USER-CENTERED DESIGN STORIES

Real-World UCD Case Studies

CAROL RIGHI JANICE JAMES

MK
MORGAN KAUFMANN

CASE 13

Isis Mobile: A Case Study in Heuristic Evaluation

Rosemary Pluchino, Texas Instruments

"Well, I'd like to be sure that people will not only like how our PDAs look but feel confident in using our products enough to consider us for a future purchase," said James. "I'm confident in the marketing strategy, but I still feel there may be something we're overlooking in terms of what the users need."

Isis Mobile's History and Direction

Isis Mobile is a successful start-up company involved in the design and production of cordless phones and venturing into the development of handheld cellular devices. James Cartwright, founder and president of the company, saw an opportunity to incorporate cell phones with more elaborate scheduling-type features found in digital organizers. The idea came to him with the surge in cell-phone popularity. If most people were carrying cell phones and depending on them throughout the day and also carrying digital organizers, why not merge the two? This product would be the first of its kind to be developed. The company could even have the first one that hit the market if they played their cards right.

James turned to his development team, which comprised industrial designers, marketing specialists, and mechanical, electrical, audio, and software engineers. He challenged them to create a product that was more than a cell phone and more than an organizer, more like a pocket-sized personal computer capable of placing and receiving calls. The team worked diligently to meet this challenge.

The marketing team was first in line, defining who the target users would be and what features would be vital to the product's success. They began with the assumption that the user base would be likely to consist of "high-end power users." The type of person who would likely benefit from a combined cell phone–digital organizer is someone in a high-profile position, who maintains a demanding schedule, who potentially travels quite a bit, and must be accessible at all times. Doctors, lawyers, and even real-estate agents would fall into their definition of a high-end power user.

Marketing then tested their assumption by hosting a set of focus groups that included a broad range of occupations to see whether the target market they defined would indeed be interested in their proposed product. Of course, people not belonging to the anticipated target market were also included for comparison. Ultimately, the results indicated that the prediction the marketing team made in terms of the target user base was in fact correct.

A few marketing representatives then recruited some participants who fit this user profile for another type of research effort. The marketing reps conducted an ethnographic study, which consisted of observing these people for several days in their work and home environments. In this way marketing learned about the participants' needs in a more contextual and concrete way. The resulting observations were used to make informed decisions about features that would be most appealing to their users and could provide the most benefit to their hectic lives.

The next step was to make a business case for this type of device. Marketing worked with engineering and finance to come up with their best estimate of how much it would cost to make the product. Once this figure was attained, they began researching how much their target users would be willing to pay for this product.

Once again a set of focus groups was run. Marketing discovered that not only were people eager to get their hands on such a device, they were willing to pay more than four times the minimum sales price they needed for a healthy profit. This product would mean absolute success for their company!

Their first concept sketches as well as the flow diagrams of their proposed user interface (UI) were near completion after 4 months. The product team was confident that their designs were well thought out and presented their proposed product line to upper management.

James Cartwright wasn't sold yet on the new concept. He had heard quite a bit on the marketing strategy and positioning of their new line of personal digital assistants (PDAs), but how could they be sure that once

people purchased the product they would be satisfied and even consider Isis for future product purchases? With increasing competition and high-profile product reviews available, James knew they ran the risk of flopping with their initial product release—a mistake that would take a huge recovery effort for their brand name.

James met with the product team directors to gain an understanding of what was driving their designs. James found extensive market analysis data but little else to really back how successful these products would be. Sure, the team had nailed down the target users, the features, and the price bracket—but nothing showed whether the design of the UI would be appealing to users, or even usable for that matter. James knew that even if the market accepted the look and cutting-edge features of their product, the new device could be headed for failure if it was too complicated to use.

Determining the Need for a Heuristic Evaluation

James turned to Michael Redding, a highly esteemed human factors engineer for a consulting firm he had read about in a number of online product reviews. He is regarded as one of the industry's best, with a particular focus on small-screen-device usability.

"Michael, how can we be confident that our proposed product strategy will not fail?" James asked.

"There are a number of ways to do that. What exactly are you trying to accomplish?" Michael replied.

"Well, I'd like to be sure that people will not only like how our PDAs look but feel confident in using our products enough to consider us for a future purchase. I am confident in the marketing strategy, but still feel there may be something we're overlooking," James clarified.

"Has there been any work done in terms of usability?" asked Michael.

"Not yet. With our budget limitations we were hoping to be able to bring someone in when our first production model was complete," James confessed.

Michael smiled, "James I'm glad you called me in to discuss this. If you wait until production units are ready, there is little likelihood that anything can be done at that point. Software will have been coded and tools will have been made—it will be very expensive to make any changes and we will be limited as to what can or cannot be improved within your release schedule."

"What do you mean exactly?" James asked.

"Think about it James," Michael continued. "If the first usability study is done with a nearly completed product, you run a high risk of users finding usability issues that will require major design changes. Any benefit that can be realized through their input may not be able to be resolved and implemented by your deadlines or within your budget."

Michael continued. "Look, what I'm about to propose will catch a number of usability issues now, so we can address those before bringing users in. Allocating a portion of your budget for an evaluation at this early stage will ensure we catch these issues before tools are made and before software is coded. This will minimize any rework that needs to be done and save cost and time too. We can still bring users in to validate the changes as we move forward."

"You mean we'd still need budget to run a study later as well? I'm not sure we're quite prepared for that," James looked sullen for a moment. "All the same, I am curious, what do you have in mind?"

"A heuristic evaluation. It's basically a sanity check where usability experts evaluate the product against several well-known design principles. It is a method that was first described by Jakob Nielsen and Robert Mack in 1994. Since then, others have contributed to elaborate on these guidelines, defining them more thoroughly and more specifically to their particular field of use. Bruce Tognazzini, for example, developed a set of heuristics geared toward interaction design of traditional global user interfaces. Think of it like preventative maintenance; typically the most obvious usability problems can be filtered out," Michael explained.

"Well, what exactly are the pros and cons to this type of evaluation?" James inquired.

"The pros of a heuristic evaluation include its cost-effectiveness and its ability to be implemented very early on in the design stages. It can be performed quickly and requires only a small number of evaluators, typically about three to five. It can be performed remotely, so you won't need to provide a lab or incur any cost for recruiting participants. Also, a heuristic evaluation can save you money by identifying and resolving major usability issues very early on in the design process."

"This sounds great so far, but what are the cons?" asked James.

Michael replied, "As for the cons, a heuristic evaluation does not leverage feedback from actual users and is subject to evaluator bias. It is also extremely dependent on the skill level of the people doing the evaluating. The sets of problems found may not have much overlap with those found by user studies,

and the same problems may not be rated at the same severity. Heuristic evaluations can produce false positives and, if not well managed, can take as many resources as a usability study. It also does not provide any quantitative data," Michael explained.

"Honestly Michael, the fact that problems found may not line up with those of our target users worries me," James confessed. "However, we are not in a position to take any chances with this release—we need this product to succeed. I am familiar with the work you've done and understand you are one of the leading experts in small-screen-device usability," James seemed torn for a moment as he thought.

"Well, I can see how getting problems out of the way early will help us later on," James agreed. "And we do want to have a usable product. Also, if spending more of our budget now will help us later on, then that makes sense. And you did say we could validate with users later, correct?"

"Yes James, absolutely, a user study can be performed to validate the product. That will provide the quantitative data that is missing from what we'll be doing up front," Michael assured him.

"Okay Michael, you've convinced me that an upfront heuristic evaluation is our best option. When can you start?" James said as he shook Michael's hand.

"Excellent!" said Michael, "I've got a couple of team members finishing up another project this week, and we can get started on Monday."

Questions

1. What factors about this situation made heuristic evaluation a better choice than user testing?

Planning and Performing a Heuristic Evaluation

After getting James' approval of his proposal for the heuristic evaluation, Michael immediately requested copies of the PDA concept sketches and flow diagrams of the proposed UI. He worked with James to identify key features that would be targeted in their marketing campaigns. This would ensure that special attention would be paid to these key features during their evaluation. Michael was eager to begin preparing for the evaluation.

Michael enlisted the help of two of his teammates, Shaun Eaves and Ivy Brenner. He knew that having multiple evaluators working separately would increase the likelihood of capturing a good portion of major usability issues.

"Shaun and Ivy, as you know, we will be performing a heuristic evaluation of Isis Mobile's PDA concept. I stress to you both the importance of not collaborating until all three of us have completed the evaluation," Michael stated as he distributed the copies of Isis Mobile's concept sketches and flow diagrams for their UI.

"Of course, Michael," Shaun interjected. "Ivy and I are aware that not collaborating prior to completing our evaluation will minimize any bias in our independent analysis."

"Also, be sure to review the documents in full before beginning your evaluations, then evaluate on the second pass," added Ivy. "It is vital to familiarize ourselves with the materials and have a good understanding of Isis Mobile's intent prior to commencing the evaluation." Michael and Shaun nodded in agreement.

Michael also distributed copies of the Usability Heuristics (Table 13.1) and the First Principles of Interaction Design (Table 13.2). First, they reviewed the heuristics together. This served to refresh their memories and to answer any questions they might have had.

Michael then reminded Shaun and Ivy to provide a written summary of their findings, including the heuristic that applies to each of the findings. He also reviewed the Severity Rating Scale (Table 13.3) that they should use to classify each issue. This would ensure that each evaluator was using the same criteria in their ratings.

Michael had prepared a list of common cell phone/organizer tasks that users would be likely to perform and key features on which Isis was focusing their market strategy (Table 13.4). He distributed these to Shaun and Ivy as well, so they would be sure to cover these in their evaluation. They reviewed these next for clarity. Michael also encouraged them to provide as many solutions as possible to any design issues they found, to ensure the design team had options from which to choose.

"In a week, we will meet again and compare notes," Michael reminded them. "I will send out an Excel spreadsheet template for you to enter your findings into so it will be easier to compile later."

"Great, that will save us a good deal of time when we meet to review our issues." Shaun was really glad to hear that Michael had already thought of that.

