per copyright restrictions.

There are too many items on the page that are animated and that the user cannot control

Hobbies and interests

- Audubon.org
- Birdwatching.com
- Cooking.com
- Genealogy.com
- Genealogy.org
- Pogo.com

Edith and Matthew used completely different Web sites in this case. Their interests and how they spend their leisure time are very different. This is due to their personalities rather than age, aptitude, or ability. (It is rather interesting, actually, that Edith's pastimes keep her online, but Matthew's pastimes don't.) All together, they visited six Web sites in this category, with Edith going to four and Matthew going to two.

The table below shows the number of observations that reviewers rated as 4, 3, 2, or 1 for each of the personas on hobby and interests sites.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	17			32		
Minor hindrance	3	13			20		
Total for 4 & 3			30	40.5		52	94.5
Serious problem	2	36			3		
Task failure	1	8			0		
Total for 2 & 1			44	59.5		3	5.5

Table 16. Hobbies and interests sites observations

Finding 10. What the numbers mean

As with other types of sites, Edith encountered more problems than Matthew did. The numbers may be skewed since Edith visited four sites and Matthew only visited two in this group, but we don't think that affects the ratio of positive to negative observations.

The range of sites here is interesting.

Although we did not know it when we selected the sites for this study, both of the genealogy sites are owned by the same "network" - a corporation that makes family tree software called MyFamily, Inc. However, the two sites are different in look and feel.

Audubon.org is owned and maintained by the Audubon Society, a large services organization. Birdwatching.com, on the other hand, seems to be owned by Michael and Diane Porter, professional birders who run tours. Audubon.org is highly polished, while Birdwatching.com seems to be a small business labor of love.

We thought at first that the sites with lower production values - less visual polish - might be harder to use. We were wrong. Birdwatching.com did as well or better than Audubon.org.

Both Genealogy.com and Audubon.org were easy to skim and to read due to their clean layouts, large type, and noticeable headings. They also did a good job of making things that were clickable obvious. It seems clear that these sites aren't counting on their audiences being super Web experts.

Edith did have some problems, however. She was surprised by having to deal with animated sequences on Genealogy.com that she couldn't control well, a disappointing site map, and language that she found difficult for someone like her, new-ish to the Web or just beginning to work on their family histories. She had also had difficulties on Pogo.com (a games Web site) and Cooking.com (where she was hoping to find recipes) with simply getting around the sites to find what she was looking for.

Matthew found few negatives on Audubon.org and Birdwatching.com - both seemed to feed his interests very well - although he grumbled about the inconsistent way in which Audubon.org partners with other sites. For example, he followed a link called Travel, which led to tours conducted by the Society. That was logical to him. There, he saw a link for Travel Clothing, which took him to a completely different site that seemed very badly set up for serving the Audubon audience. While it is not one of the sites we reviewed for this project, the partner site was very difficult to navigate, it was not clear how to identify all of the Audubon-label clothing, and pictures of the merchandise were very small, and the labeling was tiny.

What have we learned?

It is safe to say that most Web sites seeking to serve a wide audience encompassing many ages, abilities, and aptitudes fail to support older adults like Edith most of the time. For people like Matthew the sites also fall short. Matthew was more successful than Edith, but he also often met with difficulties on several types of Web sites.

In this evaluation of sites, we have counted multiple task failures in types of sites. In several cases, there were multiple task failures on a single Web site. Our personas were persistent because they were being paid. But real users probably encounter one task failure and abandon a site. They may go to a different site or to the phone. Or they may just give up.

It is difficult to imagine the pain that many older adults must tolerate to use the Web considering that Edith encountered 113 task failures and 243 serious problems in the 32 Web sites she used. That's 3.5 times on *every* site on average that Edith got stuck, met a dead end, or gave up in confusion or frustration. And 7.6 times on every site - perhaps separately from the task failures, perhaps contributing to the task failures - that Edith had serious problems.

Key findings

The Web is very difficult for the least expert among older adults. Although Edith had a high proportion of positive comments about the travel sites she used - nearly 67% of her observations about travel sites were positive - there were still three points at which she could not complete a task on these sites.

Other sites that older people often use, such as search engines and portals, shopping, news, and financial services, all rated low for Edith in terms of positive versus negative observations. It was on those sites that she had the most problems. Financial services and planning sites were the least supportive of Edith's ability and aptitude. If her experience reflects the type of issues that others like her have, it's a wonder anyone over age 70 does their banking online.

And it is still frustrating for users with more expertise. Matthew came upon 20 points of task failure and 171 serious problems over 30 sites he used. There were too many places where Matthew got stuck or gave up - almost one for each site he used (.7). And, although Matthew had more experience with computers and the Web than Edith, his lower tolerance for difficulty showed in that he met with 5 or 6 serious problems on average at each site he used (5.7).

	Edith		Matthew	
	Total task failures	Total serious problems	Total task failures	Total serious problems
	113	243	20	171
Average per site	3.5	7.6	0.7	5.7

Table 17. Task failures and problems totaled and then averaged per site visited

Although sites are doing some good things, there are still too many opportunities for serious problems and task failures. On 22 of the sites we reviewed, the personas encountered no task failures. That means that on 28 sites (56%) of the sites one or both of the personas were unable to complete what they wanted to do.

