Reprint of an article appearing in the
Society for Technical Communication's
Technical Communication, Volume 45,
Number 4, November 1998.

Measuring the Success of Visual Communication in User Interfaces

by Stephanie Rosenbaum and J.O. ("Joe") Bugental, Tec-Ed, Inc.

INTRODUCTION

Technical communicators must be visually literate in their role as information designers. Even when we have the luxury of working with professional graphic designers and artists as we produce information content, the more that technical communicators know about visual communication, the better will be the results of the collaboration.

However, most of the visual communication principles and theories in current use were originally developed for printed (hard-copy) information [Keyes 1993; Lee 1979; Ruder 1977; Smith and Ten Holt 1984; Tufte 1983 and 1990;Vignelli 1976; Wheildon 1995; White 1983 and 1984]. Professionals in many fields—including graphic design, multimedia production, user-interface design, and technical communication—are working to extend and modify print guidelines to address electronic information delivery.

Many books [Boggan 1995; Gallitz 1997; Horton 1994; Morris and Hinrichs 1996; Williams and Tollett 1998], articles, and papers address onscreen visual communication, both for user-support information (such as help systems, wizards, and online tutorials) and for product and web site user interfaces. Most of these publications describe their authors' individual experiences with graphical user interface (GUI) design and information design.

Following the experience of experts is a well-accepted way to learn new techniques, but it leaves a gap: we still need to validate the success of the results by collecting user data. Little current usability research focuses primarily on visual communication, but almost every usability study of GUIs or web sites produces some relevant results.

This article discusses three key areas of visual communication we address in user interfaces (UIs):

- Information access and navigation
- Icon recognition
- Visual appeal or "look-and-feel"

The article uses five case histories to demonstrate how usability research has helped the authors evaluate the quality of visual communication in navigation, icon recognition, and look-and-feel. It describes some of the research methodology the authors use, with examples from the case histories. For each of the three topic areas, we discuss the lessons we learned from the case histories about both usability testing methodology and visual communication guidelines. We mention, but do not concentrate on, related topics such as visual clutter.

The case histories are all from laboratory-based usability testing with representatives of the target audiences. Many other usability research methods collect valuable data and inform design insights—especially field methods such as contextual inquiry. However, because we are comparing and discussing the results of several studies in this article, we have restricted our examples to one method. The case histories still illustrate a wide variety of test designs.

This article assumes that the reader is familiar with the principles and methods of laboratory usability testing, either from working with human factors/usability specialists within your organization, or from books [Dumas and Redish 1993; Rubin 1994] and courses [Ramey 1997]. The article will be most useful to readers who either perform usability testing themselves or who want to apply the results of usability testing to improving user interfaces.

INFORMATION ACCESS AND NAVIGATION

Most current books on visual design for user interfaces recognize the importance of visual presentation to successful information access. Verplank [1988] says: "One of the most obvious contributions of good graphic design is to provide appropriate visual order and focus to the screen." Howlett [1996] says: "Functional design and visual design must work together. Like the buttons on a home appliance, controls in a user interface must look the way they work and work the way they look.

Even when visual communication isn't the primary focus of the UI design, its value is recognized. Hackos and Stevens [1997] say: "Graphics can more easily show the steps in a process or the sequence of a flow of information or product than can words."

Two kinds of usability testing yield data about the effects of visual presentation on navigation. Sometimes we perform task-completion usability tests of a single user interface. In other tests, we're asked to compare two or more user interfaces, and we have participants attempt the same tasks with each version.

Case History #1 (single UI)

In a recent usability test, system administrators used a new GUI tool to install and troubleshoot communications hardware and software on multiple systems. The GUI developers were trying to implement two jobs at once: while users installed a first system, they could simultaneously create a script on diskette for installing subsequent systems. Throughout the installation program, prompts and user selections on the right side of the screen completed the installation on the first system, and checkboxes in a small area under the illustration in the left-hand portion of each screen set parameters for subsequent installations (see Figure 1).

When participants first heard the task scenario, which talked about installing on a first computer and planning for additional computers, the test administrator asked whether that situation was typical for them. About half said, yes, it was highly typical.

However, as they performed the actual tasks, only one participant in eighteen understood how the script-creation process was working. Four others understood it after slight prompting, and for eleven, we stopped and explained it. Two never understood it at all. Some of the participants' comments revealed their frustration:

> "It needs a preamble screen explaining the process or a simple flowchart."
> "Interface ambiguous."
> "It should give you a detailed prompt as to options for creating the diskette."

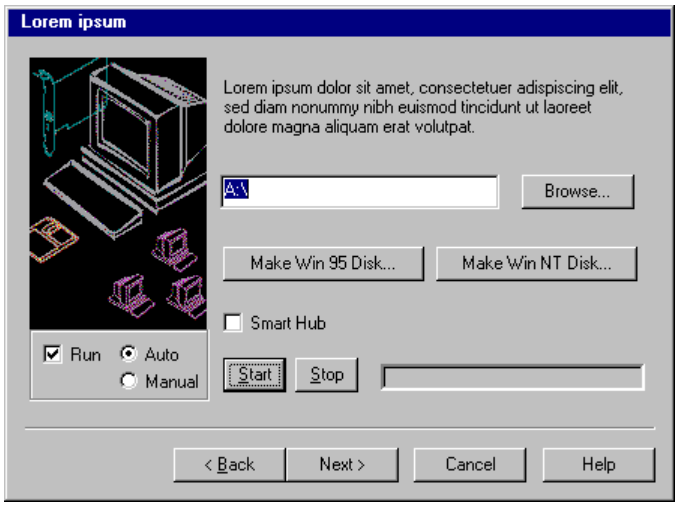

Figure 1. Sample screen from Case History #1 (the illustration is a composite of several actual screens).

