UNIT 3

## Provide Effective Feedback and

Guidance and Assistance

All user actions must be reacted to in some way. Feedback, as has been noted, shapes human performance. Without it, we cannot learn. To aid user learning and avoid frus- tration, it is also important to provide thorough and timely guidance and assistance. In this step, the following feedback topics are addressed:

* Acceptable response times
* Dealing with time delays.
* Blinking for attention
* The use of sound

This will be followed by a review of guidance and assistance, including

* Preventing errors and problem management
* The types of guidance and assistance to provide
* Instructions or prompting
* A Help facility
* Contextual Help.
* Task-oriented Help
* Reference Help
* Wizards.
* Hints or tips.

**Response Time**

**The optimum response time is dependent upon the task.** There is an optimum work pace that depends on the task being performed. Longer or shorter response times than the optimum lead to more errors. In general, response times should be geared to the user’s short-term memory load and to the way he or she has grouped the activities being performed. Intense short-term memory loads necessitate short response times. While completing chunks of work at task closures, users can with- stand longer response delays.

The human *now*, or psychological present, is under 2 to 3 seconds. This is why continuity of thinking requires a response time within this limit. Research indi- cates that for many creative tasks, response times under one second, in the range of four-tenths to nine-tenths of a second, can yield dramatic increases in produc- tivity, even greater in proportion to the decrease in response time. The probable reason is the elimination of restrictions caused by short-term memory limitations. For data entry tasks, research has found no advantages for having response times less than 1 second

If human task closures exist, high levels of concentration are not necessary, and moderate short-term memory requirements are imposed; response times of two to four seconds are acceptable. If major task closures exist, minimal short-term mem- ory requirements are imposed, and responses within 4 to 15 seconds are accept- able. As the response-time interval increases beyond 10 to 15 seconds, however, continuity of thought becomes increasingly difficult to maintain. It has been sug- gested that this happens because the sequence of actions stored in short-term memory becomes badly disrupted and must be reloaded.

The response time guidelines above, then, relate to the general tasks being per- formed. Their applicability to every situation can never be guaranteed.

**Satisfaction with response time is a function of expectations.** Expectations are based, in part, on past experiences. These experiences may be derived from work- ing with a computer, from the world in general, or from the perceived complex- ity of the task the computer is performing. These expectations vary enormously across individuals and tasks.

**Dissatisfaction with response time is a function of one’s uncertainty about delay.** The degree of frustration with delay may depend on such psychological factors as a person’s uncertainty concerning how long the delay will be, the extent to which the actual delay contradicts those expectations, and what the person thinks is causing the delay. Such uncertainty concerning how long a wait there will be for a computer’s response may in some cases be a greater source of annoyance and frustration than the delay itself.

**People will change work habits to conform to response time.** As response time in- creases, so does think time. People also work more carefully with longer response times because the time penalty for each error made increases. In some cases, more errors have been found with very short response times. This may not necessarily be bad if the errors are the result of trial-and-error learning that is enhanced by very fast response times.

**Constant delays are preferable to variable delays.** It is the variability of delays, not their length, that most frequently distresses people. From a consistency stand- point, a good rule of thumb is that response-time deviations should never exceed half the mean response time. For example, if the mean response time is 4 seconds, a 2-second deviation is permissible. Variations should range from 3 to 5 seconds. Variation should never exceed 20 percent, however, because lower response time variability has been found to yield better performance, but small variations may be tolerated.

**More experienced people prefer shorter response times.** People work faster as they gain experience, a fact that leads Shneiderman (1987) to conclude that it may be useful to let people set their own pace of interaction. He also suggests that in the absence of cost or technical feasibility constraints, people will eventually force re- sponse time to well under 1 second. In general, the longer people interact with a system, the less delay they will tolerate.

**Very fast or slow response times can lead to symptoms of stress.** There is a point at which a person can be overwhelmed by information presented more quickly than it can be comprehended. There is also some evidence indicating that when a

system responds too quickly, there is subconscious pressure on users to respond quickly also, possibly threatening their overall comfort, increasing their blood pressure, or causing them to exhibit other signs of anxious behavior. Symptoms of job burnout have been reported after substantial reductions in response time.

Slow and variable response times have also been shown to lead to a significant build-up of mood disturbances and somatic discomfort over time, culminating in symptoms of work stress, including frustration, impatience, and irritation.

**Web response time.** Bouch, Kuchinsky, and Bhatti (2000) performed three related studies. Web pages were presented at preestablished delays ranging from 2 to 73 seconds. Delay was defined as the time between a page’s request and its complete displaying on the screen. Test participants rated the response times with the fol- lowing results:

*High (Good):* Up to 5 seconds. *Average:* From 6 to 10 seconds. *Low (Poor):* Over 10 seconds.

In the second study, test participants were presented a button labeled “Increase Quality” and asked to press it when response time became unacceptable. The av- erage button-pressing time was 8.6 seconds. In the third study, the Web pages were loaded incrementally, the banner first, the text second, and the graphics last. The test participants were much more tolerant of delays under these conditions. Response times up to 39 seconds were rated as “good,“ and response times over 56 seconds were rated as “poor.”

Selvidge, Chaparro, and Bender (2000) studied the effect of downloading de- lays on user performance. For delays of 1, 30, and 60 seconds they examined task success, task efficiency, and participant frustration. They found that participants were less frustrated with the 1-second delay, but task success and efficiency were not affected by any of the response times. Ramsay, Barbesi, and Preece (1998) evaluated people’s reactions to a Web page’s quality in relation to downloading times ranging between 2 and 120 seconds. Pages with longer delays, over 41 sec- onds, were rated as less interesting and more difficult to scan.

In general, these studies seem to indicate that the same factors affecting opti- mum computer response times in general also operate in the world of the Web, al- though longer downloading times may be more readily accepted due to the graphical nature of the Web’s content. Slower response times, from a practical standpoint, however, would appear to reduce the amount of work that can be per- formed, and probably lead to increased user frustration.

##### Dealing with Time Delays

* Button click acknowledgement:
  + Acknowledge all button clicks by visual or aural feedback within one-tenth of a second.
    - Waits up to 10 seconds:
      * If an operation takes 10 seconds or less to complete, present a “busy” signal until the operation is complete.
        + Display, for example, an animated hourglass pointer.
    - Waits of 10 seconds to 1 minute:
      * If an operation takes longer than 10 seconds to complete, display:
        + A rolling barber’s pole or other large animated object.
        + Additionally, a progress indicator, percent complete message, or elapsed time message.
    - Waits over 1 minute:
      * Present an estimate of the length of the wait.
      * Display a progress indicator, percent complete message, or elapsed time message.
      * When a long operation is completed, present an acknowledgment that it is completed.
        + A significantly changed screen appearance.
        + An auditory tone.
      * If an operation is very time-consuming:
        + Consider breaking the operation into subtasks and providing progress indica- tors for each subtask.
        + Allow users to start a new activity while waiting.
    - Long, invisible operations:
      * When an operation not visible to the user is completed, present an acknowledg- ment that it is completed.
        + A message.
        + An auditory tone.
    - Progress indicator:
      * A long rectangular bar that is initially empty but filled as the operation proceeds.
        + Dynamically fill the bar.
        + Fill it with a color or shade of gray.
        + Fill it from left to right or bottom to top.
    - Percent complete message:
      * A message that indicates the percent of the operation that is complete.
      * Useful if a progress indicator takes too long to update.
    - Elapsed time message:
      * A message that shows the amount of elapsed time that the operation is consuming.
      * Useful if:
        + The length of the operation is not known in advance.
        + A particular part of the operation will take an unusually long time to complete.
    - Web page downloads:
      * For pages requiring download times greater that 5 seconds, give the user some- thing to do while waiting.
        + Quickly present, at the top of the downloading page, some text or links.

Elapsing time is in the eye of the beholder. What is important is *perceived* passing time, not actual time as measured by a clock. Dealing with time delays involves pro- viding feedback that the system is truly working, and that the system’s processing will be completed at some foreseeable and predictable time. Dealing with time delays also involves diverting people’s attention away from a delay by engaging them in some meaningful interim activities.

**Button clicks.** Acknowledge all button clicks by visual or aural feedback within one- tenth of a second. This assures the user that a requested action has been received by the system.

**Waits up to 10 seconds.** If an operation takes 10 seconds or less to complete, present a “busy” signal until the operation is finished. An hourglass pointer is the custom- ary signal. A “Please wait...” message can be presented to indicate that more com- plex processing is occurring or has been delayed. When the process is finished, provide an indication that the user may precede.

**Waits of 10 seconds to 1 minute.** If an operation takes longer than 10 seconds to com- plete, display a rolling barber’s pole or other large animated object. Additionally, present a progress indicator, percent complete message, or elapsed time message. When the process is finished, provide an indication that the user may precede.

**Waits over 1 minute.** For waits exceeding one minute, present an estimate of the length of the wait. If the length is unknown, an educated guess is better than a “Don’t Know” or no estimate at all. A time estimate allows the user to decide what to do next—wait, go to lunch, or start some other task. For these waits, display a progress indicator, percent complete message, or elapsed time message. If an op- eration is very time-consuming, consider breaking the operation into subtasks and providing progress indicators for each subtask. Also allow users to start a new activity while waiting so a delay will not be unproductive. Also, consider offering engaging text messages to keep users informed and entertained while they are waiting for process completion. Provide a clear indication of when the process is finished, significantly changing the screen’s appearance so the change may be rec- ognized from some distance away. Also include an auditory tone to attract the user’s attention back to the screen.

**Long, invisible operations.** When a long operation not visible to the user is com- pleted, present an acknowledgment message that it is completed. For example, upon completion of a search with no positive results, “Search complete, Jones not found” might be displayed. Also provide an auditory signal, since the user’s at- tention may be directed to another part of the screen, or entirely away from the screen.

**Progress indicator.** A *progress indicator* is a long rectangular bar that is initially empty but filled as the operation proceeds. Dynamically fill the bar with a color or shade of gray. Fill all bars from left to right or bottom to top. A progress indi- cator is illustrated in Figure 9.1.

**Percent complete message.** A *percent complete* message provides an indication of the percent of an operation that is complete. It is useful if a progress indicator takes too long to update. An indication of the percentage of processing that has been ac-



**Figure 9.1** Processing progress indicator.

complished can also be given through a message such as “22 of 27 transactions have been processed.”

**Elapsed time message.** An *elapsed time* message shows the amount of elapsed time the operation has consumed. It is useful if the length of the operation is not known in advance, or if a particular part of the operation will take an unusually long time to complete.

**Web page downloads.** For pages requiring download times greater that 5 seconds, give the user something to do while waiting. Quickly provide at the downloading page top some text to hold one’s interest, or some links to act upon. These diver- sions will reduce impatience while images load.

##### Blinking for Attention

* + - Attract attention by flashing an indicator when an application is inactive but must display a message to the user.
      * If a window, flash the title bar.
      * If minimized, flash its icon.
    - To provide an additional message indication, also provide an auditory signal (one or two beeps).
      * Very useful if:
        + The window or icon is hidden.
        + The user’s attention is frequently directed away from the screen.
    - Display the message:
      * When the application is activated.
      * When requested by the user.

**Attention.** Flashing an element on the screen will usually capture a person’s attention. If a window is displayed on the screen, flash its title bar. If the window is mini- mized, flash its icon.

**Auditory signal.** To provide an additional indication that a message is waiting, also provide an auditory signal (one or two beeps). This will be useful if the window or icon is hidden or the user’s attention is frequently directed away from the screen.

**Message display.** Display the message when an application is activated or when the user requests it. Displaying it when the user requests it preserves the user’s con- trol over the work environment and ensures that the message is not accidentally closed through an inadvertent key press. Finally, blinking is annoying to many people, so it should not be overused on a screen.

##### Use of Sound

* Always use in conjunction with a visual indication.
* Use no more than six different tones.
  + Ensure that people can discriminate among them.
* Do not use:
  + Jingles or tunes.
  + Loud signals.
* Use tones consistently.
  + Provide unique but similar tones for similar situations.
* Provide signal frequencies between 500 and 1,000 Hz.
* Allow the user to adjust the volume or turn the sound off altogether.
* Test the sounds with users over extended trial periods.
* Use sounds sparingly because they:
  + Are annoying to many people, including other users and nonusers in the vicinity.
  + Can easily be overused, increasing the possibility that they will be ignored.
  + Are not reliable because:
* Some people are hard of hearing.
* If they are not heard, they leave no permanent record of having occurred.
* The user can turn them off.

Sounds, sometimes called *earcons,* are useful for alerting the user: To minor and obvious mistakes.

When something unexpected happens.

Where visual attention is directed away from the screen and immediate attention is required.

When a long process is finished.

Tones used must be discriminable, nonannoying, and consistently used. Therefore, they must be thoroughly tested for discrimination and effectiveness. Brewster, Wright, and Edwards (1993) have found that high levels of recognition can be achieved by care- ful use of earcon pitch, rhythm, and timbre. They recommend:

Timbre: Use synthesized musical instrument timbres, where possible with mul- tiple harmonics.

Pitch: Do not use on its own unless there are very big differences between the pitches used.

Register: Acceptable to use alone to differentiate earcons; otherwise large differ- ences of three or more octaves should be used.

Rhythm: Make rhythms as different as possible. Including a different number of notes in each rhythm is very effective.

Intensity: 10dB to 20dB above threshold. Since sound level will be under control of the user, it should be kept in a close range.

Combinations: Leave a delay of 0.01 second between successively played earcons.

Since sounds can be annoying to some people, they should be capable of being turned down or off by the user. Playing jingles or tunes, or loud sounds, focuses atten- tion on the sound itself, which is distracting. Loud sounds can also be irritating, espe- cially to those with sensitive hearing.

Never consider sounds reliable; because they can be turned off, they leave no per- manent record of their existence, and not all users will be able to hear all tones because of hearing defects. Sounds should always be used in conjunction with a visual indica- tion of some kind.

Guidance and Assistance

We’ll begin by first looking at ways to prevent errors from occurring in the first place. Then, we will look at how to gracefully handle them when they do occur. Finally, we’ll look at guidance and assistance and its components, including the various forms of Help typically available in a system.

##### Preventing Errors

In spite of our lofty design goals, people will make errors using even the most well- designed system. When errors occur they must be properly managed. The magnitude of errors in computer systems is astounding. Studies have found error rates in com- mands, tasks, or transactions as high as 46 percent. In addition to stranding the user and wasting time, errors interrupt planning and cause deep frustration.

Errors can be classified as slips or mistakes. A *slip* is automatic behavior gone awry. One’s hands navigate the keyboard improperly and the wrong key is accidentally pressed. The wrong menu bar item is chosen because of inattentiveness. An inference error is made due to carelessness. A person often detects a slip because it is usually no- ticeable. The wrong letter appears within a word or the expected action is not performed. Slips are usually, but not always, corrected fairly easily. Slips can be reduced through proper application of human factors in design (for example, by providing adequate separation between elements to be selected).

A *mistake* results from forming a wrong model or goal and then acting on it. A mis- take may not be easily detected because the action may be proper for the perceived goal—it is the goal that is wrong. Anticipating a mistake in design is also more difficult. Mistakes can be reduced, however, by eliminating ambiguity from design. Doing us- ability testing and watching for nonsensical (to the designer) requests and actions can also detect mistake-conducive situations.

Some experts have argued that there is no such thing as an “error” in using a system; they are simply iterations toward a goal. There is much truth to that statement. It is also said “to err is human.” The corollary to that statement, at least in computer systems, might be, “. . . to forgive is good design.” Whatever we call them, errors will occur. Peo- ple should be able to correct them as soon as they are detected, as simply and easily as they were made.

MAXIM Everyone makes mistakes, so every mistake should be fixable.

##### Problem Management

* Prevention:
  + Disable inapplicable choices.
  + Use selection instead of entry controls.
  + Use aided entry.
  + Accept common misspellings, whenever possible.
  + Before an action is performed:
* Permit it to be reviewed.
* Permit it to be changed or undone.
  + Provide a common action mechanism.
  + Force confirmation of destructive actions.
* Let expert users disable this.
  + Provide an automatic and continuous Save function.
* Detection:
  + For conversational dialogs, validate entries as close to point of entry as possible.
* At character level.
* At control level.
* When the transaction is completed or the window closed.
  + For high speed, head-down data entry.
* When the transaction is completed or the window closed.
  + Leave window open.
  + Maintain the item in error on the screen.
  + Visually highlight the item in error.
  + Display an error message in a window.
* Do not obscure item in error.
  + Handle errors as gracefully as possible.
* The greater the error, the more dramatic should be the warning.
  + Use auditory signals conservatively.
* Correction:
  + Preserve as much of the user’s work as possible.
  + At window-level validation, use a modeless dialog box to display an error list.
* Highlight first error in the list.
* Place cursor at first control with error.
* Permit fixing one error and continuing to next error.
  + Always give a person something to do when an error occurs.
* Something to enter/save/reverse.
* A Help button.
* Someone to call.
* Provide a constructive correction message saying:
  + What problem was detected.
  + Which items are in error.
  + What corrective action is necessary.
* Initiate a clarification dialog if necessary.

Preventing, detecting, and correcting errors involves doing the following.

**Prevention.** It is always better to prevent errors than handle them. Errors can be re- duced in a number of ways. First, disable all choices that are not applicable at any one moment. Make improper alternatives impossible to select or activate. Next, after considering the task and user, if practical, design screens using selection in- stead of entry controls. Selection error rates tend to be lower than entry error rates. Use aided entry if at all possible. The computer has been found to be a better speller than most people. When entry is performed, human misspellings of com- mands and requests should be accepted by the system. Person-to-person commu- nication does not require perfection. Person-to-computer communication should impose no more rigor than that imposed between people. Use of the Shift key should also be discouraged, since it is such a large cause of keying errors. Actions about to be performed should also be reviewable and changeable. Keying or se- lection slips will occasionally occur.

A common send mechanism should be provided to transmit an action to the system. The existence of two or more keys to accomplish the same purpose, espe- cially if their use is mandated by different conditions, can be confusing and more prone to error. Finally, if an action causes an irreversible change and the change is critical, the user should be requested to confirm the change. A separate key should be used for this purpose, not the Send key. Let expert users disable the confirmation request, if the action is recoverable. Finally, provide an automatic continuous *Save* function so that users never lose their work due to a system (or user) malfunction.

**Detection.** Errors should be detected quickly, but without disrupting a person’s thought patterns and actions if this can be avoided. Generally, the longer the wait before editing is performed, the longer the time to accomplish the editing. So, val- idate according to how important accuracy is to the user, and the characteristics of the task being performed. It is also preferable to wait for a closure point in the dialog. For conversational dialogs, accuracy is usually more important than speed, actions are slower paced, and more closure points usually exist. In these situa- tions, validate as close to the point of input as possible: at the character level or at the control level, and when the transaction is completed or the window closed. High-speed, head-down data entry is generally fast paced. Constant interruptions for errors can be a great speed detriment. In this situation, validate when the trans- action is completed or the window closed. This is usually the only task closure point.

All errors should be maintained on the screen and identified to the user through a highlighting display technique (for example, high intensity or contrasting color). Display an error message in a dialog box and position it on the screen so it does not obscure the item in error. Handle all errors as gracefully as possible, avoiding discouraging, embarrassing, or alarming words. Words with such intent can com- pound the frustration a person already feels in having made an error. For minor problems, provide less intrusive warnings. The greater the error, the more dra- matic may be the warning.

Be cautious in using auditory signals to notify of an error. Many people, espe- cially those with status or position, do not want their inefficiencies advertised, es- pecially to peers and subordinates.

**Correction.** Preserve as much of the user’s work as possible. It can be irritating to have to reenter an entire screen when only one field is in error. At the window level of validation, use a modeless dialog box to display a list of errors. Highlight the first error in the list and place the cursor at the first control with error. Permit fixing one error, and then continuing on to the next error. If multiple errors occur, and it is impossible to display messages about all of them at one time, provide an in- dication that there are additional messages. Say, for example, “+ 2 other prob- lems.” Also, provide a distinct visual difference if the same error message is displayed more than once, because the first attempt to correct the problem failed. A person may not notice a repeated message that looks identical to the previous one. Always give a person something to do when an error occurs: something to enter, save, or reverse, or someone to call. Also provide a Help button. The Help button *must* be helpful, though.

Explicit and constructive error messages should be provided. These messages should describe what error occurred, and how it should be corrected. Corrective actions will be clearer if phrased with words like “must be” or “must have.” A study addressing restructuring messages following guidelines such as these, and others previously described, found improved success rates in fixing errors, lower overall error rates, and improved user satisfaction.

Any error ambiguities should be resolved by having the system query the user. Errors should be corrected with minimal typing. Another important error control measure is to have the system identify and store errors. This will allow tracking of common errors so that appropriate prevention programs can be implemented.

##### Providing Guidance and Assistance

New users must go through a learning process that involves developing a conceptual or mental model to explain the system’s behavior and the task being performed. Guid- ance in the form of the system’s hard-copy, online documentation, computer-based train- ing, instructional or prompting messages, and system messages serves as a cognitive development tool to aid this process. So does assistance provided by another form of online documentation, the Help function. Broadly speaking, online documentation is *every* communication provided online to help people to do their work effectively.

While it is desirable that the human-computer interface be so “self-evident” and “in- telligent” that people never experience difficulties, this goal will not be achieved in the

foreseeable future. So a great deal of emphasis should be placed on creating good guid- ance and assistance, and managing the trouble that does occur. Indeed one survey found that documentation was the second most important factor influencing the decision to purchase something. (Quality was first.)

Useful guidance and assistance answers the following questions:

What is this? What does it do?

How do I make it do it?

What is its role in the overall scheme of things?

Technical information, unlike works of fiction, is seldom read for pleasure. People turn to it only when a question has to be answered. Failure to provide the guidance and assistance needed in learning, answering questions, and problem solving makes it very difficult for the user to recover from trouble on his or her own and to avoid future trou- ble by learning from his or her mistakes. The result is most often more errors and great frustration.

###### Problems with Documentation

Wright (1991) feels that poor documentation is usually not the result of stupid and care- less writing. Most writers, professional or not, she says, try to communicate their ideas as well as they can. Poor products, however, suggest that being a native speaker of the language is not a sufficient qualification to ensure communicative success. Rather, four other factors contribute to bad design.

**Organizational factors.** First are organizational factors including management de- cisions concerning who does the writing: product developers or specialist technical authors. Product developers, by their nature, are more interested in the technical aspects and seldom have time to focus on writing. Another organizational factor is the frequency and nature of the contact between writers and developers. Suc- cessful writing requires that frequent contact be maintained between writers and developers. If not, modifications may go undocumented, and functionality may occur that is difficult to explain.

**Time scale.** Second is the time scale allocated for the writing process. Successful writ- ing also involves detailed early planning, drafting, testing, and considerable re- vising. Without adequate time being made available for the writing process, the planning, testing, and revising processes are limited, thereby increasing the po- tential for a mismatch between the final product and its documentation.

**Theoretical rationale.** Third, there is not yet a clear theoretical rationale about what content should be included in documentation and how this information should be presented. Until this is developed, one cannot be sure that the documentation being developed is the most effective that it can be.

**Resources.** Finally, Wright concludes, there are the resources. Adequate resources are needed to include people with different skills in the documentation develop- ment process. Required are people good at visual layout, writing, and test and

evaluation. Rarely does the same person possess more than one of these skills. Without the proper expertise, documentation will also suffer.

Another problem with documentation is created by the need for translation in our shrinking world. The following is found in a current user guide: “The color deviation from the original is thus resulted.” (KYE Systems, 1995) The product manufacturer is guilty of Wright’s sins number two and four above. International considerations will be presented in detail in Step 10 “Provide Effective Internationalization and Accessibility.”

###### How Users Interact with Documentation

There are three broad stages through which a reader interacts with documentation: finding information that is relevant, understanding what the documentation says, and applying that understanding to the current task in order to solve the problem that prompted turning to the documentation in the first place.

Finding information is enhanced through use of contents pages and index lists. It is also enhanced if browsing is made easy through clearly visible page headings and sub- headings. Pictures and symbols can also be used to draw the reader’s attention to par- ticular kinds of information. Understanding information is achieved through a variety of factors. Included are following good writing principles. Understanding can also be maximized through testing and revision of materials as necessary.

##### Instructions or Prompting

Instructional or prompting information is placed *within the body* of a screen. It may take the form of messages or other advice, such as the values to be keyed into a field. Prompting is provided to assist a person in providing what is necessary to complete a screen.

Inexperienced users find prompting a valuable aid in learning a system. Experienced users, however, often find prompting undesirable. It slows them down, then adds “noise” to the screen, and may reduce the amount of working information that can be displayed at one time.

Since instructions or prompting can easily create screen noise, be cautious in placing it on a screen. Use it only if all screen usage will be casual or infrequent. If people with a wide range of experience will be using a screen, it should be selectable, capable of being turned on or off as needed. As an alternative, two separate sets of screens could be made available, one with prompts and the other without. Guidelines for writing in- structions and prompts were covered in Step 8 “Write Clear Text and Messages.”

##### Help Facility

The most common form of online documentation is the Help system. The overall objec- tive of a Help facility is to assist people in remembering what to do. Its benefits include improving the usability of a system, providing insurance against design flaws that may develop, and accommodating user differences that may exist (novice versus expert). Typ- ical methods of invoking Help include: through a typed command, by pressing a Help

key or button, or by selecting a Help option from a multiple-item menu. Help may also automatically appear on the screen.

Some studies have found a Help system can aid performance, others have concluded that a Help function can impair performance if it is not task-oriented, and if it makes the interface more complex. One potential danger of the Help facility, as one study found, is that a person’s recall of command operations is related to frequency of Help facility access; fewer Help requests were associated with better command recall. The re- searchers speculate that the availability of Help may become a crutch and lead to less effective retention. People may implement a passive cognitive strategy. A Help facility, then, may influence performance in systematic and subtle ways.

The specific design characteristics that enhance an online Help are still being identi- fied. Three broad areas of Help that must be addressed in creating Help are: its content, its presentation, and its access mechanisms. Of these, presentation and access are best understood.

Elkerton and Palmiter (1991) propose that the content (and structure) of an effective online Help can be specified using the GOMS (goals, operators, methods, selection rules) model (Card, Moran, and Newell, 1983). Using GOMS, information is provided to the user on *goals* for meaningful tasks, on *operators* for actions required to be performed, on *methods* for accomplishing the goals, and where multiple interface methods exist, and on *selection rules* for choosing a specific method. Gugerty, Halgren, Gosbee, and Rudisill (1991) found that this structure was useful in remembering medical procedures. Elker- ton (1988) presents a set of suggested principles for online assistance (which he calls *on- line aiding*). These principles are reproduced in Table 9.1.

Next, we’ll look at some general Help guidelines. Then, we’ll address some specific considerations for contextual Help, task-oriented Help, reference Help, and wizards.

###### Help Facility Guidelines

* + - Kind:
      * Collect data to determine what types of Help are needed.
    - Training:
      * Inform users of availability and purpose of Help.
    - Availability:
      * Provide availability throughout the dialog.
      * If no Help is available for a specific situation, inform the user of this and provide directions to where relevant Help may exist.
    - Structure:
      * Make them as specific as possible.
      * Provide a hierarchical framework.
        + Brief operational definitions and input rules.
        + Summary explanations in text.
        + Typical task-oriented examples.
    - Interaction:
      * Provide easy accessibility.

**Table 9.1** Suggested Design Principles for Providing Online Advice Based on the GOMS Model

**USE *GOALS* IN ONLINE AIDING TO DO THE FOLLOWING:**

1. Describe what can be done in task-oriented terms (interface actions and objects) for improved initial skill learning.
2. Provide an adjustable level of detail on interface procedures for accommodating the information needs of a wide range of users.
3. Provide procedurally incomplete advice so that users can actively learn for improved long-term performance with and understanding of the interface.
4. Provide feedback to users that may help in reminding them of appropriate procedures to use, particularly when recovering from errors.
5. Develop modular assistance and instructional dialogs that can be used to describe similar and dissimilar procedural elements of the interface.

**USE *OPERATORS* IN ONLINE AIDING TO DO THE FOLLOWING:**

1. Describe simple actions, such as pressing specific keys or finding specific objects on the display, that are common to many interface procedures, to assist the user in current task performance.
2. Provide detailed information on interface procedures that inexperienced users can actively learn and that more skilled users can combine with other procedural knowledge to improve long-term performance and understanding of the interface.
3. Monitor users’ actions to provide context-sensitive Help or to diagnose user problems actively.

**USE *METHODS* IN ONLINE AIDING TO DO THE FOLLOWING:**

1. Present step-by-step interface procedures to assist the user with specific problems.
2. Improve user understanding and acceptance of online advice.
3. Decrease the cognitive load of users who are learning a new interface task by providing an explicit procedure for users to follow.
4. Provide procedural demonstrations of interface procedures so that users can quickly learn simple operations.
5. Map sequences of users’ actions to a reduced set of interface goals to help provide context-sensitive advice to users.

**USE *SELECTION RULES* IN ONLINE AIDING TO DO THE FOLLOWING:**

1. Help users choose between multiple interface methods.
2. Provide users with an understanding of representative tasks to increase their knowledge of when to apply specific interface skills.

From Elkerton (1988)

* + Leave the Help displayed until:
    - The user exits.
    - The action eliminating the need for Help is performed.
  + Provide instructions for exiting.
  + Return to original position in dialog when Help is completed.
* Location:
  + Minimize the obscuring of screen content.
  + If in a window, position priorities are: right, left, above, and below.
* Content:
  + Minimize the Help’s length.
  + Develop modular dialogs that can be used to describe similar and dissimilar
  + pro- cedural elements of the interface.
  + Provide step-by-step interface procedures to assist the user with specific problems.
  + Provide procedural demonstrations of interface procedures to aid quick learning of simple operations.
  + Provide information to help users select between multiple interface methods.
  + Provide users with an understanding of representative tasks to increase their knowledge of when to apply specific skills.
* Style:
  + Provide easy browsing and a distinctive format.
    - Contents screens and indexes.
    - Screen headings and subheadings.
    - Location indicators.
    - Descriptive words in the margin.
    - Visual differentiation of screen components.
    - Emphasized critical information.
  + Use concise, familiar, action-oriented wording.
  + Refer to other materials, when necessary.
  + Never use Help to compensate for poor interface design.
* Consistency:
  + Provide a design philosophy consistent with other parts of the system.
* Title:
  + Place the word “Help” in all Help screen titles.