	Sites <i>without</i> task failure	Number of sites in this category	
health information	3	5	
health ins and prescription drugs	2	6	
travel	4	7	
news	0	5	
financial services and planning	3	8	
shopping	6	10	
portals	1	4	
hobby & interests	3	5	
Total	22	50	44%

Table 18. Number of sites without task failures

Only three sites - all of them shopping sites - had no task failures *and* no serious problems: Amazon.com, eToys.com, and LLBean.com. Matthew used Amazon.com and LLBean.com; Edith used eToys.com. It would be interesting to see if people like Edith would also have the same level of success with Amazon.com and LLBean.com.

Many of the sites scored both positively and negatively on the same heuristic. As the next section explores further, this is mainly an issue of inconsistency within sites.

Findings by heuristic

Of Edith's observations, just over 50% were positive. Of Matthew's observations, about 73% were positive. This means that many of the sites satisfied many of the heuristics for these personas. We did, however, see issues that made sites or parts of sites difficult and in some cases impossible, for one or both of our personas to work with successfully.

Ratings		Edith			Matthew		
		observations	total	percent	observations	total	percent
No problem	4	263			343		
Minor hindrance	3	122			178		
Total for 4 & 3			385	52.0		521	73.2
Serious problem	2	243			171		
Task failure	1	113			20		
Total for 2 & 1			356	48.0		191	26.8
Total observations			741			712	

Table 19. Total observations across all sites

One of the interesting results of this study was that the same site might have both positive and negative ratings on the same heuristic. In most cases, this means the site was not consistent. It may have used industry conventions for showing links that have not yet been clicked, but then not indicate which links have been clicked. It may have used industry conventions for showing links in one part and used a different approach in another part. It may have had clear writing in one part and not in another. One section of the site may have given the user a clear pathway for the task the user is trying to do while another section stymied the user with unclear navigation.

To give some specific examples:

- On travel sites, the navigation links were well grouped and highlighted at the beginning, making it easy to get started. Further into the task, however, the presentation of search results sometimes made it difficult to distinguish content from advertising. The icons on the home page were obvious, but then later, on internal pages, there were non-obvious icons with no help or labeling.
- On financial services sites, the explanations on some pages were plain and simple. In other sections of the same site, the web writers assumed more knowledge on the part of the user about the domain than our personas had.

These inconsistencies may come from the ways that many sites are produced and managed, with different departments being permitted to develop their own web pages and with no overall editing or requirements for consistency. Because of these inconsistencies, a site that is included in the successes for a heuristic below may also be included in the failures.

Another interesting result of this study is that while many sites succeeded at a given heuristic, many did not. Because our personas did tasks at 50 web sites, we saw both successes and failures for many of the heuristics. Therefore, you will see some of the same heuristics in both the success table and the failure table below.

Successes and satisfactions. The main areas in which sites satisfied the heuristics were these:

Heuristic number	Description	Number of sites
13	Pages were easy to skim and scan because they were well organized, had clear starting points, grouped or clustered content on the page, and clearly distinguished advertising from content. Any images included were relevant to the content and animations and videos supported user goals and tasks.	33
14	Elements on the page were easy to read because the type was large and sans serif. In many cases, while there was no obvious way on the page to change the type size, changing it in the browser increased the size throughout the page and the site. Headings were large and readable.	28
18	It was easy to find things on pages quickly. The amount of text was minimal, instructions and messages were easy to recognize, and headings, lists, and links were used liberally to assist skimming. Links were labeled well, as were buttons and fields.	28
1	Links were presented conventionally and consistently through the site.	25
19	Writing was focused on the audience and purpose, using active voice, short, simple and straightforward sentences, and was directed to "you". Paragraphs were short. Headings, labels, and captions were descriptive of their associated content.	22
10	Content categories were clearly labeled to match user tasks and goals. The labels were descriptive enough to help the user predict the underlying content, and they were useful and understandable on their own. The labels also match language that most older adults are familiar with.	21
9	The structure of the Web site was highly visible. Either directory-list formats were used, or in hierarchical sites, the site implemented cross-reference and redundant links appropriately.	19
2	It was obvious what was clickable and what was not.	19
7	There was clear feedback on actions the user took. Error messages and pages were descriptive and provided solutions that were understandable. Confirmation messages were useful and clear.	17
20	Sites spoke to users using their own language, minimizing jargon and technical terms. They assumed an appropriate level of domain knowledge, explained new or technical terms, and expressed concepts or complex information in plain language.	14

Table 20. Heuristics that were most frequently satisfied

Problems and failures. The main issues causing task failures were these:

Heuristic number	Description	Number of sites
14	Type that was too small and not adjustable.	13
20	Language that was difficult or unfamiliar to this audience.	13
11	Information architectures that were too deep, causing the user to become lost or disoriented.	12
1	Treatments for links that were unconventional.	9
6	Lack of user control of animated elements or navigational menus.	8
9	Inappropriate or unhelpful labels on cross-reference links.	8
13	Pages that were difficult to skim because the content was too densely packed.	8
19	Content that was inappropriate for the audience and/or purpose of the site.	8
10	Unclear labeling of links, buttons, navigation, and fields.	6
3	Targets such as buttons or links that were difficult to click because they were too small.	6

Table 21. Heuristics on which the personas most frequently encountered task failures

Appendix

Heuristics

Here we list the heuristics that we used in the heuristic review. These heuristics are extracted from our review of the recent relevant research on older adults and Web site design⁹.