Without prompting, all participants but one followed only the dialog on the right side of the screen. The visual organization of the screen did not communicate that two parallel processes were taking place, one controlled by the left-hand checkboxes and the other by the right-hand dialog.

To help users understand how the installation was proceeding, the GUI also included a simplified system diagram, highlighted on a black background. The installation software dealt with each of the system elements in turn, and its counterpart in the picture changed color to tell the user what the program was currently doing. The interface card turned light green, the network turned light purple, and a yellow diskette outline appeared when the user could create a script diskette. However, none of the usability test participants noticed the color changes. One even asked for "more different colors and indicators to show what is happening."

At the end of the test session, when we asked people to comment on the statement, "The graphics in the displays provide useful information," comments included:

> "No information provided graphically."
> "Just the same picture over and over again."
> "That graphic on left side of the screen was just corporate image stuff."
> "No—pictures told nothing."

As is often the case, this usability test identified several problems with the visual aids to navigation. The layout of information items on the screen didn't reflect the users' work processes, and the navigation clues in the illustration—which might have helped users realize what was happening—were too subtle to be noticed. For example, stronger highlighting (perhaps beginning with a second of blinking) or momentary animation on the current system element in the illustration probably would have called the participants' attention to the element being installed, but the subtle color changes did not.

The authors—working with the graphic designer on our team—recommended a revised UI that first completed the installation process for the user's computer, then presented a dialog to create an appropriate script diskette for other computers. However, if engineering management had required a simultaneous dual process, we would have recommended a screen layout clarifying the parallel procedures, perhaps with an illustration for each.

Case History #2 (comparing three UIs)

In this project, the authors performed comparative usability testing of three product home page designs: one in current use for a product line of a large company, and two newly created designs for another company product line.

We identified several usability problems that involved visual navigation. The software developers wanted to know if participants discerned any significance in different colors of text being used as an orienting device. For example, white text on a dark green background at the left of the display provided an overview and table of contents for the whole site, with highlighting to indicate the user's current location.

Gradually escalating the level of probing in her questions, the test administrator first asked "Where are you right now? How can you tell?" If participants didn't mention the orienting device, she asked, "What clues does the screen provide?" Finally, she drew their attention to the specific items with these questions: "What does this text with the green background tell you?" and "Why do you think it's green?" None of the twelve participants commented on the orienting device without prompting.

In another aspect of visual navigation, two of the three alternative designs contained both global (corporate) and local (product line) navigation button bars. Only four of the twelve participants clearly understood the distinction between the two kinds of button bars, even after exploring both designs. More people recognized that the relationship between the button bars was hierarchical, but not where the buttons led.

Also, one of the three designs contained thin black lines connecting buttons in the two navigational button bars to show a relationship between the two bars (see Figure 2). When a participant selected a button in the top (global) row, a red line appeared connecting that button to related buttons on the bottom (local) row. To emphasize the relationship, these bottom-row buttons also became highlighted.

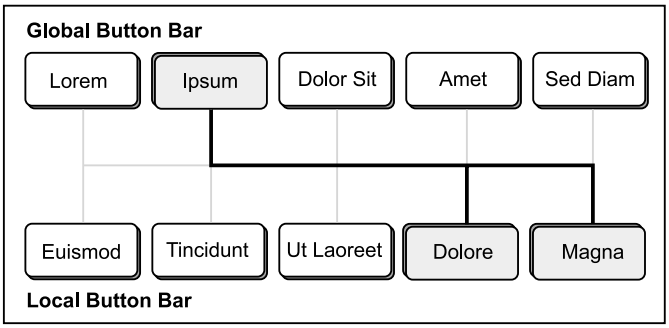

Figure 2. "Global" and "local" navigation button bars in Case History #2, connected by a red line reflecting the user's current selection in the global button bar.

Only five of the twelve participants discovered the red navigational line on their own, and the test administrator gradually increased her level of prompting (as with the green and white text) to draw the attention of the other seven to the feature. Once participants discovered the red line or were prompted to look at it, only six actively used it to orient themselves within the site.

Case History #3 (comparing two UIs)

This usability test compared two preliminary designs for a home page. The site developers had created three information "layers" in each design. One design presented the layers vertically stacked (with part of each layer showing); the other design offset them diagonally, with tabs sticking out from each layer (see Figure 3). In the stacked design, a selected layer expanded to take up more of the screen real estate. In the offset design, a selected layer moved to the top (see Figure 4).

The test participants were a mixture of business and home computer users; half were novice Web users, while the other half used the Web four to eight hours a week. During the test sessions, the participants accessed and manipulated the information layers to read their contents, then chose other information items from menus.