**Kind.** The two most common reasons people use Help are: (1) Confusion exists about something located on the screen, and (2) information about a specific function is needed. All system usability problems should be systematically identified through testing and evaluation. Monitoring users’ actions can also be a useful tool in identi- fying user problems. Online Help can then be developed to address these problems.

**Training.** Inform users of the availability and purpose of various types of Help.

Never assume that this will be obvious.

**Availability.** Make Help available at all points in the dialog. It is especially critical that Help be available consistently in all similar situations. For example, if one particular system menu has Help, ensure that all menus provide Help. If no Help

is available for a specific situation, inform the user of this and provide directions to where relevant Help may exist, including hard-copy materials.

**Structure.** The Help response should be as specific as possible, tailored to the task and the user’s current position. When accessed, the Help facility should be aware of the kind of difficulties a person is having and respond with relevant informa- tion. Only the information necessary to solve the immediate problem or to answer the immediate question should be presented. If the Help facility is unsure of the request, it should work with the user through prompts and questions to resolve the problem.

A Help facility should be multilevel, proceeding from very general to succes- sively more detailed and specific explanations to accommodate a wide range of users. The first level should provide brief definitions and rules, simple reminders, and memory joggers sufficient for skilled users. The second level should incorpo- rate more detailed explanations in a textual format. The final, and deepest, level should provide guidance in the form of task-oriented examples.

**Interaction.** A Help facility should be retrievable simply, quickly, and consistently by a key action, selection, or command. Leave the Help displayed until the user explicitly exits the Help, or performs the action eliminating the need for Help. In- structions for exiting the Help process should always be provided. These may take the form of displayed pushbuttons, function keys, or something similar.

Help should not disrupt processing. Easy return to the point of the problem should be permitted. Ideally, the problem or work should be retained on the screen when Help is accessed, but this will not always be possible unless the system pro- vides a windowing capability.

**Location.** When Help is displayed, minimize the obscuring of relevant screen con- tent. If Help is displayed within a window, position priorities are right, left, above, and below.

**Content.** Minimize the Help’s length, whenever possible. Carroll, Smith-Kerker, Ford, and Mazur (1986) recommend the development of Help text in the form of “minimal manuals.” These manuals are explicit and focus on real tasks and ac- tivities, and they have been found to be significantly better than traditional Help texts (Black, Carroll, and McGuigan, 1987; Carroll, et al., 1986).

Elkerton (1988) suggests that few Help users want detailed, fact-oriented knowledge such as a hierarchical list showing the syntax of a command. Instead, they want to know the methods to complete a task. Without knowledge of how to do things, users are left to browse through a wealth of information with little understanding of what may be useful. Hence, he recommends, among other things, providing the following:

Step-by-step interface procedures to assist the user with specific problems.

Procedural demonstrations of interface procedures to aid quick learning of sim- ple operations.

Information to help users choose between multiple interface methods.

Users with an understanding of representative tasks to increase their knowledge of when to apply specific skills.

When procedural steps are presented, consecutive numbering will make them easy to follow.

**Style.** Provide easy browsing and a distinctive format. Often the exact location of in- formation needed to answer a question cannot be definitely established. Provid- ing information in a format that can be easily skimmed aids the search process and also helps the user become familiar with the information being presented. The following techniques enhance the skimming process:

Contents screens and indexes. Screen headings and subheadings. Location indicators.

Descriptive words in the margin.

Visual differentiation of screen components. Emphasized critical information.

An index has been found to be one of the first place users turn when they have a problem. Help wording should also be concise, familiar, and action oriented. Reference to outside material may be included in the Help text, especially if the Help information cannot be provided in a concise way. Never use Help to com- pensate for poor interface design.

**Consistency.** The Help design philosophy should be consistent with the philosophy used in other parts of the system. This includes presentation techniques, style, procedures, and all other aspects.

**Title.** For easy identification, place the word “Help” in all Help screen titles.

##### Contextual Help

Contextual Help provides information within the context of a task being performed, or about a specific object being operated upon. Common kinds of contextual Help include Help command buttons, status bar messages, and ToolTips. Microsoft Windows has also introduced what is called the What’s This? Command.

###### Help Command Button

* Description:
  + A command button.
* Purpose:
  + To provide an overview of, summary assistance for, or explanatory information about the purpose or contents of a window being displayed.
* Design guidelines:
  + Present the Help in a secondary window or dialog box.

**Description and purpose.** The proper usage of a command button labeled Help, il- lustrated in Figure 9.2, is to provide supplemental Help for a secondary window,



**Figure 9.2** Help command button.

dialog box, or message box. It should provide an overview of, summary assistance for, or explanatory information about, the purpose or contents of a window.

**Design guideline.** Present this form of Help in a secondary window or dialog box. Microsoft Windows considers this Help an optional secondary form of contextual assistance, and not a substitute for the *What’s This?* command to be described shortly. The guidance and assistance provided by a Help command button differs from the “What’s This?” In that more general assistance is provided rather than information specific to the control that has the current input focus.

###### Status Bar Message

* Description:
  + An abbreviated, context-sensitive message related to the screen item with the focus.
  + Appears in window’s status bar when the primary mouse button is pressed over an item (or keyboard focus is achieved).
* Purpose:
  + To provide explanatory information about the object with the focus.
  + Use to:
* Describe the use of a control, menu item, button, or toolbar.
* Provide the context of activity within a window.
* Present a progress indicator or other forms of feedback when the view of a win- dow must not be obscured.
  + Do not use for information or access to functions essential to basic system opera- tions unless another form of Help is provided elsewhere in the Help system.
  + If extended Help is available and must be presented, place “Press F1 for Help” in bar.
* Writing guidelines:
  + Be constructive, not simply descriptive.
  + Be brief, but not cryptic.
  + Begin with a verb in the present tense.
  + If a command has multiple functions, summarize them.
  + If a command is disabled, explain why.

**Description.** An abbreviated, context-sensitive message related to the screen item with focus. The message appears in the screen’s status bar, as shown in Figure 9.3 when the primary mouse button is pressed over an item (or keyboard focus achieved).

**Purpose.** A status bar message’s purpose is to provide explanatory information about the screen object with focus. Because the user may not always notice a mes-



**Figure 9.3** Status bar message.

sage displayed in the bar, or the bar may be turned off and not displayed, it must be considered a form of secondary or supplemental assistance.

Use a status bar message to provide context for the activity being performed in window, or to describe the use of toolbars, menu items, or buttons being dis- played. When the primary mouse button is clicked over one of these items (or key- board focus achieved) display a short message describing the use of the associated command.

The bar may also be used for presentation of a progress indicator, or other forms of feedback, when the view of a window must not be obscured. Never use the bar for information or access to functions essential to basic system operations, unless another form of Help for this operation is provided elsewhere in the inter- face. If extended Help must be provided, and displaying it in the bar is not possi- ble, place “Press F1 for Help” in the bar.

**Writing.** Do not simply describe something but explain it in a constructive manner Be as brief as possible so the text can be read easily, but do not make the text so short that it is cryptic Begin all messages with a verb in the present tense. If a command with multiple functions has focus, summarize its multiple uses. If a command is disabled, explain why.

###### **ToolTip**

* + - Description:
      * A small pop-up window that appears adjacent to control.
      * Presented when the pointer remains over a control a short period of time.
    - Purpose:
      * Use to display the name of a control when the control has no text label.
    - Design guidelines:
      * Make application-specific ToolTips consistent with system-supplied ToolTips.
      * Use system color setting for ToolTips above to distinguish them.

**Description.** A ToolTip is a small pop-up window with a label that appears adjacent to a control without a label (such as a toolbar) when the pointer is positioned over the control. It is displayed after the pointer remains over the control for a short pe- riod of time. This avoids the distracting effect of a ToolTip appearing when a pointer is simply being moved past a control.

**Purpose.** To display the name of a control when the control has no text label.

**Design guidelines.** Make application-specific ToolTips consistent in size and struc- ture with system-supplied ToolTips, including using the system’s color setting to distinguish them. ToolTips were also described in Step 7 “Choose the Proper Screen-Based Controls.”

###### **What’s This? Command**

* Description:
  + A command located on the Help drop-down menu on a primary window.
  + A button on the title bar of a secondary window.
  + A command on a pop-up menu for a specific object.
  + A button on a toolbar.
* Purpose:
  + Use to provide contextual information about any screen object.
* Design guidelines:
  + Phrase to answer the question “What is this?”
  + Indicate the action associated with the item.
  + Begin the description with a verb.
  + Include “why,” if helpful.
  + Include “how to,” if task requires multiple steps.
  + For command buttons, use an imperative form: “Click this to.…”

**Description and purpose.** A *What’s This* command may take the form of a command in a menu or a button, as summarized above. It’s purpose is to provide contextual information about any screen object, including controls, dialog boxes, and prop- erty sheets.

**Design guidelines.** Phrase the label or caption to answer the question “What is this?” Indicate the action associated with the item and begin description with a verb. If helpful, include an answer to “Why?” as well. Include a “how to” if the task re- quires multiple steps. For command buttons, use an imperative form, “Click this to ”

The guidance and assistance provided by “What’s This? differs from that of a command button. With command button Help, more general assistance is provided rather than information specific to the control that has the current input focus.

##### **Task-Oriented Help**

* Description:
  + A primary window typically accessed through the Help Topics browser.
  + Includes a set of command buttons at the top; at minimum:
* A button to display the Help Topics browser dialog box.
* A Back button to return to the previous topic.
* Buttons that provide access to other functions such as Copy or Print.
  + - Purpose:
      * To describe the procedural steps for carrying out a task.
      * Focuses on *how* to do something.
    - Design guidelines:
      * Provide one procedure to complete a task, the simplest and most common.
      * Provide an explanation of the task’s goals and organizational structure at the start.
      * Divide procedural instructions into small steps.
      * Present each step in the order to be executed.
      * Label each step.
      * Explicitly state information necessary to complete each step.
      * Provide visuals that accurately depict the procedural steps.
      * Accompany visuals with some form of written or spoken instructions.
      * Begin any spoken instructions simultaneously with or slightly after a visual is presented.
      * Segment any animation to focus attention on specific parts.
      * Segment instructions.
      * Delay the opportunity to perform the procedure until all the procedure’s steps have been illustrated.
    - Presentation guidelines:
      * The window should consume a minimum amount of screen space, but be large enough to present the information without scrolling.
      * Normally, do not exceed four steps per window.
      * Use a different window color to distinguish task-oriented Help windows from other windows.
    - Writing guidelines:
      * Write simply and clearly, following all previously presented guidelines.
      * Focus on *how* information, rather than *what* or *why*.
      * Do not include introductory, conceptual, or reference material.
      * Limit steps to four or fewer to avoid scrolling or multiple windows.
      * If a control is referred to by its label, bold the label to set it off.
      * Include the topic title as part of the body.

**Description.** *Task-oriented Help,* sometimes called *procedural Help* is presented on a primary window accessed through the Help Topics browser dialog box in Mi- crosoft Windows. It includes a set of command buttons at the top, minimally, a button to display the Help Topics browser dialog box, a Back button to return to the previous topic. And buttons that provide access to other functions such as Copy or Print.

**Purpose.** Task-oriented Help details the procedural steps for carrying out a task. People prefer task-oriented Help to product-oriented Help, and research evidence shows a productivity gain using it. It is not surprising that task-oriented Help has such a preference and benefits, because people think in terms of tasks, not func- tions. This form of Help focuses on *how* to do something, rather than the *what* or *why*. Its purpose is not to document everything there is to know about a subject.

**Design guidelines.** The following guidelines are mostly derived from Harrison (1995). First, present only one procedure to complete a task, the simplest and most common. (If information about alternate methods is included, place it in a Notes or Related Topic section.)

At the beginning, provide an explanation of the task’s goals and organizational structure. Divide procedural instructions into small steps and present them in the order they are to be executed. Clearly label each step. Explicitly state what infor- mation is necessary in order to complete each step, presenting the most important information first.

Provide visuals that accurately depict the procedural steps. People prefer to follow visual examples rather than instructions, and visuals minimize orientation errors. Accompany the visuals with some form of written or, if possible, spoken instructions. Instructions provide cues as to most important aspects of the proce- dure. Begin any spoken instructions simultaneously with or slightly after a visual is presented. If animation is included, segment it to focus attention on specific parts. Segment the instructions to reinforce the concept of chunks or steps. Fi- nally, delay the opportunity to perform the procedure until all the procedure’s steps have been illustrated.

**Presentation guidelines.** A task-oriented Help window should consume a minimum amount of screen space, but be large enough to cover all the necessary informa- tion without requiring cumbersome scrolling. Normally, this means do not exceed four steps per window. To distinguish task-oriented Help windows from other windows, use a different window color to present them.

**Writing guidelines.** Write simply and clearly, following all previously presented text guidelines. Focus on *how* information, rather than *what* or *why*. Do not include in- troductory, conceptual, or reference material. If a control is referred to by its label, bold the label to set it off. Include the topic title as part of the body.

##### **Reference Help**

* Description:
  + An online reference book.
  + Typically accessed through a:
* Command in a Help drop-down menu.
* Toolbar button.
* Purpose:
  + To present reference Help information, either:
* Reference oriented.
* User guide oriented.
* Design guidelines:
  + Provide a consistent presentation style, following all previously presented guidelines.
  + Include a combination of contextual Help, and task-oriented Help, as necessary.
  + Include text, graphics, animation, video, and audio effects, as necessary.
  + Make displayed toolbar buttons contextual to the topic being viewed.
* Provide jumps, a button or interactive area that triggers an event when it is se- lected, such as:
  + Moving from one topic to another.
  + Displaying a pop-up window.
  + Carrying out a command.
* Visually distinguish a jump by:
  + Displaying it as a button.
  + Using a distinguishing color or font to identify it.
  + Changing the pointer image when it is over it.
    - Presentation guidelines:
      * Provide a nonscrolling region for long topics to keep the topic title and other key information visible.
    - Writing guidelines:
      * Write simply and clearly, following all previously presented guidelines.
      * Provide meaningful topic titles.

**Description and purpose.** *Reference Help* is another form of online documentation. Its purpose is to present Help information that may be reference-oriented, document- ing the features of a product, or it may serve as a user’s guide to a product. It is typ- ically accessed through a command in a Help drop-down menu, or a toolbar button. Reference-oriented Help is usually organized by functions and features and in- cludes more text than other types of Help. User-guide-oriented Help is usually or- ganized by tasks and may include more illustrations than other types of Help.

**Design guidelines.** Provide a consistent presentation style, following all previously presented guidelines. Include a combination of contextual Help, and task-oriented Help, as necessary. Include text, graphics, animation, video, and audio effects, as necessary and as available. Make toolbar buttons contextual to the topic being viewed in the Help window.

Provide *jumps*, a button or interactive area that triggers an event when it is se- lected. The action may be to move from one topic to another, to display a pop-up window, or to carry out a command. Jumps, when in button form are called *short- cut buttons* in Microsoft Windows. They automatically perform a task, thereby providing efficiency by reducing the amount of information necessary to present for reading by the viewer. Do not use a jump, however, if the goal is to enable the user to perform the task. Consider a balance for common tasks. Provide informa- tion that explains how to perform a task and also provide a shortcut button to ac- complish the task, making stepping through the task easier. Visually distinguish a jump by displaying it as a unique style button or using a distinguishing color or font to identify it. The system default for a textual jump in Microsoft Windows is green underlined text. Also, change the pointer image when the pointer is posi- tioned over the jump.

**Presentation guidelines.** If scrolling is necessary, provide a nonscrolling region for long topics to keep the topic title and other key information visible.

**Writing guidelines.** Write simply and clearly, following all previously presented guidelines. Also, provide meaningful topic titles.

##### **Wizards**

* Description:
  + A series of presentation pages displayed in a secondary window.
  + Include:
* Controls to collect input.
* Navigation command buttons.
  + Typically accessed through:
* Toolbar buttons.
* Icons.
* Purpose:
  + To perform a complex series of steps.
  + To perform a task that requires making several critical decisions.
  + To enter critical data and for use when the cost of errors is high.
  + To perform an infrequently accomplished task.
  + The necessary knowledge or experience to perform a task is lacking.
  + Not suited to teaching how to do something.
* Design guidelines:
  + Provide a greater number of simple screens with fewer choices, rather than a smaller number of more complex screens with too many options or too much text.
  + Provide screens of the exact same size.
  + Include on the first page:
* A graphic on the left side to establish a reference point or theme.
* A welcoming paragraph on the right side to explain what the wizard does.
  + Include on subsequent pages:
* A graphic for consistency.
* Instructional text.
* Controls for user input.
  + Maintain consistent the locations for all elements.
  + Make it visually clear that the graphic is not interactive.
* Vary from normal size or render it as an abstract representation.
  + Include default values or settings for all controls when possible.
  + For frequently used wizards, place a check box with the text “Do not show this Welcome page again” at the bottom of the Welcome page.
  + Include a Finish button at the point where the task can be completed.
  + Do not require the user to leave a wizard to complete a task.
  + Make sure the design alternatives offered yield positive results.
  + Make certain it is obvious how to proceed when the wizard has completed its process.
* Presentation guidelines:
  + Display the wizard window so it is immediately recognized as the primary point of input.
  + Present a single window at one time.
  + Do not advance pages automatically.
    - Writing guidelines:
      * Clearly identify the wizard’s purpose in title bar.
      * At the top right of the wizard window, title the Welcome page “Welcome to the

*Wizard Name* Wizard.”

* + - * + Use mixed case in headline style and no ending punctuation.
      * Write simply, concisely, and clearly, following all previously presented guidelines.
      * Use a conversational rather than instructional style.
      * Use words like “you” and “your.”
      * Start most questions with phrases like “Which option do you want . . .” or “Would you like . . .”

**Description.** A *Wizard* is a structured set of screens that guides the user through a decision-making or data entry process. Wizards are displayed in a secondary window. The screens include controls to collect input, and navigation command buttons located at the page bottom (Back, Next, Finish, and Cancel). A wizard is typically accessed through toolbar buttons or icons.

**Purpose.** A wizard’s purpose is to assist a user by automating a task and walking the user through the process. It may not appear as an explicit part of the Help inter- face. Wizards are useful for complex or infrequently occurring tasks that people may have difficulty learning or doing. Wizards are designed to hide many of the steps and much of the complexity in doing something. They are not suited to teaching how to do something, and should be considered a supplement to the ac- tual performance of the task. An experienced user who knows a process will usu- ally find a wizard inefficient or lacking access to all necessary functionality. A wizard can be accessed through toolbar buttons or icons. Microsoft (2001) suggests the following guidelines.

**Design guidelines.** Provide a greater number of simple pages with fewer choices, rather than a smaller number of more complex pages with too many options or too much text. Fewer pages will make it easier to understand the wizard and the process. Create screens of the exact same size. Include on the first page a graphic on the left side to establish a reference point or theme and a welcoming paragraph on the right side to explain what the wizard does. The graphic’s purpose is to es- tablish a reference point, or theme, or present a preview of the wizard’s result. In- clude on subsequent pages a graphic for consistency, instructional text, and the necessary controls for user input. (If screen space is critical, graphics on subsequent pages may be omitted.)

Make it visually clear that the graphic is not interactive by varying it from nor- mal size or rendering it as an abstract representation. Do not require the user to leave a wizard to complete a task. The user, often a novice, may lose context if asked to leave. Everything must be done from within the wizard. Make sure the design alternatives offered to the user yield positive results.

For frequently used wizards, place a check box with the text “Do not show this Welcome page again” at the bottom of the Welcome page. Include a Finish button at the point where the task can be completed. Make certain it is obvious how to

proceed when the wizard has completed its process by including proper closing text on the last page.

**Presentation guidelines.** Display the wizard window so it is immediately recognized as the primary point of input. Present a single window at one time, overlaying un- derlying windows so they are not visible. Do not advance pages automatically. The viewer may be unable to read all the information, and control of the dialog is removed from the user and placed in the hands of the computer.

**Writing guidelines.** Clearly identify the wizard’s purpose in the title bar. At the top right of the wizard window, title the Welcome page “Welcome to the *Wizard Name* Wizard.” Use mixed case in the headline style of presentation, and no ending punctuation. Write simply, concisely, and clearly, following all previously pres- ented guidelines. Use a conversational rather than instructional style, and words like “you” and “your.” Start most questions with phrases like “Which option do you want?” or “Would you like . . .?” People react better to phrasing that implies they are in control, rather than phrasing telling them what to do.

##### **Hints or Tips**

* Description:
  + A command button labeled Hints or Tips.
* Purpose:
  + To provide a few important contextual, but specific, items of information related to a displayed screen.
* Design guidelines:
  + Provide guidance on only two or three important points.
  + Locate the button near where its guidance applies.
  + Write concisely and to the point.

**Description.** A *Hint* or *Tip* is a command button placed on a screen and labelled as such.

**Purpose.** To provide a few important contextual, but specific, items of information related to a displayed screen. It is a supplement to the standard Help facility, but more easily accessible and relevant to the current situation. The objective is to quickly get the user back on track when disorientation or confusion occurs.

**Design guidelines.** Provide guidance on only two or three important points. Locate the button near the location where its guidance applies and write concisely and to the point.

## **Provide Effective Internationalization** **and Accessibility**

#### **International Considerations**

To create a product for use internationally may involve two steps, *internationalization*and *localization* (Russo and Boor, 1993). Internationalization is the process of isolating culturally specific elements from a product. The German text of a program developed in Germany, for example, is isolated from the program itself. This occurs in the coun- try where the product is developed. Localization is the process of infusing a specific cul- tural context into a previously internationalized product. Translating German screen components and messages into English for American users is an example of localization. Creating a product that has been properly localized and speaks fluently to another culture requires addressing a number of factors. These include text; formats for elements

such as number, date, and time; images; symbols; colors; flow; and functionality.

Localization

* When to do it:
  + When the market includes few or no English speakers.
  + When translation is required by law or by custom.
  + When the widest possible market is desired.
* When not to do it:
  + When the audience already reads English.
  + When the cost of retrofitting or rewriting the software is prohibitive.

This discussion of when, and when not, to internationalize and localize a product, is mostly based on Fowler and Stanwick (1995). Considerations include the prospective users and their English capabilities, local laws and customs, and costs associated with translation.

English is the most widely used language in the world. The current estimate for its speakers ranges from 700 million to 2 billion (Tripathi, 1992). Although many speakers of English have been taught it as a second language and may not all be facile readers and writers, they can communicate using it. The first consideration, then, is the English capabilities of the prospective user. This must be ascertained. Toward this end, both IBM (National Language Technical Center, 1991) and Apple (1992a) have documents listing the official language requirements of countries, and regions or political divi- sions. In addition, within some international business and scientific communities, En- glish is the accepted language of communication. For example, the air transportation industry uses English as the language of communication between airline pilots and flight controllers worldwide. Scientists and engineers in Japan also prefer to commu- nicate their research findings in English because of its greater precision (Kohl, Barclay, Pinelli, Keene, and Kennedy, 1993). If English is accepted in the using body, then con- cerns are only cultural.

Legal requirements may also mandate translation. For example, Canada, being com- posed of both English and French speakers, requires bilingual materials. The European Economic Community (ECC) will, at some point, mandate that all documentation shipped with imported products be written in all of the ECC languages. Whether the product will actually be used in all the countries will be immaterial.

Cost will also, of course, dictate whether a translation can, or will, be performed. Soft- ware translation rates can range from $40–80 dollars an hour or more, documentation translation $50–150 or more per page. These rates are presented for illustrative purposes only. Actual costs will be driven by many factors, including the local cost of living. For readers in need of a translation, one will be best served by getting a quote reflecting the time and locale of the translation. A translation performed in the target country often results in better quality than a translation by those who are native speakers of the pro- ducing country.

Words and Text

* + - Use very simple English.
      * Develop a restricted vocabulary.
      * Restrict the sentence structure using: noun-verb-object.
    - Avoid:
      * Acronyms and abbreviations.
      * Stringing three nouns together.
      * Local or computer jargon.
      * A telegraphic writing style.
      * An over-friendly writing style.
      * Culturally specific examples.
      * References to national, racial, religious, and sexist stereotypes.
    - Adhere to local user language idioms and cultural contexts.
    - Keep the original term for words that cannot be translated.
    - Allow additional screen space for the translation.
      * Horizontally, using Table 10.1.
      * Vertically.
    - When translating to other languages, first do:
      * European: German.
      * Middle East: Arabic.
      * Far East: Japanese.
    - Position icon captions outside of the graphic.
    - Modify mnemonics for keyboard access.
    - Adhere to local formats for date, time, money, measurements, addresses, and tele- phone numbers.

Text translation is simplified, and user interpretation errors reduced, if these guide- lines, many of which are derived from del Galdo (1990), Russo and Boor (1993) and Fowler and Stanwick (1995) are followed.

**Simple English.** Simple English text will be easier and less expensive to translate. Simple English is achieved by using a restricted vocabulary. Create a dictionary of approved terms and prohibit all synonyms, and different meanings for the same word as well. A restricted sentence structure is also necessary. Sentences meaning the same thing can be written in many ways in English. This makes text more in- teresting to look at and read. In other languages, however, word order affects the meaning. Multiple structures cause translation problems and foster errors. Follow a *noun-verb-object* structure. Another benefit of simple English: translation may not always be necessary. The number of non-native English-speaking people capable of understanding the language will increase as screen English is simplified.

**Avoid.** Do not use acronyms and abbreviations. They are difficult, and often confus- ing, to translate. A translated acronym may not be as concise, or may possess neg- ative associations. Abbreviations may also not be as concise, and they may not be understandable.

Avoid stringing three nouns together. Relationships between nouns become very explicit in many other languages, and it is difficult to determine what terms are modifying one another when three are strung together. The use of preposi- tions, such as *at*, *in, by,* and *on*, can help to clarify nouns’ relationships. Avoid local or computer jargon. Jargon is not universal and probably will not be understood.

Do not use telegraphic writing. This means a terse style where words such as, “and,” “the,” and “is” are left out. Again, translation problems can easily occur. An overly friendly style, in which the reader is addressed in the first person or in a childish manner, should also be avoided. It can be considered condescending and irritating to readers in non-English-speaking countries. Finally, avoid refer- ences to national, racial, religious, and sexist stereotypes and do not use culturally specific examples. The latter must be recreated by the translator so they are suit- able for the language and culture.

**Local Language’s Idioms.** Adhere to local language’s idioms and cultural contexts. Some words have different meanings in other languages. This is of special concern for product names. Automakers have been particular victims of this problem. Italy’s Fiat had an auto named “Uno.” They could not sell it by that name in Fin- land because uno in Finnish means “garbage.” England’s Rolls Royce planned to name a new car “Silver Mist.” Then, someone discovered that mist in German means “manure” (Taylor, 1992). Proper attention to localization can avoid some embarrassing, and costly, problems.

Some languages are not read from left to right, as English is. Arabic, for ex- ample, is read from right to left. Chinese is read from top to bottom, right to left.

**Original Terms.** Keep the original terms for words that cannot be translated. Some words do not exist in other languages. “Disk drive” and “zooming” do not exist in Thai, for example. It has been found that people often prefer the original term to a created word. Never invent words; keep the original term for nontranslatable words (Sukaviriya and Moran, 1990).

**Additional Screen Space.** Allow additional screen space for the translation. English is very concise. It usually takes less space to communicate the same word, phrase, or text than most other languages. Listed below are words with the same mean- ing from four languages. Can you translate them?

Besturingselement (Dutch) Olvadaci prvek (Czech)

Ohjausobjekti (Finnish)

Steuerelement (German)

Here is a clue. This word in English is seven characters long and has already been mentioned many, many times in this book. The Dutch version is 17 char-

acters in length, or 143 per cent longer than the English version. The others are composed of 13 characters and are 85 per cent longer. The answer will follow shortly.

Objects whose sizes are affected by translation include captions, entry areas, menu options, prompting message boxes, areas of text, and icon labels. Expan- sion room must be allowed for translation. Generally, the shorter the text the more additional room is needed. Table 10.1 (National Language Technical Center, 1991) provides some additional horizontal space guidelines. Extra vertical spac- ing may also have to be allowed. In many languages, accents and descenders fall above and below the usual ascender and descender lines. What is the English version of the above words? *Control*. Were you able to translate one or more of them?

**Translating.** When translating, start from a translator’s point of view. The language world is divided into three parts: Europe, the Middle East, and the Far East. Fowler and Stanwick (1995) report that Microsoft addresses translation in the fol- lowing manner. In Europe, where problems involve changes in words caused by gender, accented letters, and text expansion, translation begins with German. This is done because German solves for accent, gender, and expansion issues. In the Middle East difficulties in translation include bidirectional and cursive letters. To address these, Microsoft recommends starting with Arabic. When this is done, localization is accomplished for Hebrew, Farsi, Dari Persian, Pashto, and the Indian languages Sindhi and Urdu. In the Far East the main difficulty is double- byte character sets. One of the most difficult Asian languages, with ten thou- sand ideograms divided into four character sets, is Japanese. So Microsoft starts with it.

**Icon Captions.** Place icon captions outside of the graphic. Text placed within an icon may cause the icon to have to be redrawn when translation occurs. Text po- sitioned outside the icon will negate the need for redrawing.

**Table 10.1** Translation Expansion Requirements

**NUMBER OF CHARACTERS IN TEXT ADDITIONAL SPACE**

Field labels and menu options

Up to 10 100–200%

11–20 80–100%

Messages and on-screen instructions

21–30 60–80%

31–50 40–60%

Online help and documentation

51–70 31–40%

Over 70 30%

From National Language Technical Center, IBM (1991)

**Mnemonics.** Modify mnemonics used for keyboard access. Because mnemonics are established for ease of memorization, and because they are based upon a letter in a text object, when the text changes, so must the mnemonic. Attempting to create unique mnemonics may constrain the translation, but this must be addressed. Maintaining the old mnemonics will severely affect users’ ability to learn them. They will no longer be mnemonics.