Each heuristic also has a list of questions that we used to focus our observations on the Web sites.

Note that this list does not include heuristics about trust, security, credibility, privacy, and other affective issues. Although these are important issues for older adults, the research we reviewed did not touch on these issues.

Also, remember that these are not all the heuristics that you would need to follow to develop or evaluate a specific Web site. These are the elements that the research in our literature review says are particularly needed by older adults because of the special aspects of age, ability, aptitude, and attitude that come with aging.

Interaction Design: Designing the way users work with the site

- 1 Use conventional interaction elements.
 - 1.1 Does the site use standard treatments for links?
 - 1.2 Is link treatment the same from section to section within the site?
- 2 Make it obvious what is clickable and what is not.
 - 2.1 In lists of bulleted links, are the bullets clickable?
 - 2.2 Are command and action items presented as buttons?
 - 2.3 Do buttons and links show that they have been clicked?
 - 2.4 Are buttons clearly labeled?
 - 2.5 If there is an image on a button or icon, is it task-relevant?
 - 2.6 Do graphic buttons avoid symbols that will be unfamiliar to older adults who have low computer and Web expertise?
 - 2.7 Is there a visible change (other than the cursor changing) when the user “points” to something clickable with his or her mouse?
- 3 Make clickable items easy to target and hit.
 - 3.1 Are buttons large enough to easily see the image or text on them - at least 180 x 22 pixels?

⁹ Designing Web Sites for Older Adults: A Review of Recent, Relevant Research is available at www.aarp.org/olderwiserwired.

- 3.2 Is the area around buttons clickable?
 - 3.3 Is there enough space between targets to prevent hitting multiple or incorrect targets?
 - 3.4 Do buttons and links enlarge when the rest of the text size is increased?
- 4 Minimize vertical scrolling; eliminate horizontal scrolling.
- 4.1 Does the site work at the resolution that the user would typically view the site at without horizontal scrolling?
 - 4.2 Do pop-ups and secondary windows open wide and long enough to contain the content without the need for scrolling?
 - 4.3 For scrolling lists, for example, a list of all the states:
 - Are checkboxes used rather than drop-down or pull-down menus?
 - If not, are drop-down menus (a menu that drops down when requested and stays open without further action until the user closes it or chooses a menu item) used rather than pull-down menus (a menu that is pulled down and that stays available as long as the user holds it open)?¹⁰
- 5 Ensure that the Back button behaves predictably.
- 5.1 Does the Back button appear on the browser toolbar on every page?
 - 5.2 Does clicking the Back button always go back to the page that the user came from?
- 6 Let the user stay in control.
- 6.1 Is there no rolling text that goes by automatically?
 - 6.2 Does the site use static menus (a click leads to another page) rather than “walking menus” (exposing a sub-menu on hovering)?
 - 6.3 If there are walking menus, do they expand on a click (rather than a hover)?
 - 6.4 Are the sub-menus timed to stay open for 5 seconds or until they’re clicked?
- 7 Is there clear feedback on actions?
- 7.1 Are error pages descriptive and did they provide a solution to the user?
 - 7.2 Are confirmation pages clear?
- 8 Provide feedback in other modes in addition to visual.
- 8.1 Are captioning and or meaningful alternative text provided for images, video, and animation?

¹⁰ Microsoft Computer Dictionary, fourth edition. Microsoft Press. 1999.

- 8.2 Does the site support haptic (vibrating, tactile feedback) pointing devices (such as the Logitech iFeel mouse)?¹¹

Information Architecture: Organizing the content

- 9 Make the structure of the Web site as visible as possible.
 - 9.1 Does the site use a directory list format for listing topics (such as Yahoo! does or hhs.gov or firstgov.gov)?
 - 9.2 Does the site use cross-references to related topics and redundant links?
 - 9.3 Is the site hierarchy as broad and shallow as possible?
- 10 Clearly label content categories; assist recognition and retrieval rather than recall
 - 10.1 Are labels descriptive enough to make it easy to accurately predict what the content will be under each topic category?
 - 10.2 Do labels and links start with different, distinct, and relevant key words?
 - 10.3 Are labels useful and understandable each on their own?
 - 10.4 Do labels reflect language that older adults are familiar with?
- 11 Implement the shallowest possible information hierarchy.
 - 11.1 Are important, frequently needed topics and tasks closer to the surface of the Web site?
 - 11.2 Are related topics and links grouped and labeled?
 - 11.3 Do labels and category names correspond to users' tasks and goals?
 - 11.4 Do paths through the information architecture support users' tasks and goals?
 - 11.5 Is the path for any given task a reasonable length (2-5 clicks)?
 - 11.6 Is the path clear of distractors and other obstacles to reaching task goals?
 - 11.7 Are there a few, helpful cross-reference links that are related to the current task goal?
 - 11.8 Do redundant links have the same labels?
- 12 Include a site map and link to it from every page.
 - 12.1 Is there a site map?
 - 12.2 Is the site map linked from every page?
 - 12.3 Does the site map provide a quick overview of the whole site (rather than descriptions of the top level choices (see aarp.scudder.com), a rehash of the main navigation (<http://www.drugdigest.org/DD/Home/SiteMap/0,4087,,00.html>) or a list of every single topic on the site)?