Table 13.1. Usability Heuristics Developed by Jakob Nielsen and Robert Mack

Heuristic	Description
Visibility of system status	The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
Match between system and the real world	The system should speak the users' language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in natural and logical order.
User control and freedom	Users often choose system functions by mistake and need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
Consistency and standards	Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
Error prevention	Even better than good error messages is a careful design that prevents a problem from occurring in the first place.
Recognition rather than recall	Make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
Flexibility and efficiency of use	Accelerators—unseen by the novice user—may often speed up the interaction for the expert user to such an extent that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
Aesthetic and minimalist design	Dialogues should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
Help users recognize, diagnose, and recover from errors	Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
Help and documentation	Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

From Nielsen, J., and Mack, R. L. (1994). *Usability inspection methods.* John Wiley & Sons. New York, NY.

Table 13.2. First Principles of Interaction Design Developed by Bruce Tognazzini

Principle	Description
Anticipation	Applications should attempt to anticipate the user's wants and needs. Do not expect users to search for or gather information or evoke necessary tools. Bring to the user all the information and tools needed for each step of the process.
Autonomy	The computer, the interface, and the task environment all "belong" to the user, but user autonomy doesn't mean we abandon rules. Use status mechanisms to keep users aware and informed. Keep status information up to date and within easy view.
Color blindness	Any time you use color to convey information in the interface, you should also use clear secondary cues to convey the information to those who won't be experiencing any color coding.
Consistency	Levels of consistency: 1. Interpretation of user behavior (e.g., shortcut keys maintain their meanings) 2. Invisible structures 3. Small visible structures 4. The overall "look" of a single application or service (splash screens, design elements) 5. A suite of products 6. In-house consistency 7. Platform consistency
Defaults	Defaults should be easy to "blow away:" Fields containing defaults should come up selected, so users can replace the default contents with new material quickly and easily. Defaults should be "intelligent" and responsive. Do not use the word "default" in an application or service. Replace with "Standard," "Use Customary Settings," "Restore Initial Settings," or some other more specific terms describing what will actually happen.
Efficiency of the user	Look at the user's productivity, not the computer's. Keep the user occupied. Because, typically, the highest expense in a business is labor cost, any time the user must wait for the system to respond before they can proceed, money is being lost. To maximize the efficiency of a business or other organization you must maximize everyone's efficiency, not just the efficiency of a single group. The great efficiency breakthroughs in software are found in the fundamental architecture of the system, not in the surface design of the interface.

Table 13.2. *Continued*

Principle	Description
	Write help messages tightly and make them responsive to the problem: Good writing pays off big in comprehension and efficiency.
	Menu and button labels should have the key word(s) first.
Explorable interfaces	Give users well-marked roads and landmarks, then let them shift into four-wheel drive.
	Sometimes, however, you have to provide deep ruts. The closer you get to the naive end of the experience curve, the more you have to rein in your users. A single-use application for accomplishing an unknown task requires a far more directive interface than a habitual-use interface for experts.
	Offer users stable perceptual cues for a sense of "home."
	Make actions reversible.
	Always allow "undo."
	Always allow a way out.
	However, make it easier to stay in.
Fitts' law	The time to acquire a target is a function of the distance to and size of the target.
Human interface objects	Human interface objects can be seen, heard, touched, or otherwise perceived.
	Human interface objects that can be seen are quite familiar in graphic user interfaces. Objects that play to another sense such as hearing or touch are less familiar. Good work has been done in developing auditory icons.
	Human interface objects have a standard way of interacting.
	Human interface objects have standard resulting behaviors.
	Human interface objects should be understandable, self-consistent, and stable.
Latency reduction	Wherever possible, use multi-threading to push latency into the background.
	Reduce the user's experience of latency.
	Make it faster.
Learnability; limit trade-offs	Ideally, products would have no learning curve: Users would walk up to them for the very first time and achieve instant mastery. In practice, all applications and services, no matter how simple, display a learning curve.
	Limit the trade-offs. Usability and learnability are not mutually exclusive. First, decide which is the most important, then attack both with vigor. Ease of learning automatically coming at the expense of ease of use is a myth.

Table 13.2. *Continued*

Principle	Description
Metaphors	Choose metaphors well, so they will enable users to instantly grasp the finest details of the conceptual model. Bring metaphors alive by appealing to people's perceptions—sight, sound, touch, and kinesthesia—as well as triggering their memories.
Protect the user's work	Ensure that users never lose their work as a result of error on their part, the vagaries of Internet transmission, or any other reason other than the completely unavoidable, such as sudden loss of power to the client computer.
Readability	Text that must be read should have high contrast. Favor black text on white or pale yellow backgrounds. Avoid gray backgrounds. Use font sizes that are large enough to be readable on standard monitors. Favor particularly large characters for the actual data you intend to display, as opposed to labels and instructions. For example, the label, "Last Name," can afford to be somewhat small. Habitual users will learn that that two-word gray blob says "Last Name." Even new users, based on the context of the form on which it appears, will have a pretty good guess that it says "Last Name." The actual last name entered/displayed, however, must be clearly readable. This becomes even more important for numbers. Human languages are highly redundant, enabling people to "heal" garbled messages. Numbers, however, unless they follow a very strict protocol, have no redundancy, so people need the ability to examine and comprehend every single character. Pay particular attention to the needs of older people. Presbyopia, the condition of hardened less flexible lenses, coupled with reduced light transmission into the eye, affects most people over age 45. Do not trust your young eyes to make size and contrast decisions.
Track state	Because many of our browser-based products exist in a stateless environment, we have the responsibility to track state as needed. We may need to know: • Whether this is the first time the user has been in the system • Where the user is • Where the user is going • Where the user has been during this session • Where the user was when they left off in the last session and myriad other details.
Visible interfaces	Avoid invisible navigation. Most users cannot and will not build elaborate mental maps and will become lost or tired if expected to do so.

From Tognazzini, B. First principles of interaction design. Retrieved 8/27/06 from http://www.asktog.com/basics/firstPrinciples.html

Table 13.3. Severity Rating Scale Used by All Three Evaluators for Consistency

Severity Rating	Description
1	Cosmetic problem that does not impact usability; does not need to be fixed unless time permits
2	Minor usability issue; low priority; impedes usability but is easily corrected by user; likely to be a minor annoyance
3	Moderate usability issue; medium priority; impedes usability; requires effort on the part of the user to correct; some users may fail to do so
4	Major usability issue; high priority; impedes usability; limits use of certain features for most users
5	Usability disaster; problem must be fixed before product release

Table 13.4. Tasks and Features to Evaluate in the Heuristic Analysis

Powering the PDA on and off
Placing a call
Receiving a call
Storing a contact into the PDA's memory
Editing contact information
Deleting a contact from the PDA's memory
Downloading multiple contacts from your PC [Outlook] into the PDAs memory
Scheduling an appointment reminder
Retrieving a voice mail message
Sending a text message
Changing your ring tone
Setting your PDA to silent mode
Conference calling
Connecting the PDA to a wireless headset
Using speakerphone

"Okay, does anyone have any questions or concerns before we end this meeting?" Michael inquired.

Shaun shook his head no.

"All of my questions have already been answered, so I guess we're ready to go," Ivy commented.

"Great, well thanks for your time," Michael said as he began gathering up his paperwork.

Shaun and Ivy left the meeting eager to begin the evaluation. During the next week, they both worked independently to evaluate the PDA concept. Just as Michael had reminded them, they each reviewed all the documentation they had been provided and became familiar with the operation of the PDA.

Questions

2. Why does Michael have three evaluators assigned to evaluate the PDA concept?
3. Why is it important that the evaluators do not collaborate on the analysis?
4. Why does Michael stress finding as many solutions to problems as possible?

Reviewing, Comparing, and Compiling the Findings

In a week the team met to go over their findings. At this meeting each evaluator discussed each of the issues they identified. They merged the issues and their descriptions when there was more than one evaluator who identified the same problem. This allowed any additional insight into the problem to be captured for a more robust description of the issue that would be shared with Isis Mobile. Each evaluator also provided feedback and severity ratings in real time for issues not identified by one or more of the evaluators to see whether others had different perspectives unseen by the original identifier or alternate solutions.

All three evaluators identified the severity ratings they had assigned to each issue and averaged the severity scores for inclusion in the final report. They also reviewed their solutions to each issue and rank-ordered them in terms of their effectiveness in correcting the problem. The completed list of

issues found was sorted by average severity score in descending order. Table 13.5 shows an excerpt of the compiled list.

Questions

5. Why is it important that the three evaluators compare notes about each of the issues identified regardless of whether they identified the issue or not?

After compiling their findings, Michael returned to Isis Mobile to give the industrial design team and the engineering team members the evaluation report. He met with the team to review the findings and answer any questions they had about the findings or implementing the recommended changes. He urged them to contact him if any questions arose as they worked through the issues.

In a few weeks a revised copy of Isis Mobile's PDA sketches and UI was sent to Michael, Shaun, and Ivy for review. Isis Mobile had addressed all the findings and recommendations, and these were being coded into the UI software. Tool production had also begun.

Michael was pleased with the progress that had been made at this early stage of the design, and Isis Mobile was confident that some of the obvious design issues had been caught before tooling and coding. The next step would be to bring users in to validate and help further refine their design.

Validating Design Changes

As the engineering team was implementing changes from the heuristic evaluation, Michael had collaborated with one of the marketing leads at Isis to recruit participants for the user study that fit into the PDA target user base. They did this by using a series of questions that could be used to determine whether potential participants fit the particular market segment Isis was focusing on. They used these questions to screen participants and determine which would be the best representatives for the lab study.

Michael recruited 10 participants in total, 5 for each of two major east coast cities where they planned to begin their initial product release. He was sure to get an even split of men and women and an even distribution of ages within Isis's target audience.