**Local Formats.** Adhere to local formats for date, time, money, measurements, ad- dresses, and telephone numbers. A nearly infinite variety of these various units exist worldwide. They must also be localized to the exact needs of the user.

##### Images and Symbols

* Adhere to local cultural and social norms.
* Use internationally accepted symbols.
* Develop generic images.
* Be particularly careful with:
  + Religious symbols (crosses and stars).
  + The human body.
  + Women.
  + Hand gestures.
  + Flags.
  + The cross and check for check boxes.
* Review proposed graphical images early in the design cycle.

Images are the visible language of a culture and must be recognizable, meaningful, and acceptable. Like text, improper use of images, symbols, and icons can create prob- lems internationally. Social norms vary, so great variations exist in what is recognizable and acceptable throughout the world. What one culture recognizes may have no mean- ing in another. What is acceptable in one country may not be in another. The images created for graphical interfaces are particularly susceptible to these problems. To be successful internationally, images must be carefully selected and designed. The follow- ing guidelines are also derived from del Galdo (1990), Russo and Boor (1993), and Fowler and Stanwick (1995).

**Local Norms.** Adhere to local cultural and social norms. Few world travelers have not suffered embarrassment caused by failure to understand, and adhere to, local customs and mores. On an early trip to Australia, I pulled in to a service station to replenish my auto’s “petrol.” I communicated my need to the attendant through a “thumbs-up” sign, an American convention (when there were still at- tendants) meaning “fill-it-up.” The Australian attendant’s response was a stunned look and a frown. Sensing something was wrong, I hastily lowered the window and communicated my need verbally. He smiled, replying, “Ah, you’re

American, eh, OK.” It wasn’t until much later I discovered I had made a gesture considered obscene in Australia.

Examples abound in the computer literature of images that have created prob- lems internationally. The mailbox and trashcan are two examples of objects whose shape, and resulting recognizability, vary substantially around the world. A cock- tail glass used to signify an after-work appointment is a poor image to use in countries where alcohol is not associated with social activities. In the United States, a black cat is usually associated with bad luck; in the United Kingdom it means good luck. In the United States, the number 13 is considered unlucky; in Japan the number 4 is.

Images that are culturally specific must be isolated during the international- ization process. Then, proper images must be developed for use in the culture where the product will be used.

**Internationally Accepted Symbols.** Use internationally accepted symbols. Before developing an image, first determine if any international images have already been created by trade or standards organizations. The ISO (International Standards Or- ganization), for example, has developed standard shapes for a variety of pur- poses. Always consult all relevant reference books before inventing new images or modifying existing ones.

**Generic Images.** Whenever possible, create generic images that are usable in multi- ple cultures. Having different images can confuse people who may use more than one language version of a product.

**Where Caution Is Necessary.** Some topics are more susceptible to acceptability problems than others. Inappropriate presentation can result in the viewer’s being offended or insulted. Be particularly careful when using religious symbols such as crosses or stars. Also be wary of images depicting a human body, particularly the female. In some cultures simply revealing a woman’s arms and legs is unaccept- able. During the World Cup soccer tournament in the United States in 1994, one Middle East country televised the soccer games using a several-second tape delay. This was done so that stadium crowd pictures potentially containing pictures of women dressed to accommodate the USA’s summer heat would not be shown on local television. What crowd scenes were substituted instead? Pictures of people attending an American football game in December in a northern city (like Green Bay, Wisconsin) when the temperature hovered around zero degrees Fahrenheit. Needless to say, *all* the fans were well covered from toes to top of head. While tele- vision viewers in this middle-eastern country saw spectators with skin barely visi- ble, the soccer players cavorted in shorts and jerseys.

Also, be wary of hand gestures, as my Australian experience illustrates. Actu-

ally, I’m in pretty good company in committing this kind of faux pas. A former American president departed Air Force One on a visit to Germany exhibiting his customary hand wave to the welcoming people. Unfortunately, his Protocol Offi- cer neglected to inform him that his wave had a vulgar connotation in Germany. We can only hope that German viewers of this action interpreted what he felt in is heart, not what he indicated with his hands. Also exercise caution in using a

country’s flag for a language icon. Many countries are multilingual, Canada, South Africa, and Switzerland, for example. Their flags may not be associated with any one language. Be more generic in nature, a word (such as French, Spanish, Italian) encompasses many countries and cultures.

Last, the X and check mark used for check boxes do not have meaning uni- versally. It has also been found that they do not have universal meaning in the United States. In recent years, various graphical systems have moved away from X to the check mark as the symbol to indicate an active or set check box control. Why? In an engineering environment an X in a check box means the choice is not applicable or not set, a check means it is applicable, or it is set. Thus, an “X” was found to be confusing to some people when it meant active or set. Research has also indicated that when people complete a form with check boxes, the symbol most often used is the check mark.

MAXIM Perception = Reality!

**Review Images Early.** Review proposed graphical images early in the design cycle. Creating acceptable images can be a time-consuming process*.* Start developing them early so ample time exists for extensive testing and modification.

##### Color, Sequence, and Functionality

* Adhere to local color connotations and conventions.
* Provide the proper information sequence.
* Provide the proper functionality.
* Remove all references to features not supported.

Other international considerations include the following.

**Local Color Connotations.** Color associations also differ among cultures. In the United States mailboxes are blue; in England they are red; in Greece they are yel- low. In the United States red is associated with danger or stop, green with OK or go. This red-green association does not exist everywhere in the world. Table 10.2, derived from Russo and Boor (1993), lists some common cultural color associations. Colors used on screens must also reflect the color expectancies of its viewers.

**Information Sequence.** Information within a screen will be arranged to reflect the logical flow of information. In many cultures, including those we are most famil- iar with, it will be from left to right for text and from top to bottom, left to right for ease of scanning. Some cultures, however, read from right to left. For these, in- formation sequence must be reorganized to reflect this right-to-left sequence. Sim- ilarly, cascaded windows for left-to-right readers are usually presented in an upper-left to lower-right structure. These will have to be reorganized to reflect dif- ferent reading patterns.

**Table 10.2** Some Cultural Color Associations

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **RED** | **YELLOW** | **GREEN** | **BLUE** | **WHITE** |
| China | Happiness | Birth | Ming Dynasty | Heavens | Death |
|  |  | Wealth | Heavens | Clouds | Purity |
|  |  | Power | Clouds |  |  |
| Egypt | Death | Happiness | Fertility | Virtue | Joy |
|  |  | Prosperity | Strength | Faith |  |
|  |  |  |  | Truth |  |
| France | Aristocracy | Temporary | Criminality | Freedom | Neutrality |
|  |  |  |  | Peace |  |
| India | Life | Success | Prosperity |  | Death |
|  | Creativity |  | Fertility |  | Purity |
| Japan | Anger | Grace | Future | Villainy | Death |
|  | Danger | Nobility | Youth |  |  |
|  |  |  | Energy |  |  |
| United | Danger | Cowardice | Safety | Masculinity | Purity |
| States | Stop | Caution | Go |  |  |

**Proper Functionality.** Product features developed for one culture may not be ap- propriate for all cultures. Nielsen (1990), for example, describes a school hypertext product developed in France. During requirements determination it was estab- lished that only the schoolteacher should be able to add comments and view- points to the screens, not the students. This was a socially acceptable practice in France. Later, when the product was marketed in Sweden this aspect created problems. In Sweden independent discovery is greatly valued, and the inability of the students to add comments and viewpoints was unacceptable. All interna- tional products have to be reviewed for functionality as well and may require multiple versions to reflect the individual needs of cultures.

**Features Not Supported.** All aspects of a product not supported internationally should be removed from the system. Any references to features not supported should also be eliminated from all documentation. To leave this information in cre- ates visual noise and will be confusing.

##### Requirements Determination and Testing

* Establish international requirements at the beginning of product development.
* Establish a relationship within the target culture.
* Test the product as if it were new.

When a product is translated for a new culture, it becomes a new product. Russo and Boor (1993) suggest the following should be accomplished.

**Establish Requirements at Beginning.** Developers must establish in what cultures the product will be used in at the start of the development cycle. Then, differing product requirements must be established, reflecting the differing needs of the various users. This permits localization issues to be addressed throughout the de- velopment process.

**Relationship with Target Culture.** A close working relationship with *natives* from all using cultures during requirements and development will permit local, culturally specific feedback to be obtained in a timely manner. A close working relationship will also educate the designers about the culture where their product will be used.

**Testing.** When a product is translated for a new culture, it is a new product and it should be subjected to a normal testing during the development cycle. If interna- tional testing is delayed until after the product is released to the domestic market, problems may be difficult, if not impossible, to address.

#### Accessibility

*Accessibility*, in a general sense, means a system must be designed to be usable by an al- most unlimited range of people, essentially anyone who desires to use it. In a narrower sense, accessibility can be defined as providing easy access to a system for people with disabilities. We’ll focus on this aspect of accessibility in the following paragraphs. De- sign objectives in creating accessibility for users with disabilities are:

Minimize all barriers that make a system difficult, or impossible, to use. Provide compatibility with installed accessibility utilities.

Many governments have passed laws requiring that most employers provide rea- sonable accommodation for workers with disabilities. In the United States, one piece of legislation with this intent is the Americans with Disabilities Act. Accessible system de- sign, then, seeks to ensure that no one with a disability is denied access to computer technology.

##### **Types of Disabilities**

Worldwide, a significant number of people have disabilities of one form or another. Disabilities may be temporary or permanent, or simply the result of aging. Disabilities can be grouped into several broad categories: visual, hearing, physical movement, speech or language impairments, cognitive disorders, and seizure disorders.

*Visual* disabilities can range from slightly reduced visual acuity to total blindness.

*Hearing* disabilities range from an inability to detect certain sounds to total deafness.

*Physical movement* disabilities include difficulties in, or an inability to, perform certain physical tasks such as moving a mouse, pressing two keyboard keys simultaneously, or accurately striking a single keyboard key. People with *speech or language* disabilities may find it difficult to read and write (as with dyslexia). *Cognitive* disabilities include memory impairments and perceptual problems. People with *seizure* disorders are sen- sitive to visual flash rates, certain rates triggering seizures.

##### Accessibility Design

* Consider accessibility issues during system planning, design, and testing.
* Provide compatibility with installed accessibility utilities.
* Provide a customizable interface.
* Follow standard Windows conventions.
* Use standard Windows controls.

Accessibility issues should be considered throughout the entire system development cycle. Costs of retrofitting after the design is completed are always much higher than costs associated proper design itself. Unlike internationalization, where design costs are weighed against potential benefits, designing for accessibility may be required because of federal laws. All accessibility issues and requirements must be understood in system planning so that they may be incorporated within the design and testing processes.

Provide compatibility with accessibility utilities installed by users (screen review and voice input, for example). Also provide a customizable interface to accommodate the widest variety of user needs and preferences. Users are then free to choose an array of properties most satisfying to their viewing and usage needs. Whenever possible follow standard Windows conventions and use standard Windows controls in design. Most accessibility aids work best with applications that follow standard system conventions. Custom controls may not be usable by screen-review utilities.

###### Visual Disabilities

* Utilities:
  + Ensure compatibility with screen-review utilities.
  + Ensure compatibility with screen-enlargement utilities.
* Screen components:
  + Include meaningful screen and window titles.
  + Provide associated captions or labels for all controls, objects, icons, and graphics.
    - Including graphical menu choices.
  + Provide a textual summary for each statistical graphic.
  + Allow for screen element scalability.
  + Support system settings for high contrast for all user interface controls and client area content.
* When a “high contrast” setting is established, hide any images drawn behind the text to maintain screen information legibility.
  + Avoid displaying or hiding information based on the movement of the pointer.
* Exception: Unless it’s part of the standard interface (Example: ToolTips).
* Keyboard:
  + Provide a complete keyboard interface.
  + Provide a logical order of screen navigation.
* Color:
  + Use color as an enhancing design characteristic.
  + Provide a variety of color selections capable of producing a range of contrast levels.
* Create the color combinations based on the system colors for window components.
* Do not define specific colors.

Visual disabilities range from impaired visual acuity, often resulting from aging; de- creased sensitivity to a specific color or colors; partial blindness, or total blindness. Mod- erately impaired vision may simply require the availability of larger fonts or restrictions in the use of colors. Severe impairments, such as blindness, may require compatibility with speech or Braille utilities.

**Utilities.** For people who cannot use a screen’s visual content, a s*creen-review* utility will be necessary. These utilities, also called *screen-reader* programs or *speech access* utilities, take the displayed information being focused on and direct it through another medium. Alternate media include synthesized speech and refreshable Braille displays. *Screen enlargement* utilities enable the user to enlarge a portion of the screen, the monitor becoming a viewport that displays only a section of an en- larged display. These programs, also referred to as *screen magnification* utilities or *large-print* programs, track the user’s use of a keyboard or mouse, moving the view- port to different areas of the screen as the user navigates within it.

**Screen components.** Meaningful, specific, and unique screen and window *titles* will assist the user in differentiating between these, especially when using a screen- review review utility. When using a reader, content must be addressed sepa- rately, so it will not be available with the title to aid in comprehension of what is presented. Provide associated *labels or captions* for all controls, objects, icons, and graphics, since all screen information must be presented as text by a screen re- viewer. These labels must also be located in close proximity to the screen elements they refer to. A screen reviewer will relate the label to its associated screen ele- ment by its physical proximity, if it is not related programmatically. In rare situ- ations, where the caption may be visually distracting (display-only data on inquiry screens, for example), provide a label but do not make it visible. Follow all the conventions presented in Step 3 “Understand the Principles of Good Screen

Design” for caption and label placement. *Graphical menu* choices, such as illus- trated colors, shades, and patterns, must also possess textual labels.

Also provide a textual summary for each *statistical graphic*. Statistical graphics are images containing detailed information and, because of their graphic nature, their contents cannot be conveyed by a screen reader. The textual summary should include all information available to a sighted user.

Support screen element *scalability*, the presentation of larger text and graphics for people with only slight or moderate vision impairment. Also consider pro- viding a “Zoom” command that scales the information displayed within a win- dow. Support system settings for *high contrast* for all user interface controls and client area content. Users with visual impairments require a high contrast be- tween foreground and background elements for best text legibility. Poor contrasts may result severely degraded legibility because the background may “bleed” into the foreground. When a “high contrast” setting is established, hide any images drawn behind text (watermarks, logos, patterns, and so on) to maintain screen in- formation legibility. Monochrome versions of graphics and icons can also be pre- sented using an appropriate foreground color for the displayed background color. In general, use black text on a white background to achieve the best foreground- background contrast. While some softer colors may be more attractive to look at, black on white always yields the best legibility.

Finally, avoid displaying or hiding information based on the movement of the pointer, unless it is part of the standard interface (a ToolTip, for example). These techniques may not be available to screen-review utilities. If these techniques are used, however, allow them to be turned on or off if a screen-review utility is used.

**Keyboard.** Provide a thorough and *complete keyboard interface*. Blind users cannot use a mouse to navigate because the pointer’s location is unknown. All mouse actions, therefore, must be available through the keyboard using keyboard equivalents and keyboard accelerators. A logical order of screen *element navigation* is also a re- quirement for blind users. While this principle is standard for all screen users, a failure to adhere to it can be especially confusing for the blind because, when using a screen-review utility, they must navigate a screen sequentially in the predeter- mined navigation order. Their ability to scan the entire contents of a control or screen to establish context is simply not possible.

**Color.** Color must always be used as a supplemental or *enhancing* design charac- teristic. Users with a color-viewing deficiency may not be able to discriminate certain colors, and, consequently, they may be unable to understand that an ac- tion is required if the action is based upon an element’s color alone. Provide a *variety* of color selections capable of producing a range of contrast levels. Create these combinations based on the system colors for window components. Never define and use specific colors. With a selection variety, the user may then cus- tomize the interface, choosing the best combination for his or her visual needs. Use of color is much more thoroughly described in Step 12 “Select the Proper Colors.”

###### Hearing Disabilities

* Provide captions or transcripts of important audio content.
* Provide an option to display a visual cue for all audio alerts.
* Provide an option to adjust the volume.
* Use audio as an enhancing design characteristic.
* Provide a spell-check or grammar-check utility.

Hearing disabilities range from an inability to detect or interpret auditory output at normal or maximum levels certain sounds, to total deafness. A noisy work environment may also disrupt hearing, or sound may be turned off to avoid annoying neighboring workers.

Because audio may be missed or not understood, provide *captions or transcripts* of all important audio content. Also provide an option to display a *visual cue* for all audio alerts. Methods include displaying the alert in a message box or within the status bar. Provide an option to *adjust the volume* so that auditory content may be turned louder or off as necessary. A volume control may also benefit the vision-impaired user, who re- lies on a speech access utility to understand the screen. Like color, always use audio as an *enhancing* design characteristic; never rely on it as the sole means of communicating with the user.

Many people who are deaf, and whose language is American Sign Language, can be helped by a *spell-* or *grammar-check* utility. Uses of and problems with sound were de- scribed more fully in Step 9 “Provide Effective Feedback and Guidance and Assistance.”

###### Physical Movement Disabilities

* Provide voice-input systems.
* Provide a complete and simple keyboard interface.
* Provide a simple mouse interface.
* Provide on-screen keyboards.
* Provide keyboard filters.

**Voice input.** People who have difficulty typing should have the option of using a voice-input system. Voice-input systems, also called *speech recognition* systems, permit the user to control software by voice instead of by mouse or keyboard. In a voice-input system, captions or labels are used to identify manipulable screen objects. Speaking the object’s label then activates the object.

**Keyboard interface.** People with limited use of their hands may not be able to effec- tively use a mouse because of the fine motor movements necessary to control it. All mouse actions, therefore, must also be available through the keyboard using keyboard equivalents and keyboard accelerators for people with this physical movement disability. Accessibility Options in the Windows Control Panel also pro-

vide a setting to allow the mouse pointer to be controlled through the numeric key- pad. Some people may have difficulty pressing two keys at the same time. This can be remedied by ensuring that the keyboard interface is simple.

**Mouse interface.** Pointing devices may actually be more efficient for some users with physical movement disabilities. Therefore, a simple mouse interface is also important. As is done for the entire range of system users, do not require basic system functions to be performed through double-clicks, drag-and-drop manipu- lation, and keyboard-modified mouse actions. These are shortcut techniques for advanced users.

**On-screen keyboards.** Some people cannot even use a standard keyboard. Keyboards can be presented on the screen and activated through special switches, a special mouse, or a headpointer, a device used to manipulate a pointer through head motion.

**Keyboard filters.** People with erratic motion, tremors, or slow responses often make incorrect keystrokes. Keyboard filters can be used to ignore brief or repeated key- strokes. Accessibility Options in the Windows Control Panel provide a range of keyboard filtering options.

###### Speech or Language Disabilities

* Provide a spell-check or grammar-check utility.
* Limit the use of time-based interfaces.
  + Never briefly display critical feedback or messages and then automatically re- move them.
  + Provide an option to permit the user to adjust the length of the time-out.

**Spell-checker.** People with language disabilities, such as dyslexia, find it difficult to read and write. A spell-checker or grammar-checker can help these users, as well as or people with writing impairments, and people whose first language is not English.

**Time-based interfaces.** Limit the use of interface techniques that “time-out” and are removed after a prescribed period of time. People with some speech and lan- guage disabilities may not be able to react, either by reading text or pressing keys, within the allotted time period. Again, it is helpful to provide an option to permit the user to adjust and extend the time-out period.

###### Cognitive Disabilities

* Permit modification and simplification of the interface.
* Limit the use of time-based interfaces.
  + Do not briefly display critical feedback or messages and then automatically re- move them.
  + Provide an option to permit the user to adjust the length of the time-out.

**Interface modification and simplification.** People with memory or perceptual prob- lems can often be aided by a simplified interface. Allowing modification of the in- terface, customization of menus, customization of dialog boxes, or hiding graphics, for example, should be permitted. Conversely, some people with cognitive diffi- culties can be assisted by more extensive use of icons and graphics to illustrate ob- jects and choices. Permitting modifications of this sort is also beneficial.

**Time-based interfaces.** People with cognitive impairments may also not be able to react to some situations in a timely manner. Again, for these people limit use of interface techniques that “time-out” and are removed after a prescribed period of time. It is also helpful to provide an option to permit the user to adjust and extend the time-out period.

###### Seizure Disorders

* Use elements that do not blink or flicker at rates between frequency ranges of 2 Hz and 55 Hz.
* Minimize the area of the screen that is flashing.
* Avoid flashing that has a high level of contrast between states.
* Provide an option to enable users to slow down or disable screen flashing.

People with seizure disorders may experience photosensitive epileptic seizures when exposed to certain visual flicker or flash rates of screen elements. In general, the higher the intensity of the flash, the larger area of the flash, or the faster the frequency of the flash, the greater the problem may be. Screen elements particularly susceptible to this phenomenon are flashing text, graphics that repeatedly turn on and off, and screen im- ages that repeatedly change.

**2Hz to 55Hz flicker.** Use screen elements that do not blink or flicker at rates between 2Hz and 55Hz.

**Flashing area.** Smaller areas of flicker are less likely to cause seizures than larger areas. Minimize the area of the screen that is flashing.

**High contrast.** Avoid flashing that has a high level of contrast between states. Some people are more susceptible to high-intensity flashing.

**Slow down or disable.** Provide an option to enable users to slow down or disable screen flashing. In Windows, the Keyboard option in the Control Panel permits adjustment of the cursor blink rate. When set to slow, the cursor will flash 1.2 times per second. The rate increases 100 milliseconds for each notch up to a maximum of five times per second.

###### Web Page Accessibility Design

* Provide a “Skip to Main Content” link at the top of each page.
* Structure articles with two or three levels of headings.
* For all images provide associated text.
  + For lengthy descriptions, provide a link to a separate page.
* For all audio content include one or more of the following:
  + A caption or pop-up text window.
  + A textual transcript.
  + A textual description.
    - For lengthy transcripts or descriptions, provide a link to a separate page.
* For all video content include one or more of the following in both a textual and audio format:
  + A transcript.
  + A description.
    - For lengthy transcripts or descriptions, provide a link to a separate page.
* For image maps, provide equivalent text menus.
* Provide alternate ways to access items contained within tables.
* For online forms that cannot be read by screen utilities, provide alternate methods of communication.
* Provide an option to display animation in a nonanimated presentation mode.
* If accessibility cannot be accomplished in any other way, provide a text-only page with equivalent information and functionality.
* Follow the standards set by the World Wide Web Consortium for accessibility of Web content.

In addition to the previously described guidelines, Web applications require some additional considerations.

**Skip to Main Content link.** When a navigation bar is located at the top or left side of a page, a user using speech synthesis must listen to all navigation links before arriving at the main page content. This can become especially cumbersome if the links are consistently repeated on successive pages. (A sighted user can eas- ily ignore the links.) To by-pass these links, and other elements such as tables of contents, provide a “Skip to Main Content” link attached to an unimportant image at the beginning of each page. The user can activate this link when it is presented, and the focus will move directly to the start of the page’s content. (The user always has the choice to pass over this link and continue through the nav- igation links.)

**Headings.** Structure articles with two or three levels of headings. Nested headings facilitate access using screen-review utilities.

**Images.** Screen-review utilities cannot reveal images to visually impaired users. There- fore it is important to associate text with active images, particularly links or com- mand buttons. When an image is not active, whether to provide associated text must be determined based upon the situation. The visually impaired user will not be able to ignore this audio text as a sighted user can if text is included. Include, then, short textual description for all important images. To include a lengthy text description, provide a link to a separate page that contains a complete description.

**Audio.** For hearing-impaired users, include one of the following textual alternatives. For short audio pieces, provide a *caption or small pop-up* window describing the audio. For longer audio, consider providing a *textual transcript,* an exact word-for- word version of the audio. Give the user the choice of reading the transcript, lis- tening to the audio, or both reading and listening. Also consider a *textual description,* a longer and more extensive audio presentation than a transcript. This kind of de- scription can be both subjective and artistic, depending on the needs of the use. Governmental legislation in some countries requires that these textual alternatives be synchronized with the audio presentation. For a very lengthy transcript or de- scription, consider providing a link to a separate page containing a complete tran- script or description.

**Video.** For all video content include a transcript or description in both a textual and audio format. As mentioned above, a t*ranscript is* an exact word-for-word version of the video. A *description,* also both subjective and artistic*,* is a longer and more extensive summarization of the video. It generally includes actions, settings, body language, and scene changes necessary to fully understand the video. Again, gov- ernmental legislation in some countries requires that these alternatives be syn- chronized with the video presentation.

**Image maps.** For content embedded in image maps provide equivalent textual menus because their content may not be accessible to screen-review utilities.

**Tables.** For tables, provide alternate ways to access their content. It may be difficult for sight-impaired users to navigate within a table using screen-review utilities.

**Online forms.** For online forms that cannot be read by screen utilities, provide alter- nate methods of communication. For example, provide instructions for supplying needed information by telephone, regular mail, or e-mail.

**Animation.** Screen reviewers cannot read information that is animated. Provide an option that enables users to stop animation. Also, ensure that the information conveyed by the animation is available in an alternate format.

**Text-only pages.** If accessibility cannot be accomplished in any other way, provide a text-only page with equivalent information and functionality. These pages must be maintained and updated in conjunction with the primary Web page.

**WWW Consortium.** The World Wide Web Consortium (www.w3.org) is establish- ing guidelines for Web content accessibility. Follow the existing guidelines, and new guidelines as they are approved and become available.

Other useful and thorough references for creating accessible systems are found in the Web sites of Microsoft [(www.msdn.microsoft.com/library/books)](http://www.msdn.microsoft.com/library/books)) and IBM (www-3.ibm.com).

###### Documentation

* Provide documentation on all accessible features.
* Provide documentation in alternate formats.
* Provide online documentation for people who have difficulty reading or handling printed material.

All accessibility features must be documented for the user. Much standard documen- tation does not address keyboard access as thoroughly as is required by disabled people. Also, some people have difficulties in reading or handling printed material. Documen- tation in alternate formats, such as audio or Braille, may be required.

###### Testing

* Test all aspects of accessibility as part of the normal system testing process.

Testing for accessibility must be part of the normal testing process.

The design of graphical systems and Web pages, and their screens, is a complicated process. As has been shown, in both a host of factors must be considered. In graphical systems among the many design elements are the types of windows used, the way the windows are organized, what controls are selected to collect and present information, and the way the controls are organized within one window and between several win- dows. Web page design factors include the proper integration of text, graphics, naviga- tion links, and controls, page size, writing for simplicity and clarity, the characteristics of browsers and monitors, and accessibility requirements. In both design processes nu- merous design trade-offs will be made. Also, some design decisions may be based on skimpy data and reflect the most educated guess possible at the moment. Finally, the implications for some design decisions may not be fully appreciated until the results can be seen.

To wait until after a system has been implemented to uncover and correct any system usability deficiencies can be aggravating, costly, and time-consuming for both users and developers. Indeed, after implementation many problems may never be corrected because of time constraints and costs. To minimize these kinds of problems, and ensure usability, interfaces must be continually tested and refined before they are implemented. What follows is an overview of the testing process and the role it plays in design. Its purpose is to provide an awareness of the testing procedures and methods, and to

summarize some basic testing guidelines. Testing steps to be reviewed are:

Identifying the purpose and scope of testing. Understanding the importance of testing.

Developing a prototype.

Developing the right kind of test plan. Designing a test to yield relevant data.

Soliciting, selecting, and scheduling users to participate. Providing the proper test facility.

Conducting tests and collecting data.

Analyzing the data and generating design recommendations. Modifying the prototype as necessary.

Testing the system again. Evaluating the working system.

The Purpose of Usability Testing

Usability testing serves a twofold purpose. First, it establishes a communication bridge between developers and users. Through testing, the developer learns about the user’s goals, perceptions, questions, and problems. Through testing, the user is exposed to the capabilities of the system early on, before design is solidified.

Second, testing is used to evaluate a product. It validates design decisions. It also can identify potential problems in design at a point in the development process where they can be more easily addressed. Testing also enables comparison of alternate versions of a design element, when a clear direction is not immediately evident. How well the in- terface and screens meet user needs and expectations can also be assessed.

Thorough testing also has one other benefit for the developer. It can prevent the massive embarrassment that often results from letting things “slip through the cracks.”

#### The Importance of Usability Testing

A thorough usability testing process is important for many reasons, including all of the following.

**Developers and users possess different models.** As discussed earlier, developers and users have different expectations and levels of knowledge. Specialized knowl- edge possessed by the developers enables them to deal with complex or ambigu- ous situations on the basis of context cues not visible to the users. Developers also frequently use terminology that does not always match that of the users.

**Developer’s intuitions are not always correct.** The intuition of designers, or anyone for that matter, no matter how good or bad they may be at what they do, is error prone. This is illustrated by the previously reported Tullis and Kodimer (1992) study evaluating several screen-based controls. They found that programmers’ predictions of control usage speed correlated only .07 with actual measured speeds. They also found that programmers’ predictions of user control preferences corre- lated only .31 with actuality. Intuition is too shallow a foundation on which to base design decisions.

**There is no average user.** We all differ—in looks, feelings, motor abilities, intellectual abilities, learning abilities and speeds, device-based control preferences, and so forth. In a keyboard data entry task, for example, the best operators will probably be twice as fast as the poorest and make 10 times fewer errors. The design must permit people with widely varying characteristics to satisfactorily and comfort- ably learn and perform the task or job.

**It’s impossible to predict usability from appearance.** Just as it is impossible to judge a person’s personality from his or her looks, it’s impossible to predict a system’s usability from its appearance.

**Design standards and guidelines are not sufficient.** Design standards and guidelines are an important component of good design, laying the foundation for consis- tency. But design standards and guidelines often fall victim to trade-offs. They also cannot address all the countless design element interactions that occur within a completed system.

**Informal feedback is inadequate.** Informal feedback is a hit-and-miss proposition. Parts of the system may be completely overlooked; significant problems in other parts may never be documented.