At the end of the task sequences with both UI designs, we asked the participants which UI design they preferred and why. Their reasons gave us insight into the visual navigation issues:

> "It's easier than the other to move around ... the tabs give a better idea of what to expect." [Offset design]
> "I can click on one thing and see just that." [Offset design]
> "I'm less likely to put something where I can't get to it again." [Stacked design]
> "Can see both [layers] at the same time." [Stacked design]

Note that participants had reasons why each of the UI designs was preferable for information access. The Case History #3 "reprise" later in this article discusses methodology for user preferences.

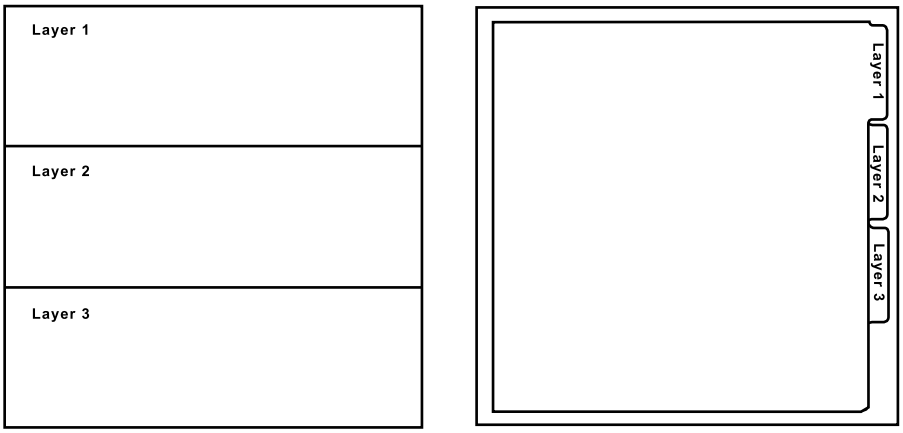

Figure 3. Two presentations of information layers in home page design from Case History #3, showing the layers as they first appear to the user.

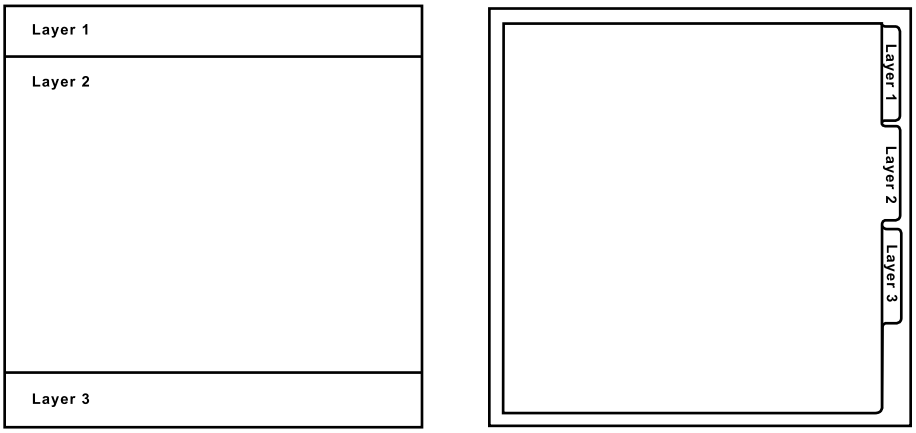

Figure 4. Information layers in Case History #3 after Layer 2 was selected.

As the test participants moved deeper into the site, both designs enabled them to choose product information from menus and submenus. When a menu item was selected, it turned blue, and the menu item became the title of the displayed product information page. We asked the participants why some submenu items were colored; only three of eight understood that blue indicated the displayed pages.

Usability methodology guidelines for information access

Not all usability practitioners are trained in how to prompt (or hint or remediate) when participants become lost or confused. Ideally, the test administrator describes a test task in the most general terms, and the participant proceeds independently. For example, you can give someone a pie chart and say, "Create something that looks like this. [Ramey 1997]" Wixon and Wilson [1997] call this approach a results-based test, in contrast with process-based tests where the user follows a prescribed set of tasks; but process-based tests should also not over-specify tasks.

But what should we do when a participant doesn't "get it" and can't figure out what to do next? What should we do when a participant chooses an incorrect path and thus doesn't encounter some of the UI elements we want to evaluate? While developing the details of the test script, the usability team should also decide and practice what to say when participants make serious errors or come to a standstill.

First, wait longer than feels comfortable; many participants won't help themselves until they're convinced you won't help them. Then offer the smallest suggestion from your script of prompts, and wait again. The key is to give the least amount of prompting that will keep the participant moving through the task—and to give consistent prompts to everyone, as illustrated in Case History #2.

A useful approach is for the usability team to record not only whether a participant needed prompting at any point, but also how many prompts were required before the participant "got it." The results report can then include tables or charts showing how many participants needed each level of prompt.

Although most usability tests include an interview component, the primary value of usability testing is to observe participants' behavior as they perform tasks. If we prompt too early or too often, the actual behavior is masked. In all the case histories in this article, for each subtask, we first observed behavior without intervening, prompted only as much as necessary, and only then asked participants' opinions.

Visual communication guidelines for information access

A key user question is always, "Where do I go next in the UI to accomplish my goals?" In Case History #1, there were not enough cues; in Case History #2, too many alternatives were offered at once without clear distinctions.