¹¹ We tried out an iFeel mouse, and as far as we can tell, if the mouse drivers and software are properly installed, there is nothing special to code into Web sites to support tactile feedback.

Visual Design: Designing the pages

13 Make pages easy to skim or scan.

- 13.1 Are pages clean looking and well organized (versus cluttered or busy)?
- 13.2 Is there a clear visual “starting point” to the page?
- 13.3 If pages are dense with content, is content grouped or otherwise clustered to show what is related?
- 13.4 Is it easy to tell what is content and what is advertising?
- 13.5 Do task-supporting keywords stand out?
- 13.6 Are images relevant to, and supportive of, the text content?
- 13.7 Are there videos or animated sequences? If so, do they support specific goals or tasks?

14 Make elements on the page easy to read.

- 14.1 Is the default type size 12-point or larger?
- 14.2 If not, is there an obvious way on the page to increase the type size?
- 14.3 If not, does changing the type size in the browser enlarge all of the text?
- 14.4 Is the type size on pull-downs and drop-down menus the same size as the text content? Does it change when the user increases the type size?
- 14.5 Are headings noticeably larger than body content (18- or 24-point)?
- 14.6 Is sans serif type used for body content?
- 14.7 Are headings set in a typeface that is easy to read?
- 14.8 Are there visual cues to direct users’ attention to important items that are in the left and right columns?

15 Visually group related topics.

- 15.1 Are pages dense with information, or sparse, or in between? Is the amount appropriate for the audience and type of site?
- 15.2 Are the most important and frequently used topics, features, and functions, close to the center of the page rather than in the far left or right margins?
- 15.3 Are task-related topics grouped together?
- 15.4 Are frequently used topics, actions, and links “above the fold”?

16 Make sure text and background colors contrast.

- 16.1 Are text and interaction elements a different color from the background (not just a different hue)?
- 16.2 Do the colors that are used together make information easy to see and find?
- 16.3 Are clickable items highlighted differently from other non-clickable highlighted items?

- 16.4 Are multiple types of highlighting minimized on each page?
- 17 Use adequate white space.
 - 17.1 Are there visual cues in the layout of the page that help users know there is more content “below the fold”?
 - 17.2 Is there line space between clickable items? (at least 2 pixels)
 - 17.3 Is body text broken up with appropriate and obvious headings?

Information Design: Writing and formatting the content

- 18 Make it easy to find things on the page quickly.
 - 18.1 Is the amount of text minimized; is only necessary information present?
 - 18.2 If there are introduction paragraphs, are they necessary?
 - 18.3 Are instructions and messages easy to recognize?
 - 18.4 Is there liberal use of headings, bulleted lists, and links to assist skimming?
 - 18.5 Do bulleted lists have the main points and important keywords at the beginning of each item?
 - 18.6 Do links have meaningful labels?
 - 18.7 Are buttons labeled clearly and unambiguously?
 - 18.8 Do button and link labels start with action words?
- 19 Focus the writing on audience and purpose.
 - 19.1 Is the content written in active voice, directed to “you”?
 - 19.2 Are sentences short, simple, and straightforward?
 - 19.3 Are paragraphs short?
 - 19.4 Is humor used appropriately, if at all?
 - 19.5 Are headings, labels, and captions descriptive of associated content?
 - 19.6 Are conclusions and implications at the top of a body of text, with supporting content after? (inverted pyramid)
- 20 Use the users’ language; minimize jargon and technical terms
 - 20.1 Does the site use words that most older adults know?
 - 20.2 If there are technical words or jargon, are they appropriate for the level of domain expertise that the audience has?
 - 20.3 If there are new or technical terms, does the site help users learn what the terms mean?
 - 20.4 Are concepts and technical information (such as safety and effectiveness information about a prescription drugs) written in plain language?
 - 20.5 Are instructions written in plain language?

- 20.6 Is the reading level appropriate for the capabilities of the audience and their literacy in the topic area? Is it easy to draw inferences and to understand the implications of text?

Tasks

With the goal in mind of getting a measure of how well Web sites support older adults, we want to make our task situations as realistic as possible. So, rather than starting with dozens of predetermined tasks for each of the sites, we started with high-level scenarios that are more reflective of how people form tasks. Here are the five high-level tasks that we used:

Task	Types of sites covered
1. Start the day	Portals and search engines; news
2. Do some research about options my doctor has given me	Health information; health insurance and prescription drugs
3. Plan a trip	Travel
4. Do the monthly household bookkeeping	Financial services and planning
5. Leisure activities or hobbies, shopping	Shopping; hobbies and interests

Sites

We generated the list of sites to evaluate from sites that were mentioned during AARP's 2003 usability studies and from doing searches ourselves on topics that are important to older adults. We adjusted the list as our personas directed. Dana reviewed the sites shaded blue-green; Ginny reviewed the sites shaded yellow. The task number in each table corresponds to the list of tasks in the table above. As planned for in the project schedule and budget, Dana did more of the tasks for both personas than Ginny did.

Health Information

All sites were used as part of high-level task 2, Do some research about options my doctor has given me.