Table 13.5. Excerpt of the Compiled List of Usability Issues Found, Listing Applicable Heuristic(s) and Average Severity Scores

			Severity Ratings				
Issue	**Description**	**Heuristic**	**Michael**	**Shaun**	**Ivy**	**Average**	**Solutions**
No tactile indicator on the number 5 key	This is a standard accessibility requirement and must be included on all telecommunication devices.	Consistency and standards	5	5	5	5.00	Add a raised tactile indicator on the 5 key that complies with standard.
Blank screen in battery save mode	After 10 minutes of inactivity, the product enters battery save mode. In this state, any button press would "wake" the PDA. However, there is no indication of the mode on the screen, which may lead users to believe their PDA was powered off by an inadvertent button press or the battery has failed. A user's first instinct may be to press power, which wakes up the PDA, but initiates power down.	Visibility of system status	3	4	4	3.67	1. Allow power key to wake PDA first and then power down if user presses it again. 2. Create a small battery save icon that remains blinking on a corner of the screen.
Ambiguous option labels when entering a contact	After entering a name and number into the address list, the options presented to the user are "Done" on the left and	Consistency and standards	3	3	4	3.33	Change "Done" to "Cancel," "Exit," or something that implies the number

Table 13.5. *Continued*

			Severity Ratings				
Issue	**Description**	**Heuristic**	**Michael**	**Shaun**	**Ivy**	**Average**	**Solutions**
	"Save" on the right. Users read left to right and may not see the "Save" option before selecting "Done." When a user presses "Done," the address list entry is not stored. It seems logical that users, upon completion of the entry, would assume they were done and not expect this choice would exit them without saving.						will not be stored in memory.
Unintuitive error message in calendar	The calendar allows a maximum of 50 appointment reminders. Once this amount is reached, if the user enters another reminder, a pop-up with the following message appears: "Error—Buffer Overload." This is likely to confuse users because it is not plain language, and it does not help users understand how to correct the problem.	Error prevention	3	3	4	3.33	Change error message to say, "Memory full. Delete expired reminders?" Include "Yes" and "No" button options for users to select.

Table 13.5. *Continued*

			Severity Ratings				
Issue	**Description**	**Heuristic**	**Michael**	**Shaun**	**Ivy**	**Average**	**Solutions**
Unintuitive error message when searching for a contact not listed in phonebook	When searching for a contact in the address list that has not been saved previously, the following error appears: "Error 7x65b; Exit Application and Enter New Information." The error number holds no meaning to users and although some direction is given, it is not clear to users what the problem is. Also, "Enter New Information" is confusing in that it does not specify whether users should search for the contact again or whether they need to add this person to their contact list.	Aesthetic and minimalist design; error prevention	3	4	3	3.33	Reword error message to exclude error number and include clear explanation of problem, consider, "Contact Not Found. Add New Contact?" Have "Yes" and "No" button options for the user to select.
Inconsistent exit labels throughout UI	To exit some menus, the left option reads "back" and for others it reads "cancel," "done," or "exit" but the function is the same.	Consistency and standards	2	3	2	2.33	Choose a term and use it consistently throughout the interface.
No visual indicator for	A press and hold on the number 1	Flexibility and efficiency	3	2	2	2.333333	Add a voice mail icon onto

Table 13.5. *Continued*

			Severity Ratings				
Issue	**Description**	**Heuristic**	**Michael**	**Shaun**	**Ivy**	**Average**	**Solutions**
voice mail shortcut	key automatically dials voice mail as an alternative to accessing it through menus. However, there is no visual cue to this shortcut on the 1 key itself.	of use; recognition rather than recall					the number 1 key.
Multiple labels for calendar feature	In different areas of the User Interface and User Manual, the Calendar feature is referred to as "Calendar," "Date Book," or "Organizer."	Consistency and standards	2	2	3	2.33	Choose one term and use it consistently throughout both the interface and manual.
User manual too wordy and large	Although the information in the manual is thorough, it can be distilled down to contain clear concise steps to accessing features; this reduces the manual size and increase its usefulness.	Help and documentation	2	2	2	2.00	1. Use a bullet format for feature access instructions. 2. Consider providing a separate "quick start" guide that outlines how to get started and two or three major features to get users up and running with their new product.

Michael spent 2 weeks running the test sessions and 1 week preparing the data for the report. He confirmed a high success rate for completing tasks and that the majority of users reported the PDA was extremely intuitive to use. A few minor changes, such as changing some labels, were included as recommendations. Overall, these were easy fixes, and Michael was confident that the PDAs were ready to make their release on schedule.

Isis Mobile benefited from planning and incorporating a heuristic evaluation early on in the design phase of their new product line. They were able to prevent usability problems from entering later phases of the project. They minimized cost in terms of time and rework and validated their design with users before release to ensure their product debut was not greeted with bad press and low sales, thus ensuring that end users would consider Isis Mobile for future purchases.

James Cartwright was not only pleased with the outcome of their initial release, he was also inspired by the possibility of being able to further improve products in a way that would positively impact those who buy and use them. He believed he may have found a way to truly differentiate their brand in the marketplace.

Going forward, as a direct result of the heuristic evaluation, James decided to hire a team of user-centered design personnel as part of the design team to help drive concepts into usable form and provide guidance throughout the product development process—not just at a point when concerns arise. He also mandated that a portion of the product development budget be allocated for usability studies and forward-looking research. He was convinced that the company would flourish because they were focusing on their customer's experience as a top priority.

Summary

The steps involved in planning and performing the heuristic evaluation included the following:

- Identifying the evaluators; three to five is typical for a heuristic evaluation.
- Ensuring that evaluators do not collaborate before completion; this minimizes bias.
- Using a common rating scale for severity of issues; this ensures consistency.

- Reviewing the entire interface at least twice; first to familiarize yourself with it and then to scrub for issues once a good understanding of the design intent has been realized.
- Documenting issues, applicable heuristics, severity, and possible solutions using a common template and/or checklist; this allows for ease of sorting and organizing issues for reporting.
- Comparing findings, severities, and proposed solutions after all evaluators are done with their evaluations; this distills the information into a single report.
- Validating design changes with users and refining the design where necessary.

The major benefits of the heuristic evaluation included the following:

- Can be implemented very early in the design process; it does not require functional products.
- Requires only a small number of evaluators and does not require the additional cost of recruiting participants or booking a lab.
- Can be done in a short amount of time.
- Provides a cost-effective way to prevent major known usability issues from making their way into later stages of the design process.
- Saves both time and money by eliminating problems very early and minimizing any last-minute rework.

A heuristic evaluation does have some shortcomings:

- It does not leverage feedback from actual users.
- It is subject to evaluator bias.
- It is extremely dependent on the skill level of the people doing the evaluating.
- It does not provide any quantitative data.
- A user study is recommended as a follow-up to validate the design.

Conducting up-front heuristic evaluations of design concepts can minimize rework later in the development process. It is a healthy way to validate whether initial designs are on the right track. Following up with user studies is highly recommended. Testing early and often can mean the difference between product success and failure.

Further Reading

Jeffries, R., Miller, J. R., Wharton, C., and Uyeda, K. M. (1991). *User interface evaluation in the real world: a comparison of four techniques*. Presented at CHI '91 Proceedings, (New Orleans, LA, April 28–May 2): pp. 119–124.

John, B. E. (1997). Tracking the effectiveness of usability evaluation methods. *Behavior & Information Technology,* 16:188–202.

Nielsen, J. (1992). *Finding usability problems through heuristic evaluation*. Presented at CHI '92, Proceedings (Monterey, CA, May 3–7), pp. 373–380.

Nielsen, J. (1994). *Enhancing the explanatory power of usability heuristics*. Presented at CHI '94 Proceedings, (Boston, MA, April 24–28), pp. 152–158.

Nielsen, J., and Landauer, T. K. (1993). *A mathematical model of the finding of usability problems*. Presented at INTERCHI '93 Proceedings, Amsterdam, April 1993 pp. 206–213.

Nielsen, J., and Mack, R. L. (1994). *Usability inspection methods*. John Wiley & Sons New York, NY.

Nielsen, J., and Molich, R. (1990). *Heuristic evaluation of user interfaces*. Presented at ACM CHI '90 Proceedings, (Seattle, WA, April 1–5), pp. 249–256.

Tognazzizi, B. *First principles of interaction design*. Retrieved 8/27/06 from http://www.asktog.com/basics/firstPrinciples.html

Vijavan, R. (1997). *CS6751: Topic report, heuristic evaluation*. Retrieved from http://www-static.cc.gatech.edu/classes/cs6751_97_winter/Topics/heur-eval/

CASE 21

Aikot Corporation: A Case Study in Qualitative/Quantitative Remote Evaluation

Carol Farnsworth, Keynote Systems

"Look, here's the thing," said Mark. "Truth is, we focus most of our attention on in-store retail, advertising, and packaging. We only built the website because everyone has one. But now all this troubles me. There is a possibility that our website might not be just increasing market share but may actually be contributing to the decline in our market share."

Decline of a Strong Brand

Aikot Corporation manufactures and sells electronic products, such as televisions, video cassette recorders, video players, radios, stereos, and mobile phones, which are used by both consumers and businesses. Aikot Corporation grew quickly during the 1980s and 1990s, beginning with thousands of retail stores throughout the United States. In the 1990s Aikot expanded to Canada, Central America, South America, and Europe. During the 1990s Aikot was a popular brand and was ranked in the top 100 Global Brands in 1996 by the Business Week/Interbrand report.

Aikot Corporation was creative about marketing their products and used a variety of marketing strategies. They attempted to establish product branding through a 360-degree strategy that included multimedia advertising (print, television, radio), product packaging, and more.

In the mid-1990s the Internet became a viable marketing channel. As part of the overall corporate marketing strategy, Aikot created an online Internet marketing team whose main mission was to develop the corporate

website. The first website was launched for the U.S. market, and within a few years the European and Latin American sites were launched.

For years Aikot Corporation enjoyed continual growth in market share and revenue, which resulted in climbing stock prices. But beginning in the late 1990s and in the early 2000s this growth began to steadily and slowly decline. Market share was falling and the value of the brand decreasing. With the rising anxiety there were threats of layoffs, and the Board of Directors contemplated selling off a portion of the business.

Mark, the chief marketing officer, conducts marketing campaigns that drive consumers to the retail stores and the websites. He always makes sure that the in-store teams know when the campaigns are in progress so they are prepared for customers. He also provides the website teams with the content required for each marketing campaign.

In spite of all the careful planning, Mark did not understand why the company was experiencing the slow, yet steady, decline. During a weekly executive meeting, Mark responded to his peers' questions regarding this issue. "The public perception and feedback we received from our customers is that they recognize Aikot to be a leader in innovative design and that we have competitive pricing."

To try to get a better handle on the declining market share problem, Mark sought guidance from a variety of market and consumer research firms to obtain a better understanding of the purchase decision-making process and the marketplace as a whole.

From this consumer purchase-behavior research, Mark validated his hypothesis that the web is a key channel in the purchase decision process for both consumers and businesses. He knows that if a website does not meet visitors' goals, then the visitors move on to other websites and brands. Visit-and-purchase success on a website increases the likelihood consumers will return to the site, recommend the brand to others, and purchase additional products either online or off-line.