The participants' comments in Case History #3 illustrate other information access challenges in UI visual design. Users want the object of their immediate interest to be large and centrally located; yet they also want all their choices to be visible at all times. A successful UI design must achieve a balance between these user goals. Mullet and Sano [1995] say:

> "Organization and visual structure are the staples of successful communication-oriented design…Organization begins with classification, which involves *grouping* related elements and establishing a *hierarchy* of importance for elements and groups. When this hierarchy is clear, the display itself can be structured to reflect the *relationships* between the elements while maintaining a pleasing *balance* in the resulting composition."

They then list seven common errors involving organization and visual structure: haphazard layout, conflicting symmetries, ambiguous internal relationships, aligning labels but not controls, alignment within but not across controls, false structure, and excessive display density. The problems the participants encountered in Case Histories #1, #2, and #3 illustrate ambiguous internal relationships and false structure; however, other usability tests the authors have performed have shown problems caused by all of these errors.

Although written about magazine design, Jan White's comment [1982] about the purpose of layout applies well to user interfaces: "Layout is the specific, individually tailored solution to a particular story. It is the day-to-day manner in which the tools provided by the magazine's styling are put to use by the designer for journalistic purposes, to make the story clear, interesting, and memorable." Just as magazine layout must support the story, UI layout must support the user's tasks.

Interestingly, although most guidelines to color in UIs caution against overuse of strong colors [Marcus 1992; Schneiderman 1998], the navigation problems in our case histories were exacerbated by color cues that may have been too subtle. In Case History #1, participants didn't notice the color changes in the illustration. In Case History #2, they didn't notice either of two color navigation aids. And in Case History #3, they didn't recognize a color cue connecting a menu list with the related content.

This experience makes us point out the distinction between the use of color for style and corporate identity—where subtle colors are preferred—and its use to prompt specific user actions. When we want to ensure successful user behavior, UI designers shouldn't depend on subtle color changes to make critical distinctions for information access. Bright colors may be needed for navigation cues, especially if the object is small, and redundant cues—using motion or blinking in addition to color—are even better.

ICON RECOGNITION

It's not necessary for a successful UI that every icon be correctly recognized by its target users the first time they see it. Immediately recognizable icons are desirable, but very difficult to create for some applications. However, many products are used only occasionally, by a broad audience. Kiosk-like interfaces geared to successful first use are becoming more widespread. Because interfaces that are usable without training require recognizable icons, usability testing of icon recognition is also increasing.

Case History #4 (icon recognition, 40 participants)

In an extensive usability test of icon recognition, we tested 27 different icons for a new product, with three alternative designs for each icon, so participants saw 81 different icon designs. Using automated test administration and data collection through an interactive slide-show created by the software developers, we tested 40 participants in four sessions. Ten participants used individual workstations in each session, while the usability team walked among the workstations, observing participants' behavior.

Because users normally see icons in the context of an interface, we grouped the icons according to the six windows in which they would appear. We began each group of designs by showing participants the window and a one-sentence description of it. The slide-show then displayed in random order all the alternative icon designs associated with that window.

For each alternative icon design, participants first typed what they thought that icon would do (or let them do) if they clicked on it. Then they chose from a multiple-choice list of five possible meanings for that icon, including one that was the graphic designer's intent. The final choice in each list was Other (fill in), although that was never the intended choice.

The slide-show software prevented back-tracking, so participants couldn't change their free-form entries after seeing the multiple choices. The software also randomized the display order separately for each participant, both ensuring counterbalancing (that is, keeping participants' learning curves from affecting the results) and preventing participant cooperation in the group setting.

At the conclusion of the test sessions, the usability team rated the participants' free-form responses according to the following criteria:

- Right. Both the object and the action match the intended meaning, if not the exact words.
- Partly right. Either the object or the action matches the intended meaning, or, if the participant provided multiple answers, one of them is right.
- Wrong. Neither object nor action match, or the answer is completely different from the intended meaning.

There were 3220 opportunities to provide a free-form meaning for an icon; 33% of these answers were right, another 17% were partly right, and 47% were scored as wrong. ("No answer" texts accounted for 3% of the responses.) Participants screened as Intermediate and Advanced were right about 7% more often than were Novices.

When we looked at individual icon designs, the percentage of right answers varied from 6% to 90%. This variation was statistically significant at a level beyond one in a million: it is virtually certain that there were real differences in recognition of different icons. We also found 23 icon designs (out of the 81) where participant recognition exceeded 50%; we called these "high-recognition" icons.

Figure 5 shows the high-recognition icons, and Figure 6 shows the icons with the lowest recognition. For an icon to be included in either of the illustrated high-recognition groups, all answers (free-form and multiple-choice; by novice, intermediate, and advanced participants) had to meet the specified recognition level. Similarly, in Figure 6, all answers for these icons met the specified recognition level.

When we looked at the designs grouped by the icon meaning they represented, we found at least one of the three designs for each of 10 icon meanings was high-recognition. Sometimes all three alternative designs for an icon were high-recognition. For nine more icon meanings, the variations in recognition percentages were large enough for us to make a recommendation, even though none of the designs was classified as high-recognition.

Thus the usability test gave us quantitative data to decide on a higher-recognition icon for 19 of the 27 icon meanings. We were able to make recommendations about design directions for most of the remaining icon meanings by referring to the participants' free-form text responses.