**Products’ built-in pieces almost always have system-level inconsistencies.** This is a normal and expected result when different developers work on different aspects of a system. We might also say that developers differ—there is no average developer.

**Problems found late are more difficult and expensive to fix.** Unless they’re really severe, they may never be fixed.

**Problems fixed during development mean reduced support costs later.** Support costs are directly proportional to the usability problems that remain after develop- ment. The more problems, the higher the support costs.

**Advantages over a competitive product can be achieved.** Many products can do something. The most successful products are those that permit doing something easily.

#### Scope of Testing

Testing should begin in the earliest stages of product development and continue throughout the development process. It should include as many of the user’s tasks, and as many of the product’s components, as reasonably possible. Always involve all mem- bers of the design team in the testing to ensure a common reference point for all. Involv- ing all also permits multiple insights into the test results from the different perspectives of team members.

#### Prototypes

A prototype is primarily a vehicle for exploration, communication, and evaluation. Its purpose is to obtain user input in design, and to provide feedback to designers. Its major

function is the communicative role it plays, not accuracy or thoroughness. A prototype enables a design to be better visualized and provides insights into how the software will look and work. It also aids in defining tasks, their flow, the interface itself, and its screens.

A prototype is a simulation of an actual system that can be quickly created. A proto- type may be a rough approximation, such as a simple hand-drawn sketch, or it may be interactive, allowing the user to key or select data using controls, navigate through menus, retrieve displays of data, and perform basic system functions. A prototype need not be perfectly realistic, but it must be reasonably accurate and legible. A prototype also need not be functionally complete, possessing actual files or processing data. Today, many software support tools for prototyping are available that permit the prototype to be integrated directly into the application code.

A prototype may have great breadth, including as many features as possible to pre- sent concepts and overall organization, or it might have more depth, including more detail on a given feature or task to focus on individual design aspects. By nature, a pro- totype cannot be used to exercise all of a system’s functions, just those that are notable in one manner or another.

Particularly useful early in design, a prototype should be capable of being rapidly changed as testing is performed. A prototype is characterized by its fidelity, the exact- ness and thoroughness of its replication of a system’s screens and user interaction. Prototypes range in fidelity from low to high, from rough hand-drawn sketches to fully functioning software (Microsoft, 1995; Weinschenk, 1995; Winograd, 1995). Various kinds of prototypes, in general order of increased fidelity, are as follows.

##### Hand Sketches and Scenarios

* Description:
  + Screen sketches created by hand.
  + Focus is on the design, not the interface mechanics.
  + A low-fidelity prototype.
* Advantages:
  + Can be used very early in the development process.
  + Suited for use by entire design team.
  + No large investment of time and cost.
  + No programming skill needed.
  + Easily portable.
  + Fast to modify and iterate.
  + A rough approximation often yields more substantive critical comments.
  + Easier to comprehend than functional specifications.
  + Can be used to define requirements.
* Disadvantages:
  + Only a rough approximation.
  + Limited in providing an understanding of navigation and flow.
  + A demonstration, not an exercise.
  + Driven by a facilitator, not the user.
    - Limited usefulness for a usability test.
    - A poor detailed specification for writing the code.
    - Usually restricted to most common tasks.

**Description.** The first, and simplest, prototype is a low-fidelity rough hand-drawn sketch, or mock-up, of the screens. These can start early in the design process and before any attempt is made to create a prototype using an available toolkit or in- terface builder. With sketches, the focus is on the design of individual screens, not the interface mechanics. The hand sketch should be an entity that has enough of a general look to suggest the functionality, interaction, and layout of screens. The goal is a rough vision, not a polished work of art. This sketch will be useful in defin- ing and refining task organization, conceptual ideas, and the general layout of screens.

MYTH The design is finished.

**Advantages.** Hand-drawn sketches of screens can be easily developed and used very early in the development process. Many usability problems can then be identified and corrected quickly. Sketches are also suitable for use by the entire develop- ment team, giving everyone a sense of what the real design issues are. Sketches re- quire no large investment of time and money, and they are portable, placing few restrictions on where the testing may occur. Sketches can also be quickly modified and iterated, as many times as necessary. Because there has been no emotional investment in “code” and the status quo, there is no necessity for team members to defend something already created from hard work. Screen sketches are rough approximations, and rough approximations often yield more substantive sugges- tions or critical comments than actual screen-drawn versions. Their “draft” or “un- polished” look greatly softens the attitude that everything is cast in concrete. Sketches can also be used to define a system’s requirements.

**Disadvantages**. Since hand-drawn sketches are rough approximations, they can only suggest the final layout of the interface. They are limited in helping understand system navigation and flow, and are a demonstration device driven by a facilita- tor, with the user assuming a more passive role. They are usually restricted to the most common user tasks. As a result, they possess limited usefulness for usability testing.

###### Sketch Creation Process\*

* Sketch (storyboard) the screens while determining:
  + The source of the screen’s information.
  + The content and structure of individual screens.
  + The overall order of screens and windows.

\* Based on Weinschenk (1995).

* Use an erasable medium.
* Sketch the screens needed to complete each workflow task.
* Try out selected metaphors and change them as necessary.
* First, storyboard common/critical/frequent scenarios.
  + Follow them from beginning to end.
  + Then, go back and build in exceptions.
* Don’t get too detailed; exact control positioning is not important, just overall order and flow.
* Storyboard as a team, including at least one user.
* Only develop online prototypes when everyone agrees that a complete set of screens has been satisfactorily sketched.

Sketch the screens while determining the source of the screen’s information, the con- tent and structure of the individual screens, and the overall flow of the screens. Use an erasable medium so that as ideas are explored and changed, the modifications will be easy to make. Sketch the screens needed to complete each task in the workflow. First, sketch the most common, critical, or frequent activities, following them from beginning to end. Then, go back and build in the exceptions. Try out all selected metaphors and modify them as necessary. Make sure the major user objects are very obvious. Avoid get- ting too detailed. Most important is the overall screen flow, order, and structure. Ap- proximate the positioning of controls simply to verify fit. Exact positioning will come later. Sketch the screens as a team, including at least one user. To avoid solidifying the product too soon, develop online prototypes only when everyone agrees a complete set of screens have been satisfactorily sketched.

##### Interactive Paper Prototypes

* Description:
  + Interface components (menus, windows, and screens) constructed of common paper technologies (Post-It notes, transparencies, and so on).
  + The components are manually manipulated to reflect the dynamics of the software.
  + A low-fidelity prototype.
* Advantages:
  + More illustrative of program dynamics than sketches.
  + Can be used to demonstrate the interaction.
  + Otherwise, generally the same as for hand-drawn sketches and scenarios.
* Disadvantages:
  + Only a rough approximation.
  + A demonstration, not an exercise.
  + Driven by a facilitator, not the user.
  + Limited usefulness for usability testing.

**Description.** Another simple low-fidelity prototype involves use of common office supplies such as Post-It notes, transparencies, markers, and scissors. Menu bars, pull-down menus, pop-up windows, screen bodies, and so on reflecting system tasks are created using these media. Then, the components are manually manip- ulated to illustrate the dynamics of the software. The purpose of this kind of pro- totype is to provide a sense of the dynamics of a program without actually having to build a functional version of it. The objective, again, is to create a rough vision, not a polished piece of art, in order to elicit substantive suggestions and critical comments. The goal is not to show how the system or Web site will look, but to as- sess if people can easily use it.

**Advantages.** Interactive paper prototypes are more illustrative of program dynam- ics than simple screen sketches. System components can be manipulated to demonstrate the interactive nature of the system. The other paper prototype ad- vantages are generally the same as those described for hand-drawn sketches and scenarios.

**Disadvantages.** The disadvantages are similar to those for hand-drawn sketches. They are only rough approximations, are demonstrations and not exercises, are driven by a facilitator not the user, and possess limited usefulness for usability testing.

##### Programmed Facades

* Description:
  + Examples of finished dialogs and screens for some important aspects of the system.
  + Created by prototyping tools.
  + Medium-fidelity to high-fidelity prototypes.
* Advantages:
  + Provide a good detailed specification for writing code.
  + A vehicle for data collection.
* Disadvantages:
  + May solidify the design too soon.
  + May create the false expectation that the “real thing” is only a short time away.
  + More expensive to develop.
  + More time-consuming to create.
  + Not effective for requirements gathering.
  + Not all of the functions demonstrated may be used because of cost, schedule lim- itations, or lack of user interest.
  + Not practical for investigating more than two or three approaches.

**Description.** To provide a realistic surface view of a real program and to illustrate some of the program’s functioning, a programmed facade can be created. Using prototyping tools such as Hypercard, Supercard, and Toolbook, examples of fin- ished dialogs and screens for some important aspects of the system are con- structed and viewed. A facade is very shallow, duplicating only a small portion of

what is ultimately intended in both depth and breadth. Because of this shallow nature, it is best described as a medium-fidelity to high-fidelity prototype.

**Advantages.** While much is missing underneath, what is visible can provide a good impression of finished design. Programmed facades also provide a good detailed specification for writing code, and can be a vehicle for the actual collection of data.

**Disadvantages.** First, a highly polished product can foster a sense of finality because of its appearance. Significant reorganization or restructuring suggestions may be inhibited, with the focus simply being on screen cosmetics. Second, the false ex- pectation that the real thing is only a short time away may easily be created, even though much work still remains to be done. Programmed facades are also much more expensive to develop than paper-based prototypes, and they take much longer to create. Also not all of the functions demonstrated may eventually be used because of cost, schedule limitations, or lack of user interest, and they are not practical for investigating more than two or three alternate design approaches.

##### Prototype-Oriented Languages

* Description:
  + An example of finished dialogs and screens for some important aspects of the system.
  + Created through programming languages that support the actual programming process.
  + A high-fidelity prototype.
* Advantages:
  + May include the final code.
  + Otherwise, generally the same as those of programmed facades.
* Disadvantages:
  + Generally the same as for programmed facades.

**Description.** To present an example of completed dialogs and screens for some parts of the system, prototypes can be constructed using programming languages that support the actual programming process. Examples include Power Builder, Visual Basic, and so on. Using these tools, a high-fidelity, real program can be created to illustrate some of the program’s functioning and the mechanics of the interface. One consideration to be decided up front: Will the prototype software be a “throw- away,” or the actual system software? This will have implications concerning the amount of programming effort expended on the prototype.

**Advantages.** The greatest advantage of this kind of prototype is that it may include the actual code needed for the system. Otherwise, advantages are generally the same as those of programmed facades.

**Disadvantages.** Like a programmed facade, a danger is that the highly polished prod- uct can foster a sense of finality because of its appearance, inhibiting reorganization or restructuring suggestions, the focus simply being on screen cosmetics. Other disadvantages are also similar to those of programmed facades.

##### Comparisons of Prototypes

The fidelity of the prototypes described above varies from low to high. Does fidelity af- fect a prototype’s usefulness as a testing tool? Two recent studies have addressed this issue. The first study, by Catani and Biers (1998), examined prototypes created at three fidelity levels: low (paper), (medium) screen shots, and high (using a prototype-oriented language—Visual Basic). There were no significant differences in the number and sever- ity of problems identified with each kind of prototype. There was also a high degree of commonality in the specific problems uncovered.

The second study, reported by Uceta, Dixon, and Resnick (1998), compared a paper- based prototype with a computer-based prototype. Both interfaces were identical except for the medium of presentation. Most of the usability problems were found using both approaches, the results being statistically comparable. Identifying problems using the paper prototype, however, took about 30 percent longer than using the computer-based prototypes.

The results of these studies indicate that prototype fidelity seems to have no impact on the identification of usability problems. Other prototyping issues (prototype creation time and cost, testing time, and so on) remain important issues in usability testing, how- ever. It seems reasonable that any system development effort should use combinations of these prototyping techniques throughout the entire design cycle in order to visualize the design, solicit users’ input, and obtain needed feedback for the developers. Succes- sively increasing the complexity of the prototypes as development moves toward the final solution, allows users to continually get a better idea of how the system will actu- ally look and work. This will give them greater and greater insights into how the system may be improved.

#### Kinds of Tests

A test is a tool that is used to measure something. The “something” may be: Conformance with a requirement.

Conformance with guidelines for good design. Identification of design problems.

Ease of system learning. Retention of learning over time. Speed of task completion.

Speed of need fulfillment. Error rates.

Subjective user satisfaction.

A test is usually formal; it is created and applied intentionally and with a purpose. It is usually based upon some kind of criteria, an understanding of what a good result would be. Several testing techniques, at varying levels of sophistication and cost, are available to exercise the system.

##### Guidelines Review

* Description:
  + A review of the interface in terms of an organization’s standards and design guidelines.
* Advantages:
  + Can be performed by developers.
  + Low cost.
  + Can identify general and recurring problems
  + Particularly useful for identifying screen design and layout problems.
* Disadvantages:
  + May miss severe conceptual, navigation, and operational problems.

**Description.** A *guidelines review* is an inspection of an interface’s navigation and screen design and layout in the context of an organization’s standards and design guide- lines. A checklist summarizing a system’s standard or guideline document is prepared and is used as the basis for the comparison. Failure to comply with a guideline or standard indicates that a design modification may be necessary.

**Advantages.** Software developers can perform this kind of test. Its advantages in- clude its low cost and its ability to identify recurring and general problems. It is particularly useful in identifying screen design and layout problems.

**Disadvantage.** Because this review tends to be static in nature, it may miss severe conceptual, navigation, and operational problems.

##### Heuristic Evaluation

* Description:
  + A detailed evaluation of a system by interface design specialists to identify problems.
* Advantages:
  + Easy to do.
  + Relatively low cost.
  + Does not waste user’s time.
  + Can identify many problems.
* Disadvantages:
  + Evaluators must possess interface design expertise.
  + Evaluators may not possess an adequate understanding of the tasks and user communities.
  + Difficult to identify systemwide structural problems.
  + Difficult to uncover missing exits and interface elements.
  + Difficult to identify the most important problems among all problems uncovered.
  + Does not provide any systematic way to generate solutions to the problems uncovered.
* Guidelines:
  + Use 3 to 5 expert evaluators.
  + Choose knowledgeable people:
    - Familiar with the project situation.
    - Possessing a long-term relationship with the organization.

**Description.** In a *heuristic evaluation*, interface specialists study a system in depth and look for properties they know, from experience, will lead to problems. The interface is judged for its compliance with recognized usability principles, the heuristics*.*

**Advantages.** A heuristic evaluation is relatively easy to do and its cost is fairly low (See “Heuristic Evaluation Process” below). It does not waste the user’s valuable time. It also can identify many usability problems (See the research studies below).

**Disadvantages.** The evaluators should possess user interface expertise; this greatly re- stricts the population from which evaluators can be are chosen. Because of the small size of this available group, evaluators may not possess an adequate understanding of the system tasks and the user communities. (Even expert reviewers have great difficulty knowing how typical users, especially first-time users, will really behave.) With this method of evaluation, it is difficult to identify deeper design prob- lems, including systemwide structural problems, missing exits, and missing inter- face elements or features. It is also difficult to identify the most important problems among those documented. Finally, this method does not provide any systematic

way to generate solutions to the problems uncovered.

**Guidelines.** Based upon a study, Nielsen (1992) suggests that the optimum expert group size to satisfactorily perform a heuristic evaluation is 3 to 5 people. He found that different evaluators tend to find different problems, and one person will never be able to find the number of usability problems that several people will. Nielsen also found that including more than five evaluators does not gain enough additional usability information for the extra cost involved. Ideally, evaluators used should be familiar with the project situation and possess a long-term relationship with the developing organization. This suggestion is often difficult to achieve, however.

###### Heuristic Evaluation Process

* Preparing the session:
  + Select evaluators.
  + Prepare or assemble:
    - A project overview.
    - A checklist of heuristics.
  + Provide briefing to evaluators to:
    - Review the purpose of the evaluation session.
    - Preview the evaluation process.
    - Present the project overview and heuristics.
    - Answer any evaluator questions.
    - Provide any special evaluator training that may be necessary.
* Conducting the session:
  + Have each evaluator review the system alone.
  + The evaluator should:
* Establish own process or method of reviewing the system.

— Provide usage scenarios, if necessary.

* Compare his or her findings with the list of usability principles.
* Identify any other relevant problems or issues.
* Make at least two passes through the system.
  + Detected problems should be related to the specific heuristics they violate.
  + Comments are recorded either:
* By the evaluator.
* By an observer.

— The observer may answer questions and provide hints.

* + Restrict the length of the session to no more than 2 hours.
* After the session:
  + Hold a debriefing session including observers and design team members where:
* Each evaluator presents problems detected and the heuristic it violated.
* A composite problem listing is assembled.
* Design suggestions for improving the problematic aspects of the system are discussed.
  + After the debriefing session:
* Generate a composite list of violations as a ratings form.
* Request evaluators to assign severity ratings to each violation.
* Analyze results and establish a program to correct violations and deficiencies.

**Preparing the session**. First, as described above, *select* 3 to 5 evaluators. Ideally, the evaluators used should be familiar with the project situation and possess a long- term relationship with the developing organization. Then, *prepare or assemble* a proj- ect overview and a checklist of heuristics. A useful checklist may be available from one or more of the evaluators. Examples of checklists are found in Tables 14.2 and

14.3. Finally, provide a *briefing* to the evaluators to review the purpose of the eval- uation session, to preview the entire evaluation process that is being undertaken, to present the project overview and heuristics, and to answer any questions the eval- uators may have. Any special evaluator training that may be necessary can also be provided at this time. This briefing session will normally consume 45 to 60 minutes.

**Conducting the session**. Each evaluator should inspect the system *alone*, not with or in the presence of other evaluators. This is to ensure independent and unbiased evaluations from each evaluator. Allow the evaluators to establish their *own process or method* of reviewing the system. Ideally, let the evaluator use the system without procedural assistance, browsing as is felt necessary. If, however, the eval- uators are not familiar with the system’s content and purpose, they may be pro- vided with scenarios listing the steps a user would take to perform a sample set of realistic user tasks.

During the session, the evaluators will compare their findings with the list of usability principles. *Detected problems* should be related to the specific heuristics

they violate. Multiple heuristics can be linked to any one identified problem. Other relevant problems or issues can also be noted. Two or more passes should be made through the system.

Evaluator *comments* can be recorded either by the evaluator or by an observer. Evaluator-written comments require additional effort by the evaluator during the review process and can break the continuity of the evaluation process. Having an observer record the evaluator’s verbal comments adds to the overhead of the ses- sion, but reduces the evaluator’s workload and allows greater focusing on the review process. If the same observer is used for all evaluation sessions, session re- sults should be available faster since the observer only needs to understand and organize one set of personal notes, not multiple reports or reviews written by all session participants. The observer may answer questions and provide hints to the evaluator, as necessary. Evaluators not familiar with the system’s content will occasionally need advice. Also precious time will not be wasted by the evaluator’s struggling with the interface’s mechanics. Observer comments should be restricted to situations where the evaluator is clearly confused or in trouble.

MAXIM Not even the most brilliantly conceived and ingenious computer system can do all that it was designed to do—or even a small part of what it was designed to do—unless the brilliance of its operation and purpose is matched by the cunning simplicity of its user interface. (Hicks and Essinger, 1991)

Finally, to minimize evaluator fatigue, restrict the *length of a session* to about 2 hours. For large or complicated interfaces that require more time to evaluate, it is better to split the session into several sessions, each concentrating on a part of the interface.

**After the session.** When all evaluator and/or observer notes have been complied, hold a *debriefing session,* no more than 2 hours in length. Include all observers and design team members. Have each evaluator present the problems detected and the heuristic each violated. Assemble a composite list of usability problems (if one has not yet been compiled). Solicit and discuss design suggestions for improving the problematic aspects of the system.

*After the debriefing session,* form the composite list of violations into a ratings form. Request the evaluators to assign severity ratings to each violation on a scale ranging from “usability catastrophe” to “not a usability problem,” as shown in Table 14.1. It is difficult to obtain these estimates during the evaluation process, where the focus is on finding problems. Then, analyze the results and establish a program to correct the necessary violations and deficiencies. The ratings will identify the most serious problems, the first in need of attention.

###### Heuristic Evaluation Effectiveness

One of the earliest papers addressing the effectiveness of heuristic evaluations was by Nielsen (1992). He reported that the probability of finding a *major* usability problem av- eraged 42 percent for single evaluators in six case studies. The corresponding probabil- ity for uncovering a *minor* problem was only 32 percent. He also found that the actual

**Table 14.1** Severity Ratings in Heuristic Evaluation

0 = I don’t agree that this is a usability problem at all.

1 = A cosmetic problem only. Need not be fixed unless extra time is available. 2 = A minor usability problem. Fixing should be given a low priority.

3 = A major usability problem. Important to fix and should be given a high priority. 4 = A usability catastrophe. Imperative to fix before the product can be released. From useit.com

number of located minor problems exceeded the number of major problems uncovered by about a 3:1 ratio (152 minor problems versus 59 major problems). Minor design problems are normally more prevalent in design (for example, inconsistent use of typefaces) and, consequently, show up in greater numbers in this kind of evaluation. Minor problems, such as inconsistencies, are more susceptible to identification by inspection, and may not be noticed in testing with actual users, who are focusing on a task or procedure. Nielsen suggests that severity ratings are especially useful in prioritizing minor problems.

For a number of years, Bailey (2001) has questioned the effectiveness of heuristic evaluations, as now conducted. In a 1992 article (Bailey, Allan, and Raiello, 1992) he sug- gested that many of the “problems” identified by heuristic evaluations were really not problems at all. He recently bolstered his argument by reporting on three studies look- ing at their effectiveness (Catani and Biers, 1998; Rooden, Green, and Kanis, 1999; Stan- ton and Stevenage, 1998). In each study, he determined what the heuristic evaluators thought the problems were, and then he compared these determinations with the prob- lems users actually had in performance testing. The results showed that about one-third (36 percent) of identified problems were actually usability problems, and almost one- half (43 percent) of the problems identified by evaluators *did not* turn out to be problems at all. The evaluators also missed about one-fifth (21 percent) of the problems users ac- tually had. He concluded that when a heuristic evaluation is conducted, about one-half of the problems identified will be problems, and one-half will not be problems.

Bailey ends his article not by suggesting that heuristic evaluations are bad, but that the heuristics themselves must be greatly improved. (The messenger is all right—the problem is the message.) He recommends establishing more solidly research-based heuristics. The most used set of heuristics can be traced back to a paper by Molich and Nielsen in 1990, and revised by Nielsen in 1994 (Bailey, 1999b). The research foundation of these papers is somewhat suspect. Bailey suggests a better research-based set of heuristics will lead to improved evaluation results, for example, those proposed by Gerhardt-Powals (1996). This set of heuristics is summarized in Table 14.2.

Web heuristics are still an evolving entity and must also be validated through re- search. The set proposed by Levi and Conrad (1996), and summarized in Table 14.3, seem a good place to start.

In conclusion, heuristic evaluations are useful in identifying many usability prob- lems and should be part of the testing arsenal. Performing this kind of evaluation before beginning actual testing with users will eliminate a number of design problems, and is but one step along the path toward a very usable system.

**Table 14.2** Research-Based Set of Heuristics

1. Automate unwanted workload.
   * Free cognitive resources for high-level tasks.
   * Eliminate mental calculations, estimations, comparisons, and unnecessary thinking.
2. Reduce uncertainty.
   * Display data in a manner that is clear and obvious.
3. Fuse data.
   * Reduce cognitive load by bringing together lower-level data into a higher-level summation.
4. Present new information with meaningful aids to interpretation.
   * Use a familiar framework, making it easier to absorb.
   * Use everyday terms, metaphors, and so on.
5. Use names that are conceptually related to functions.
   * Context-dependent.
   * Attempt to improve recall and recognition.
6. Group data in consistently meaningful ways to decrease search time.
7. Limit data-driven tasks.
   * Reduce the time needed to assimilate raw data.
   * Make appropriate use of color and graphics.
8. Include in the displays only that information needed by a user at a given time.
   * Allow users to remain focused on critical data.
   * Exclude extraneous information that is not relevant to current tasks.
9. Provide multiple coding of data where appropriate.
10. Practice judicious redundancy.

* To resolve the conflict between heuristics 6 and 8.

From Gerhardt-Powals (1996).

**Table 14.3** Possible Web Page Heuristics

1. Speak the user’s language.
   * Use familiar words, phrases, and concepts.
   * Present information in a logical and natural order.
2. Be consistent.
   * Indicate similar concepts through identical terminology and graphics.
   * Adhere to uniform conventions for layout, formatting, typefaces, labeling, and so on.
3. Minimize the user’s memory load.
   * Take advantage of recognition rather than recall.
   * Do not force users to remember key information across documents.
4. Build flexible and efficient systems.
   * Accommodate a range of user sophistication and diverse user goals.
   * Provide instructions where useful.
   * Lay out screens so that frequently accessed information is easily found.

(*continues*)

**Table 14.3** Continued

1. Design aesthetic and minimalist systems.

* Create visually pleasing displays.
* Eliminate information that is irrelevant or distracting.

1. Use chunking.

* Write materials so that documents are short and contain only one topic.
* Do not force the user to access multiple documents to complete a single thought.

1. Provide progressive levels of detail.

* Organize information hierarchically, with more general information appearing before more specific detail.
* Encourage the user to delve as deeply as needed, but to stop whenever sufficient information has been obtained.

1. Give navigational feedback.

* Facilitate jumping between related topics.
* Allow the user to determine his/her current position in the document structure.
* Make it easy to return to an initial state.

1. Don’t lie to the user.

* Eliminate erroneous or misleading links.
* Do not refer to missing information.

From Levi and Conrad (1996).

##### Cognitive Walkthroughs

* Description:
  + Reviews of the interface in the context of tasks users perform.
* Advantages:
  + Allow a clear evaluation of the task flow early in the design process.
  + Do not require a functioning prototype.
  + Low cost.
  + Can be used to evaluate alternate solutions.
  + Can be performed by developers.
  + More structured than a heuristic evaluation.
  + Useful for assessing “exploratory learning.”
* Disadvantages:
  + Tedious to perform.
  + May miss inconsistencies and general and recurring problems.
* Guidelines:
  + Needed to conduct the walkthrough are:
* A general description of proposed system users and what relevant knowledge they possess.
* A specific description of one or more core or representative tasks to be performed.
* A list of the correct actions required to complete each of the tasks.
  + Review:
    - Several core or representative tasks across a range of functions.
    - Proposed tasks of particular concern.
  + Developers must be assigned roles of:
    - Scribe to record results of the action.
    - Facilitator to keep the evaluation moving.
  + Start with simple tasks.
  + Don’t get bogged down demanding solutions.
  + Limit session to 60 to 90 minutes.

**Description.** In a *cognitive walkthrough*, developers walk through an interface in the context of representative user tasks. Individual task actions are examined and the evaluators try to establish a logical reason why the user would perform each examined action. Actions are compared to the user’s goals and knowledge. Dis- crepancies and problems are noted and analyzed and the tasks modified as neces- sary. Walkthroughs require that the task definition methodology must have been properly accomplished in the system requirements stage. The user’s goals and as- sumptions must also be clearly defined before the walkthrough is performed.

**Advantages.** Walkthroughs permit a clear evaluation of the task flow early in the de- sign process, before empirical user testing is possible. The earlier a design flaw can be detected, the easier it is to fix it. They can also be used to evaluate alternate design solutions. Walkthroughs are of low cost and can be performed by develop- ers. They are also more structured than a heuristic evaluation, being less likely to suffer from subjectivity because of the emphasis on user tasks. Walkthroughs are very useful for assessing “exploratory learning,” first-time use of a system without formal training.

**Disadvantages.** A cognitive walkthrough is tedious to perform, and it may miss in- consistencies and general and recurring problems.

**Guidelines.** Needed to conduct the walkthrough are a general description of pro- posed system users and what relevant knowledge they possess, a specific descrip- tion of one or more core or representative tasks to be performed, and a listing of the correct actions required to complete each of the tasks. The walkthrough should review several core or representative tasks across a range of functions, as well as proposed tasks of particular concern to the developers. Start with simple tasks to get the brain juices flowing. Don’t get bogged down looking for solutions to prob- lems. Avoid detailed screen designs; they often inhibit assessment of the big pic- ture. Limit a session to 60 to 90 minutes.

##### Think-Aloud Evaluations

* Description:
  + Users perform specific tasks while thinking out LOUD.
  + Comments are recorded and analyzed.
* Advantages:
  + Utilizes actual representative tasks.
  + Provides insights into the user’s reasoning.
* Disadvantages:
  + May be difficult to get users to think out loud.
* Guidelines:
  + Develop:
* Several core or representative tasks.
* Tasks of particular concern.
  + Limit session to 60 to 90 minutes.

**Description.** In a *think-aloud evaluation*, users perform specific tasks while thinking out load. The objective is to get the user to talk continuously. All comments are recorded so all thoughts are captured and subtle points are not missed when analysis occurs.

**Advantages.** This kind of evaluation utilizes actual representative tasks. Valuable in- sights into why the user does things are obtained.

**Disadvantages.** It may be difficult to get all people to think out loud.

**Guidelines.** Develop core or representative task scenarios, or scenarios of proposed tasks of particular concern. Limit a session to 60 to 90 minutes.

##### Usability Test

* Description:
  + An interface evaluation under real-world or controlled conditions.
  + Measures of performance are derived for specific tasks.
  + Problems are identified.
* Advantages:
  + Utilizes an actual work environment.
  + Identifies serious or recurring problems.
* Disadvantages:
  + High cost for establishing facility.
  + Requires a test conductor with user interface expertise.
  + Emphasizes first-time system usage.
  + Poorly suited for detecting inconsistency problems.

**Description.** A *usability test* evaluates an interface under real-world or controlled con- ditions. Specific tasks are performed by users, measures of performance taken, and the results compared with the previously defined performance goals. Evaluators also gather data on problems that arise. Errors, confusion, frustrations, and com- plaints can be noted and discussed with the user. It is also useful to have the user talk aloud about what he or she is doing. Failure to meet these usability design ob- jectives will indicate that redesign is necessary.