Mark learned of the growing trend for consumers to transact online with suppliers directly rather than with resellers, such as Best Buy and Office Depot. The web is a key online channel for making purchase decisions and transacting, and until recently consumers have been shopping more on reseller sites than on supplier sites. Consumers use online portals and search engine results to find resellers when beginning their product research. The reseller websites are generally comprehensive and convenient one-stop shops with multiple products and suppliers available on one website, making comparison shopping possible and convenient.

Resellers have made it easy for consumers to make the move from offline to online transactions. They were the first to develop websites with online shopping and purchase transaction functionality. They offer low discount pricing because they buy large quantities from suppliers. Consumers can comparison shop without leaving the website. Resellers have invested in understanding the user experience on their websites and have made continual improvements to the user experience over the years. Generally, consumers find what they are looking for, at a price that is acceptable, and complete the purchase process online.

Mark also learned that consumers are spending more time researching products and looking for information on the web. To keep up, companies need to understand what is driving and influencing the purchase decision process and give consumers what they want—good quality products at a fair price, using a website that is easy to find and purchase products.

Using this research data, Mark began to take a good look at Aikot's online initiatives. Aikot's products are excellent, innovative, and seem to meet customers' needs. They have strong competitors, but Mark believes Aikot's marketing continues to be superior. They have the majority of the market share, although those numbers are decreasing.

Mark understands that competition is tough everywhere and comes from unexpected places. Website visitors researching products can browse products and services from multiple companies at one time by opening multiple web browsers. People will keep looking until they find exactly what they want at the price they are willing to pay. Even auction sites, like eBay, are strong competitors for many businesses. Other websites serve as conduits for many other companies, such as BizRate.com and Buy.com. Through these single websites potential customers can see many companies' products. With stores, customers must travel to get there. So if they don't find the perfect product in the store, sometimes they will take the next best thing. On the web, visitors must be able to find and learn about products. If you provide a way for people to buy products on your site, then you have to make sure people can actually complete the full transaction.

Aikot's website has a pretty good e-commerce application, and most of the customers who do want to purchase online are able to complete the process. However, Mark realizes there may be other difficulties of which he is not aware. For example, he knows that an important aspect of the consumer product research process is finding and comparing prices for products. Some products on the Aikot site do not have prices, and he does not know what impact this has on the customer experience.

Mark reflected that from a business perspective, it is very difficult to provide pricing information for cell phones. Cell phone sales depend on available wireless phone service within the location of the buyer. As a result, cell phones are sold by service providers, such as Sprint and Verizon. The "service package" includes cell phones and, depending on what the current promotion is, the phone unit itself might be free or cost several hundred dollars. Mark knows consumers want to see prices and will continue to go to the site for this purpose.

Aikot doesn't offer pricing on most of their sites because the list price may not be the price available for any given individual consumer. There are so many dependencies based on location alone. Cell phones may not be available for some consumers if wireless providers don't have service in their area. Cell phone units may be free of charge with some service contracts, whereas with others the consumer may have to pay full price. If consumers want to upgrade their phone before the service contract is complete, they may have to pay a penalty fee on top of the cost of the phone unit. For these reasons, Aikot doesn't publish cell phone pricing on the sites that do not have online stores.

Mark realizes that he does not know very much about the experiences visitors have on the Aikot sites. Specifically, he doesn't know if visitors can find what they are looking for on the site or if they are successful and satisfied. He doesn't know if the experience on the website is having a positive or negative impact on the impression of the Aikot brand. He doesn't know when in the purchase process people are visiting the site. He has some assumptions about who the visitors are, but he does not have any data to support his assumptions.

The data Mark currently has available are the traffic pattern data, such as the number of unique visitors to each page, number of pages for an average site visit, and time spent on each page, provided by a product called "Hit Box." To capture the data using the Hit Box application, pages are tagged with code that tracks the number of people who visit that page. For multistep task flows, or process funnels, such as the registration process and online store checkout for purchases, Hit Box provides information about how many people begin the process funnel, from which pages in the process funnel people drop out, and how many people complete the process funnel all the way to the confirmation page. This provides information about conversion rates—the actual number of people completing the checkout process resulting in a purchase divided by the number who begin the checkout process. The fundamental questions for Mark are "why are visitors dropping out of the

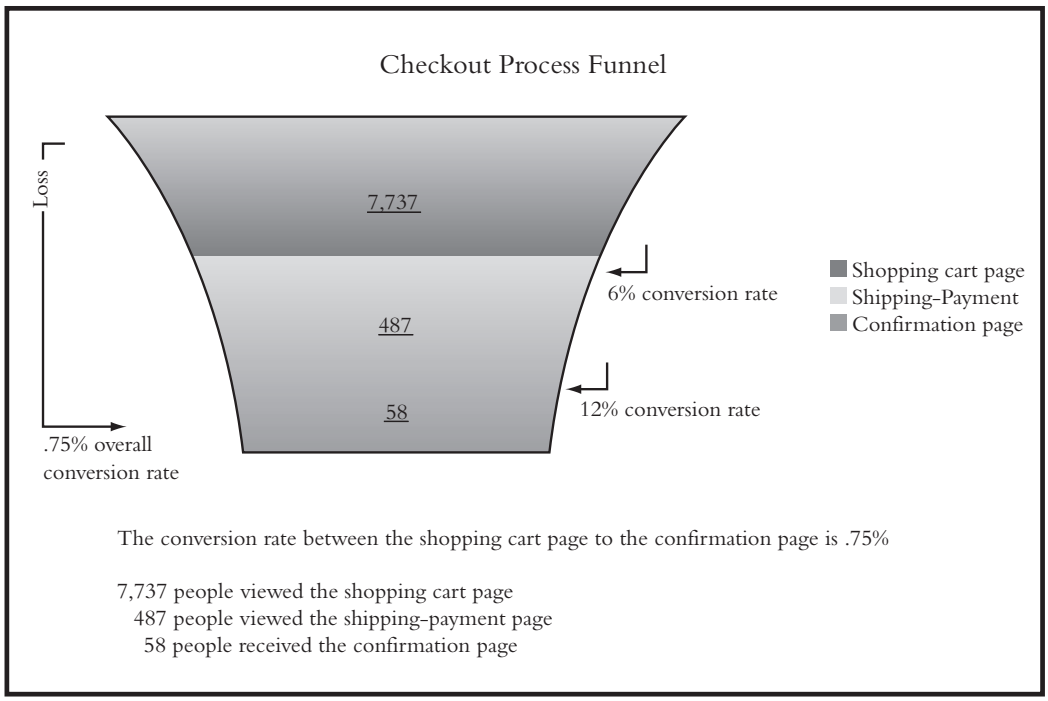


Figure 21.1. A funnel graph depicting the checkout process.

purchase process funnel?" and "why are they not completing the purchase transaction?" (See Figure 21.1.)

At lunch one day with his friend Anthony, whose company does business exclusively on the web selling beauty and cleaning supplies primarily to small businesses with fewer than 200 employees, Mark was talking about the website and what he knows and what he needs to learn about the user experience on the Aikot website. He asked Anthony for suggestions.

"I have several people on my website development team who are responsible for understanding our site visitors," Anthony responded. "The data they collect drives our development efforts. I have learned to pay attention to site visitors. If they are not successful in finding what they want, they will not buy our products, visit the site in the future, or refer the site to their friends. If people don't come back to the site, I lose money."

"Look, here's the thing," said Mark. "Truth is, we focus most of our attention on in-store retail, advertising, and packaging. We only built the website because everyone has one. But now all this troubles me. Based on everything I know and have learned from the consumer research, and

assuming that our product offering is better or as good as our competitors, there is a possibility that our website might actually be contributing to the decline in market share."

"Yes!" said Anthony. "A website is almost never 'neutral'—it either helps or hurts a business. You need to find out what's going on with the site and have a better understanding of visitors' experience. Right now, it sounds to me like you have no idea what impact your website is having on your company."

Anthony continued, "Members of our web team have compensation incentives for hitting certain success and satisfaction targets we have identified for the site. Do you have anyone on your web team responsible for the visitor experience on the sites? Have you established a set of performance metrics that can be tracked over time?"

"We don't have a specific person on the team responsible for the visitor experience. The web teams consist of a manager who is responsible for executing and managing the day-to-day details of hosting the website, product managers, a website manager, a web content manager, and a third-party design agency," responded Mark. "We also don't have specific goals identified for the site other than to have a corporate online presence."

"You need to have people on your team who understand the site visitors and the impact the site is having on your brand. You also need to identify the goals and objectives of the site and establish targets for measuring the visitor experience," said Anthony.

The course was clear to Mark: Aikot needed to review and renew their overall marketing plan, establish goals for the site, hire team members who can manage the visitor experience, and understand the experience on the websites from a visitor's point of view.

Questions

1. What types of questions should Mark ask himself about his website to give him clues about how it is affecting sales?
2. What methods of customer feedback does Aikot use today, and why are they not enough?
3. What about the company structure, team makeup, and process contributed to the lack of knowledge about the influence of the website on the company's bottom line?
4. What should Mark do now that he understands the source of the problem may be the website?

The Nature of Quantitative/Qualitative Remote Website Visitor Research

Staffing the Customer Experience Team

Mark knew he had his work cut out for him. Changes would have to be made to the overall marketing plan, recognizing that the website plays a critical role in preserving market share. Aikot's business managers' perceptions of the role of the website in the overall marketing plan will need to change as well. Mark decided to first focus on team development, because any marketing plan needs a strong online marketing team. Anthony had told him about a new role many of his peers in other corporations had created, Director of Customer Experience. Mark decided it was time to hire one for Aikot.

After an intensive search, Mark hired Anne to lead the customer experience team. She has both an MBA and a Master's degree in Human Factors. Throughout her career, Anne has helped software application and website development teams make the experiences of the people who use the applications and visit the websites more enjoyable and successful. In short, she helps teams to focus on the customers' needs. At Aikot, Anne's role would be to listen to the customer and to communicate customer requirements and the strengths and weaknesses of the website to the design and development team.

Anne understands that the Internet plays an important role in the overall customer experience, especially in the purchase decision-making process. Each year, customers spend an increasing amount of time on the Internet researching products, finding information, and making purchases. Success in her new role at Aikot would be accomplished as Anne discovers what drives and influences customers' purchase decision process and decision to make future purchases with the company. She must translate this knowledge into actionable development plans for the web team.