Usability methodology guidelines for icon recognition

When beginning a usability test, we often take participants on a tour of the opening screen or other high-level screens. Unlike most tours, we ask questions instead of demonstrating. For example, we say, "What do you think will happen if you click on each of these buttons?" Case History #4 is based entirely on collecting people's expectations about icons.

This approach has several benefits. It's excellent for testing preliminary paper-and-pencil prototypes, when none of the interface elements is functional yet (and you don't have follow-on screen drawings for every button). Even for functional products, these questions enable the usability team to record participants' expectations before a sequence of rapid mouse-clicks renders them moot.

The tour questions can be repeated every time a new key screen appears, although once the participants begin interacting with a product, it's often tricky to slow them down enough to collect expectations. We often start by saying, "Without clicking on anything yet, tell me [for example] what you think will happen when you click on each of these menu items."

Visual communication guidelines for icon recognition

What guidelines help us create recognizable icons? Marcus [1992] says: "An icon is something that looks like what it means; it is representational and easy to understand. ... However, icons and symbols are not a panacea for electronic publishing or user-interface design; they can not completely replace words in some complex situations."

Horton [1994] says: “Every pixel contributes either message, enabler, or noise. The *message* is .. the concept you are trying to represent. The *enablers* help the user get the message reliably and accurately. *Noise* interferes with that process.” He describes 40 techniques for encoding meaning in icons, grouped in five categories:

- The subject directly
- A related or associated object
- Textual identifiers
- Spatial arrangement
- Geometric elements and properties

Although it was not an explicit part of the usability test in Case History #4, the test results seem to support the findings of Byrne [1993]: complex icons are less successful than simple ones. Byrne defines “simple icons” as “those discriminable based on a few features.” For the most part, the low-recognition icon designs in Figure 6 are more complex than the high-recognition icons in Figure 5.

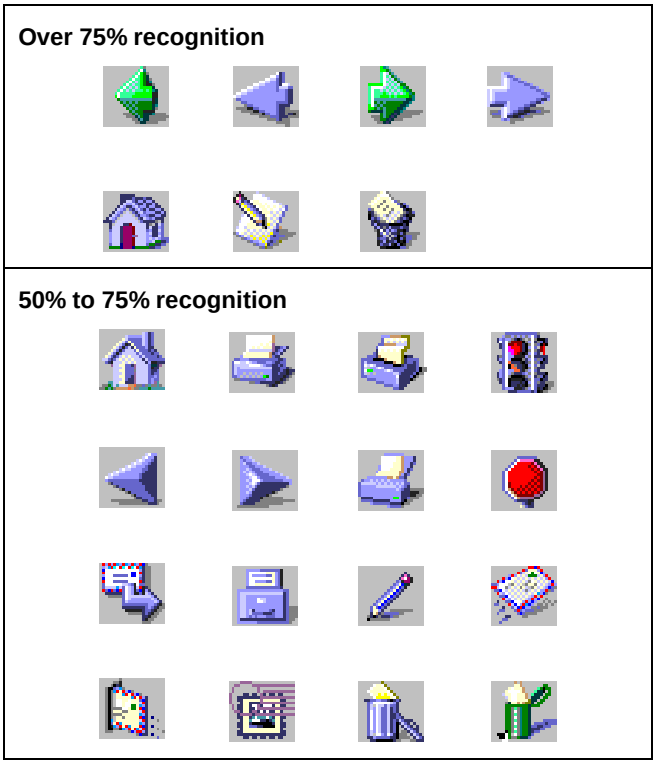

Figure 5. High-recognition icons from Case History #4.

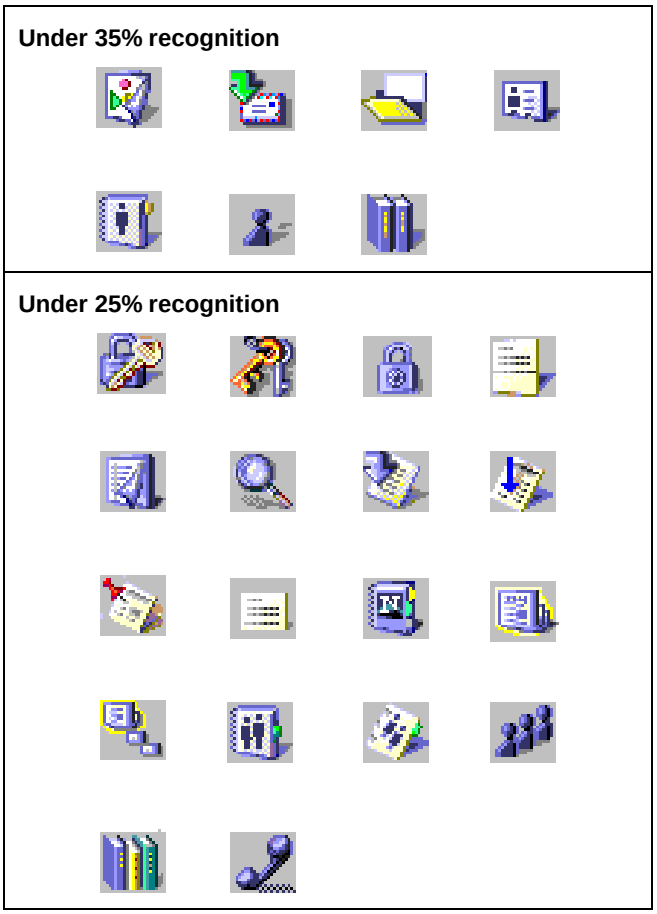

Figure 6. Low-recognition icons from Case History #4.