**Advantages.** Usability tests incorporate a realistic work environment. Tasks are per- formed in an actual work setting, either a usability laboratory or another controlled setting. A usability test can identify serious or recurring problems, avoiding low- priority items

**Disadvantages.** Its most serious problem is the high cost associated with establishing a facility to perform the testing. To effectively perform usability test also requires a test conductor with user interface expertise. A usability test also emphasizes first- time or early system use, collecting little data concerning use of a system by expe- rienced system users. These tests are also poorly suited to detecting problems with consistency.

##### Choosing a Testing Method

Unfortunately, there is little published detailed advice on which tests to use, when to use them, and which tests work best together. Beer, Anodenko, and Sears (1997) suggest a good pairing is cognitive walkthroughs followed by think-aloud evaluations. Using cognitive walkthroughs early in the development process permits the identification and correction of the most serious problems. Later, when a functioning prototype is avail- able, the remaining problems can be identified using a think-aloud evaluation.

A substantial leap forward in the testing process would be the creation of a software tool simulating the behavior of people. This will allow usability tests to be performed without requiring real users to perform the necessary tasks. One such example is a system, described by Hornof and Kieras (1997), called Executive Process Interactive Control (EPIC). Formal evaluations by a tool such as this have the potential to greatly improve the quality of many user interfaces.

In conclusion, each testing method has strengths and weaknesses. A well-rounded testing program will use a combination of some, or all, of these methods to guarantee the usability of its created product. It is very important that testing start as early as pos- sible in the design process and, continue through all developmental stages.

#### Developing and Conducting the Test

A usability test requires developing a test plan, selecting test participants, conducting the test, and analyzing the test results.

##### The Test Plan

* Define the scope of the test.
* Define the purpose of the test.
  + Performance goals.
  + What the test is intended to accomplish.
* Define the test methodology.
  + Type of test to be performed.
  + Test limitations.
  + Developer participants.
* Identify and schedule the test facility or location.
* Develop scenarios to satisfy the test’s purpose.

An inadequately planned test can waste time and money, and provide flawed or misleading information.

**Scope.** How extensive and detailed will the test be? A test’s scope will be influenced by a variety of factors. Determinants include the following issues: (1) The *design stage:* early, middle, or late—the stage of design influences the kinds of prototypes that may exist for the test, (2) the *time available* for the test—this may range from just a few days to a year or more, (3) *finances allocated* for testing—money allocated may range from one percent of a project’s cost to more than 10 percent, (4) the project’s *novelty* (well defined or exploratory)—this will influence the kinds of tests feasible to conduct, (5) expected *user numbers* (few or many) and *interface criticality* (life-critical medical system or informational exhibit)—much more testing depth and length will be needed for systems with greater human impact, and (6) finally, the development team’s *experience* and testing knowledge will also affect the kinds of tests that can be conducted.

**Purpose.** Define the purpose of the test in simple statements, including performance goals, what the test is expected to accomplish in system learning, use of screens, system navigation, and efficiency of operation. Learning issues will center on ef- fectiveness in starting to use the system, recognizing system features, and explor- ing system features. Screen issues will be directed toward general screen layout, including efficiency and aesthetics, and recognition and understanding of screen- based controls, icons, and messages. Navigational issues involve use of system navigation methods, including menus, buttons, and icons. Operational issues in- clude how many steps are needed to accomplish a task, system responsiveness, and forms of system feedback provided.

**Methodology.** Define the specific *type* of test to be carried out. Also identify any test *limitations*. Set reasonable expectations. If the test is measuring only a part of a sys- tem or its behavior, the results must be interpreted with this limitation in mind. Identify all *participants* from the development team.

**Test Facility or Location.** Select the location where the test will be conducted, in a usability lab, a conference room, or some other location. The location should be away from distractions and disturbances. If the test is being held in a usability lab- oratory, the test facility should resemble the location where the system will be used. It may be an actual office designated for the purpose of testing, or it may be a laboratory specially designed and fitted for conducting tests. The advantage of a laboratory from a data collection perspective is that it can be designed with one- way mirrors for observers and easily equipped with recording devices such as video cameras. The physical environment—lighting, temperature, and so on—can also be more precisely controlled.

MYTH Testing would be nice, but it is costly and we don’t have time for it.

**Scenarios.** Task scenarios to adequately satisfy the test’s purpose must be identified and constructed. Ideally, the entire system will be tested, but time and costs often limit what can actually be done. When time or cost constraints exist, good candi- dates for testing include the user’s most important tasks or the most representa- tive tasks. Always test the functions or features whose design foundation is not as solid as desired. These are features where the trade-off issues were not clear-cut, not strongly pointing to one design alternative over other possible alternatives.

After preparing task scenarios, try them out and refine them as necessary. Make sure they are clearly written and capable of being performed in the allo- cated testing time.

The testing of Web sites will follow the same methodology as that for graphical sys- tems. Web sites, however, present some unique issues that must also be addressed. Many of these important issues are summarized in Table 14.4. This listing is not intended to be all-inclusive. It is simply presented as a reminder of “what not to forget to do.”

**Table 14.4** Some Things to Remember to Test in Web Site Design

**TEST: ENSURE THAT:**

All the browsers, servers, and monitors Different servers and browsers don’t used. introduce adverse new behaviors.

Different monitors don’t negatively affect color appearance, legibility, and page size.

At different dial-up speeds. Download times are satisfactory for all users.

The navigation design. Users know how to find information.

The navigation sequence through related information makes sense.

Users know where they are in the Web site’s structure.

Users know how to return to visited points when they are browsing.

All links are working properly. Unnecessary links are found and removed.

The visual design style. The style reflects the site’s purpose.

The style creates a positive and proper impression of the organization owning the site.

The style is compatible with user tasks and needs.

Content legibility and readability. The design succeeds for all default fonts.

All grammar and spelling is correct.

## Create Meaningful Graphics,

Icons and Images

Icons

Icons are most often used to represent objects and actions with which users can interact with or that they can manipulate. These types of icons may stand alone on a desktop or in a window, or be grouped together in a toolbar. A secondary use of an icon is to re- inforce important information, a warning icon in a dialog message box, for example.

##### Kinds of Icons

The use of icons to reflect objects, ideas, and actions is not new to mankind. We’ve been there before. Early humans (100,000 years or so ago) used pictographs and then ideo- graphs to communicate. Some of these early communications can still be found today on rock walls and in caves around the world. Until recent times, this was also the only way to communicate in some cultures (Native Americans and Australian aborigines, for example).

Word writing is traced back to Chinese writing from about 6000 BC and Egyptian hi- eroglyphics from about 3000 BC. This was followed by cuneiform (Babylonia and Assyria) from about 1900 BC, and the contemporary Chinese vocabulary (numbering about 50,000) around 1500 BC. In 1000 BC the Phoenicians developed a 22-sign alphabet that the Greeks adopted about 800–600 BC. The Greeks passed this alphabet on to the Romans about 400 BC, who then developed a 23-character alphabet. This alphabet has been modified and embellished but has remained essentially the same for the last 2000 years. Pictorial representations, then, have played a prominent role in mankind’s history.

Word writing, however, unleashed much more flexibility and richness in communica- tion. This has caused some skeptics to wonder why, after taking 2500 years to get rid of iconic shapes, we have now revived them on screens.

Whatever the past, today, objects or actions *are* depicted on screens by icons. The term icon by itself, however, is not very specific and can actually represent very different things. An attempt has been made by some to define the actual types of icons that do exist. Marcus (1984) suggests icons fall into these categories:

**Icon.** Something that looks like what it means.

**Index.** A sign that was caused by the thing to which it refers.

**Symbol.** A sign that may be completely arbitrary in appearance.

He states that what are commonly referred to as icons may really be indexes or sym- bols. A true icon is something that looks like what it means. It is representational and easy to understand. A picture of a telephone or a clock on a screen is a true icon. An index is a sign caused by the thing to which it refers. An open door with a broken window in- dicates the possible presence of a burglar. The meaning of an index may or may not be clear, depending upon one’s past experiences. A symbol is a sign that may be completely arbitrary in appearance and whose meaning must be learned. The menu and sizing icons

Resemblance—An image that looks like what it means. Symbolic—An abstract image representing something.

Exemplar—An image illustrating an example or characteristic of something.

Arbitrary—An image completely arbitrary in appearance whose meaning must be learned.

Analogy—An image physically or semantically associated with something.

She suggests that an icon is *used* in a number of different ways: for *objects* such as a document, *object attributes* such as a color or fill pattern, *actions* such as to paste, *system* *states* such as ready or busy, and *message types* like critical or warning.

The different ways icons are used may then be represented by different design schemes. A *resemblance* icon is an image that looks like what it means—a book, for ex- ample, to represent a dictionary. This is equivalent to Marcus’s icon. A *symbolic* icon is an abstract image that represents something. A cracked glass, for example, can represent something fragile. Marcus’s symbol would be similar. An *exemplar* icon represents an example or characteristic of something. A sign at a freeway exit picturing a knife and fork has come to indicate a restaurant. An *arbitrary* icon is not directly related in any way and must be learned. Marcus’s symbol would be an equivalent. Finally, an *analogy* icon is an image physically or semantically associated with something—a wheelbarrow full of bricks for the move command, for example. Marcus’s symbol would also be similar. In a study looking at various kinds of icons, Rogers found that those depicting both an action and an object were quite effective. For example, a drawing of a page and an arrow pointing up means “go to the top of the page.” She also found that arbitrary icons were only meaningful in very small sets, and that icons based on analogies were rela-

tively ineffective.

##### Characteristics of Icons

An icon possesses the technical qualities of syntactics, semantics, and pragmatics (Mar- cus, 1984). *Syntactics* refers to an icon’s physical structure. Is it square, round, red, green, big, small? Are the similarities and differences obvious? Similar shapes and colors can be used to classify a group of related icons, communicating a common relationship. *Se- mantics* is the icon’s meaning. To what does it refer—a file, a wastebasket, or some other object? Is this clear? *Pragmatics* is how the icons are physically produced and depicted. Is the screen resolution sufficient to illustrate the icon clearly? Syntactics, semantics, and pragmatics determine an icon’s effectiveness and usability.

##### Influences on Icon Usability

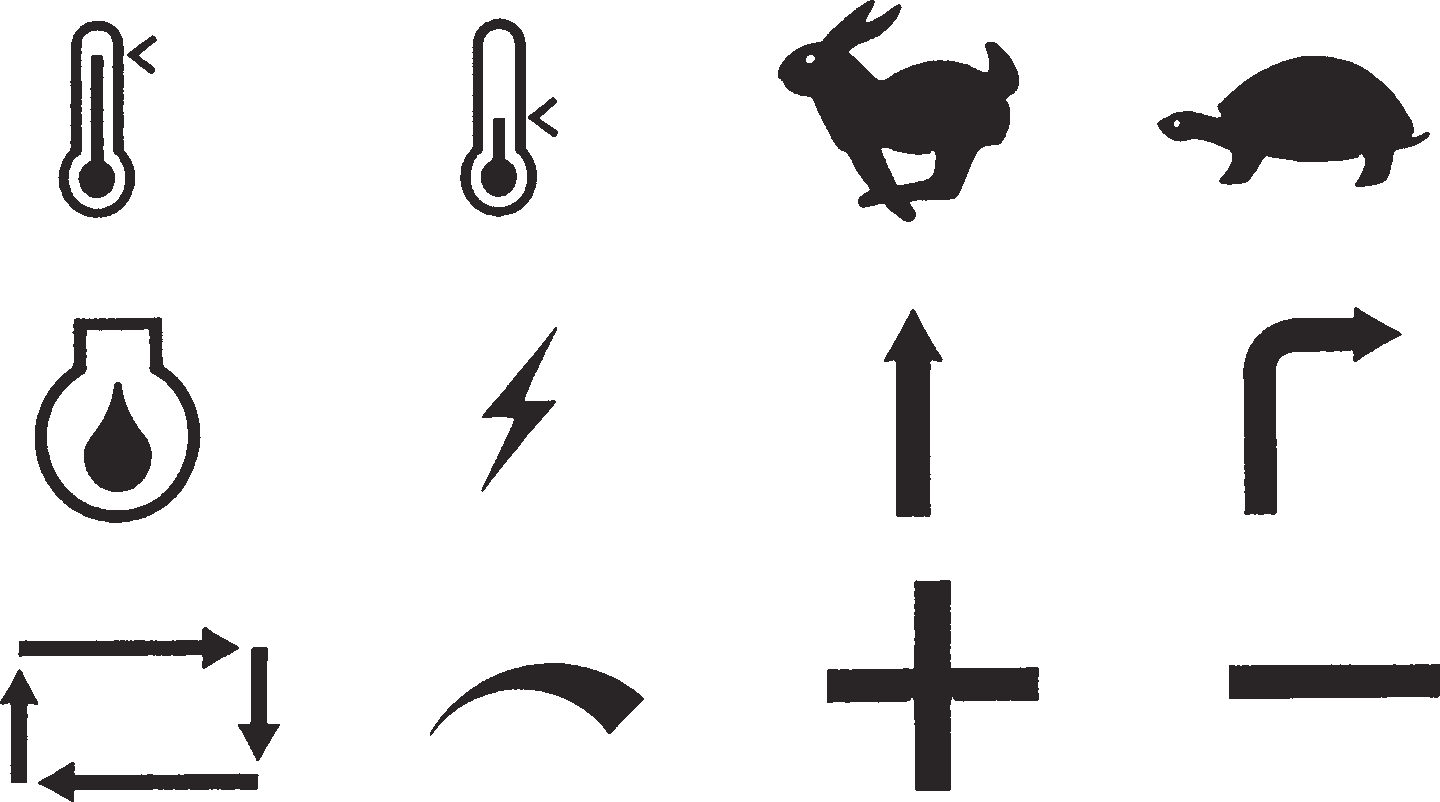
Simply providing an icon on a screen does the user no particular favor, unless it is care- fully designed to present a natural and meaningful association between the icon itself and what it stands for. Unfortunately, a sampling of many current systems finds icons

that do not achieve this objective. Icons are included because “this is the thing to do” in a graphical system today. Little concern is given to effectiveness. The result is too often a cluttered and confusing screen that is visually overwhelming. So, proper icon design is important from an acceptance, learning, and productivity perspective. The following factors influence an icon’s usability:

* Provide icons that are:
  + Familiar.
  + Clear and Legible.
  + Simple.
  + Consistent.
  + Direct.
  + Efficient.
  + Discriminable.
* Also consider the:
  + Context in which the icon is used.
  + Expectancies of users.
  + Complexity of task.

**Familiarity.** How familiar is the object being depicted? Familiarity will reduce learn- ing time. How familiar are the commonly seen icons in Figure 11.1? Lack of famil- iarity requires learning the icons’ meanings. Very unfamiliar icons require a great deal of learning.

Experience often makes words and numbers more familiar to a person than symbols. Confusion matrices have been developed through extensive research for alphanumeric data (0 versus O, 1 versus I). Graphic symbols may be more visually similar to each other.



**Figure 11.1** Some common icons. What do they stand for? Answers are on the next page. From Micro Switch (1984)

|  |  |  |  |
| --- | --- | --- | --- |
| The icons depicted in Figure 11.1 have the following meanings: | | | |
| Hot | Cold | Fast | Slow |
| Engine Oil | Ammeter/Generator | Straight | Turn |
| Automatic | Variable Regulation (Increase/Decrease) | Plus/Positive | Minus/Negative |

**Clarity.** Is the icon legible? Does the shape, structure, and formation technique on the screen permit a clear and unambiguous depiction of what it is? Screen resolu- tion should be sufficiently fine to establish clear differences of form at the normal working distance. The resolution and pixel shapes for screens differ from one an- other. Icons must appear correctly and consistently no matter what kind of screen is used. If color is used, it should contrast well with the background. Poor clarity will lead to identification errors and slower performance.

**Simplicity.** Is the icon simple? Is the shape clean and devoid of unnecessary em- bellishments? Too many parts will only confuse the screen viewer.

**Consistency.** Are families of icons consistent in structure and shape? Are the same icons displayed on different screens consistent in shape and structure? Are the same icons displayed in different sizes also consistent in structure and shape?

**Directness.** How “sign-like” is the icon; how well does it convey its intended meaning? For concrete objects and actions, direct links are more easily established. Adjectives, adverbs, conjunctions, and prepositions can cause problems, however. Also, how does one easily convey concepts such as bigger, smaller, wider, or narrower?

**Efficiency.** In some situations, a graphics screen may be less efficient, consuming more screen display space than a word or requiring more physical actions by the user than text. A telephone directory of 50 names and numbers listed on an alphanumeric screen may consume the same screen space required for, and manipulation of, 15 file cards. Raising an arm or moving a mouse may be slower than simply typing. In other situations, icons can be more effective than words in communicating concepts in a smaller area of space. Icons’ strength lies in situations where this occurs.

**Discriminability.** The symbols chosen must be visually distinguishable from other symbols. A person’s powers of differentiation for shapes and other forms of codes have been experimentally determined over the years. The maximum number of codes that can be effectively differentiated by a human being, including geomet- ric shapes, is summarized in Table 11.1. A person’s ability to discriminate alpha- betic or alphanumeric information is much more potent.

**Context.** The context of a symbol may change its meaning. Does the rabbit symbol illustrated in Figure 11.1, if seen on a road sign in a national park, mean, “go faster”? From this contextual perspective, icons are similar to words.

**Expectancies.** The symbol may be comprehended, but a false conclusion may be reached about the desired action because of an incorrect expectancy. A study of international road signs found that eight percent of all drivers never saw the “do not do” slash through a symbol on a road sign. Their expectancy was that they could do it, not “not do it.”

**Table 11.1** Maximum Number of Codes for Effective Human Differentiation

|  |  |  |
| --- | --- | --- |
| **ENCODING METHOD** | **RECOMMENDED MAXIMUM** | **COMMENTS** |
| Alphanumerics | Unlimited | Highly versatile. |
|  |  | Meaning usually self-evident. |
|  |  | Location time may be longer than for |
|  |  | graphic coding. |
| Geometric Shapes | 10–20 | High mnemonic value. |
|  |  | Very effective if shape relates to object |
|  |  | or operation being represented. |
| Size | 3–5 | Fair. |
|  |  | Considerable space required. |
|  |  | Location time longer than for colors and |
|  |  | shapes. |
| Line Length | 3–4 | Will clutter the display if many are used. |
| Line Width | 2–3 | Good. |
| Line Style | 5–9 | Good. |
| Line Angle | 8–11 | Good in special cases (such as wind |
|  |  | direction). |
| Solid and Broken Lines | 3–4 | Good. |
| Number of Dots or Marks | 5 | Minimize number for quick assimilation. |
| Brightness | 2–3 | Creates problems on screens with poor |
|  |  | contrast. |
| Flashing/Blinking | 2–3 | Confusing for general encoding but the |
|  |  | best way to attract attention. |
|  |  | Interacts poorly with other codes. |
|  |  | Annoying if overused. |
|  |  | Limit to small fields. |
| Underlining | No data | Useful but can reduce text legibility. |
| Reverse Video | No data | Effective for making data stand out. |
|  |  | Flicker easily perceived in large areas, |
|  |  | however. |
| Orientation (location on | 4–8 | — |
| display surface) |  |  |
| Color | 6–8 | Attractive and efficient. |
|  |  | Short location time. |
|  |  | Excessive use confusing. |
|  |  | Poor for the color blind. |
| Combinations of Codes | Unlimited | Can reinforce coding but complex |
|  |  | combinations can be confusing. |

Data derived from Martin, 1973; Barmack and Sinaiko, 1966; Mallory et al., 1980; Damodaran et al., 1980; and Maguire, 1985.

**Complexity of task.** The more abstract or complex the symbol, the more difficult it is to extract or interpret its intended meaning. It has been found that more concrete graphic messages are easier to comprehend than the more abstract. Icons, there- fore, cannot completely replace words in more complex situations.

##### Choosing Icons

Icon design is an important process. Meaningful and recognizable icons will speed learn- ing and recall and yield a much more effective system. Poor design will lead to errors, delays, and confusion. While the art of icon design is still evolving, it is agreed that the usability of a system is aided by adhering to the following icon design guidelines.

###### A Successful Icon

* + - Looks different from all other icons.
    - Is obvious what it does or represents.
    - Is recognizable when no larger than 16 pixels square.
    - Looks as good in black and white as in color.

Fowler and Stanwick (1995) provide these general guidelines. An icon must look dif- ferent from all other product icons, making it discriminable and differentiable. What it does or represents must also be obvious so it is interpretable. It must be recognizable when no larger than 16 pixels square. Finally, it must look as good in black and white as in color. Color is always an enhancing quality of an icon.

###### Size

* + - Supply in all standard sizes.
      * 16  16 pixels.
        + 16- and 256-color versions.
      * 32  32 pixels
        + 16- and 256-color versions.
        + Effective: 24  24 or 26  26 in 32  32 icon.
      * 48  48 pixels
        + 16- and 256-color versions.
    - Use colors from the system palette.
    - Use an odd number of pixels along each side.
      * Provides center pixel around which to focus design.
    - Minimum sizes for easy selection:
      * With stylus or pen: 15 pixels square.
      * With mouse: 20 pixels square.
      * With finger: 40 pixels square.
    - Provide as large a hot zone as possible.

**Size.** Typically, icons come in three standard sizes, 16, 32 and 48 pixels square. For clarity, 16 x 16 should be an icon’s minimum size. An effective combination for an image is a 24  24 or 26  26 in a 32-pixel square icon.

**Colors.** Microsoft suggests that while 256 colors may be used in the smaller sizes than 48  48 pixels, to do so increases icon storage requirements, and they may not be displayable on all computer configurations. If 256 colors are used for icons, they suggest that the standard 16-color format should *always* be provided. Also, use colors from the system palette to ensure that the icons look correct in all color configurations.

**Odd number of pixels.** Horton (1994) recommends using an odd number of pixels along each side of the matrix. This provides a center pixel around which to focus, thus simplifying the design process.

**Icon selection.** For easy selection the following are minimum icon sizes: with a sty- lus or pen, 15 pixels square; with a mouse, 20 pixels square; with one’s finger, 40 pixels square.

**Hot zone.** An icon’s hot zone, the area within it that allows it to be selected, should be as large as possible, preferably the entire size of the icon. This allows easier selection.

##### Choosing Images

* Use existing icons when available.
* Use images for nouns, not verbs.
* Use traditional images.
* Consider user cultural and social norms.

**Existing icons.** Many standard icons have already been developed for graphical sys- tems. Use these standard icons where they are available. This will promote con- sistency across systems, yielding all the performance benefits that consistency provides. Where standard icons are not available, determine if any applicable icons have already been developed by industries and trade or standards organizations. The International Standards Organization (ISO), for example, has developed stan- dard shapes for a variety of purposes. Always consult all relevant reference books before inventing new symbols or modifying existing ones.

**Nouns.** An object, or noun, is much easier to represent pictorially than an action or verb. Choose nouns for icons whenever possible.

**Traditional images.** Old-fashioned, traditional images often work better than newer ones. They have been around longer, and more people recognize them.

**Cultural and social norms.** Consider users’ cultural and social norms. Improper de- sign of icons can create problems internationally. Social norms vary, so great vari- ations exist in what is recognizable and acceptable throughout the world. What one culture recognizes may have no meaning in another. What is acceptable in one

country may not be in another. International considerations were thoroughly de- scribed in Step 10 “Provide Acceptable Internationalization and Accessibility.”

##### Creating Images

* Create familiar and concrete shapes.
* Create visually and conceptually distinct shapes.
  + Incorporate unique features of an object.
  + Do not display within a border.
* Clearly reflect objects represented.
* Simply reflect objects represented, avoiding excessive detail.
* Create as a set, communicating relationships to one another through common shapes.
* Provide consistency in icon type.
* Create shapes of the proper emotional tone.

**Concrete and familiar shapes.** Ideally, an icon’s meaning should be self-evident. This is enhanced when concrete shapes are provided, those that look like what they are. An icon should also be intuitive or obvious, based upon a person’s preexisting knowledge. Familiar shapes are those images that are well learned. Figure 11.2 il- lustrates concrete and familiar icons for a file folder, book, and telephone as well as images for the same objects that are more abstract and unfamiliar. A study found that concrete, familiar icons were preferred to abstract, unfamiliar ones.

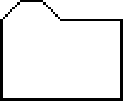
Keep in mind, however that familiarity is in the eye of the viewer. The concrete images pictured may be familiar to us, readers of this book, but not to a tribal chief living in a remote area of the world where these objects do not exist. Similarly, items familiar to those working on the factory floor may not be at all familiar in the office or in the home, and vice versa. Mayhew (1992) also cautions that some

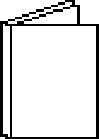
File Folder

Book

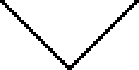
Telephone

Concrete/Familiar

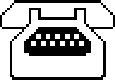
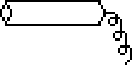




Abstract/Unfamiliar





**Figure 11.2** Concrete and familiar shapes.

abstract images should not be discounted because they have become familiar, in spite of their being abstract. On a road sign, for example, an angled red bar in- scribed over an object means do not do what is pictured beneath (at least to most people, as described earlier). While abstract, it is a very familiar shape today. If an abstract image must be used, it should be capable of being learned quickly and eas- ily recalled. Familiarity can only be determined through knowing one’s user.

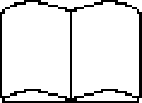
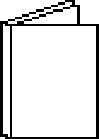
**Visually and conceptually distinct shapes.** It must be easy to tell icons apart so the chances of confusing them are minimized. Differentiation is aided when icons are visually different from one another. It is also aided when icons are conceptually different—that is, when they portray specific features of an object that are rela- tively unique within the entire set of objects to be displayed. Figure 11.3, based upon Mayhew (1992), illustrates how distinctiveness may be achieved for two sim- ilar items, a dictionary and a telephone book. Visual distinctiveness is achieved by incorporating unique features of each: for the dictionary, it is its content of letters and words, for the telephone book, numbers and the telephone bell.

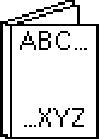
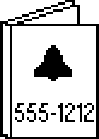
Visual distinctiveness is degraded when borders are placed around icons, as illustrated in Figure 11.4. Borders tend to obscure the shape of the object being displayed.

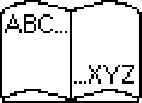
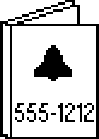
**Clearly reflect objects.** The characteristics of the display itself should permit drawings of adequate quality. Poorly formed or fuzzy shapes will inhibit recognition.

DICTIONARY TELEPHONE BOOK

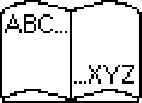
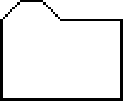
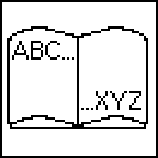
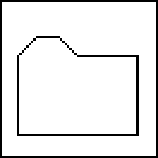
Conceptually Visually

Similar Distinct  

Distinct Similar  

Distinct Distinct

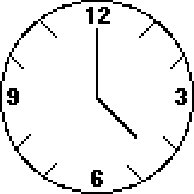
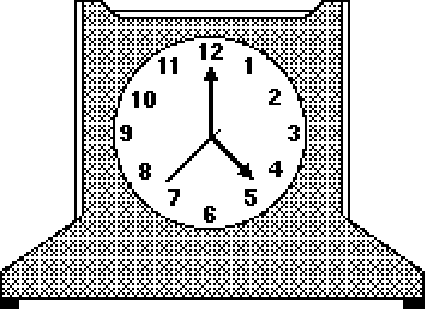
**Figure 11.3** Visually and conceptually distinct shapes.



**Figure 11.4** Borders degrading icon distinctiveness.

**Simply reflect objects.** Construct icons with as few graphical components as neces- sary, using no more than two or three, if possible. Also, use simple, clean lines, avoiding ornamentation. Byrne (1993) found that simple icons, icons containing fewer graphical elements, were located faster in a visual search task than complex icons, icons with more components. He concluded that complex icons seemed to clutter a screen with information that people were unable to employ to their ad- vantage. Too much detail inhibits rather than facilitates perception, as illustrated in Figure 11.5. For real-world objects, use only enough detail to permit recognition and recall.

**Design as a set.** Do not design icons in isolation, but as a family considering their re- lationships to each other and the user’s tasks. Provide a common style. When icons are part of an overall related set, create shapes that visually communicate these re- lationships. Objects within a class, for example, may possess the same overall shape but vary in their other design details, as illustrated in Figure 11.6. Color may also be used to achieve this design goal. In creating sets, always avoid repeating un- related elements.



POOR GOOD

**Figure 11.5** Avoid excessive detail in icon design.

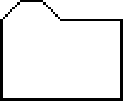
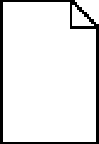
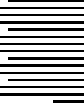


Chart Letter Document File

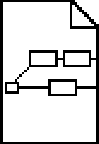
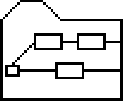
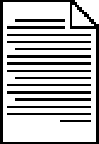
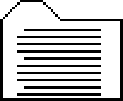
 

Chart in Document



Letter in Document

Chart in File



Letter in File

**Figure 11.6** Communication relationships in icons.

**Consistency in icon type.** As previously noted, there are many different kinds of de- sign schemes for icons (resemblance, symbolic, arbitrary, and so on). All these schemes might be used to create a meaningful family of icons for an application. Learning the meaning of icons and searching for the right icon, however, will be aided if the same design scheme is used for all icons within a family. In present- ing a series of icons for actions such as paint, cut, and so on, one could, for exam- ple, (1) depict a before-after representation of the action, (2) depict the action itself being performed, or (3) picture the tool to perform the action. While a series of meaningful icons could be developed using each scheme, the best approach would be to use only one of these schemes to develop the entire family of icons.

**Proper emotional tone.** The icon should appropriately reflect the environment in which it is used. A sewage disposal system would be an inappropriate metaphor for an electronic mail system wastebasket.

##### Drawing Images

* Provide consistency in shape over varying sizes.
* Do not use triangular arrows in design to avoid confusion with other system symbols.
* When icons are used to reflect varying attributes, express these attributes as mean- ingfully as possible.
* Provide proper scale and orientation.
* Use perspective and dimension whenever possible.
* Accompany icon with a label to assure intended meaning.