Anne's compensation plan for the next several years is based on improving the online customer experience and increasing brand value. To do this, Anne needs to have information on the current customer purchasing experience on the website. Anne will focus on the prepurchase, purchase, and postpurchase processes on the Aikot website.

To be successful in achieving Aikot's customer experience goals, Anne needs people on her team with usability and user-centered design skills. Anne's team comprises the website managers for three countries: David from Germany, Tomas from Mexico, and Jennifer from the United States.

Anne was concerned she wouldn't be able to do what was needed to manage the customer experience on the websites with the existing team whose primary focus is on executing the day-to-day activities required to manage the websites. Anne believed that having another person design and execute the user research activities was critical to the success of the new plan. Anne called this the "customer experience manager" role. The person hired in this role would have a background in usability and user-centered design.

Anne wanted to hire an external person for the customer experience manager role, but Mark insisted that there could not be any new head count positions this year. He did give Anne the approval to hire a contractor on a part-time basis. So, Anne hired Cynthia Wilson, a user-centered design/usability consultant who has been working with global corporations to improve customer experience on websites. Cynthia has extensive experience conducting U.S. and international focus groups, in-lab, one-on-one usability studies, and remote qualitative/quantitative research.

Reviewing the Consumer Purchase Process

Anne met with her team—David, Tomas, Jennifer, and Cynthia—to review key aspects of the online purchase process and to set goals for improving the online customer experience. The meeting started first with a high-level review of the consumer purchase process.

After introductions, Anne asked David to summarize the consumer prepurchase process. David explained, "The customer experience during the prepurchase process involves all activities consumers complete before making a decision to buy a product. This stage of the buying process includes

- Receiving communication from the company about their products through multimedia advertising, such as television, radio, newspaper, magazines, billboards, etc.;
- Learning about products from family and friends; and
- Visiting the stores and websites."

Cynthia interjected, "We have to make it easier for visitors to find information in the prepurchase stage since learning about the products is the primary customer goal at this stage. Competitive intelligence research conducted by Keynote Systems, Inc. indicates that good websites typically have visit success rates of 80% or more."

Next, Tomas described the purchase process on the Aikot sites. "Customers are able to purchase products online on the U.S. and Germany websites.

On the Mexico website, visitors are only able to shop for and learn about the products and locate stores to purchase the products. We do not have an online store on this site but plan to add this functionality in the next 6 to 12 months. To add online shopping to the Mexico site we need make sure the site has everything required to make the purchase online, including product displays, product descriptions, shopping baskets, registration, billing, delivery, shipping, payment, and confirmation notifications."

Jennifer cautioned that other considerations would have to be included in the online shopping experience. "Once products are purchased, customers often need additional support, such as learning how to set-up, fix, and/or learn more about how to use the products. The Aikot website has an extensive customer support section and is maintained by the corporate marketing web development team."

Identifying Aikot Website Visitors

"Who visits the website?" Cynthia asked.

Being the most familiar with the web analytics data that reports site traffic, David was quick to respond. "We have 387,328 unique visits per month and, of those, 9.5% are returning visitors."

Cynthia said, "Traffic data do provide important information. But what are the profiles of the people who visit the site?"

"I am sure they are the same people who visit and purchase products in our stores," said Jennifer. "We have four main customer profiles—enthusiasts, conservatives, stylists, and low budget—developed from market research studies conducted for our off-line stores and sales. The enthusiasts are early adopter consumers who like to be the first to have the latest products, and their purchase cycles are somewhat short. Conservatives do everything they can to extend the life of their product, and they spend a very long time in the prepurchase decision phase. The stylists care most about the impressions they make on their friends and would rather have a product with a lot of style instead of one with the most functionality. Finally, the low budget consumer wants the best product for the lowest cost. The low budget consumer buys only products on sale or last year's models. They also purchase refurbished products when we offer them."

"We use content, product messaging, and advertising for the websites similar to what we use in the stores. It makes the production process more efficient," said Tomas.

"Do you have any idea who the actual people are who visit the website?" asked Cynthia.

"Look, Cynthia, the fact is, we have not conducted any research on the sites. For development purposes we have been making the assumption that the people who visit the website are the same as those who visit our stores," replied Anne. "After all, why would we think otherwise? Besides, not all of our websites have online stores, so I would think that the profiles of our brick-and-mortar stores are more important than the website visitor profiles."

Cynthia folded her hands on the table. "This is important because this means that you also do not know why visitors are coming to the site and whether or not they are successful. It also means that we have no idea what impact the website has on visitors' impression of Aikot the company and decisions to purchase Aikot products or not. Do you have key web metrics in place for the website so you know how well the site is performing?"

"No," responded Anne. "We need to develop key web metrics at some point during this year, but in the meantime it sounds like we need to learn more about the visitors of our website. Cynthia, you have conducted research all over the world for large international corporations, what methodologies do you use to learn more about site visitors?"

Cynthia explained, "Quantitative/qualitative online remote research is most effective for answering the types of questions we have for this site. The research I am proposing is quantitative in that we invite a large number of people to participate. Generally, 200 to 400 participants is a good number, but if we want to segment the data by certain criteria such as customer profile or visit purpose, we would want 400 to 800 participants. The research is qualitative in that we capture behavioral data as the participants are navigating the site and we can ask open-ended questions about why they are doing what they are doing. Some of the behavioral data that are captured includes navigation paths, time on each page, time on each task, and number of pages visited."

"I thought six to eight people were enough to obtain insights into the usability of a website or design," interjected Tomas.

Cynthia responded, "Traditional usability studies where we ask six to eight people to interact with an existing website or a new design for a website will not provide us with the necessary visitor data to answer the questions to help us understand who is visiting our website, why they come, and are they successful. Traditional usability studies rely on a handful of people who match a particular user profile to interact with a website, application, or device to help us understand if there are any problems with the design. The usability sessions provide good indications of the problems people may encounter, but they don't tell you how often they happen to real customers and visitors to your website."

"Quantitative research is well suited for answering a number of questions organizations may have about their website and help business managers make decisions regarding development efforts that may cost hundreds of thousands and even millions of dollars to build," continued Cynthia. "Using statistically significant sample sizes of 200 to 400, business managers have more confidence in understanding the strengths and weaknesses of a website than on the responses of six to eight people."

"For the type of exploratory research we need for the Aikot website, we will rely on people naturally visiting the website and in statistically significant quantities—around 200 to 600 completed studies for this type of exploratory research. The more participants we have, the more segmenting we can do and still have large numbers in the sample size. The sample size we need for this study is partly dependent on how many segments we want to analyze after we collect all the data. For a study with no segmentation, 200 participants is a good number and will provide statistically significant data at a high confidence rate. To answer our questions, we need to listen to people who are visiting our websites," replied Cynthia.

"I've seen pop-up surveys on a few of the websites I surf. I completed one survey and they asked a zillion questions about my experience," added Jennifer.

Cynthia replied, "The problem with a 'survey' is that people have to *remember* an experience and self-report data of an event that happened some time in the past. People do not always report information about a past experience accurately. If we conduct quantitative/qualitative studies online, we can obtain both self-reported and behavioral data. We can also capture how long it takes visitors to complete their tasks, how long they spend on each page, and their navigation paths during their visit. It is a known fact that self-reported information does not always match actual behaviors. This is why we need research methodologies that help us gather as much information as possible. The quantitative results give significance to the self-reported information and behaviors, whereas the qualitative observations deliver insights into the visitors' thoughts and actions that help us understand the strengths and weaknesses of the site."

"How can we do that?" asked David.

"With applications that allow us to conduct remote evaluations. Some people call this 'unmoderated' research because the application does it automatically, participants are in their own environment, using any computer available to them at that time, and a researcher is not present during the evaluation. The data analysis from this type of research is much more robust

than analysis of pure survey questions as well," responded Cynthia. "We will be able to analyze the navigation path, correlate actions/factors with positive and negative experiences, and segment the data to identify differences between needs of visitors based on what they came to the site to do. We will also be able to conduct an explicit or implicit needs assessment, and . . ."

Tomas interrupted Cynthia. "I don't understand how you are going to get enough data to do all this analysis and what data will be collected."

Cynthia responded, "There are a number of companies, such as Keynote Systems, Inc., with software applications that allow teams like ours to conduct quantitative/qualitative research online with 200, 400, 600, or more website visitors. Since a large number of people naturally visiting the site are invited to participate in the study, we will have confidence in the findings regarding the strengths and weaknesses of the site. In addition, the large sample sizes will help us understand who is visiting the site and provide data to be able to create profiles of these site visitors."

"The remote evaluation applications allow us to ask participants questions while behind the scenes the application captures data, such as time on task, page load time, URLs visited, and words entered in search entry fields," continued Cynthia. "The steps for conducting this type of evaluation include designing the research activities and questions, programming the study using an online remote application, adding code to invite participants to the Aikot website, collecting data, analyzing the data, and reporting findings and recommendations."

Anne nodded her head. "It's pretty clear that we need to first identify our goals for this year. Second, we need to identify key metrics for the website and, finally, learn more about conducting the quantitative/qualitative remote online exploratory research so we can learn more about our visitors."

Questions

5. Based on the discussion of the team members, what should some of the overall goals for the Aikot user experience team be?
6. What are some of the advantages of remote testing that Anne and her team can exploit?
7. What information should the team be able to collect by conducting an online remote research study?
8. Anne and her team want to ensure that the visit experience is as natural and unobtrusive as possible. How could they best invite and entice visitors to participate?
9. How could the team structure the study itself to be as natural and unobtrusive as possible?

Planning and Conducting the Quantitative/ Qualitative Remote Study

"Looks like we are ready to begin planning the study," said Cynthia. "In this quantitative/qualitative remote evaluation, we will invite 200 to 800 website visitors to participate. We will invite people who are naturally visiting the website by intercepting them with a pop-up, or 'floater image,' as they are called today. This floater image contains a link to the research study and is almost like an advertisement asking the visitor to give their opinions and feedback about their experience on the website. This type of evaluation is great because people complete the study using their own computers at home or work or any place they have access to a computer. They can also take as much time as they want to explore your website and answer the questions. Giving people this type of freedom results in a richness of data that we are not always able to get in more traditional user research."