Site number	Site name	Edith	Matthew
1.	WebMD	X	
2.	Dr. Koop	X	
3.	Cancer.org (American Cancer Society)		X
4.	nhlbi.nih.gov/index.htm (National Heart, Lung, and Blood Institute)		X
5.	americanheart.org (American Heart Association)	X	

Health insurance and prescription drugs

All sites were used as part of high-level task 2, Do some research about options my doctor has given me.

Site number	Site name	Edith	Matthew
6.	Medicare.org	X	
7.	Medicare.gov for Edith; aetna.com for Matthew	X	X
8.	Medlineplus.gov, drug information	X	X
9.	Drugs.com	X	X
10.	purplepill.com		X

Travel

All sites were used as part of high-level task 3, Plan a trip.

Site number	Site name	Edith	Matthew
11.	Frommers.com		X
12.	Expedia.com	X	X
13.	AAA.com	X	
14.	Mapquest.com		X
15.	Elderhostel.org	X	
16.	Fodors.com		X
17.	Travelocity.com	X	

Financial services and planning

All sites were used as part of high-level task 4, Do the monthly bookkeeping.

Site number	Site name	Edith	Matthew
18.	Citibank	X	X
19.	Wells Fargo Bank		X
20.	Chevy Chase Bank	X	
21.	www.smartmoney.com	X	X
22.	Charles Schwab		X
23.	Americanfunds.com	X	X
24.	tiaacreff.com	X	

Site number	Site name	Edith	Matthew
25.	Vanguard.com		X

Shopping

All but two shopping Web sites were part of the high-level task 5, Leisure activities or hobbies, shopping. Matthew used LLBean.com and Travelsmith.com as part of high-level task 3, Plan a trip.

Site number	Site name	Edith	Matthew
26.	eBay	X	X
27.	Carmax	X	
28.	Amazon		X
29.	Barnes & Noble		X
30.	L L Bean		X
31.	Fisher-Price	X	
32.	eToys	X	
33.	Buy.com		X
34.	Travelsmith		X
35.	Land's End	X	

News

All sites were used as part of high-level task 1, Start the day.

Site number	Site name	Edith	Matthew
36.	CNN	X	
37.	BBC	X	
38.	New York Times online	X	X
39.	Weather.com		X
40.	ESPN	X	X

Hobbies and interests

All sites were used as part of high-level task 5, Leisure activities or hobbies, shopping.

Site number	Site name	Edith	Matthew
41.	genealogy.com	X	

42.	genealogy.org	X	
43.	audubon.org		X
44.	birdwatching.com		X
45.	Pogo.com	X	
46.	Cooking.com	X	

Portals and search engines

In task 1, Start the day, Edith used AOL, Matthew used Yahoo.com. Edith and Matthew used Google.com in tasks 2, 3, and 5. Edith used iwon.com in task 5.

Site number	Site name	Edith	Matthew
47.	AOL	X	
48.	Google	X	X
49.	Yahoo		X
50.	iwon.com	X	
		32	30
		11 sites in both columns, used by both personas	

Example evaluation worksheet

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Heuristic review																						
2	DEC 11/3/04																						
3	Task scenario: HEALTH INFO ONLY																						
4	Persona: Matthew																						
5		Site type	Site number	Interaction Design					Information Architecture					Visual Design					Information Design				
6				conventi ons H1	obviously clickable H2	ease of hitting targets H3	minimize scrolling H4	back button H5	user control H6	clear feedback H7	multi- modal feedback H8	visible structure H9	retrieval vs recall H10	broad + shallow H11	site map H12	ease of skimming H13	ease of reading H14	related topics grouped H15	contrast H16	white space H17	ease of finding H18	audience & purpose H19	appropria te language H20
7	Comments & Observations																						
39	So the blurb says that this is part of a set. But I don't want the set. I wonder if I can download or order just the one? I click on the picture of the booklet, which starts up Acrobat. I wish they had told me that was going to happen. I just took a chance that clicking the picture would lead somewhere. There weren't any instructions or anything.	health info	NIH			3				3											3	3	
40	This booklet is just for African Americans. That's good! I guess I can save this to my computer and print it out later or email it to Lynn. But does it answer my question about counts?	health info	NIH																				
41	Yes, it does, in a nice table on page 3.	health info	NIH												4		4			4	4	4	
42	I'm going back to the main site, now, using the Back button.	health info	NIH					4															
43	It shows that I've been to the PDF (the link color changed).	health info	NIH	4																			
44	I've just downloaded a copy for free, but if you want printed copies, they cost money!	health info	NIH																		4		
45	Just for the heck of it, I clicked the Order Online button. The page has instructions for ordering materials and a bar across the top with nice icons showing what you can do now.	health info	NIH								4				4								
46	I can click on those words or icons, although that's not obvious. I was sort of expecting that they would tell me where I was in the process by changing color or something, but they don't.	health info	NIH		2																		
47	Hmmm. That's odd. On one page there are two icons for checking out. But it's the same thing, right?	health info	NIH												3								
48	Unfortunately, you have to register once you get this far, and I'm not interested enough to do that. I have already found what I wanted.	health info	NIH								1												
56	<i>no problem -- satisfies the heuristic</i>	4	23	3	0	0	0	1	0	0	0	3	1	0	0	3	0	3	0	3	2	2	2
57	<i>minor hindrance -- possible issue, but probably will not hinder this persona/user</i>	?	19	1	1	1	1	0	0	1	0	3	2	1	0	3	0	0	1	0	1	3	0
58	<i>serious problem -- may hinder this persona/user</i>	?	9	0	1	0	0	0	0	0	0	2	0	2	0	1	0	0	0	0	0	1	2
59	<i>task failure -- prevents this persona/user going further</i>	?	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Ratings:

- 4 = No problem - satisfies the heuristic
- 3 = Minor hindrance - possible issue, but probably will not hinder this persona/user
- 2 = Serious problem - may hinder this persona/user
- 1 = Task failure - prevents this persona/user going further

Here's a close-up example of some observations and ratings for Matthew.

Comments & Observations	Site type	Site number	Interaction Design								Information Architecture				Visual Design					Information Design		
			conventions H1	obviously clickable H2	ease of hitting targets H3	minimize scrolling H4	back button H5	user control H6	clear feedback H7	multi-modal feedback H8	visible structure H9	retrieval vs recall H10	broad + shallow H11	site map H12	ease of skimming H13	ease of reading H14	related topics grouped H15	contrast H16	white space H17	ease of finding H18	audience & purpose H19	appropriate language H20
So the blurb says that this is part of a set. But I don't want the set. I wonder if I can download or order just the one? I click on the picture of the booklet, which starts up Acrobat. I wish they had told me that was going to happen. I just took a chance that clicking the picture would lead somewhere. There weren't any instructions or anything.	health info	NIH			3														3	3		
This booklet is just for African Americans. That's good! I guess I can save this to my computer and print it out later or email it to Lynn. But does it answer my question about counts?	health info	NIH																	4	4		
Yes, it does, in a nice table on page 3.	health info	NIH												4		4		4	4	4		
I'm going back to the main site, now, using the Back button.	health info	NIH					4															
It shows that I've been to the PDF (the link color changed).	health info	NIH	4																			
I've just downloaded a copy for free, but if you want printed copies, they cost money!	health info	NIH																	4			
Just for the heck of it, I clicked the Order Online button. The page has instructions for ordering materials and a bar across the top with nice icons showing what you can do now.	health info	NIH									4			4								
I can click on those words or icons, although that's not obvious. I was sort of expecting that they would tell me where I was in the process by changing color or something, but they don't.	health info	NIH												4								
Hmmm. That's odd. On one page there are two icons for checking out. But it's the same thing, right?	health info	NIH			2									3								
Unfortunately, you have to register once you get this far, and I'm not interested enough to do that. I have already found what I wanted.	health info	NIH										1										

How to analyze the data

Because we evaluated so many sites, we had to find a relatively quick and straightforward way to pick out which supported older adults well and which did not. When using the basic methodology to evaluate any one site, you may not need to perform the steps below.

1. Skim through the totals on the main task spreadsheets to see where the high numbers of 1, 2, 3, 4s are. See the sample spreadsheet on the previous page.
2. In a Summary Data spreadsheet, list the heuristics that got the top numbers for one persona.
3. In the main task sheet, filter the data to show only the rows with that rating for that heuristic

	A	B	C	D	E	F	G	
1	Shopping							
2								
3	SUCCESSSES	(4/3)						
4	Heuristic	Edith incidents	Matthew incidents	Edith sites	Matthew Sites	4s	3s	
5	visible structure	6/1	12/2	carmax 1; landsend 1; ebay 1; fisher 2; etoys 1 / fisher	travelsmith 2; bean 3; ebay; buy 2; amazon 4 / travelsmith 2	18	3	
6	link conventions	9/0	6/1	carmax 3; landsend 2; ebay 2; etoys 2	bean; ebay 2; buy; amazon; b&n / travelsmith	15	1	
7	clear feedback	7/0	8/1	carmax 3; landsend 2; ebay 1; fisher 1	ebay; buy; amazon 5; b&n / b&n	15	1	
8	ease of skimming	1/1	12/8	landsend; etoys	bean; ebay 3; buy; amazon 5; b&n 2 / travelsmith; bean 2; ebay 3; buy; b&n	13	9	
9	ease of reading	0/0	8/5		travelsmith 2; bean; ebay; buy 2; amazon 2 / travelsmith; buy; amazon 2; b&n	8	5	
10	audience & purpose	0/3	8/1	carmax; fisher 2	travelsmith; bean; ebay 2; amazon 4 / b&n	8	4	
11	ease of finding	0/0	8/3		travelsmith 2; ebay 2; amazon 4 / amazon 2; b&n 1	8	3	
12	related topics groupe	0/0	8/1		ebay; buy; amazon 4; b&n 2 / ebay	8	1	
13	ease of hitting	2/0	4/2	carmax; fisher	travelsmith; bean; amazon 2 / buy; b&n	6	2	
14	retrieval v recall	0/0	4/3		bean; ebay; amazon 2 / travelsmith; buy; b&n	4	3	
15	obviously clickable	0/0	4/2		travelsmith; bean; amazon 2 / buy; b&n	4	2	
16	white space	0/0	3/2		bean; buy; amazon / buy; b&n	3	2	
17								
18								
19								
20	FAILURES	(2/1)						
21	Heuristic	Edith incidents	Matthew incidents	Edith sites	Matthew Sites	1s	2s	
22	audience & purpose	3/5	1/0	carmax 2; ebay / ebay 4; fisher	b&n	5	4	
23	ease of hitting	0/3	3/0	/ landsend 2; ebay	travelsmith 2; buy	3	3	
24	broad + shallow	0/3	1/0	/ carmax; fisher 2	b&n	3	1	
25	link conventions	2/4	1/0	landsend	buy	1	1	

Part of a summary spreadsheet

so you can see what the site and the observation were.