**Consistency.** When drawing images, create consistency in shapes for identical icons of differing sizing. Preserve the general shape and any distinctive detail. Consis- tency is achieved through limiting the variations of angles, line thicknesses, shapes, and amount of empty space.

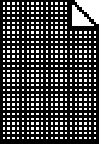
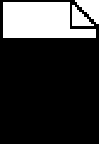
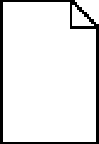
**Triangular arrows.** Avoid using a triangular graphic similar to that used as a cascade symbol for menus, a drop-down button for controls, and scroll arrows. The simi- larity may cause confusion.

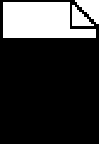
**Meaningful attributes.** When an icon is also used to express an attribute of an object, do this as meaningfully as possible. The status of a document, for example, might be represented by displaying it in a different shade, but would be more effectively illustrated by filling it in, as illustrated in Figure 11.7. Shading requires remem- bering what each specific type of shading stands for; the filled-in proportion is more intuitively obvious.

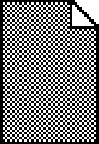
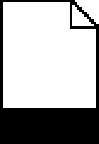
**Scale and orientation.** Ensure that the size and orientation are consistent with other related objects. Also ensure that they fit well on the screen.

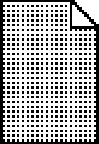
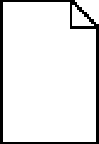
**Perspective and dimension.** Use lighting and shadow to more accurately reflect the real-world experiences of people. When a light source is used, it must be located upper left, as is done with other screen elements.

POOR GOOD

Report Finished  Report Mostly Completed



Report Started

Report Planned

**Figure 11.7** Expressing attributes in icon design.

**Caption or label.** Since icons may not be used often, the ability to comprehend, learn, and recall an icon’s meaning can be greatly improved by attaching textual cap- tions or labels to them The preferred label location is directly beneath the icon, not within it, because of the international considerations discussed in Step 10. Labels beneath the icon also provide a larger target, speeding selection. Labels should al- ways be related to icons in a consistent positional way. “Mystery icons,” icons with no caption or label to explain them, lead to a user guessing game and many errors. While ToolTips can be used to present labels, they are time-consuming to present, taking about two-thirds of a second to appear and be comprehended. Scanning an entire row of 15 icons with ToolTips, therefore, will consume about 10 extra seconds.

MAXIM If people must remember hieroglyphics, they won’t stick around long.

##### Icon Animation and Audition

* Animation:
  + Use:
    - To provide feedback.
    - For visual interest.
  + Make it interruptible or independent of user’s primary interaction.
  + Do not use it for decoration.
  + Permit it to be turned off by the user.
  + For fluid animation, present images at 16 or more frames per second.
* Audition:
  + Consider auditory icons.

**Animation.** Recent research has explored the use of bringing to life on screens the icons representing objects and actions. An animated icon appears to move instead of maintaining a static position on the screen. Animation can take two forms, best described as static and dynamic. A *static* icon’s appearance is unchanged over a period of time and changes only at the moment that a system event occurs. An ex- ample would be the open door of a mailbox shutting when an electronic message is received. A *dynamic* icon’s movement is independent of a system event, chang- ing appearance to represent functions, processes, states, and state transitions. An example is an icon that begins movement to illustrate an action when a pointer is moved close to it.

Animation can be used to provide feedback and to create visual interest. Re- searchers caution, however, that there are many outstanding issues. Among them are that few animation creation rules exist, prototyping is difficult, a scheme for how they fit into a larger system is lacking, and whether they can be made useful for more complex and abstract concepts is not known. Morimoto, Kurokawa, and Nishimura, (1993) found that dynamic animation of the type in the example above did not increase the comprehensibility of icons. Its only advantage was its enter- tainment value.

Some general guidelines, however, seem appropriate. First, do not prevent the user from interacting with the system while the animation is performed. Un- less the animation is part of a process, it should be independent of what the user is doing. It should also be interruptible. Be conservative in its use; do not use ani- mation simply for decoration. It can be very distracting or annoying. Finally, pro- vide the user with the option of turning it on or off, as desired

Microsoft recommends that to achieve fluidity in movement, images should be presented at a speed of at least 16 frames per second. The reader interested in more information on animation is referred to Baecker and Small (1990).

**Audition.** Objects make sounds as they are touched, dragged, bumped against one another, opened, activated, and thrown away. Auditory icons are computer sounds replicating everyday sound-producing events. When a printer near one’s desk begins printing, the sound of the printing mechanism is heard. This provides auditory feedback that a print operation one has just asked for has successfully started. An auditory icon would be the same sound, generated by the computer. Another example would be to convey information about an object’s dimensions. If a file is large, it can sound large. If an object is dragged over a new surface, the new surface is heard. If an ongoing process starts running more quickly, it sounds quicker.

Sounds can convey information about many events in computer systems, per- mitting people to listen to computers as we do in the everyday world. It may be well suited to providing information:

About previous and possible interactions. Indicating ongoing processes and modes. Useful for navigation.

To support collaboration.

*Auditory icons* are distinct from *earcons*, abstract synthetic tones used in struc- tured combinations to create sound messages. Auditory icons may also be sus- ceptible to the distracting influences that sounds can cause to listeners, especially others. The use of sound was more fully discussed in Step 9 “Provide Effective Feedback and Guidance and Assistance.” The reader in need of more information on auditory icons is referred to Garver (1993).

##### The Design Process

* Define the icon’s purpose and use.
* Collect, evaluate, and sketch ideas.
* Draw in black and white.
* Draw using an icon-editing utility or drawing package.
* Test for user:
  + Expectations.
  + Recognition.
  + Learning.
* Test for legibility.
* Register new icons in the system’s registry.

**Define purpose.** To begin the design process, first define the icon’s purpose and use. Have the design team brainstorm about possible ideas, considering real-world metaphors. Simple metaphors, analogies, or models with a minimal set of con- cepts are the best places to start in developing icons.

**Collect, evaluate, and sketch ideas.** Start by designing on paper, not on the computer (Fowler and Stanwick, 1995). Ask everyone to sketch his or her ideas. Do not worry about too much detail; exact pixel requirements are not necessary at this time.

**Draw in black and white.** Many icons will be displayed in monochrome. Color is an enhancing property; consider it as such.

**Test for expectation, recognition, and learning.** Choosing the objects and actions, and the icons to represent them, is not a precise process, and will not be easy. So, as in any screen design activity, adequate testing and possible refinement of de- veloped images must be built into the design process. Icon recognition and learn- ing should both be measured as part of the normal testing process.

**Test for legibility.** Verify the legibility and clarity of the icons in general. Also, ver- ify the legibility of the icons on the screen backgrounds chosen. White or gray backgrounds may create difficulties. An icon mapped in color, then displayed on a monochrome screen, may not present itself satisfactorily. Be prepared to re- draw it in black and white, if necessary.

**Register new icons in the system’s registry.** Create and maintain a registry of all system icons. Provide a detailed and distinctive description of all new icons.

##### Screen Presentation

* Follow all relevant general guidelines for screen design.
* Limit the number of symbols to 12, if possible, and at most 20.
* Arrange icons:
  + In a meaningful way, reflecting the organization of the real world.
  + To facilitate visual scanning.
  + Consistently.
* Place object and action icons in different groups.
* Present an interactive icon as a raised screen element.
* Ensure that a selected icon is differentiable from unselected icons.
* Permit arrangement of icons by the user.
* Permit the user to choose between iconic and text display of objects and actions.

In designing, or establishing, screen layout rules; adhere to the following presenta- tion rules.

**General guidelines.** Follow all relevant general guidelines for screen design. Icons are but one part of a larger picture.

**Number of icons.** A person’s ability to identify shapes is limited (see Figure 11.1). A literature review, suggest using no more than eight to twelve or so functions that require icons at one time. At most, present no more than 20. If labels are attached to icons, however, the meaning of the icon is greatly clarified. Too many icons on a screen, though, will greatly increase screen clutter and create confusion. In gen- eral, fewer are better.

**Arranging icons.** Organize icons in a way that reflects the *real-world* organization of the user. Place object icons and action icons within different groupings. *Visual scanning* studies, in a non-iconic world, universally find that a top-to-bottom scan of columnar-oriented information is fastest. Generalization of these findings to an icon screen may not necessarily be warranted if icons have attached labels. Columnar orientation icons (with labels below the icons) will separate the labels from one another by the icons themselves. The labels will be farther apart and fewer icons will fit in a column than in a horizontal or row orientation. A row ori- entation would seem to be more efficient in many cases, as adjacent icons will be in closer physical proximity. Until research evidence is established to the contrary, organizing icons either in a column or a row seems appropriate. In either case, a *consistent* straight eye movement must be maintained through the icons.

**Object and action icons.** Conceptually similar items should always be arrayed to- gether. Locating them will be easier.

**Interactive icons.** To provide a visual indication that an icon is interactive or “click- able,” present it in a three-dimensional state raised from the screen background.

**Selected icon.** Ensure a selected icon is visually differentiable from unselected icons. One common method to achieve this is to present the selected icon in a three- dimensional “pressed” state.

**User arrangement.** Allow the user to arrange the icons in a manner that is meaning- ful for the task. A default arrangement should be provided, however.

**Iconic or text display.** In some situations, and for some users, pure text labels may be more meaningful than icons. The option to display text only should always be provided.

#### Multimedia

The graphical flexibility of the Web permits inclusion of other media on a screen, in- cluding images, photographs, video, diagrams, drawings, and spoken audio. The avail- ability of these additional interface elements has, however, been a double-edged sword. On the one hand, the various media can be powerful communication and attention-getting techniques. Multimedia can hold the user’s attention, add interest to a screen, entertain, and quickly convey information that is more difficult to present tex- tually. It can also make the Web much more accessible to people with disabilities. On the other hand, effective use of multimedia in design has been hindered by a lack of knowledge concerning how the various media may best be used, and a scarcity of ap- plied design guidelines. (GUI guidelines relevant to Web page design have been avail- able for years, but their existence was either unknown or ignored.) Effective multimedia use has also been hindered by the “let’s use it because we have it” attitude exhibited by many designers. (To be fair, early GUI design has suffered from the same problem.) The resulting usability problems, user confusion and frustration, poor screen legibility, and slow downloads, and so on have created situations where the user was too often denied an efficient and meaningful Web experience.

As a result, recent studies (Spool et al., 1997, for example) have found that the most

difficult to use Web sites were those that were graphically intense, and the top Web sites were characterized by little, if any, multimedia. Studies have also found that for users, text is currently a much more important Web site component than graphics. (At least at this stage in Web evolution.) Today, consequently, good interface design em- ploys multimedia in a conservative and appropriate manner. The objective is good in- teraction design, not “sparkle.” In the future, experts say, multimedia elements will be much better integrated with browsers, alleviating many of today’s usability problems.

##### Graphics

* Use graphics to:
  + Supplement the textual content, not as a substitute for it.
  + Convey information that can’t be effectively accomplished using text.
  + Enhance navigation through:
* Presenting a site overview
* Identifying site pages.
* Identifying content areas.
* Limit the use of graphics that take a long time to load.
* Coordinate the graphics with all other page elements.

Graphics contained in Web pages serve several distinct purposes, which can be clas- sified as follows:

**Navigational.** To identify links that may be followed. **Representational**. To illustrate items mentioned in the text. **Organizational.** To depict relationships among items mentioned in text. **Explanative.** To show how things or processes work.

**Decorative.** To provide visual appeal and emphasis.

Graphics must always be used for a specific purpose. This purpose must be deter- mined before designing or choosing the graphic itself. Graphics should only be used when they add to a Web site’s message. Graphics that do not relate to a Web site’s pur- pose, and do not strengthen the Web site’s message should never be used.

**Supplement textual content.** Use graphics to supplement text, not as a substitute for it. Graphics are not easily accessible to search facilities and screen reviewers, and are slower to download than text. As studies have shown, people prefer textual page content to graphical content. So, never use graphics when text will do the job. If a graphic will help people understand the text they are reading, then certainly use it.

**Convey information not possible using text.** Use graphics to convey information that can’t be effectively conveyed using text. In some cases the old adage “a pic- ture is worth a thousand words” is indeed true. Photographs, for example, can be used to communicate the exact appearance of objects. Video is useful for showing objects or things that move. Diagrams can be used to present an object’s structure. Drawings are useful when selected parts of an object need to be emphasized or represented. If a graphic does a better job of communicating an idea or concept than text, then use it. (Remember, however, text descriptions or transcripts of the graphic will always be necessary for accessibility reasons.)

**Enhance navigation.** Graphics can be used to enhance navigation. A graphical *overview* of a site’s organizational scheme will enable the user to conceptualize and learn the site’s structure faster than can be done through textual overviews. *Site pages* can be related through a consistent graphical theme carried from page to page. This will reinforce the browsing user’s sense of place. Graphics can also be used to identify and represent major site *content areas*. The experienced user will locate and identify the content areas faster using meaningful graphical iden- tifiers rather than text.

**Limit long-loading graphics.** Limit the use of graphics that take a long time to load In general, all graphics must be smaller on the Web than on the printed page. Large graphics take longer to download testing the user’s patience. If a large graphic is needed, present a small version and link it to a page containing the large version. Richly colored graphics and pages containing numerous graphics are also slower to load.

**Coordinate graphics.** Graphics are only one component of a Web page. The graph- ics must fit in with the style of typography used, the colors used, and the page lay- out itself. Use of plain and simple fonts are best coordinated with simple graphics. Realistic graphics work best with elements like three-dimensional effects and more complex typography.

##### Images

* General:
  + Use standard images.
  + Use images consistently.
  + Produce legible images.
  + Provide descriptive text or labels with all images.
  + Distinguish navigational images from decorative images.
  + Minimize:
    - The number of presented images.
    - The size of presented images.
      * Restrict single images to 5K.
      * Restrict page images to 20K.
      * Provide thumbnail size images.
* Image animation.
  + Avoid extraneous or gratuitous images.
* Color:
  + Minimize the number of colors in an image.
* Format:
  + Produce images in the most appropriate format.
    - GIF.
    - JPEG.
* Internationalization:
  + Provide for image internationalization.
* Screen design:
  + Reuse images on multiple pages.

**Standard images.** Whenever possible, use standard images that have already been developed and tested. This will promote consistency across systems, yielding all the performance benefits that consistency provides. These standard images may be found in guideline books, company or organizational documentation, or in

industry, trade, or standards organizations’ documentation. The International Standards Organization (ISO), for example, has developed standard image shapes for a variety of purposes. Always consult all relevant reference books before in- venting new images or modifying existing ones.

**Consistency.** Use an image consistently throughout an application or Web site. Mul- tiple images with the same meaning will be difficult to learn.

**Legibility.** Create legible images, images that are easy to identify from a variety of viewing distances and angles. Legibility is affected by a number of factors, in- cluding contrast with the background, image complexity, and image size. Images with a minimum amount of detail are usually easier to comprehend and faster to load. If an image with more detail is needed, provide a link to a page containing the detailed version. An image that is perfectly legible when it is drawn or ren- dered large may, when shrunk for placing on a page, become incomprehensible.

**Descriptive text or labels.** Many images are not immediately clear, even if well de- signed. The ability to comprehend, learn, and recall an image’s meaning, espe- cially if it is used for navigation, can be greatly improved by providing images with descriptive text or labels. Also, many people browse the Web with their graphics turned off. Without alternate, text an image’s purpose and function will not be known. Alternate text for an image also provides the following benefits:

It provides vision-impaired users with access to content through a screen- review utility.

It helps sighted users determine whether they want to wait for the image to fully load.

It enables users to read a description of a linked image and activate the link before image fully loads.

**Navigational and decorative images.** Clearly indicate which graphical images on the screen are used for navigation by providing a visual indication that an image is interactive or “clickable.” Possibilities include giving the image a raised or three- dimensional appearance (like a navigational icon) or underlining any descriptive text contained within it (like a textual link). Navigational images that cannot be dis- tinguished from decorative images force users to “mouse over” each image to de- termine which are interactive. (Once they are over their initial state of confusion.) This is time-consuming, and important navigation links may be missed.

**Minimize number of images.** The more images presented on a Web page, the slower the download time. Use text whenever possible

**Minimize size of images.** Oversized images also take a long time to load. Slow- loading graphics rarely add value to text, and people often don’t bother to stick around for them. The design goal is to produce images that load quickly. Make the graphic as small as possible while still retaining sufficient image quality. In general, restrict *single images* to 5K, *page images* to no more than 20K. A 200K file can take several minutes to load. Never put borders around an image with a drawing program because this also adds to the file size.

**Thumbnail size.** A thumbnail is a small version of an image, usually fairly low in quality. This small image will load quickly because of its small file size. Link this

thumbnail image to a large high-quality version of the image. Users can then de- cide whether or not they want to retrieve and view the full-size version. Always let the user know the size of the full-size image. Thumbnails are especially useful when several images, or a collection of images, must be displayed on a Web page.

**Minimize animation.** Animated images take a long time to load and are distracting to many people. Only use animation when it serves a useful purpose.

**Extraneous or gratuitous images.** Similarly, do not present extraneous or gratuitous images. Images take longer to load than text, and Web users prefer text. Images must always serve a useful purpose.

**Minimize the number of colors.** To reduce the size of image files, reduce the size of the color palette and the number of colors in the image. Color-rich images tend to be large. If the image color palette is too small, however, the image will be de- graded. The objective is to retain sufficient image quality while making the file as small as possible. To create images of sufficient color quality while at the same time reducing file size, begin with a high-quality image and create versions using successively smaller color palettes. Stop when the image degradation becomes ap- parent. (Guidelines for the use of color in screen design are addressed in Step 12 “Choose the Proper Colors*.*”)

**Appropriate format.** Produce images in the most appropriate format, GIF or JPEG. CompuServe developed the GIF format (Graphics Interchange Format) in 1987. The JPEG (Joint Photographic Experts Group) was developed for the transfer of photographic images over the Internet.

**GIF.** Most Web color images and backgrounds are GIF files. They are usually smaller and load faster than JPEGs. They are particularly useful for images that contain flat areas of color. Since GIFs are limited to 256 colors, they are ideal for graphics that use only a few colors. GIFs exist in either a *dithered* or *nondithered* format. Dithering is the color-mixing process a computer goes through when it encounters a color not in its palette. In this process palette colors are mixed to approximate the appearance of the desired color. The resulting color may be grainy or unac- ceptable. The dithering will be most apparent in gradations, shadows, and feath- ered edges. A nondithered GIF attempts to match the closest colors from the palette to the image. This is referred to as *banding.* This banding may also create an un- acceptable image.

One way to control the dithering process is to create images that only use non- dithering colors. The 216 colors that are shared by PCs and Macintoshes are called the Web palette or browser-safe colors. These colors display properly across all platforms without dithering.

GIFs may also be *interlaced.* Interlacing is the gradual display of an image in a series of passes on the screen. The first pass displays a low-resolution out-of focus image and each succeeding pass creates a clearer view until finally a complete image is displayed. With interlacing, users see a complete, although not clear, image much more quickly. An impression that the image is loading much faster is achieved, and users can quickly determine if they are interested in the image. With a *noninterlaced* GIF, the graphic unfolds more slowly one row at a time. Use interlaced GIFs to give users a preview of graphics while they unfold.

Most Web servers call up to four GIFs at a time for display. Limiting GIF im- ages on a page to four will allow pages to load much faster.

**JPEG.** JPEG formats are superior for images such as photographs that contain nu- merous changes in color tonality. They look best on monitors capable of display- ing 16 million colors. A JEPG’s range of colors cannot be produced in monitors displaying 256 or fewer colors. Images that contain flat areas of color may also find that JEPGs introduce unwanted artifacts. JPEGs usually take longer to down- load than GIFs.

JPEGs may be displayed as progressive or standard. *Progressive* images gradu- ally fade into view like interlaced GIFs, each pass an increasingly higher quality scan. *Standard* images are drawn from top to bottom like noninterlaced GIFs. Use progressive JPEGs to give users a preview of the graphics while they are unloaded.

**Internationalization.** When designing for international or multilingual users, using images may eliminate the need for translating words. All images, however, must comply with the internationalization design guidelines covered in Step 10.

**Reuse images.** Repeat the same images on multiple pages. Repeated images will be stored in a *cache*, the browser’s temporary storage area. Loading an image from cache significantly reduces an image’s downloading time.

###### Image Maps

* Use:
  + To provide navigation links to other content.
* Advantages:
  + Can be arrayed in a meaningful and obvious structure.
  + Faster to load than separate images.
* Disadvantages:
  + Consume a significant amount of screen space.
  + “Hot spots” not always obvious.
  + One’s location within image map is not always obvious.
* Guidelines:
  + Use with caution.
  + Provide effective visual cues and emphasis to make it easy to identify link boundaries.
  + Ensure image maps are accessible to the vision impaired.

**Use**. An *image map* is a complete image containing individual segments with naviga- tion links to other content. It primary use is to present a meaningfully structured image within which the links are contained.

**Advantages/disadvantages.** An *advantage* of an image map is its meaningful and ob- vious structure. It can reflect the user’s mental model of an object, minimizing or- ganizational learning requirements. An image map may be a map of a country, for

example, with areas reflecting regions that can be selected as links to more detailed content. An image map can also be an image reflecting a site’s organization. Image maps, because of their graphical nature, can aid conceptualization of a Web site and how it is organized. Another image map advantage is that they are faster to load than individual images, at least for users accessing the Web through a modem. There are several *disadvantages* of image maps. First, they are quite wasteful of screen space. Providing large enough “hot spots” or “clickable” areas for each ele- ment often necessitates creating very large maps. Within the maps, clickable re- gions are also not always obvious because they cannot be seen. Whether to click on the map, or where to click, is not always known. This can be confusing for the new user. Unclear or poorly designed image maps can cost users a great deal of time when they make erroneous navigation selections. Selected image map links are also not obvious to the user. A link just selected may be again selected, direct- ing the user right back to the page displayed with no indication that anything has changed. User confusion can again exist. Another disadvantage is that search fa-

cilities may not be able to index an image map.

MYTH Cool = Usable

**Guidelines.** Because of these disadvantages, be cautious in the use of image maps. Some experts recommend not using them at all. If used, provide effective visual cues and emphasis to make it easy to identify individual selectable segments and where link boundaries exist. Consider supplementing the image map graphic with text to inform users what they will see when they select a particular area. Fi- nally, ensure that image maps are accessible to vision-impaired users.

##### Photographs/Pictures

* + - Use:
      * When every aspect of the image is relevant.
    - Guidelines:
      * Use JPEG format.
      * On the initial page:
        + Display a small version.

A thumbnail size image.

Zoom-in on most relevant detail.

* + - * + Link to larger photos showing as much detail as needed.

**Use.** When every aspect of an image or object is relevant, present a picture or pho- tograph of it. A photo or picture will capture all visible aspects.

**Guidelines.** The *JPEG* format was developed for presenting photographs that con- tain numerous changes in color tonality. Pictures or photos look best on monitors capable of displaying 16 million colors.

A large photo will have an excessively long downloading time. To minimize this time, on the initial page display a small version of the photo and provide a link to a larger, high-quality, complete photo on another page. The small version may be a *thumbnail* image, a complete miniature photograph, usually fairly low in quality. Because of the complexity of a photographic image, a thumbnail may not always be legible. When legibility is a problem, instead of resizing the image to a minia- ture photo, provide a *zoom-in* on the most relevant photo detail, cropping and re- sizing as necessary to provide a meaningful and legible image.

For linked full-size photographs provide as much detail as the users need and always inform the users of the image’s size. Also, if necessary, provide a zoom or rotation capability for the photograph on the linked page.

##### Video

* Uses:
  + To show things that move or change over time.
  + To show the proper way to perform a task.
  + To provide a personal message.
  + To grab attention.
* Disadvantages:
  + Expensive to produce.
  + Slow to download.
  + Small and difficult to discern detail.
* Guidelines:
  + Never automatically download a video into a page.
  + Create short segments.
  + Provide controls, including those for playing, pausing, and stopping.
  + Consider using:
* Existing video.
* Audio only.
* A slide show with audio.

**Uses.** Video is especially suited to showing things that move or change over time. Examples include product demonstrations, how to repair a piece of equipment, or how to perform a dance step. Videos can also be used to present personal messages, although the speaker’s “presence” may not always have the desired emotional ef- fect. Because of their animation, videos can also be used to grab attention.

**Disadvantages.** Videos are expensive to produce and slow to download and play. They are also small and limited in the detail they can present. Always inform the user of a video’s size so a choice of whether or not to download it can be made. Depend- ing on a video’s purpose, its animation may also be distracting to the user.

**Guidelines.** Do not *automatically download* a video into a loading Web page. Create *short segments*. There are many distractions people may encounter while using a video (the telephone or interruptions by people, and so on), so long segments

should be avoided. A 60 to 90 second video is considered long so keep a video’s length well within these limits. For all playable files provide the following *controls*: Play, Pause/Resume, Stop, Rewind, Fast Forward, and Volume.

Because of a video’s disadvantages consider using existing videos, audio alone, or a slide show with audio. Reusing an *existing video* will save production time and money. A new voice-over may be all that is necessary. *Audio alone* may be as pow- erful a tool as a video, since the human voice is an important aspect of all videos. Determine whether audio alone will accomplish the video’s objectives. An *audio slide show* may also be a good substitute for a video. The impression of movement is still achieved as the slides change, but they are quicker and easier to create and download.

##### Diagrams

* + - Uses:
      * To show the structure of objects.
      * To show the relationship of objects.
      * To show the flow of a process or task.
      * To reveal a temporal or spatial order.
    - Guidelines:
      * Provide simple diagrams.
      * Provide cutaway diagrams or exploded views to illustrate key points.

**Uses.** Diagrams are useful for illustrating the structure of an object, its key parts and how they are related to each other. Diagrams are also useful for illustrating the re- lationships of objects, the structure of an organization, or the structure of a Web site. Other uses are to illustrate the flow of a process or task, a software program, or an airline passenger check-in sequence, for example. (Guidelines for displaying flow charts were presented in Step 3 “Understand the Principles of Good Screen Design.”) Diagrams can also be used to reveal temporal or spatial order, includ- ing activities such as the sequence in which an object’s parts should be assembled.

**Guidelines.** Provide simple diagrams showing only as much detail as necessary to clearly illustrate the diagram’s objective. Simpler diagrams will also load faster on a Web page. To illustrate key points, provide cutaway diagrams or exploded dia- gram views.

##### Drawings

* + - Use:
      * When selective parts need to be emphasized or represented.
    - Guidelines:
      * Provide simple drawings showing minimal detail.
      * Provide a link to a complete drawing.

**Use.** Use a drawing when only certain parts of an image are of relevance, and these parts must be emphasized or clearly represented. If the working of a specific ob- ject is to be described, a diagram illustrating its relevant parts should be used.

**Guidelines.** Provide simple drawings showing minimal detail. They are easier to view and understand and they also load more quickly. Photographs are likely to be less effective since they contain information that is not relevant, they lack clar- ity, and they take longer to load on a Web page. If the user is also in need of a de- tailed drawing, provide a link to a page containing a complete drawing.

##### Animation

* Uses:
  + To explain ideas involving a change in:
* Time.
* Position.
  + To illustrate the location or state of a process.
  + To show continuity in transitions.
  + To enrich graphical representations.
  + To aid visualization of three-dimensional structures.
  + To attract attention.
* Disadvantages:
  + Very distracting.
  + Slow loading.
* Guidelines:
  + Use only when an integral part of the content.
  + Create short segments.
  + Provide a freeze frame and stop mode.
  + Avoid distracting animation.

**Uses**. Only use animation when it serves a useful purpose. Animations can be used to enhance textual explanations of objects changing over *time.* A map illustrating population growth can be animated to illustrate population densities and patterns over a sequence of years or centuries. Proper sequential body p*ositions* needed to skillfully perform a sport can also be illustrated as they are textually described. The current *location* within, or the state of, a process can be highlighted through animating flow arrows or process steps.

*Continuity* in transitions can also be illustrated. The changing of states of an el- ement with two or more states will be easier to understand if the transitions are animated instead of being instantaneous. In Windows, actually seeing an icon mov- ing as it is dragged from a desktop to the Recycle Bin or the My Documents file strengthens one’s understanding of the task and the results. *Graphical representa- tions* can also be enriched. Some kinds of information are easier to visualize with movement rather than with still pictures. *Visualization* of three-dimensional struc- tures can also be aided. While a two-dimensional screen can never provide a full understanding of a three-dimensional element, animating the element by slowly

turning it aids in understanding its structure. Animation can also be used to *at- tract attention*. The user’s attention can be directed to an important screen element or alerted to an important condition.

MAXIM *Content* is always more important than graphics.

**Disadvantages.** Any discussion of screen image animation includes a strong caution concerning animation’s side effects. Screen animation is difficult to ignore, often overpowering a person’s peripheral vision. As was discussed in Step 1 “Know Your User or Client,” peripheral vision competes with foveal vision for a person’s attention. That sensed in the periphery is passed on to our information-processing system along with what is actively being viewed foveally. It is, in a sense, visual noise. Mori and Hayashi (1993) experimentally evaluated the effect of windows in both a foveal and peripheral relationship and found that performance on a foveal window deteriorates when there are peripheral windows, and the performance degradation is even *greater* if the information in the peripheral is dynamic or mov- ing. Reeves and Nass (1996) measured brain waves with an EEG and found that at- tention increased every time motion appeared on a screen. Permanently moving animation on a screen makes it very hard for people to concentrate on reading text, if the brain wants to attend to the motion. Animation can also be very annoying.

Another current negative side effect of Web page animation is its close associa- tion with advertising. Animation, including scrolling text, is frequently being used by advertisers to try and gather the users’ attention. Studies suggest that people have started equating animation with advertising, so animation as a screen element is being routinely ignored. Important animation may, therefore, be missed. Ani- mated images also take longer to load.

**Guidelines.** Use animation sparingly. Only use it when it is an *integral part* of the textual content, or reinforces the content. Create *short segments*. There are many distractions people may encounter while watching animation, so long segments should be avoided. Animation, when used, should be capable of being *stopped* by the user so an image may be studied in detail. It should also be capable of being ended entirely so it is eliminated as a visual distraction. In conclusion, always avoid animation or special effects that detract from the screen’s message.