"Aikot has three websites—United States, Germany, and Mexico," Cynthia continued. "Since this is the first user research to be conducted on the site, I recommend that we focus this study on the U.S. site. Later we can have the study script translated into German and Spanish and conduct the study on the other sites independent of the U.S. evaluation."

"I agree," said Anne.

Cynthia worked with Anne and the team to finalize the study questions. In addition, she arranged for a one-time use license for the online remote application for the Aikot U.S. website. Cynthia programmed the study, sent the study URL to the team, and asked the team to review the programmed study and send her feedback.

Here is the flow of the study from the participants' point of view:

- Visitors arrive at the home page of the website.
- Visitors are intercepted with an invitation to participate in the study.
- Participants answer a few key questions before completing their self-directed task.
    - How did you hear about Aikot?
    - How many times have you visited the Aikot website in the past?
    - What is the primary purpose of your visit to the Aikot website today? Please select the most appropriate answer from the list below.
    - Which Aikot products are you mainly interested in? Please select one.
    - Would you be interested in buying any of the following online? Please select all that apply.

- Before beginning the next part of the study, participants are asked to download a small application. This application collects behavioral data such as URLs visited, time on task, and so on. These data can also be collected through a proxy server if the download option is not desirable.
- For this part of the study, participants interact with the Aikot website, completing the tasks they originally intended to complete when they came to the site. This is the only task in the study, and it is self-directed. As soon as the participant is finished their task, they click on the "answer" button.
    - Here's how we phrase the self-directed task: "Now please continue doing what you originally came to the Aikot site to do today. Spend as much time and effort as you typically would on any site visit. When you are finished, please click the 'Answer' button."
- After participants complete the self-directed task, we ask them to answer questions designed to capture the overall experience while completing their intended task.
    - Which of the following main content on the Aikot site did you visit today?
    - Did you find the information you were looking for?
    - How successful were you in achieving your primary goal for today's visit? (This is a rating question followed by free-form questions: Why were you successful? or Why were you not successful?)
    - How satisfied are you with the website overall? (This is a rating question followed by free-form questions: Why were you satisfied? or Why were you not satisfied?)
    - What, if any, difficulties or frustrations did you encounter? Please select all that apply. (This is a multiple-select question and includes responses that describe common difficulties that might be encountered, no difficulties encountered, and Other—please specify.)
- In the next part of the study, participants answer questions regarding their current Aikot product ownership. This section also includes questions regarding how much influence—positive or negative—the Aikot website had on the previous purchase of the products.
- Next, participants are asked several questions about their future intent to purchase new products and which brand they intend to select for purchase.
- After the current and future purchasing behavior questions, we ask them questions regarding brand value.

For example, we ask, "How did your visit today change your opinion of Aikot in general? Ratings 1, 2 much worse opinion, 3, 4, 5 no impact to my opinion, and 6, 7 much better opinion.
We also ask participants to rate future calls to action such as likelihood to return to the site, likelihood to purchase from the site, likelihood to purchase in a store as a result of this visit, and likelihood to recommend this site to other people.
When you consider everything surrounding your interaction with Aikot, what is your overall satisfaction with the company itself?

- In the final section of the study, we ask a variety of demographic questions targeted to help us understand the profile of the visitors. This section helps us identify and validate our four main off-line profiles: enthusiasts, conservatives, stylists, and low budget customers.

"This is great to see the programmed study. Are you sure it is not too long?" asked Jennifer.

Cynthia shook her head. "I've conducted hundreds of these exploratory studies and have found that 30 to 40 questions is an acceptable length for a quantitative/qualitative remote study, especially since we are not going to give these participants an incentive and that this study should take 15 to 20 minutes longer than the participants' normal visit to the website. If we planned to conduct a regular directed task-based online study so we can understand the strengths and weaknesses of the site in more detail, we would need a longer study. When we ask participants to spend more than 20 to 25 minutes (but never longer than 45 minutes) on a study, we reward their efforts with a gift certificate."

The team was full of questions. Jennifer wanted to know, now that the questionnaire was programmed into the application, how to launch the study and what the next steps were, such as how they'd get participants. David wondered whether they would be able to see the data while the study was in progress. Tomas wanted to know how participants in non-English language tests would be able to see the study in their own language.

Cynthia was glad that the team had a lot of questions about conducting the research.

"How are we going to get people to participate in this study?" asked Jennifer.

Cynthia responded, "This is an exploratory study, so we need people who are naturally visiting our website. Intercepting people from the home page is the best way to invite people who are visiting a website. We might

use different invitation methods for studies with different goals. For example, if we want to conduct a directed task-based usability study or competitive comparison study, we would use a panel of participants. We would recruit the participants from a market research vendor and invite them through an e-mail invitation. Embedded within the invitation would be a link to our study."

"That makes sense. So we are planning to intercept people on our site for this study. How exactly do we do this?" ask Jennifer.

Cynthia replied, "Potential participants can be recruited using pop-up Javascript code, but with pop-up blockers so prevalent these days, most website visitors won't see the invitation. A more effective alternative is to add layered-based 'floater code' to the home page HTML. This is also Javascript code that is inserted into the header of the home page, but instead of a pop-up an image appears on the page. The image contains an invitation to participate in the study and can be set to appear at specific intervals. For example, if a website has high traffic, the floater image may be set to appear at a frequency of 1 out of every 50 visitors and for a website with very few unique visitors the floater image may be set to appear at frequency of 1 out of every 2 visitors."

David, the team member most familiar with the site traffic, suggested, "We should use floater code and intercept at a rate of one out of three visitors. Here's the message we can present to our visitors. . . ." (See Figure 21.2.)

Before adding the code to the website, everyone reviewed the programmed study and gave approvals for launch. All final modifications were made. The study was ready to be launched. Jennifer, the U.S. website manager, had her web development team add the floater code to the home page header. The floater code is the Javascript code that is added to the home page of the website to intercept people naturally visiting the website. The

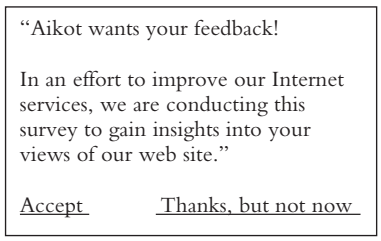


Figure 21.2. Intercept floater image.

```
<script language="JavaScript1.2"><!--
 var InterceptLikelihood = 0.01;
 var InterceptTaskKey = 'A1B23C4D5E6F';
 var InterceptType = 'Layer';
 //Height and Width in pixels
 var InterceptWidth = '400';
 var InterceptHeight = '450';
 function Handlentercept()
 {
  try {
   if (Math.random() >= (InterceptLikelihood*5)) return;
   var s = document.createElement('script');
   s.src = 'http://company.com/applications/intercept/filter_page.asp?inv=' +
InterceptTaskKey + '&type=' + InterceptType + '&rate=' +
InterceptLikelihood + '&max=5' + '&width=' + InterceptWidth + '&height='
+ InterceptHeight;
    document.body.insertBefore(s, document.body.firstChild);
    window.Window = 'primary';
  }
  catch(e){}
 }
 if (window.attachEvent) window.attachEvent('onload',HandleIntercept);
  else window.addEventListener('load',HandleIntercept,false);
//--></script>
```

Figure 21.3. Example intercept floater code.

team inserted the code between the <HEAD> and </HEAD> tags on the web page except for the <BODY> tag. The <BODY> code either needs to be combined with the page's <BODY> tag or to replace it. Figure 21.3 shows an example of what the code looked like.

The website managers sent a note to the team letting them know the study had begun. The floater code was removed from the web page as soon as the team's goal of 300 completed studies was met.

Analysis of the Data

As soon as data were collected from the desired number of participants, Cynthia began the analysis of the data, documenting the key findings and recommendations in a report. Cynthia's final report would be in the form of an in-depth PowerPoint presentation that she would deliver to Anne and the rest of the team.

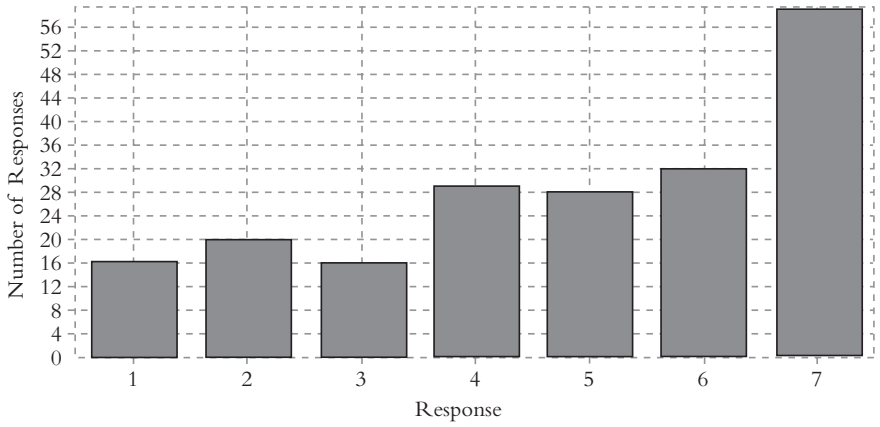
"Cynthia, now that we have all the data collected, what is the next step?" asked Tomas.

"To analyze the data, we use the remote online research application, WebEffective from Keynote Systems, Inc., which collects all the data from the participants," responded Cynthia. "The application makes all the responses to the questions available in percent and number of participants format. We will also be able to segment, or filter, the data by various criteria that are important to us, such as visit intent, success, satisfaction, and other key metrics." (See Figure 21.4.)

Question #92
How difficult or easy was it to complete this task?

Q #92 SHARE...

Mean = 4.8 Standard Deviation = 2.0



COPY

Response	#	%
1	16	8.0
2	20	10.0
3	16	8.0
4	29	14.5
5	28	14.0
6	32	16.0
7	59	29.5
Total number of respondents	200	100.0

Figure 21.4. Aggregated question responses inside the evaluation tool.

Cynthia continued, "The key aspects of this phase of the project include an analysis of

- Qualitative and quantitative feedback associated with the participant's own task;
- Behavioral data using click streams, which shows the navigational paths visitors took during their visit, the total time taken to complete the primary task, and number of pages visitors viewed during the task completion;
- Visitors' attitudes, preferences, and motivations by key visit intent segments, such as research product information, see what's new, and compare products; and
- Visitors' overall satisfaction, their ratings for site organization, and frustration while completing tasks on the website.