However, other issues besides visual complexity were also involved here. Some quite simple visual images received low recognition scores, which we judged were due to the complexity of the concept represented. For example, the first three icons with "Under 25% recognition" in Figure 6 all represented "Choose Encryption Method"; the last icon with "Under 25% recognition" represented "Call a Data Phone Number"; and the last icon with "Under 35% recognition" represented "Locate Address on Server."

If Case History #4 had been an academic thesis rather than a time-constrained applied research project in commercial product development, it would have been interesting to identify implementations of Horton's 40 techniques in the successful and unsuccessful icons. We hope that describing a methodology for testing recognition of icon designs will help other researchers to work with graphic designers to validate and prioritize specific techniques for design of icons.

VISUAL APPEAL ("LOOK AND FEEL")

Usability testing provides both performance and preference data about product use. We observe the kinds of problems participants have, and we hear what they think about the product. During usability testing of a UI, the authors are usually asked to learn the participants' preferences about its visual appeal or look and feel.

Case History #5 (testing product packaging)

In this usability test, we extended our view of the user interface to include the physical parts of a software product that users encountered on their way to the installation process: packaging box, license, CD containing the software, registration cards, and documentation. We tested 13 participants who represented three major characteristics: prior experience with the products, size of organization, and whether the participants were primarily developers or system administrators (used the product themselves or supported others' use).

The target users for this software product were highly technical people: engineers, programmers, and system administrators. We expected that the test participants might resist expressing preferences about visual appeal—and they did, making comments like:

"This is the first time I'm reacting to logos and graphics and all that stuff."
"I usually don't pay any attention [to packaging]; I just pop it open."
"It's hard to get that excited about packaging."
"I don't pay attention to that, I just use it."
"It's attractive. I'm an engineer, though; I don't care."

In anticipation of such attitudes, we developed our test administrator's script to include more probing questions than we would in a situation where we expected participants' behavior to communicate the desired information more easily. This approach was successful; we were able to collect comments relating specifically to the product's look and feel. For example, when we asked participants what drew their eyes to various parts of the packaging and how they would compare the new packaging to the former packaging, they said:

"New and fresh"
"More current"
"Much more colorful"
"Vibrant, pretty confident, primary colors, enthusiastic"
"A catchy package is always nice."

On the negative side, one participant said, "As a piece of artwork, I'm not that keen on it... It's not very restful; it kind of glares at me a bit." Another said, "They're spending too much time designing fancy graphics." Comparing the product's new look to previous versions, several current users made disparaging comments on the previous purple color scheme; one participant said, "Purple is a little girl's bicycle color."

Even when we asked specifically about the product's look and feel, we also heard about information access. For example, one participant wanted the company to change the accent color with each new product release, to help him keep track of versions.

Reprise of Case Histories #2 and #3 (comparing UIs)

In Case History #2, participants compared two prototype home pages with a current home page. Despite the participants' repeated navigational problems and criticisms that the usability team observed during the test task with the current home page, most participants said that information was easiest to find in that design. Its visual appeal overcame the performance issues; the participants' overall positive impression of that UI probably caused them to feel it offered better navigation aids as well.

In questionnaires accompanying the test, the usability team attempted to distinguish between home pages that were simply pleasant to look at and those whose appeal made the participants want to explore further. Participant responses, however, tended to remain consistently positive or negative in response to both kinds of questions; either they liked the page or they didn't.

In Case History #3, after the participants saw and used each home page design, we asked them to agree or disagree with the statement, "The visual design of this page was appealing to me." For the stacked design, six (of eight) participants agreed or were neutral. Their reasons why or why not included:

> "Very busy" [Neutral]
> "I like cool colors (blue, purple)." [Strongly agree]
> "It appeared logically laid out." [Agree]
> "I would probably strongly agree, but initially I was overwhelmed by the busyness of the page." [Agree]

For the offset design, five participants agreed. Their comments included:

> "Pleasing colors." [Strongly agree]
> "Nice looking, softer." [Agree]
> "Liked the look of the icons and how the different texts slid into view." [Agree]
> "Detail is very useful—the tabs." [Strongly agree]
> "Screen was too busy. Too much information/options." [Somewhat disagree]

We also asked the participants to name their most-liked and least-liked features for either design. One most-liked feature was the tabs on the offset design; and three participants mentioned "busyness" as a least-liked feature. The comments about busyness reflect another key area of visual communication—clutter—which we didn't specifically address in this article.

Usability methodology guidelines for visual appeal

In comparative usability testing, the authors not only observe participants' behavior with two or more UIs, we are also usually asked to find out which UI design the participants prefer. Almost invariably, we learn more from the participants' comments and behavior about the various interface elements than from their overall choice of UI. Most participants like (and dislike) elements from all the UIs being tested; these preferences—and the reasons for them—provide valuable input for design decisions.

Simply asking participants about their preferences isn't enough; questionnaire data and self-reporting in general have many limitations. Also, as described in Case History #5, there is a population of users—primarily technical people like engineers, programmers, and system administrators—who believe visual appeal is irrelevant to product quality.