##### Audition

* + - Uses:
      * As a supplement to text and graphics.
      * To establish atmosphere.
      * To create a sense of place.
      * To teach.
      * To sample.
    - Advantages:
      * Does not obscure information on the screen.
      * Shorter downloading time than video.
* Disadvantages:
  + Is annoying to many people, including users and nonusers in the vicinity.
  + Can easily be overused, increasing the possibility that it will be ignored.
  + Is not reliable because:
* Some people are hard of hearing.
* If it is not heard, it may leave no permanent record of having occurred.
* The user can turn it off.
* Audio capability may not exist for the user.
* Guidelines:
  + When words are spoken:
* The content should be simple.
* The speed of narration should be about 160 words per minute.
  + When used to introduce new ideas or concepts the narration should be slowed.
  + Off-screen narration should be used rather than on-screen narration.
* Unless the narrator is a recognized authority on the topic.
  + Create short segments.
  + Provide segments of high quality.
  + Provide audio controls.
  + Play background audio softly.

This discussion of audition focuses on sound as a communication medium for pres- enting meaningful information, words, music, and so on A discussion of sounds used to alert the user is found in Step 9*.*

**Uses.** Use audio as a *supplement* to text and graphics and only to reinforce visual con- tent. Audio should never be used alone because of the disadvantages listed above. Audio can also be used to establish *atmosphere.* A particular type of music reflect- ing a Web site’s content can be played to establish ambience and also to create ori- entation signposts fostering a *sense of place*. Audio can also be used to *teach* word pronunciation or to provide *samples* of music.

**Advantages.** An advantage of audio is its ability to offering commentary or help for a visual display. Audio does not obscure information on the screen, and it down- loads faster than most other types of graphics.

**Disadvantages.** Audio’s disadvantages are similar to those of sounds described in Step 9. Audio can be annoying to many people, including users and nonusers in the vicinity. It can easily be overused, increasing the possibility that it will be ignored. Audio is also not reliable because some people are hard of hearing, it may leave no permanent record of having occurred, it may not be available to the user, or it may be turned off. Loud audio can also be irritating, especially to those with sen- sitive hearing.

**Guidelines.** Williams (1998) in a multimedia literature review extracted most of the following guidelines. When words are spoken, the content should be simple, and the speed of narration should be about 160 words per minute. When the narration is used to introduce new ideas or new concepts the narration should be slowed. Off-screen (invisible) narration should be used rather than on-screen narration.

On-screen narration is acceptable, however if the narrator is a recognized author- ity on the topic being presented.

Other audition guidelines include these: Create *short segments*. There are many distractions people may encounter while listening to audio, so long segments should be avoided. Always provide audio segments of *high quality*. Research has found (Reeves and Nass, 1996) that while people will accept poor video, they are very affected by poor audio. Let users control the playing of audio. Provide the fol- lowing *controls*: Play, Pause/Resume, Stop, Rewind, Fast Forward, and Volume. Any *background* audio should be subdued so it does not interfere with main in- formation being presented on the screen.

##### Combining Mediums

* + - Combinations:
      * Use sensory combinations that work best together:
        + Auditory text with visual graphics.
        + Screen text with visual graphics.
    - Integration:
      * Closely integrate screen text with graphics.
    - Relevance:
      * Both the visual and auditory information should be totally relevant to the task being performed.
    - Presentation:
      * Visual and auditory textual narrative should be presented simultaneously, or the visuals should precede the narrative by no more than 7 seconds.
      * To control attention, reveal information systematically.
        + Limit elements revealed to one item at a time and use sequential revelations for related elements.
      * Animation must show action initiation as well as the action’s result.
      * Avoid animation that distracts from other more important information.
    - Downloading times:
      * Consider downloading times when choosing a media.
    - Testing:
      * Thoroughly test all graphics for:
        + Legibility.
        + Comprehensibility.
        + Acceptance.

Interface technology encourages inclusion of the various graphical media (images, photos, video, diagrams, drawings, and audio) along with text on a screen. The design issue is which mediums work best with other mediums, and which mediums should not be employed together. Before reviewing research on this topic, which does find per- formance advantages for certain combinations of multimedia, theories for why this may happen will be summarized.

The first theory is called the *dual code* theory. It proposes that people store informa- tion in two ways in memory: verbally and pictorially. This theory postulates that, be- cause of this dual-storage capability, information communicated to a person in both a verbal and pictorial manner has a greater likelihood of being remembered than infor- mation arriving in only one format. Also postulated is that too much information ar- riving in one format can overtax that particular memory. Combining verbal audio with displayed text is one such overtaxing combination.

The second theory also proposes two independent working memories, but is slightly different in concept. The first type of memory is a visual-spatial sketchpad in which in- formation accumulated visually is stored. This visual information may be graphical or textual in nature. The second type of working memory is a phonological loop for deal- ing with and storing auditory information. This theory postulates that performance may be improved for certain more complex tasks because working memory is expanded through the application of two senses. The general conclusion is that combining visual and verbal auditory information can lead to enhanced comprehension, when com- pared to relying on one sense alone.

The two theories diverge on the storage of audio. The former suggests that verbal audio and displayed text is stored together; the latter suggests that they are stored separately.

**Combinations and integration.** Williams (1998) in a literature review found that combining visual and verbal auditory information in multimedia design can lead to enhanced comprehension, when compared to use of these medias alone. Sev- eral recent studies have also explored the effects of various media, or combina- tions of media, on user performance. One such study is that of Lee and Bowers (1997). These researchers evaluated various mediums to see which yielded the best learning. The results, summarized in Table 11.2, compared a control group to groups learning material by the various methods described.

Another series of three studies were those of Tindall-Ford, Chandler, and Sweller (1997). They compared combinations of the following multimedia conditions for learning and performance:

|  |  |
| --- | --- |
| **Table 11.2** Learning Improvements for Various Media |  |
| **MEDIUM** | **PERCENT MORE LEARNING** |
| Hearing spoken text and viewing graphics | 91% |
| Viewing graphics alone | 63% |
| Viewing text and viewing graphics | 56% |
| Hearing spoken text, viewing text, and viewing graphics | 46% |
| Hearing spoken text and viewing text | 32% |
| Viewing text alone | 12% |
| Hearing spoken text alone | 7% |
| From Lee and Bowers (1997). |  |

* A visual diagram or table and separated visual text.
* A visual diagram and integrated visual text.
* A visual diagram or table and spoken instructions.

They found that the visual-audio combinations yielded reliably better perfor- mance for complex tasks, but no differences were found for easy tasks. They also found that visual text integrated into a diagram yielded better performance than separated visual text. They attributed the better results for the audiovisual combi- nation and the integrated text and diagram alternative to reduced demands on working memory.

What can we conclude from these studies?

* **The proper multimedia combinations can improve learning and perfor- mance.** Hearing spoken text combined with a visual graphic is an espe- cially useful combination, especially for complex tasks. All studies found this pairing useful.
* **Visual graphics do enhance learning and performance.** In the Bowers and Lee study, the various graphical combinations yielded the higher learning rates.
* **Single-dimensional textual media are not as successful when used alone.** In the Bowers and Lee study, viewing text or hearing spoken text alone yielded the lowest learning rates.
* **Hearing spoken text and viewing text at the same time may not be great, but it may not be terrible, either.** This combination yielded “middle-of- the-road” results in the Bowers and Lee study. The dual code theory would suggest, however, that its use be minimized. Exercise caution in this area.
* **Visual text should always be integrated with related visual graphics.** Tindall-Ford et al. found much better user performance when visual text was closely integrated with, or adjacent to, related visual graphics. It will be much easier for user to coordinate and integrate the visual materials. Pre- senting spatially separated text and related graphics places greater de- mands on working memory.

**Relevance.** Both the visual and auditory information should be totally relevant to the task being performed. All spoken text should reinforce presented graphics.

**Presentation.** Faraday and Sutcliffe (1997) also conducted a series of studies address- ing multimedia design. Like the above-mentioned studies, they found displayed graphics (images and animation) improved user performance, specifically the recall of information. Based upon these studies, they developed the following guidelines. Provide sufficient *time* for reading screen graphic captions. Present si- multaneously all visual and auditory *narrative information* to the user, or have the visual information precede the auditory narrative by no more than 7 seconds. To control the users’ attention, *reveal* or expose information systematically on the screen, either from left to right or from top to bottom. Limit the information re- vealed to one item at a time, and sequentially reveal related elements. Finally, any *animation* must show an action being initiated as well as the action’s result, and avoid any animation that distracts from other more important screen information.

**Downloading times.** Consider times in choosing a graphical medium. In general, downloading times range from the fastest, audio, to the slowest, video.

**Testing.** Thoroughly test all graphics for *legibility*. Make sure visual graphics are easy to see from a variety of viewing distances. Also test them for *comprehensibil- ity*. Are visual graphics and related audio clear and understandable? Are the graphics *acceptable* to the using audience? This is especially critical if the users are multicultural. Always test graphics with all representative user groups.

## **Choose the Proper Colors**

Color adds dimension, or realism, to screen usability. Color draws attention because it attracts a person’s eye. If used properly, it can emphasize the logical organization of in- formation, facilitate the discrimination of screen components, accentuate differences among elements, and make displays more interesting. If used improperly, color can be distracting and possibly visually fatiguing, impairing the system’s usability. In this step we will:

Come to understand color’s characteristics.

What color is. The uses of color.

Possible problems and cautions when working with color. The results of color research.

Color and human vision.

Come to understand how to use color.

Choose the proper colors for textual graphic screens. Choose the proper colors for statistical graphics screens. Choose the proper color for Web screen text and images.

Effective use of color in screen design has taken great steps forward in the last four decades. Earlier text-based displays could only present a few colors, and many of the colors were not very legible. Color was often overused in combinations that reminded one more of a Christmas tree than of an effective source of communication. The evolu- tion to graphical screens expanded the use of color, but did not immediately eliminate

some of the color problems. Today, because technology has improved, as well as our understanding of what constitutes good design, colors in screens are being used much more effectively. Pastels have replaced bright reds and dark blues, and the number of colors presented at one time on a screen has been reduced, dramatically in some cases. This is not to say, however, that all the problems have been solved. A tour around the office will usually uncover some questionable, or awful, uses of color. Two of two most common problems are screen backgrounds being more attention-grabbing than the screen data (which is the most important element of a screen), and overuse of color as a graphic language or code (the color itself meaning something to the screen viewer). This latter kind of use forces the user to interpret a color’s meaning *before* the message it is communicating can be reacted to.

In recent years, the development of the Web and the availability of monitors with significantly expanded color capability have initiated a replay of the early color-use prob- lems that surfaced in both text-based and graphical systems. Infatuated with the almost unlimited supply of available colors, developers have eagerly raced to include a multi- tude of colors on Web pages, with too little thought given to the consequences for users. The “Christmas tree effect” has lived on as users struggled with illegible text and numerous visual distractions. Today, the use of color in Web pages has improved some- what. Too many site designers still, however, associate good design with “splashy” color. The discussion to follow begins by defining color. Next is a review of how color may be used in screen design and some critical cautions on its use. Then, the human visual system and its implications for color are discussed. Continuing, a series of general screen guidelines are presented for choosing and using colors. This is followed by a compilation of guidelines for specific kinds of screens: textual and graphical, statistical

graphics, and Web screens and their associated graphical elements.

Color—What Is It?

Wavelengths of light themselves are not colored. What is perceived as actual color re- sults from the stimulation of the proper receptor in the eye by a received light wave. The name that a color is given is a learned phenomenon, based on previous experiences and associations of specific visual sensations with color names. Therefore, a color can only be described in terms of a person’s report of his or her perceptions. The visual spectrum of wavelengths to which the eye is sensitive ranges from about 400 to 700 mil- limicrons. Objects in the visual environment often emit or reflect light waves in a lim- ited area of this visual spectrum, absorbing light waves in other areas of the spectrum. The dominant wavelength being “seen” is the one that we come to associate with a spe- cific color name. The visible color spectrum and the names commonly associated with the various light wavelengths are shown in Table 12.1.

To describe a color, it is useful to refer to the three properties it possesses: hue, chroma or saturation, and value or intensity, as illustrated in Figure 12.1. *Hue* is the spectral wave- length composition of a color. It is to this we attach a meaning such as green or red. *Chroma* or *saturation* is the purity of a color in a scale from gray to the most vivid version of the color. The more saturated a hue is, the more visible it is at a distance. The less

**Table 12.1** The Visible Spectrum

**APPROXIMATE COLOR WAVELENGTHS IN MILLIMICRONS**

|  |  |
| --- | --- |
| Red | 700 |
| Orange | 600 |
| Yellow | 570 |
| Yellow-green | 535 |
| Green | 500 |
| Blue-green | 493 |
| Blue | 470 |
| Violet | 400 |

**Figure 12.1** The relationship of hue, chroma, and value.

Light



VALUE

Dark

Violet Blue

Red

Green

HUE HUE

Orange Yellow

saturated, the less visible it is. *Value* or *intensity* is the relative lightness or darkness of a color in a range from black to white. Two-word descriptors, such as light red or dark blue, are usually used to describe lightness differences. Some hues are inherently lighter or darker than others, for example, yellow is very light and violet is very dark.

The primary colors of illuminated light are red, green, and blue, whose wavelengths additively combine in pairs to produce magenta, cyan, and yellow, and all the other visible colors in the spectrum. The three primary colors additively combine to produce white. The long-wavelength colors (red) are commonly referred to as warm, and short- wavelength colors (blue) as cool.

Color, then, is a combination of hue, chroma, and value. In any one instance what we call “blue” may actually be one of several hundred thousand “blues.” This problem has confounded color research over the years. A “blue” may be unacceptable in one situation because it is highly saturated and dark, but perfectly acceptable in another, being less sat- urated and light. The exact measures of a color are rarely reported in the literature.

##### RGB

Many color monitors use the three primary colors of light, in various combinations, to create the many colors we see on screens. By adjusting the amounts of red, green, and blue light presented in a pixel, millions of colors can be generated. Hence, color palette editors exist with labels R, G, and B (or the words spelled out).

##### HSV

Some palette editors use a convention based on the Munsell method of color notation called HSV, for hue, saturation, and value (or HSL for hue, saturation, and lightness). Again, various combinations produce different colors.

##### Dithering

The eye is never steady, instead trembling slightly as we see. If pixels of different colors are placed next to each other, this tremor combines the two colors into a third color. This is referred to as *dithering*, and sometimes *texture mapping*. Taking advantage of this phenomena, an optical illusion, a third color can be created on a screen. Dithering is often used to create a gray scale when only black and white pixels are available to work with. A difference of opinion exists on whether dithering should, or should not, be used on a screen*. The Macintosh Human Interface Guidelines* (Apple Computer, 1992b) discour- age its use, stating it creates unnecessary visual clutter.

In systems containing large palettes of colors, the color-mixing process a computer goes through when it encounters a color not in its palette is also called dithering. In this process palette colors are mixed to approximate the appearance of the desired color.

#### Color Uses

* Use color to assist in formatting a screen:
  + Relating or tying elements into groupings.
  + Breaking apart separate groupings of information.
  + Associating information that is widely separated on the screen.
  + Highlighting or calling attention to important information by setting it off from the other information.
* Use color as a visual code to identify:
  + Screen components.
  + The logical structure of ideas, processes, or sequences.
  + Sources of information.
  + Status of information.
* Use color to:
  + Realistically portray natural objects.
  + Increase screen appeal.

Color may be used as a formatting aid in structuring a screen, or it may be used as a visual code to categorize and identify information or data. It may also be used to por- tray objects naturally and make a screen more appealing to look at.

##### Color as a Formatting Aid

As a formatting aid, color can provide better structure and meaning to a screen. It is es- pecially useful when large amounts of information must be included on a screen and spacing to differentiate components is problematic. For example, displaying groupings of information in different colors can enhance differentiation of the groupings. Spa- tially separated but related fields can also be tied together through a color scheme.

Color can also replace highlighting as a means of calling attention to information. Color is much more flexible than other techniques because of the number of colors that are available. Color, as an attention-getting mechanism must, however, be chosen in light of the psychological and physiological considerations to be described shortly.

##### Color as a Visual Code

A color code indicates what category the information being displayed falls into. It has meaning to the screen’s user. A properly selected color-coding scheme permits a per- son to identify a relevant information category quickly, without having to read its con- tents first. This permits focusing on this category, while the remaining information is excluded from attention.

One common color-coding scheme used to differentiate screen components is to dis- play screen captions and data in different colors. Another is to identify information from different sources—information added to a process from different locations, or text added

to a message from different departments, may be colored differently. Using color-coding to convey status might involve displaying, in a different color, information that passed or failed system edits. Color can also be used as a prompt, guiding a person through a complex transaction.

To be effective, color as a visual code must be relevant and known. Relevance is achieved when the color enables a person to attend to only the information that is needed, and easily exclude that which is not needed. A relevant code, however, will be useless unless its meaning is also understood by the person who must use it. If a color’s meaning is not known, one must first interpret its meaning. This can place burdens on a person’s memory. It can also impede performance, requiring one to consult a manual or a legend in order to understand it.

##### Other Color Uses

Color can also be used to more realistically portray objects in the world around us that must be displayed on a screen. It is also thought that the addition of color increases a screen’s appeal and makes working with a display more pleasant.

#### Possible Problems with Color

The simple addition of color to a screen will not guarantee improved performance. What may have been a poorly designed product will simply become a colorful poorly designed product. When used improperly, color may even impair performance by dis- tracting the viewer and interfering with the handling of information. Possible problems may be caused by the perceptual system itself or the physiological characteristics of the human eye.

##### High Attention-Getting Capacity

Color has an extremely high attention-getting capacity. This quality causes the screen viewer to associate, or tie together, screen elements of the same color, whether or not such an association should be made. A person might search for relationships and dif- ferences that do not exist, or that are not valid. The result is often bewilderment, con- fusion, and slower reading. The effect achieved is often described as the Christmas tree mentioned earlier.

##### Interference with Use of Other Screens

Indiscriminate or poor use of color on some screens will diminish the effectiveness of color on other screens. The rationale for color will be difficult to understand and its attention-getting capacity severely restricted.

##### Varying Sensitivity of the Eye to Different Colors

All colors are not equal in the eye of the viewer. The eye is more sensitive to those in the middle of the visual spectrum (yellow and green), which appear brighter than those at the extremes (blue and red). Thus, text composed of colors at the extremes is thought to be more difficult to read. Research evidence on this topic is mixed. Several studies have found that acuity, contrast sensitivity, target recognition, legibility, and perfor- mance were not influenced by color. On the other hand, other studies have found ad- vantages for central spectral colors in reaction times, resolution, and error rates.

Also, it is thought that some combinations of screen colors can strain the eye’s ac- commodation mechanism. The wavelengths of light that produce blue are normally focused in front of the eye’s retina, the red wavelengths behind it. Simultaneous or se- quential viewing of red and blue causes the eye to continually refocus to bring the image directly onto the retina, thereby increasing the potential for eye fatigue. Again, the research evidence is mixed. Some studies have found this a problem while others have not.

What does one conclude after looking at the research addressing the above prob- lems? The reasonable assumption is that they have neither been proved nor disproved. We have not properly defined all the terminal-based tasks being performed, and the exact qualities of the colors being studied. Also, the studies have used only a few of the many devices in existence. And, a firm definition of “visual fatigue” remains elusive. Finally, none of the studies have addressed extended terminal viewing. The prudent course is to be cautious and avoid using colors and combinations that color theory sug- gests can create problems.

The perceived appearance of a color is also affected by a variety of other factors, in- cluding the size of the area of color, the ambient illumination level, and other colors in the viewing area. Also, larger changes in wavelength are needed in some areas of the vi- sual spectrum for a color change to be noticed by the eye. Small changes in extreme reds and purples are more difficult to detect than small changes in yellow and blue-green. Failure to consider the eye and how it handles color, then, can also lead to mistakes in color identification, misinterpretations, slower reading, and, perhaps, visual fatigue.

##### Color-Viewing Deficiencies

Another disadvantage of color is that about 8 percent of males and 0.4 percent of females have some form of color-perception deficiency—color blindness, as it is commonly called. In actuality, very few people are truly color-blind; most of those with problems simply have difficulties discriminating certain colors. A red viewing deficiency is called *protanopia*, a green deficiency is called *deuteranopia*, and a blue deficiency is called *tri- tanopia*. These common color deficiencies, their results, and the percentage of people who experience these problems, are summarized in Table 12.2. Total color blindness af- fects no more than 0.005 percent of both sexes. For an individual with color-perception deficiency, all the normal colors may not be discernible, but often differences in light- ness or intensity can be seen. The important point: a person experiencing any form of color deficiency must not be prohibited from effectively using a screen.

**Table 12.2** Results of Color-Defective Vision

|  |  |  |  |
| --- | --- | --- | --- |
| **ACTUAL COLOR** | **COL**  **RED-VIEWING DEFICIENCY (2.04%)** | **OR SEEN WITH: GREEN-VIEWING**  **DEFICIENCY (6.39%)** | **BLUE-VIEWING DEFICIENCY (0.003%)** |
| Red | Brown | — | — |
| Yellow | Greenish-Yellow | Orange | Deeper Yellow |
| Purple | Dark Blue | Red | Deep Red |
| Green | — | Light Brown | — |
| Brown | — | Reddish-Brown | — |
| Blue | — | — | Green |

From Barnett (1993); Fowler and Stanwick (1995).

##### Cross-Disciplinary and Cross-Cultural Differences

Colors can have different meanings in different situations to different people. A color used in an unexpected way can cause confusion. An error signaled in green would con- tradict the expected association of red with stop or danger. The same color may also have a different connotation, depending upon its viewer. The color blue has the follow- ing quite different meanings:

For financial managers—Corporate qualities or reliability. For health care professionals—Death.

For nuclear reactor monitors—Coolness or water.

For American movie audiences—Tenderness or pornography.

Differences in color connotations also exist between cultures. Red, for example, in the United States is associated with danger, in Egypt with death, and in India with life. In- correct use in a different culture may cause severe problems. A listing of some common cultural associations with color was found in Table 10.2 in Step 10 “Provide Effective Internationalization and Accessibility.”

Color appeal is also subjective. People have different tastes in color, what is pleasing to one person may be distasteful or unusable by someone else.

The proper use of color, then, requires an analysis of the expectations and experiences of the screen viewer. The use of color in design must always keep these possible prob- lems clearly in focus. The designer must work to minimize their disruptive and destruc- tive effects.

MAXIM Poor use of color is worse than not using it at all.

#### Color—What the Research Shows

The effectiveness of color in improving the usability of a display has been studied for many years. The research results have been mixed. To illustrate, on a positive note, color has been shown to improve performance (Kopala, 1981; Nagy and Sanchez, 1992; Sidorsky, 1982), to improve visual search time (Christ, 1975; Carter, 1982), to be useful for organizing information (Engel, 1980), to aid memory (Marcus, 1986b), and to demar- cate a portion of a screen (as opposed to lines or type font, Wopking, Pastoor, and Beldie 1985). Color has also created positive user reactions (Tullis, 1981), was preferred to monochromatic screens for being less monotonous and reducing eyestrain and fa- tigue (Christ, 1975), and is more enjoyable (Marcus, 1986b).

On the other hand, it has also been shown that color does not improve performance (Tullis, 1981), does not have much of an effect on reading text (Legge and Rubin, 1986), may impair performance (Christ and Teichner, 1973; Christ, 1975), and is less important than display spacing (Haubner and Benz, 1983). It has also been demonstrated that poor character-background color combinations lead to poorer performance (McTyre and Frommer, 1985). Finally, no evidence has been produced that color, as compared to black and white, can significantly improve aesthetics or legibility or reduce eyestrain (Pas- toor, 1990).

Research has found, moreover, that as the number of colors on a display increases, the time to respond to a single color increases, and the probability of color confusion in- creases (Luria, Neri, and Jacobsen 1986). Many studies have found that the maximum number of colors that a person can handle is in the range of 4 to 10, with emphasis on the lower numbers (for example, Brooks, 1965; Halsey and Chapanis, 1951; Luria, Neri, and Jacobsen, 1986).

The conclusion to be derived from these studies is that for simple displays, color may have no dramatic impact. Indeed, a monochromatic display may serve the purpose just as well. As display complexity increases, however, so does the value of color. A second conclusion is that people like using color and think it has a positive influence on their productivity, even though it may not.

To be effective, color must be properly used. Poor use of color will actually impair performance, not help it. When using color, keep in mind its value will be dependent upon the task being performed, the colors selected, the number of colors used, and the viewing environment.

#### Color and Human Vision

To understand how color should be used on a screen, it is helpful to know something of the physiology of the human eye.

##### The Lens

Muscles control the lens of the eye. These muscles focus received wavelengths of light on the retina. The lens itself is not color corrected. The wavelengths of light that create

different colors are focused at different distances behind the lens, the longer wavelengths (red) being focused farther back than the shorter wavelengths (blue). The result is that colors of a different wavelength from the color actually being focused by the lens will appear out of focus. To create a sharp image of the out-of-focus colors requires a refo- cusing of the eye. Excessive refocusing (such as between red and blue) can lead to eye fatigue.

The effect of this focusing for most people is that blues appear more distant and reds appear closer. This can give a three-dimensional appearance to what is being viewed. A critical problem is that the wavelength of light that creates blues can never be brought into focus on the retina but is always focused in front of it. A sharp blue image is im- possible to obtain. Very pure or saturated colors require more refocusing than less pure or unsaturated colors. Therefore, a color with a large white component will require less refocusing.

The lens does not transmit all light wavelengths equally. It absorbs more wavelengths in the blue region of the spectrum than those in the other regions. Additionally, as the lens ages, it tends to yellow, filtering out the shorter blue wavelengths. Thus, as people get older, their sensitivity to blue decreases. The lens also refocuses for light waves of different brightness. Sharp contrasts in brightness in things being viewed can lead to visual fatigue as the eye continually makes muscular adjustments. Driving an auto- mobile through a forest of trees on a bright sunny day illustrates this effect. The eye con- tinually adjusts as the auto sequentially moves through areas of bright sunlight and patches of shadows.

##### The Retina

The retina is the light-sensitive surface of the eye. It comprises two kinds of receptors, rods and cones, which translate the incoming light into nervous impulses. Rods are sensitive to lower light levels and function primarily at night. Cones are stimulated by higher light levels and react to color. The sensitivity of cones to colors varies, different cones possessing maximum sensitivity to different wavelengths of light. About two- thirds (64 percent) of the cones are maximally sensitive to longer light wavelengths, showing a peak response at about 575 millimicrons. These cones have traditionally been referred to as “red” sensitive cones. In actuality, however, the peak sensitivity is in the yellow portion of the visual spectrum (see Table 12.1). About one-third (32 percent) of the cones achieve maximum sensitivity at about 535 millimicrons and are commonly referred to as “green” sensitive cones. The remainder (2 percent) primarily react to short light wavelengths, achieving maximum sensitivity at about 445 millimicrons. These are known as “blue” sensitive cones. Any light wave impinging on the retina evokes a re- sponse, to a greater or lesser degree, from most or all of these cones. A perceived “color” results from the proportion of stimulation of the various cone kinds.

Rods and cones vary in distribution across the retina. The center is tightly packed

with cones and has no rods. Toward the periphery of the retina, rods increase and cones decrease. Thus, color sensitivity does not exist at the retina’s outer edges, although yel- lows and blues can be detected further into the periphery than reds and greens. The very center of the retina is devoid of blue cones, creating a “blue-blindness” for small objects being looked at.

The receptors in the eye also adjust, or adapt, their level of sensitivity to the overall light level and the color being viewed. Adaptation to increases in brightness improves color sensitivity. Color adaptation “softens” colors.

The brightness sensitivity of the eye to different colors also varies. It is governed by output from the red and green cones. The greater the output, the higher the brightness, which results in the eye being most sensitive to colors in the middle of the visual spec- trum and less sensitive to colors at the extremes. A blue or red must be of a much greater intensity than a green or yellow even to be perceived.

The ability of the eye to detect a form is accomplished by focusing the viewed image on the body of receptors to establish edges. Distinct edges yield distinct images. Edges formed by color differences alone cannot be accurately brought into focus and thus cre- ate fuzzy and indistinct images. A clear, sharp image requires a difference in brightness between adjacent objects, as well as differences in color.

The components of the eye—the lens and retina—govern the choices, and combina- tions, of colors to be displayed on a screen. The proper colors will enhance performance; improper colors will have the opposite effect, as well as increase the probability of vi- sual fatigue.

MYTH If we can’t do it right, do it big. If we can’t do it big, do it in color.

#### Choosing Colors

When choosing colors for display, one must consider these factors: the human visual system, the possible problems that the colors’ use may cause, the viewing environment in which the display is used, the task of the user, how the colors will be used, and the hardware on which the colors will be displayed. The primary objective in using color is communication, to aid the transfer of information from the screen to the user.

##### Choosing Colors for Categories of Information

* Choosing colors for categories of information requires a clear understanding of how the information will be used.
* Some examples:
  + If different parts of the screen are attended to separately, color-code the different parts to focus selective attention on each in turn.
  + If decisions are made based on the status of certain types of information on the screen, color-code the types of status that the information may possess.
  + If screen searching is performed to locate information of a particular kind or qual- ity, color-code these kinds or qualities for contrast.
  + If the sequence of information use is constrained or ordered, use color to identify the sequence.
  + If the information displayed on a screen is packed or crowded, use color to pro- vide visual groupings.
* Use color as a redundant screen code.

**Categories.** Color chosen to organize information or data on a screen must aid the transfer of information from the display to the user. This requires a clear under- standing of how the information is selected and used. The examples above describe some common ways of classifying information for color-coding purposes.

**Redundancy.** Never rely on color as the only way of identifying a screen element or process. Users with a color-viewing deficiency may not be able identify a specific color when it is important. It is also important to remember that information on one screen may be used in more than one way. What is useful in one context may not be in another and may only cause interference. Therefore, when developing a color strategy, always consider how spatial formatting, highlighting, and mes- sages may also be useful and employ these structural and coding methods as well.