"As I analyze the data, I make recommendations for improving the customer experience based on usability best practices and visitor insights," Cynthia explained.

"Here's a specific analysis example. One of the goals of the study was to understand how successful and satisfied visitors are with their experiences on the website," Cynthia continued. "I want to know if success and satisfaction ratings differ based on the visit intent. Visit intent is the main reason why people come to the website and answers the question 'why did you come here today?'"

"How do we know the main reasons why people come to our site?" asked Jennifer.

Cynthia answered, "The third question in our study asked people about their primary purpose of their visit. Here's the exact question we asked." Why are people coming to the Website?

"When we look at the online report that aggregates all participants' responses, we can see that visitors come to the website primarily to research product information, see what is new, compare products, and look for pricing," continued Cynthia. "Since we cannot predict all of the reasons why visitors come to a website, we include a response called 'Other (Please specify).' This gives the participant an opportunity to tell us what they intend to do on the website if their visit intent is not listed. The 'Other' responses are stored as freeform responses, which we read and categorize. Sometimes the responses are variations of the list of tasks presented in the question. This response type is a great way to learn new reasons why people come to a

website as well. For example, in another study I conducted I learned that people came to the website to find the address of local stores. The website did not have this information and it was later added to the website."

"To find out if there is a difference between success, satisfaction, and brand impact ratings based on different reasons why people visit the website, I used the filter tool in the application to look at the data segmented by each of the four main reasons why people visit the site. I found that satisfaction rates varied slightly depending on visit intent. Success and brand impact had higher variances depending on why people are visiting the website. In other words, visitors were least successful when they were researching product information and consequently, their website visit had a lower impact on the 'brand' or their image of Aikot as a company. To measure the brand impact we asked 'How did your visit today change your opinion of Aikot in general?'" (See Figure 21.6.)

"What were the rating questions asked for these three metrics—success, satisfaction, and brand impact," asked Tomas.

Cynthia responded, "Here are the three 7-point rating questions we asked after participants completed their own task on the site. Remember the task they completed was self-directed—it was the task they originally came to the website to complete. Ratings 6 to 7 are the scores we use to determine success, satisfaction, and brand value." (See Sidebar.)

Visit Success

In thinking about your own expectations for the Aikot site how successful were you in achieving your goals for today's visit?

1 = not at all successful

2 3 4 5 6 7 = extremely successful

Satisfaction

Considering your experience today, how satisfied are you with the Aikot website?

1 = not at all satisfied

2 3 4 5 6 7 = extremely satisfied

Brand Impact

How did your visit today change your opinion of Aikot in general?

1 = Much worse opinion of Aikot

2

3

4 = No impact to my opinion of Aikot

5

6

7 = Much better opinion of Aikot

Cynthia continued with the analysis explanation, "Ninety participants' primary visit intent was to research product information. After filtering the data by 'research product information,' we see that 34% of the 90 participants rated their success as a 6 or 7 on a 7-point scale, 48% of the 90 participants rated their satisfaction as a 6 or 7 on a 7-point scale, and 21% of the 90 participants said they had a much better opinion of Aikot as a result of their visit. These scores are low compared with the 59 participants who came to the website to see what's new. The 'see what's new' group reported a success rate of 67%, satisfaction at 53%, and brand impact (influence of the website on the opinion of Aikot, the company) at 58%." (See Figure 21.5.)

"How can we learn more about what's going on for each of these groups of participants," asked Jennifer.

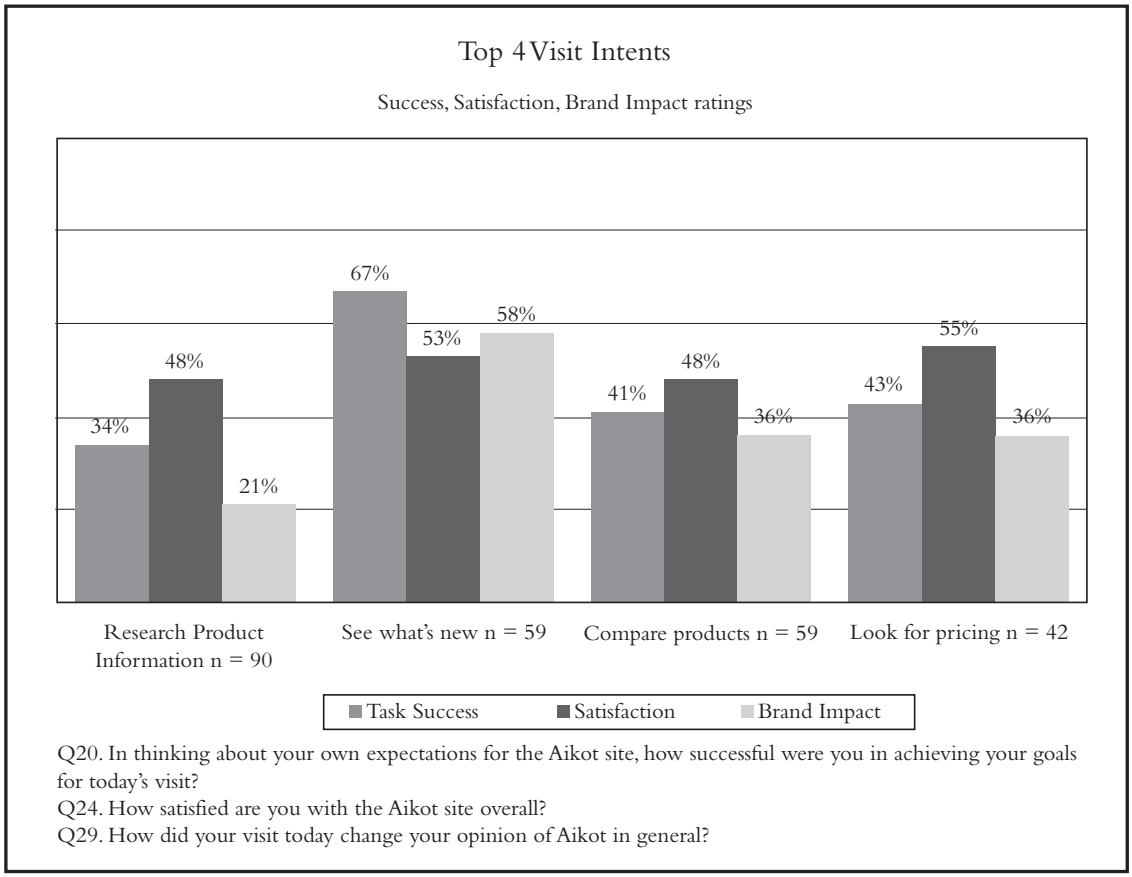


Figure 21.5. Top four visit intents: task success, satisfaction, and brand impact.

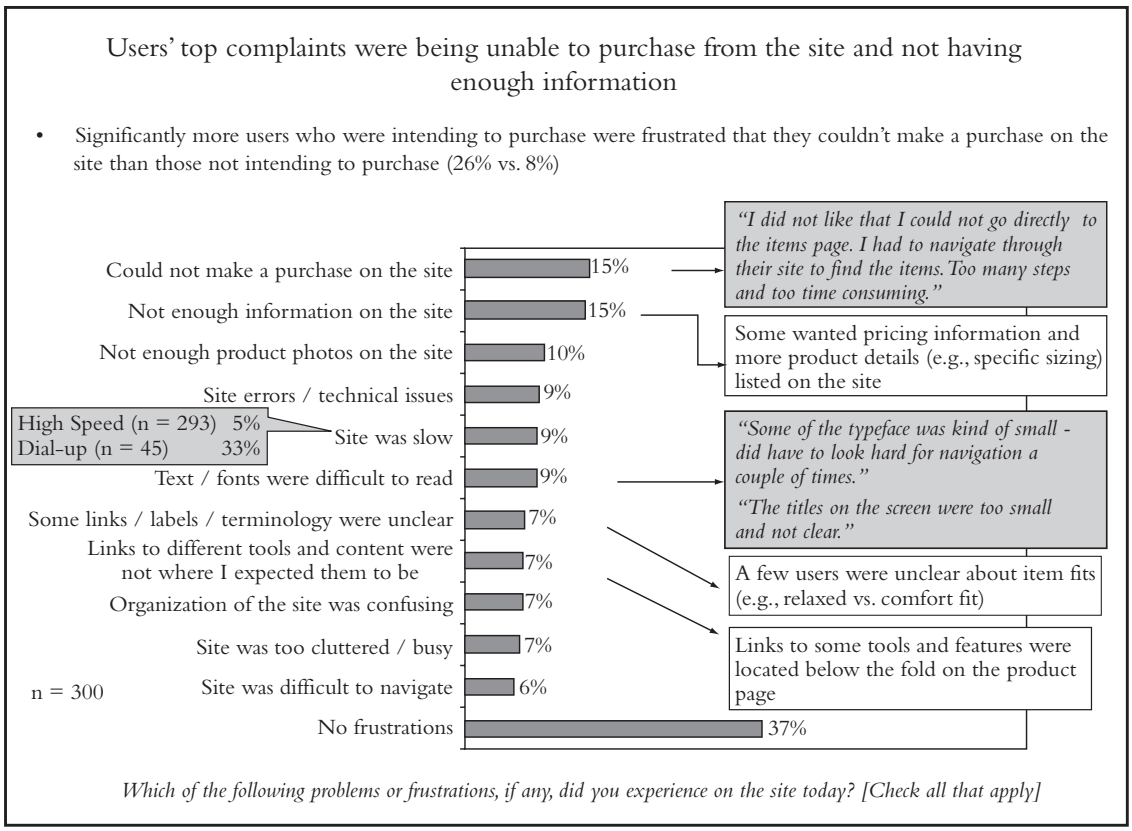


Figure 21.6. Main difficulties encountered while completing the task.

Cynthia responded, "We have to dig deeper into the data, especially the freeform responses from the open-ended questions, the difficulties encountered while completing their own task, and review the navigation paths. Here is a summary of the main difficulties visitors encountered." (See Figure 21.6.)

In the navigation paths, Cynthia was interested in visitors' back-button usage, first-page clicks, time spent on each page, and the average number of pages visited to complete a task. She used a small section of the clickstream navigation to summarize this information in the PowerPoint presentation (Figure 21.7). In Figure 21.7, the yellow rectangles represent web pages, the black lines indicate the navigation paths, the magenta lines represent back-button use, and the number on the lines and rectangles represent the number of people following this path and visiting a particular page.