1	Heuristic review		
91	The left side of this page (the map page) has a list of other links but I don't know why the first part is Investor Information and the second one is Job Openings. I would think the Questions? section would be first. It is nice that the links are still underlined blue even on the side. Oh, I see, when I went to the locations, it took me to the Company Info section. I did not realize that. I'm not sure that makes sense.	carmax	4
01	They don't actually say you don't have to buy before it is transferred but they do say that if we buy it part of the transfer fee applies to the cost so I guess you don't have to buy it, but you pay for getting it to your location if you don't buy it. The information is pretty clear. It's in numbered steps with big numbers, imperative verbs, clear language. There's more about transferring, so I'll look at that. The link is underlined blue, it's good that it's so consistent.	carmax	4
12	The links are underlined blue and clear.	landsend	4
18	The language is clear. They say "we have 17 offerings. View them all." And the links are underlined blue.	landsend	4
49	I chose Musical Instruments that time and there's an option on that page for Equipment and Metronomes. The pages in the categories have lists by topic. It's all out there on the page. That's good and the links are in underlined blue.	ebay	4
50	If I go to a choice (Metronomes) and then go back to the Musical Instruments page, the choice I've been to is now in red. That's clear.	ebay	4
09	The links in the left are underlined blue.	etogs	4
27	I went to Help at the top to see if it said anything there about returning items. It has a good help page. A few have good pictures with them. There are sets of questions. All are underlined blue links. One of the most common questions is How do I return an item?	etogs	4

Example of a data spreadsheet filtered on heuristic 1 for observations rated 4

- List in the Summary Data sheet which sites the persona had the difficulty on along with the number of times that site got that rating.
- Keep going like that for all the main successes and failures. Don't look at one-off issues unless they are the only ones (that is, there are only 3s but no 4s, for example).
- Do the same thing for the other persona.
- Match up the incidents by heuristic if possible. That is, when one persona has multiple incidents of the same heuristic, go back and look at what happened on the same heuristic for the other persona and note that in the Summary Data spreadsheet, too.
- Add the 4s and 3s together, the 2s and the 1s together.
Re-order the lists of successes and failures to see where sites did well and did badly.

Afterword

Evaluating 50 Web sites was a huge task that we sometimes feared we would not be able to complete within the schedule in a way that would produce results that AARP and the rest of the user research and Web design community would find useful. Our special methodology was the key to reaching those goals.

While we had each used parts of the methodology in different contexts before this project, when the parts came together in this particular package, we were pleasantly surprised with how well they worked together. Creating methodologies to suit particular projects isn't that unusual - practitioners do it all the time to meet demands of budget, schedule, and research objectives. Our original goal was to do just that. We first talked about the project in July 2004 and we knew it had to be completed by the end of December.

At first, it seemed radical to "channel" personas while "they" performed tasks at the same time that we as usability practitioners kept the heuristics in mind. We had several concerns:

- Had anyone else done anything like this? Would the professional communities we were interested in serving accept the methodology as legitimate and useful?
- Would the reviews be thorough enough on each of the sites to give us useful, defensible data?
- Would we be able to get into and then "stay in character" without projecting our usability practitioner selves into the reviews?
- What if the personas did not remark on all of the heuristics on a given site? What would that absence of data mean?
- Would the personas actually find different issues, and if so, what would that mean?

Published, working precedents

Using the personas was not that much of an extension of the work that both of us have done in the past. Ginny has always used user profiles and task scenarios in doing heuristic evaluations. Tec-Ed (where Dana used to work) has also performed heuristic evaluations based on task scenarios for many years.

After we started reviewing sites, we also saw Kathy Gill's article in the November - December 2004 issue of *Interactions* magazine in which she relates how she reviewed sites from "two perspectives," which she characterizes as two types of Web site users. We found this example to be comforting. It showed that we weren't alone in thinking that there were ways to approach evaluating user interfaces that might be effective expansions of and possible improvements on the original heuristic evaluation methodology.

Depth and breadth of review

We were surprised at how doing even one short (and sometimes unsuccessful) task on a Web site gave us good sense of the site and how well it supported (or did not support) the persona and task in relation to the set of heuristics that we were using. The number of observations we made during any given task on any given Web site ranged from 5 to 50.

The number of observations varied naturally. We purposely did not use the heuristics as a checklist but, instead, had the personas do tasks that were realistic to them. Later, we reviewed the persona's observations against the heuristics.

The persona's observations were much like the think aloud commentary during an exploratory (diagnostic) usability test. Some people are more observant than others; some are more aware of Web

design issues than others; and we also know that there is great variation in performance using any user interface across a group of users.