However, everyone exposed to the mass communication and advertising in today's society is influenced by visual presentation quality in products, even if they're not consciously aware of this influence. Thus usability test protocols should be designed to elicit behavior and comments relating to product look and feel.

In Case History #5, the administrator's early questions focused on information access (see Figure 7). First, we expected resistance if we explicitly asked these engineers about the product's look and feel. Second, we wanted to collect unsolicited comments about visual appeal before asking specific questions, and certainly before asking for high-level preferences. While performing the task and responding to the questions, the participants did volunteer some comments about visual appeal.

In the hierarchy of opinion-collecting, we place the most credence in user "thinking aloud" and unsolicited comments made during task performance, as well as physical user behaviors: grimaces, shrugging, smiles, and other body language. When we move from behavior observations to direct interviewing, we have more confidence in the answers if we begin with small, specific questions. Only at the end of an interview do we ask general questions.

Task 1.4 Examining the License Pack and Media Kit (combined package), expected duration: 2-3 min.	
Topic: Finding the combined package	
Administrator Script	**Observer Notes**
[When they go to it:] What were you looking for when you picked this up? What is there about this piece that drew your eye to it?	Wanted this? Y / N If not, what? _____ Recognition by: Size? Shape? Materials? Color? Other: _____ Other issues: Quotes:
Topic: Expectations before opening the combined package	
Administrator Script	**Observer Notes**
What do you expect to find inside this piece of the package? What else? How can you tell what you've got here?	Media? Install manual? Other doc? _____ License? Registration? Marketing material? Other: _____ Version no. an issue? Y / N Other issues: Quotes:

Figure 7. A page from the administrator's script in Case History #5.

For example, in Case History #3, the last question we asked participants was which home page version they preferred overall. That way, we learned their opinions of the individual features of each design on a feature-by-feature basis. Had we asked their overall preference first, the participants would have tended to make their specific answers consistent with their high-level preference.

Similarly, in Case History #5, the last questions the administrator asked the participants were:

> "If you'd never heard of [company name] and someone handed you this package, what would it make you think about the company?"
> "What did you especially like about the packaging you examined today?"
> "What did you especially dislike about the packaging you examined today?"

Finally, unless user preferences are overwhelmingly for one design, which rarely occurs, a somewhat greater preference for one UI will not be statistically significant. For example, in Case History #3, five of eight participants preferred the offset design; one cannot reliably predict from this data that more users will prefer the offset design (and the authors had to emphasize this concern to a product development group eager to make decisions).

Visual communication guidelines for visual appeal

The lesson we learned most clearly when attempting to test the look and feel of a user interface is that it's difficult to assess visual appeal alone, independent of visual navigation and information access issues. Users themselves rarely separate usability and appeal. In Case History #3, when we asked the participants why (or why not) each home page design was visually appealing, many of the comments referred to information access characteristics: logical layout, display density, and tabs.

Therefore, it's important to make the effort to ask explicit questions probing look and feel, especially at the conclusion of test sessions. In addition to the Case History #5 questions listed above, other probing questions we've found helpful include:

- List three words that describe the appearance of this product [Wilson 1998]
- If this product were a person, what would he or she be like?

Sometimes, as we saw in Case History #2, there is a conflict between usability and visual appeal; participants preferred an interface they had difficulty using. Although this can happen even when all alternatives being tested are equally "polished," it's especially likely if one prototype is online and another paper-and-pencil, or if an existing product is being compared with a prototype.

Ideally, we should test equally polished alternatives, but resource and schedule constraints can interfere. If resources permit, one solution is to create sketches or prototypes of an existing product, for a valid comparison with proposed alternatives. Otherwise, bear any prototype unevenness carefully in mind when making UI design decisions based on usability testing.

CONCLUSIONS: USABILITY TESTING OF VISUAL COMMUNICATION

The authors are often asked about usability testing of documentation and how it can be combined with "rest-of-product" testing [Rosenbaum and Anschuetz 1996]. We recommend that during the product development cycle, some tests—or some tasks within tests—be designed to focus on documentation. Otherwise we miss collecting data from those users who don't turn to documentation without prompting.

We see some parallels in this experience with studying visual communication during usability research into user interfaces. However, the visual communication of a user interface is more closely coupled with product functionality than are even online documentation tools such as help systems and wizards. It's therefore harder to decouple testing of visual communication from rest-of-product testing.

Certainly, usability testing focused on visual issues is possible; Case History #4 is an example of such a test. Other test design possibilities include:

- Preparing screen layouts with "greeked" or unreadable content, so that users can focus entirely on visual appeal
- Preparing prototypes with information access tools but greeked content, so that users can focus on the visual navigation issues rather than the content
- Showing users examples of existing or imaginary product screens that are not for the product being designed, and asking which screen looks better and why
- Giving users a short tutorial on graphic design principles before conducting design walkthroughs or usability tests [Wilson 1998]

In the press of most software product development, it's unlikely that we'll have many opportunities to evaluate visual communication independent of other aspects of product usability. Rather, we'll need to continue to learn as much as we can about visual communication during general-purpose usability studies.

Therefore, it is critically important that we measure the success of visual communication in UIs. Our entire development teams—including our graphic designers—should be involved in studies of users' behavior. As we observe people accomplishing typical tasks, we can also probe to learn their opinions about visual appeal.