"What can we learn from these data?" asked Tomas.

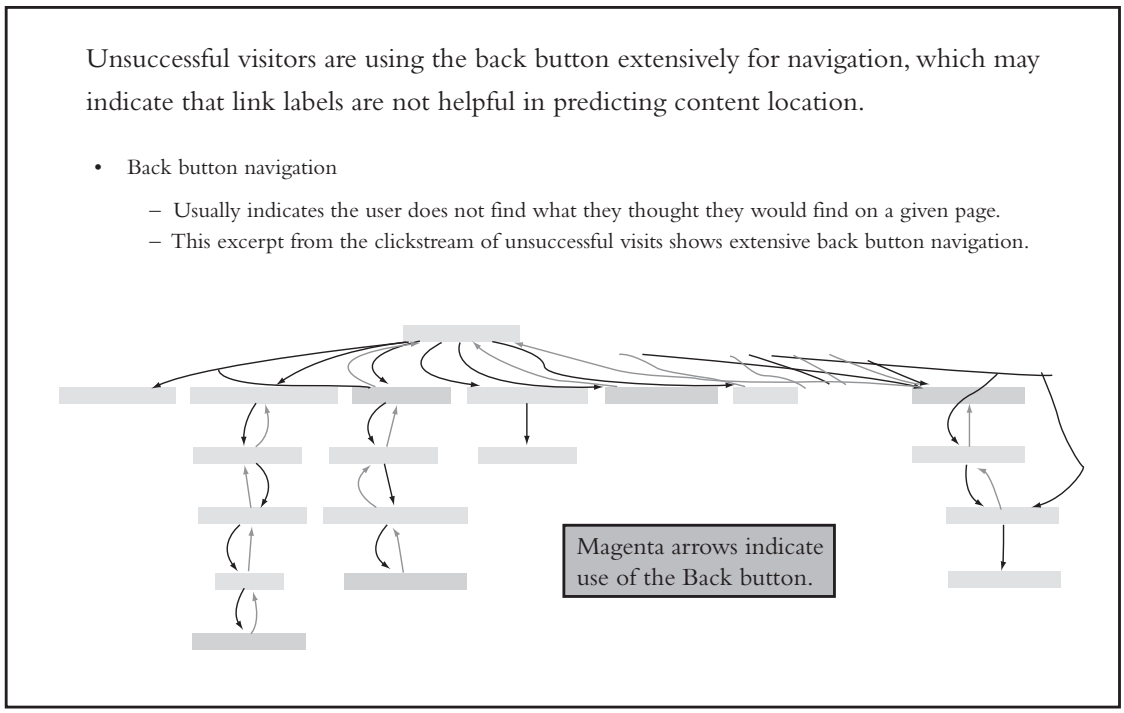


Figure 21.7. Click-stream navigation paths.

Cynthia responded, "We know that Aikot visitors want to be able to view all products within a particular category so they can see the similarities and differences in the products at a glance. They also want help in selecting the right products that meet their needs. Once they own the product, visitors expect help using the product and want to get involved with Aikot."

Jennifer wanted to know more, "What other key findings do you have?"

Cynthia continued, "We learned about who is visiting the website, their profile, reasons for visiting, and impact of the site on their image of the company. I summarized most of this information on one slide." (See Figure 21.8.) "Other slides in the report contain additional information. In summary, visitors had most difficulty with

- Particular products are not available,
- Too many clicks to find something,
- Pricing information not available,

USA Visitor Profile Summary
Aikot's USA site visitors can be described as:

- ***Age/gender/household***
  - **67%** of the site visitors are between the ages of **25–49 years**; **71%** are **male**; **28%** live in a **single household with no children**.
- ***Brand loyal***
  - **58%** currently own an Aikot product – **Aikot 3329** is one of the **most popular** phone models.
- ***New to the Aikot website***
  - **1 out of 4** are visiting the site for the **first time; slightly fewer than 1 out of 6** have visited the site **2 to 5 times**.
- ***Likely to shop online***
  - **1 out of 3** are interested in **buying a product online**.
- ***Site impact on Aikot brand***
  - **Almost 1 out of 3** visitors had a **positive** change in their general opinion of Aikot as a result of their visit.
- ***Primary reason for visit***
  - Visitors are coming to the site largely to **research product info (28%)**.
  - **3 out of 5** visitors are interested in products for **personal use**.
- ***Re-visits and referral***
  - **2 out of 5** are likely to **re-visit** the site and **less than one-third** are likely to recommend the site to others.

Figure 21.8. U.S. visitor profile summary.

- Poor search functionality, and
- Information about products was hard to find.

"In addition, visitors had the lowest satisfaction with

- Support area,
- Help choosing a product,
- Look and feel, and
- Finding product information.

"In interpreting the data from this study, we can say that most of the factors causing difficulty related to visitors not easily finding what they were looking for are caused by one or more of the following website design issues:

- Poor site architecture (how the site is organized),
- Lack of pertinent content,
- Poor search functionality, and
- The 'look and feel' of the website was also singled out as a dissatisfier."

"What should we do to improve the website," asked Tomas.

"We should think about rearchitecting the site. To begin this process we should perform a card sort exercise to determine how visitors perceive the site content should be organized, ensure that the main horizontal navigation bar is consistent across the site, and ensure that navigation choices are presented in a way that reflect the way the site is organized," Cynthia answered. "We should also use 'bread crumbs' to show the current path through the site hierarchy, create a site map to depict the site hierarchy, and provide easy-to-understand 'deep links' to popular areas."

Cynthia's complete presentation entailed 63 slides. She was able to develop a story about who is visiting the website, why they are coming to the site, and whether or not they are successful and satisfied—all the goals the team set out to accomplish with the online study. In her executive summary, Cynthia summarized the problems visitors had with the site (Figure 21.9).

The team was surprised at how much they learned about their website visitors. They now have a better understanding of why people come to the website, whether or not they are successful in completing their tasks, what works well on the site, what needs to be improved, and what needs to be added.

Executive Summary

- **Overall: visitors *love* the Aikot brand, and they *tolerate* the Aikot site.**
  - ➢ The site does not have a high influence on visitors because information is buried deep in the site or release dates are inaccurate, but they still come often because they're **eager to find out more about Aikot products**.
    - As a matter of fact, visitors who love the latest technology, come to the web site specifically to find out about products and they are very interested in new products and when they are available
  - ➢ The **main reason** behind visitors' lack of enthusiasm for the Aikot site is the site's poor organization which makes it difficult to find what they are looking for.
    - Visitors wish the site would be much more **structured and organized**, with **better search capabilities**.
    - They also wish that the information about dates for new product releases would be **more accurate** by providing new models when promised.
  - ➢ Visitors don't want to share their experience of the site with others and they are not inspired to purchase an Aikot product based on their site visits. However, they **do** want to revisit the site.
    - They come to the site to try to find information. **They love the Aikot brand regardless of their experience on the site**.

Figure 21.9. Executive summary.

The results of this study were a particular revelation for Mark, the chief marketing officer. He had no idea how influential the website is on consumers' purchase decision behavior. The role of the web is critical during the active decision-making stages in consumers' purchase decision behavior. With this in mind, he also realized that the Aikot websites have some serious problems, especially because fewer than 50% are successful in completing their primary visit task. The Aikot website is not positively impacting visitors' opinion of the brand and, in fact, may be one of the causes of the slow, yet steady, decline in market share.

Mark was grateful for all the work Anne and her team did to conduct this research, and he realized they need to evaluate their websites on an ongoing basis to understand the website's impact and success. As a result of this study, Mark and Anne spearheaded the following customer-centered design plan for the Aikot websites:

- Conduct semiannual exploratory studies where participants are instructed to complete self-directed tasks. Questions should include the following:
  - Who is visiting the site?
  - Why are they coming to the site?
  - Are they successful?
  - Are they looking for information that is not available on the site? If yes, what information are they looking for that they cannot find?
  - Are they satisfied with the website and with the company?
  - What are the trends over time?
- Conduct competitive comparison studies at least once a year to better understand best practices and other competitive information. This research helps answer the question, "How does Aikot stack up against its competitors?"
- Conduct quantitative/qualitative studies during the development cycles for targeted areas of the website. Evaluating prototype designs is more cost effective than evaluating a newly launched website that has no previous user feedback.

Summary

The web is becoming an increasingly important medium for consumer purchase decision-making processes, and customer visits to company websites are

steadily increasing. Now more than ever before, websites must provide excellent customer experiences or market share will decline.

Remote online testing can accomplish the following goals:

- Evaluates websites using large sample sizes, giving user experience teams more confidence in the results.
- Large sample sizes also allow the user experience teams to segment the data by criteria of interest to them.
- Participants complete the study when they want to, using their own computers in their own environments.
- During the data collection phase of the project, the user experience team can work on other projects because evaluators are not moderating the study sessions.
- Both self-reported data and behavioral data can be captured about a website experience.
- This methodology combines market research, usability research, and web analytics into one study.

The steps for conducting remote evaluations are as follows:

1. Design the research activities and questions.
2. Program the study using an online remote application such as Keynote Systems's WebEffective application.
3. Add code to invite participants to the website.
4. Collect data.
5. Analyze the data.
6. Report findings and recommendations.

Results from a remote online evaluation include the following:

- Qualitative and quantitative feedback associated with the participant's own task
- Behavioral data using clickstreams (this shows the navigational paths visitors took during their visit), the total time taken to complete the primary task, and number of pages visitors viewed during the task completion
- Visitors' attitudes, preferences, and motivations by key visitor segments, such as frequent visitors versus infrequent visitors
- Visitors' overall satisfaction, their ratings for site organization, and frustrations experienced while completing tasks on the website

Further Reading

Cooper, A. (1999). *The inmates are running the asylum*. Indianapolis, IN: Sams.

Dumas, J., and Redish, J. (1993). *A practical guide to usability testing*. Intellect Books.

Van Duyne, D., Landay, J., and Hong, J. (2002). *The design of sites: patterns, principles, and processes for crafting a customer-centered web experience*. Boston, San Francisco, New York: Addison-Wesley.

Weinstein, A. (2004). *Handbook of market segmentation: strategic targeting for business and technology firms*. Haworth Series in Segmented, Targeted, and Customized Market. New York: The Haworth Press, Inc.