In the end, the observations and ratings that came directly from working through the personas generated about 90 percent of the data used in our analysis. The other 10 percent came from the practitioner side of us as we reviewed the specifics of issues that the personas had.

Channeling personas

The personas we used in this study became close friends to us, but as close friends can be, we are very different from each of them, as we are different from each other.

It surprised us both how easy it was, based on the sort of character sketches that we started out with, to get into the fictional heads of personas. Quickly, we empathized with their abilities (and disabilities), technical expertise (and lack of it), life experiences, preferences, attitudes, even memories. In the course of deciding on specific tasks and making observations, we expanded on the original sketches quite extensively. We started with Matthew and Edith's first names and a few demographic characteristics and ended up with characters with relationships, habits, and emotions - real personalities.

It *was* sometimes difficult to stay in character in the same way that a study moderator might momentarily drop out of a neutral tone or impartial posture. We were careful, though, to separate what we as practitioners wanted to cover from what we felt was realistic for the personas to do.

Coverage of the heuristics

Letting go this way meant that we wouldn't hit all of the heuristics on all of the tasks on all of the Web sites. While we struggled with this issue, we believe that there's a parallel in our approach to conducting usability sessions in which there are no directed tasks. It produced the kind of data we hoped and planned for, although we were not sure at the beginning of the evaluation that it would.

Different "people," different issues

Our personas did find many of the same types of things to be difficult or problematic on Web sites. But they also found different things. In performing usability studies, the accepted wisdom (supported by research), is that 6 to 8 participants who have similar characteristics can reveal a large majority of usability problems for that user group with any given user interface. We practitioners expect that different types of users will have different types of problems and successes. And that was true in this case, too.

	Persona/Evaluator	
	Edith	Matthew
start the day	Dana	Ginny
research about medical / health options	Dana	Dana
plan a trip	Dana	Dana
do monthly bookkeeping	Ginny	Dana
leisure activities or hobbies, shopping	Ginny	Dana

Table 22. Who did what. One evaluator performed three of the high-level tasks (18 sites), the other did seven (46 sites).

Interpreting the data

Measurements in this method are different from other methods, and it took practice and discipline to keep our heads around what we were really counting. It would be very easy to slip into counting the observations as points where there were issues or usability incidents, but that isn't what they represent. The only thing they represent is a point at which the persona noticed something or did something. The rest is classifying that observation within the set of heuristics we used. Some of the observations did not fit within our heuristics, but almost all did. Still, we had to discuss and agree in several cases what the observation meant and which heuristic applied.

A challenging method

It is difficult to know whether our new method for heuristic evaluation takes more time than the classic checklist method, for two reasons. First, we created this method because we had so much to cover - 50 sites with two personas in just a few weeks. Second, we found that we got tired. Channeling was very much like *being* the personas, and it was exhausting, especially to be Edith. Informally, we limited the time we worked through a task to two and a half or three hours at a shot. Think of participants in a usability test. After about 90 minutes, it is difficult to get productive data. Older adults can tire out even sooner than younger people.

It was also challenging because we were actually working as the participant and as the usability specialist. The cognitive load was pretty heavy, so we had to take breaks after gathering the data. Analysis usually happened a day or two later.

We look forward to using this method in evaluations with more personas but fewer (or only one) sites - a situation that we know most practitioners are more likely to have.

About the Authors

Dana Chisnell was the co-principal investigator and lead researcher on the project. She conducted an extensive review of the literature about older users and the Web. She developed the heuristics and conducted many of the persona- and scenario-based reviews. Earlier, during 2003 and 2004, she had

conducted exploratory studies with 70 older adults in three U.S. cities in addition to several heuristic reviews and usability tests on Web sites covering consumer information about prescription drugs, general health information, mutual funds and retirement accounts, and Medicare supplemental insurance - all studies specifically with older adults in mind or as participants.

As an independent user researcher and usability consultant operating UsabilityWorks in San Francisco, California, Dana has been doing usability, user interface design, and technical communications consulting and development since 1982. Her focus is on helping companies build products that real users can use. Dana is a senior member of the usability community who has presented at various international conferences on usability and human-computer interaction.

Dr. Janice (Ginny) Redish was co-principal investigator on the project that this report covers, helping to develop the new model and the new methodology, taking the personas through some of the sites, and contributing through several rounds of writing and revising the report. Ginny is one of the most senior people in the usability and user experience community. As President of Redish & Associates, Inc., in Bethesda, Maryland, Ginny helps colleagues and clients solve problems in usability, web design, and clear communication. She has been working with AARP since 2001, consulting on many different issues and projects related to usability of web sites.

Ginny is co-author of two of the major books on usability techniques, *A Practical Guide to Usability Testing*, with Joseph Dumas, and *User and Task Analysis for Interface Design*, with JoAnn Hackos. She is an internationally respected consultant in all aspects of user centered design as well as clear communication and information design. Her more than 25 years of experience includes work with many government agencies, organizations, and commercial companies. She is a sought-after seminar speaker and workshop leader and has presented at conferences in many countries.

Other AARP Publications Related To This Report

For an up-to-date list of publications related to older adults and the Internet, please visit AARP's Older Wiser Wired site at <http://www.aarp.org/olderwiserwired> or AARP's extensive offerings at <http://www.aarp.org/research>.