We should endeavor to perform both performance and preference studies on all user interfaces. Depending on studies of preferences only, distinct from task-based usability testing, is likely to be misleading. Yet the majority of decisions about visual communication in UI design—when user data is collected at all—are made by asking for user preferences.

The case histories and guidelines in this article may help others refine their techniques for collecting both behavior and opinion data from target users about user interfaces.

ACKNOWLEDGMENTS

Many people contributed to the experiences and information in this article. We particularly wish to thank Janice James and Tracy Halgren for their contributions to the design and implementation of Case History #4 at Netscape Communications. Other people who worked with us on these case histories included Michael Nico and Janice Rohn. Robert Farrell, Joan Hantz, and Chauncey Wilson read early drafts of the article and made valuable suggestions. Tec-Ed staff members who helped with the article include William Cotter, Natalie Grandinetti, Deborah Hinderer, John Nelson, April Pichlik, Larry Rusinsky, and Jacie Wedberg.

REFERENCES

Boggan, Scott. 1995. *Developing Online Help for Windows*. Carmel, IN: Sams Publishing.

Byrne, Michael D. 1993. "Searching: Tools and Strategies," in *Proceedings of INTERCHI '93 Conference on Human Factors in Computing Systems* (Reading, MA: Addison-Wesley Publishing Company), 446-453.

Dondis, Donis and Peter Dondis. 1973. *A Primer of Visual Literacy*. Cambridge: MIT Press.

Dumas, Joseph S. and Redish, Janice. 1993. *A Practical Guide to Usability Testing*. Norwood, NJ: Ablex Publishing Corp.

Gallitz, Wilbert O. 1997. *The Essential Guide to User Interface Design.* New York: John Wiley & Sons, Inc.

Hackos, JoAnn and Dawn Stevens. 1997. *Standards for Online Communication*. New York: John Wiley & Sons, Inc.

Horton, William. 1994. *Designing and Writing Online Documentation,* 2d ed. New York: John Wiley & Sons, Inc.

———. 1994. *The Icon Book: Visual Symbols for Computer Systems and Documentation*. New York: John Wiley & Sons, Inc.

Howlett, Virginia. 1996. *Visual Interface Design for Windows*. New York: John Wiley & Sons, Inc.

Keyes, Elizabeth. 1993. "Typography, Color, and Information Structure." *Journal of the Society of Technical Communication*, 40 No. 4 (November): 638-654.

Lee, Marshall. 1979. *Bookmaking*, 2d ed. New York: R.R. Bowker Company.

Marcus, Aaron. 1992. *Graphic Design for Electronic Documents and User Interfaces*. New York: ACM Press.

Morris, Mary and Randy Hinrichs. 1996. *Web Page Design: A Different Multimedia*. Mountain View: Sunsoft Press.

Mullet, Kevin and Darrell Sano. 1995. *Designing Visual Interfaces: Communication Oriented Techniques*. Mountain View: SunSoft Press.

Ramey, Judith. 1998. *Usability Testing of Product User Interfaces and Documentation*. Course sponsored by Tec-Ed, Inc.—Ann Arbor, Michigan. Sunnyvale, CA (August).

Rosenbaum, Stephanie and Lori Anschuetz. 1996. "Whole-Product Usability: Integrating Documentation and Rest-of-Product Usability Testing," in *Proceedings of the Silicon Valley Ergonomics Conference & Exposition, ErgoCon'96* (San Jose: Silicon Valley Ergonomics Institute), 17-21.

Rubin, Jeffrey. 1994. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. New York: John Wiley & Sons, Inc.

Ruder, Emil. 1977. *Typography: A Manual of Design*. Niederteufen, Switzerland: Verlag Arthur Niggli.

Shneiderman, Ben. 1998. *Designing the User Interface,* 3d ed. Menlo Park: Addison-Wesley Longman, Inc.

Smith, Stan and Professor H.F.Ten Holt. 1984. *The Designer's Handbook*. New York: Gallery Books.

Travis, David. 1991. *Effective Color Displays: Theory and Practice*. New York: Academic Press.

Tufte, Edward. 1990. *Envisioning Information*. Cheshire, CT: Graphics Press.

———. 1983. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Verplank, William L. 1988. "Graphic Challenges in Designing Object-oriented User Interfaces." Chap. 17 in *Handbook of Human-Computer Interaction*. North Holland: Elsevier Science Publishers B.V.

Vignelli, Massimo. 1976. *Grids: Their Meaning and Use for Federal Designers*, No. 036-000-00038-4. Washington: National Endowment for the Arts.

Wheildon, Colin. 1995. *Type & Layout*. Berkeley: Strathmoor Press.

White, Jan V. 1982. *Editing by Design*, 2d ed. New York: R.R. Bowker Co.

———. 1983. *Mastering Graphics*. New York: R.R. Bowker Co.

———. 1984. *Using Charts and Graphs*. New York: R.R. Bowker Co.

Williams, Robin and John Tollett. 1998. *The Non-Designer's Web Book*. Berkeley: Peachpit Press.

Wilson, Chauncey. 1998. Telephone conversation with author.

Wixon, Dennis and Chauncey Wilson. 1997. "The Usability Engineering Framework for Product Design and Evaluation." Chap. 27 in *Handbook of Human-Computer Interaction*, 2d ed. North Holland: Elsevier Science, Ltd.