##### Colors in Context

Colors are subject to contextual effects. The size of a colored image, the color of images adjacent to it, and the ambient illumination all exert an influence on what is actually perceived. At the normal viewing distance for a screen, maximal color sensitivity is not reached until the size of a colored area exceeds about a 3-inch square. Smaller images become desaturated (having a greater white component) and change slightly in color. Also, small differences in actual color may not be discernible. Blues and yellows are particularly susceptible to difficulties in detecting slight changes. Finally, small adja- cent colored images may appear to the eye to merge or mix. Red and green, for example, might appear as yellow.

Adjacent images can influence the perceived color. A color on a dark background will look lighter and brighter than the same color on a light background, for example. A color can be *induced* into a neutral foreground area (gray) by the presence of a colored background. A red background can change a gray into a green. Induced colors are the complement of the inducing color. Looking at a saturated color for a period of time can also induce complementary afterimages. Colors also change as light levels change. Higher levels of ambient light tend to desaturate colors. Saturated colors will also ap- pear larger than desaturated colors.

##### Usage

* Design for monochrome first.
* Use colors conservatively.
  + Do not use color where other identification techniques, such as location, are avail- able.

**Monochrome.** Design for monochrome first, or in shades of black, white and gray. A screen should be as capable of being effectively used as if it were located in a monochrome environment. Spatial formatting, consistent locations, and display techniques such as highlighting and multiple font styles, should all be utilized to give information a structure independent of the color. Doing this will permit the screen to be effectively used:

By people with a color-viewing deficiency. On monochrome displays.

In conditions where ambient lighting distorts the perceived color. If the color ever fails.

**Conservative Use.** Use color sparingly since it has such a high attention-getting qual- ity. Only enough colors to achieve the design objective should be used. More col- ors increase response times, increase the chance of errors due to color confusion, and increase the chance of the Christmas tree effect. If two colors serve the need, use two colors. If three colors are needed, by all means use three. A way to mini- mize the need for too many colors is not to use color in situations where other identification methods are available. A menu bar, for example, will always be lo- cated at the top of the screen. Its position and structure will identify it as a menu bar. To color-code it would be redundant.

##### Discrimination and Harmony

* + - For best absolute discrimination, select no more than four or five colors widely spaced on the color spectrum.
      * Good colors: red, yellow, green, blue, and brown.
    - For best comparative discrimination, select no more than six or seven colors widely spaced on the color spectrum.
      * Other acceptable colors: orange, yellow-green, cyan, violet, and magenta.
    - Choose harmonious colors.
      * One color plus two colors on either side of its complement.
      * Three colors at equidistant points around the color circle.
    - For extended viewing or older viewers, use brighter colors.

**Absolute discrimination.** The population of measurable colors is about 7.5 million. From this vast number, the eye cannot effectively distinguish many more than a handful. If color memorization and absolute discrimination is necessary (a color must be correctly identified while no other color is in the field of vision), select no more than 4 to 5 colors widely spaced along the color spectrum. Selecting widely spaced colors will maximize the probability of their being correctly identified. Good choices are red, yellow, green, blue, and brown.

Two good color opponent pairs are red/green and yellow/blue. All of these colors except blue are easy to resolve visually. Again, be cautious in using blue for data, text, or small symbols on screens because it may not always be legible. If the meaning for each of more than five colors is absolutely necessary, a legend should be provided illustrating the colors and describing their associated meanings.

**Comparative discrimination.** If comparative discrimination will be performed (a color must be correctly identified while other colors are in the field of vision), se- lect no more than six or seven colors widely spaced along the visual spectrum. In

addition to those above, other colors could be chosen from orange, yellow-green, cyan, violet, and magenta. Again, be cautious of using blue for data, text, or small symbols. If the intent is to portray natural objects realistically, the use of more col- ors is acceptable.

**Harmony.** Choose harmonious colors. Harmonious colors are those that work well together or meet without sharp contrast. Harmony is most easily achieved with a monochromatic palette. For each background color, different lightness or values are established through mixing it with black and white. Marcus (1986a) suggests a minimum of three values should be obtained.

Harmonious combinations in a multicolor environment are more difficult to obtain. Marcus recommends avoiding complementary colors—those at opposite sides of the circle of hues in the Munsell color system, a standard commercial color system. He suggests using split complements, one color plus two colors on either side of its complement, or choosing three colors at equidistant points around the color circle.

**Extended viewing.** For older viewers or extended viewing, use bright colors. As eye capacity diminishes with age, data, text, and symbols in the less bright colors may be harder to read. Distinguishing colors may also become more difficult. For any viewer, long viewing periods result in the eye adapting to the brightness level. Brighter colors will be needed if either of these conditions exists.

##### Emphasis

* To draw attention or to emphasize elements, use bright or highlighted colors. To deemphasize elements, use less bright colors.
  + The perceived brightness of colors from most to least is white, yellow, green, blue, red.
* To emphasize separation, use contrasting colors.
  + Red and green, blue and yellow.
* To convey similarity, use similar colors.
  + Orange and yellow, blue and violet.

**Drawing attention.** To draw attention or emphasize, use bright colors. The eye is drawn to brighter or highlighted colors, so use them for the more important screen components. The data or text is the most important component on most screens, so it is a good candidate for highlighting or the brightest color. Danger signals should also be brighter or highlighted. The perceived brightness of colors, from most to least, is white, yellow, green, blue, and red.

Keep in mind, however, that under levels of high ambient illumination, colors frequently appear washed out or unsaturated. If some means of light attenuation is not possible, or if the colors chosen are not bright enough to counter the illu- mination, color should be used with caution.

**Emphasizing separation.** Use contrasting colors to emphasize separation. The greater the contrast, the better the visibility of adjacent elements. To emphasize the sepa- ration of screen components, use contrasting colors. Possible pairs would be red/ green and blue/yellow.

**Similarity.** Use similar colors to convey similarity. Displaying elements in a similar color can bring related screen components together. Blue and green, for example, are more closely related than red and green.

##### Common Meanings

* + - To indicate that actions are necessary, use warm colors.
      * Red, orange, yellow.
    - To provide status or background information, use cool colors.
      * Green, blue, violet, purple.
    - Conform to human expectations.
      * In the job.
      * In the world at large.

**Actions.** The warm colors, red, yellow, and orange, imply active situations or that actions are necessary. Warm colors advance, forcing attention.

**Status or background.** The cool colors, green, blue, violet, and purple, imply back- ground or status information. Cool colors recede or draw away.

**Expectations.** Conform to human expectations. Use color meanings that already exist in a person’s job or the world at large. They are ingrained in behavior and difficult to unlearn. Some common color associations are the following:

Red—Stop, fire, hot, danger. Yellow—Caution, slow, test.

Green—Go, OK, clear, vegetation, safety. Blue—Cold, water, calm, sky, neutrality. Gray—Neutrality.

White—Neutrality.

Warm colors—Action, response required, spatial closeness.

Cool colors—Status, background information, spatial remoteness.

Some typical implications of color with dramatic portrayal are: High illumination—Hot, active, comic situations.

Low illumination—Emotional, tense, tragic, melodramatic, romantic situations. High saturation—Emotional, tense, hot, melodramatic, comic situations.

Warm colors—Active, leisure, recreation, comic situations. Cool colors—Efficiency, work, tragic and romantic situations.

Proper use of color also requires consideration of the experiences and expec- tations of the screen viewers.

##### Location

* In the center of the visual field, use red and green.
* For peripheral viewing, use blue, yellow, black, and white.
* Use adjacent colors that differ by hue and value or lightness.

**Central vision.** The eye is most sensitive to red and green in the center of the visual field. The edges of the retina are not sensitive to these colors. If used in the view- ing periphery, some other attention-getting method such as blinking must also be used.

**Peripheral vision.** For peripheral viewing, use blue, yellow, black, or white. The retina is most sensitive to these colors at its periphery.

**Adjacent colors.** Colors that appear adjacent to one another should differ in hue and lightness for a sharp edge and maximum differentiation. Also, adjacent colors that differ only in their blue component should not be used so that differentiation is possible. The eye is poorly suited for dealing with blue.

##### Ordering

* Order colors by their spectral position.
  + Red, orange, yellow, green, blue, indigo, violet.

If an ordering of colors is needed, such as from high to low, by levels of depth, and so on, arrange colors by their spectral position. There is evidence that people see the spectral order as a natural one. The spectral order is red, orange, yellow, green, blue, in- digo, and violet, most easily remembered as “ROY G. BIV.”

##### Foregrounds and Backgrounds

* Foregrounds:
  + Use colors that highly contrast with the background color.
  + For text or data, use:
* Black.
* Desaturated or spectrum center colors such as white, yellow, or green.
* Warmer more active colors.
  + Use colors that possess the same saturation and lightness.
  + To emphasize an element, highlight it in a light value of the foreground color, pure white, or yellow.
  + To deemphasize an element, lowlight it in a dark value of the foreground color.
    - Backgrounds:
      * Use a background color to organize a group of elements into a unified whole.
      * Use colors that do not compete with the foreground.
      * Use:
        + Light-colored backgrounds of low intensity: Off-white or light gray.
        + Desaturated colors.
        + Cool, dark colors such as blue or black.
        + Colors on the spectral extremes.

###### Foregrounds

Foreground colors should be as different as possible from background colors. A widely different foreground will maximize text legibility. With today’s high-resolution moni- tors, the most recommended text color is black presented on a light-colored background of low intensity, either off-white or light gray. Bright white backgrounds should be avoided because of the harsh contrast between the dark text and the background.

Desaturated spectrum center colors, such white, yellow, or green, on dark back- grounds also work well. These colors do not excessively stimulate the eye and appear brighter to the eye. Saturated colors excessively stimulate the eye. Color theory also sug- gests using warmer, more active colors for foregrounds. Warmer colors advance, forc- ing attention. Exercise caution in using more fully saturated red and orange, however. They may be difficult to distinguish from one another.

Use foreground colors that possess the same saturation and lightness. Highlight el- ements in a light value of the foreground color. If off-white is the foreground color, high- light elements in pure white. Yellow can also be used to highlight elements. To deemphasize an element, lowlight it in a darker value of the foreground color. When lowlighting, a strong enough contrast with both the background and the non- lowlighted element must be maintained so that legibility and visual differentiation is possible.

The simultaneous use of highlighting and lowlighting should be avoided. Used to- gether they may confuse the viewer. Also, as with other display techniques, be conserv- ative in using highlighting and lowlighting, so that simplicity and clarity are maintained.

###### Backgrounds

A background color should organize a group of elements into a unified whole, isolating them from the remainder of the screen. Use colors that do not compete with the fore- ground. A background must be subtle and subservient to the data, text, or symbols on top of it.

As mentioned above, with today’s high-resolution monitors, the most recommended background color is a low-intensity off-white or gray with black text. Pastoor (1990), in a study to be described shortly, also found that desaturated backgrounds in almost any color work well. Foreground colors must be chosen with consideration of the back- ground color, however.

For dark backgrounds, use the cool, dark colors. Cool, dark colors visually recede, providing good contrast to the advancing lighter, foreground colors. Blue is especially good because of the eye’s lack of sensitivity to it in the retina’s central area and in- creased sensitivity to it in the periphery. Lalomia and Happ (1987), also in a study to be described shortly, found the best background colors to be black and blue. In a similar study, Pastoor (1990) found that cool colors, blue and bluish cyan, were preferred for dark background screens.

Also consider colors at the extreme end of the color spectrum. Marcus (1986a) recom- mends, in order of priority, the following background colors: blue, black, gray, brown, red, green, and purple.

##### Three-Dimensional Look

* Use at least five colors or color values to create a 3-D look on a screen.
  + Background: The control itself and the window on which it appears.
  + Foreground: Captions and lines for buttons, icons, and other objects.
* Usually black or white.
  + Selected mode: The color used when the item is selected.
  + Top shadow: The bezel on the top and left of the control.
  + Bottom shadow: The bezel on the bottom and right of the control.

At least five colors or color values are needed to create a three-dimensional look on a screen (Fowler and Stanwick, 1995): the backgrounds of the control and the surface on which it is placed, the foreground (captions, lines, and so on), the selected mode, and the top and bottom shadows of the controls. These shadows assume an upper-left light source. Motif has created an algorithm to automatically calculate the top and bottom shadows, and the select color based upon the background (Kobara, 1991). Briefly, it rec- ommends the following:

**Background.** Midrange colors, 155–175 on the RGB scale.

**Foreground.** Black or white, depending on the lightness or darkness of the back- ground.

**Selected mode.** About 15 percent darker than the background color, halfway be- tween the background and bottom shadow. (Calculate this by multiplying the background color’s RGB value by 0.85.)

**Top shadow.** About 40 to 50 percent brighter than the background color. (Cal- culate this by multiplying the background color’s RGB by 1.50.)

**Bottom shadow.** About 45 to 60 percent darker than the background color. (Cal- culate this by multiplying the background’s RGB values by 0.50.)

One reminder: A raised look should only be used on operable controls.

##### Color Palette, Defaults, and Customization

* + - Permit users to customize their colors.
    - Provide a default set of colors for all screen components.
    - Provide a palette of six or seven foreground colors.
      * Provide 2 to 5 values or lightness shades for each foreground color.
    - Provide a palette of six or seven background colors.
    - Never refer to a screen element by its color.

**Customization.** Because color preference is subjective, permit users to customize their displayed colors. While little research has been performed on color customization, Familant and Detweiler (1995) have measured the frequency of color changes by users. Compared were displayed color combinations that were judged to be “good” or “poor.” They found that users with the poorer color combinations changed their screen colors more often than those with good combinations. Color satis- faction for those with “poor” color combinations must be fleeting. When color customization is permitted, whenever possible allow users to see the results of their color choices before they are applied. Include a sample screen in a preview function within the customization process.

**Default set.** While some users experiment with different color combinations, many others take what is provided them and never attempt to change it. Actually, many people do not know how to apply color to create a clear and appealing screen. Others may have the talent and skills but not the time to choose a proper combina- tion. For these users, a preselected set of default colors should be developed for all screen elements.

Both the Macintosh and Microsoft Windows provide standard, well thought out color schemes. While thousands of colors may be available for display on a screen, most platforms recommend the use of restricted palettes. This is actually a good thing, reducing the probability of very poor color combinations and Christmas trees. Most Macintosh colors are subdued to avoid a “circus” effect on the screen (Apple, 1992b). Microsoft offers a number of predefined schemes such as “Arizona.”

Do not provide the color spectrum; limit the number of choices available. A maximum of six or seven foreground and background colors will provide the necessary variety. It is also worthwhile to note that 2 to 5 values or lightnesses for each foreground color be developed.

With these palettes, however, some sort of guidance concerning maximum number of colors to use and what are good and poor combinations should be pro- vided. Macintosh, for example, suggests that, if you create your own color schemes, colors compatible with the ones on the Color Control Panel be used. Guidelines will make the color selection process more efficient and reduce the likelihood of visually straining conditions developing.

**Color reference.** Finally, never refer to a screen element by its designed color. What was originally on the screen in yellow may not now be so on some users’ screens.

##### Gray Scale

* For fine discriminations use a black-gray-white scale.
  + Recommended values are white, light gray, medium gray, dark gray, black.

The perception of fine detail is poor with color. The eye resolves fine detail much better on a black-white scale. Marcus (1986b) recommends five tonal values for black and white, higher-resolution screens: black, dark gray, medium gray, light gray, and white. He suggests the following general uses:

White: Screen background.

Text located in any black area.

Light Gray: Pushbutton background area.

Medium Gray: Icon background area. Menu drop shadow. Window drop shadow. Inside area of system icons. Filename bar.

Dark gray: Window border.

Black: Text.

Window title bar. Icon border.

Icon elements.

Ruled lines.

Motif presents the following scheme for designing windows in a gray scale (Kobara, 1991).

Background: A 30 percent light gray (RGB 79, 79, 79).

Foreground: White (RGB 0, 0, 0).

Selected mode: A 70 percent dark gray (RGB 181, 181, 181). Top Shadow: White.

Bottom Shadow: Black (RGB 255, 255, 255).

Gray scale values must differ by at least 20 to 30 percent (White, 1990).

##### Text in Color

* When switching text from black to color:
  + Double the width of lines.
  + Use bold or larger type:
  + If originally 8 to 12 points, increase by 1 to 2 points.
  + If originally 14 to 24 points, increase by 2 to 4 points.
* Check legibility by squinting at text.
  + Too-light type will recede or even disappear.

Text in color is not as visible as it is in black. Fowler and Stanwick (1995) report that the size of text has to be increased to maintain legibility when the text is switched from black to color. Lines should be doubled in width and type made larger or bolder. If the existing type ranges from 8 to 12 points, increase it one or two points. If the existing type ranges from 14 to 24 points, increase it by 2 to 4 points. They suggest that by squinting at it, you can check the legibility of type. A type that is too light will recede, or even disappear, from view.

##### Monochromatic Screens

* + - At the standard viewing distance, white, orange, or green are acceptable colors.
    - At a far viewing distance, white is the best choice.
    - Over all viewing distances, from near to far, white is the best choice.

Monochromatic, or one-color, screens are still found in graphical systems, most fre- quently on notebook PCs. In a study by Hewlett-Packard (Wichansky, 1986) white, or- ange, and green monochrome desktop display device screens were evaluated for performance and readability at various viewing distances. At the standard screen view- ing distance (18 to 24 inches), no performance differences were found between white, or- ange, and green phosphor in either polarity (light characters on a dark background, or dark characters on a light background). Subjective ratings of ease of reading were high- est for green and orange light background screens as compared to dark background screens, while no differences in ease of reading were found for either polarity with white phosphor at this distance. At a far viewing distance (4 to 5 feet), orange and green light background screens could be seen more clearly than dark background screens, while white screens were equally legible in either polarity. More errors were found with green than the other two colors.

A green screen yielded red or pink afterimages for 35 percent of the screen viewers; orange, blue afterimages for 20 percent; white yielded afterimages for 5 percent. A 35 percent pink afterimage rate for green screen viewing was also found by Galitz (1968).

Some conclusions:

At standard viewing distances, no significant performance differences exist for white, orange, or green. All are acceptable. Subjective preferences may vary, however, so providing the viewer a choice of any of these colors is desirable.

At far viewing distances, white is the more legible color and therefore the best choice.

Over all viewing distances, white is the best choice.

White has the lowest probability for creating visual afterimages.

##### Consistency

* + - Be consistent in color use.

Consistency in color usage should exist within a screen, a set of screens, and a system. A person can sense the relatedness of color in space and over time, thereby linking el- ements not seen immediately together. An identical background color in windows on different screens, for example, will be seen as related. Changing specific color meanings must be avoided. It will lead to difficulties in interpretation, confusion, and errors. In general, broadly defined meanings (such as red indicating a problem) permit more scope for variations without inconsistency.

##### Cultural, Disciplinary, and Accessibility Considerations

* Consider the impact of specific colors on:
  + Users of various cultures.
  + Users of various disciplines.
  + Users with color-viewing deficiencies.
  + Users relying on accessibility utilities.

As previously described, colors may possess different meanings and interpretations in different cultures and disciplines. Where applicable, color choices for screen elements should reflect these differences and not be offensive. See Table 10.2 in Step 10 for a sum- mary of some cultural color differences. Also consider users with a color-viewing defi- ciency. If color is used as a code, it must be recognizable by all users. It is best to use color as a supplement to other coding methods such as location, size, or element orien- tation. See Table 12.2 for a summary of problem-creating colors for people with defective color vision. Colors chosen should also consider the impact of users relying on accessi- bility utilities. Some utilities, such as the Magnifier accessory included with Windows, alter the colors displayed on a screen.

#### Choosing Colors for Textual Graphic Screens

For displaying data, text, and symbols on a textual graphical screen (as opposed to sta- tistical graphics screens to be described shortly) colors selected should have adequate visibility, meaning, contrast, and harmony.

Use effective foreground/background combinations. Use effective foreground combinations.

Choose the background color first.

Display no more than four colors at one time. Use colors in toolbars sparingly.

Test the chosen colors.

##### Effective Foreground/Background Combinations

Lalomia and Happ (1987) established effective foreground/background color combi- nations for the IBM 5153 Color Display. From a color set of 16 different foregrounds and 8 different backgrounds, 120 color combinations were evaluated for (1) response time to identify characters, and (2) subjective preferences of users. The results from each measure were ranked and combined to derive an overall measure of color combi- nation effectiveness. The best and poorest color combinations are summarized in Table

* 1. In this table “Best” means the specified combination was in the top 20 percent for overall effectiveness; “Poor” means it was in the bottom 20 percent. Those combina- tions composing the “middle” 60 percent are not marked.

The results yield some interesting conclusions:

The majority of good combinations possess a bright or high-intensity color as the foreground color.

The majority of poor combinations are those with low contrast. The best overall color is black.

**Table 12.3** Effective Foreground/Background Combinations

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BACKGROUND**  **FOREGROUND BLACK BLUE GREEN CYAN RED MAGENTA BROWN WHITE** | | | | | | | | |
| BLACK | x |  |  | Good |  | Good |  | Good |
| BLUE |  | x |  |  | Poor |  |  | Good |
| H.I. BLUE |  |  | Poor | Poor |  |  | Poor | Poor |
| CYAN | Good |  | Poor | x |  |  | Poor |  |
| H.I. CYAN | Good | Good |  | Good | Good | Good |  |  |
| GREEN | Good | Good | x | Poor | Good |  | Poor | Poor |
| H.I. GREEN |  | Good |  |  |  |  |  |  |
| YELLOW | Good | Good |  | Good |  | Good |  |  |
| RED |  |  | Poor |  | x | Poor | Poor |  |
| H.I. RED |  |  | Poor |  |  |  |  |  |
| MAGENTA |  |  | Poor |  | Poor | x | Poor |  |
| H.I. MAGENTA | Good |  | Good |  |  | Poor |  |  |
| BROWN |  |  | Poor |  |  | Poor | x |  |
| GRAY |  | Poor |  |  | Poor |  | Poor |  |
| WHITE |  | Good |  | Poor |  |  |  | x |
| H.I. WHITE | Good |  | Good | Good |  |  |  |  |

(H.I. = High Intensity)

From Lalomia and Happ (1987).

The poorest overall color is brown.

Maximum flexibility and variety in choosing a foreground color exists with black or blue backgrounds. (These backgrounds account for almost one-half of the good combinations.)

Brown and green are the poorest background choices.

Bailey and Bailey (1989), in their screen creation utility Protoscreens, have a table summarizing research-derived good foreground/background combinations. This table, which uses the results of the Lalomia and Happ study plus some others, is shown in modified form in Table 12.4.

The studies referenced above did not control character-background luminance- contrast ratios. Because of the characteristics of the eye, some colors appear brighter to it than others. A conclusion of the Lalomia and Happ study was that good combina- tions usually possessed a bright or high-intensity foreground color.

**Table 12.4** Preferred Foreground/Background Combinations from Protoscreens

**BACKGROUNDS ACCEPTABLE FOREGROUNDS**

|  |  |  |
| --- | --- | --- |
| Black | Dark Cyan Dark Yellow Dark White | Light Green Light Cyan Light Magenta Light Yellow Light White |
| Blue | Dark Green Dark Yellow Dark White | Light Green Light Cyan Light Yellow Light White |
| Green | Black Dark Blue | Light Yellow Light White |
| Cyan | Black Dark Blue | Light Yellow Light White |
| Red |  | Light Green Light Cyan Light Yellow Light White |
| Magenta | Black | Light Cyan Light Yellow Light White |
| Yellow | Black Dark Blue Dark Red |  |
| White | Black Dark Blue |  |

Pastoor (1990) equalized luminance-contrast ratios at preoptimized levels for about 800 foreground/background color combinations. For light foregrounds and dark back- grounds, the ratio was 10:1; for light backgrounds and dark foregrounds, 1:6.5. He then had the combinations rated with the following results:

For dark on light polarity:

Any foreground color is acceptable if the background color is chosen properly.

Increased saturation of the foreground only marginally affected ratings, im- plying that any dark, saturated, foreground color is satisfactory.

Saturated backgrounds yield unsatisfactory ratings.

Less saturated backgrounds generally receive high ratings with any fore- ground color.

For light on dark polarity:

Combinations involving saturated colors tend to be unsatisfactory.

As foreground color saturation increases, the number of background colors yielding high ratings diminishes.

Generally, desaturated foreground/background color combinations yielded the best ratings.

Short wavelength, cool colors were preferred for backgrounds (blue, bluish cyan, cyan).

In general, Pastoor concluded that: (1) there was no evidence suggesting a differential effect of color on subjective ratings or performance (except that for light on dark polar- ity, blue, bluish cyan, or cyan were preferred as backgrounds), and (2) overall, desatu- rated color combinations yielded the best results.

##### Choose the Background First

When choosing colors to display, it is best to select the background color first. Then, choose acceptable foreground colors.

##### Maximum of Four Colors

While not experimentally verified, experience indicates that displaying more than four colors at one time on a textual screen gives rise to a feeling of “too much.” Marcus (1986a) suggests an even more conservative approach, a maximum of three foreground colors and, even better, only two. An application of good use of color can often be viewed in one’s living room. Note the use of color by the television networks when textual or tab- ular information is presented (for example, sport scores, news highlights, and so on). The use of only two, or sometimes three, colors is most commonly seen.

So, while more than four colors may be displayed over a period of time or on a se- ries of screens, do not display more than four colors at one time on a single screen. For most cases, restrict the number of colors to two or three.

##### Use Colors in Toolbars Sparingly

Toolbar icons are usually small in size. Presenting them in color is rarely useful, most often disrupting legibility. Use color in toolbar icons simply and conservatively, and only if the color aids icon identification, makes it easier to distinguish icons, or adds meaning. A file folder in yellow or a “stop” icon in red are examples of good uses of color.

##### Test the Colors

Because color is such a complex phenomenon, because definitions of a color can vary, and because the hardware on which a color is used can affect its look, always test all chosen colors as part of the system testing process (see Step 14 “Test, Test, and Retest”).

#### Choosing Colors for Statistical Graphics Screens

The visual, spatial, or physical representation of information—as opposed to numeric, alphanumeric, textual, or symbol representation—is known as *statistical* or data graph- ics. Common kinds of statistical graphics include bar graphs, line graphs, scatterplots, and pie charts. Color can also be used to render a statistical graphic screen more legi- ble and meaningful.

##### Emphasis

* Emphasize the graphic’s data.

The main emphasis of color in a statistical graphics screen should be in the data area. Brighter colors and highlighting should attract the eye to the presented data so that trends and conclusions can be quickly perceived. Supporting text, numbers, and legends should receive slightly less emphasis. Aids in data interpretation such as grids should receive the least emphasis.

##### Number of Colors

* Use no more than six colors at one time.
* Use one color of five values or lightness.

Experience indicates that displaying more than six colors at one time on statistical graphics screens is “too much.” Even five or six colors, however, may be distracting or confusing if they are not properly chosen or are not harmonious. Marcus (1986a) suggests a more pleasing arrangement can often be achieved for graphics with five or less segments by using one color and displaying each segment in a different value or lightness.

##### Backgrounds

* Surround images:
  + In a neutral color.
  + In a color complementary to the main image.

A neutral background will help set off a full color. A background in the comple- mentary color of the main image will minimize visual afterimages.

##### Size

* Provide images of an adequate size for the task.
* If the image changes in size, use colors that exhibit a minimum shift in hue or lightness.
  + White, yellow, and red on dark backgrounds.

As color areas decrease in size, they appear to change in lightness and saturation. Similar colors may look different, and different colors may look similar. Interactions with the background color also increase. Thin gray images (lines or borders, for example) appear as a desaturated color complement of their background.

Provide *adequately sized* images. Where color identification is important, an image should be large enough to eliminate these distortions. For images *changing in size*, use colors that exhibit minimal hue or lightness shifts. Marcus (1986b) recommends that white, yellow, and red be used for light text, thin lines, and small shapes on dark back- grounds (blue, green, red, light gray).

##### Status

* To indicate a status, use the following colors:
  + *Proper, normal, or OK:* Green, white, or blue.
  + *Caution:* Yellow or gold.
  + *Emergency or abnormal:* Red.

To indicate a status, use green, white, or blue to indicate OK; yellow or gold for cau- tion; and red for emergency or abnormal. The use of red, yellow, and green are well- learned color conventions.

##### Measurements and Area-Fill Patterns

* Display measurements in the following colors:
  + Grids: Gray
  + Data points: Yellow
  + Variance or error bars: Blue
  + Out of specified range data: Red
  + Captions and labels: Lavender, lime green, or cyan
* Display area-fill patterns in the following colors:
  + Widely spaced dots: Red
  + Closely spaced dots: Green
  + Wide dashed lines: Magenta
  + Narrow dashed lines: Cyan
  + Wide crosshatch: Blue
  + Narrow crosshatch: Yellow

For *measurements,* Smith (1986) recommends the above. They balance emphasis con- siderations (gray for grids, yellow for data points, lavender, lime green, or cyan for labels) and human expectations (red for out-of-specified range). Marcus (1986a) recommends that all text and the horizontal and vertical axis lines of a statistical graphic should be off-white. This will aid in focusing users’ main attention on the colored data. To ensure that fill-in area patterns are identifiable, discriminable, and free from unintended brightness effects, Smith (1988) recommends the above.

##### Physical Impressions

* Size:
  + To convey an impression of:
* Larger: Use bright or saturated colors.
* Smaller: Use dark or desaturated colors.
* Similar: Use colors of equal lightness.
* Weight:
  + To convey an impression of:
* Heavy: Use dark, saturated colors.
* Light: Use light, desaturated colors.
* Distance:
  + To convey an impression of:
* Close: Use saturated, bright, long-wavelength (red) colors.
* Far: Use saturated, dark, short-wavelength (blue) colors.
* Height:
  + To convey an impression of height, use desaturated, light colors.
* Depth:
  + To convey an impression of depth, use saturated, dark colors.
* Concentration level:
  + To convey an impression of concentration level, use:
* High: Saturated colors.
* Low: Desaturated colors.
  + - Magnitude of change:
      * To convey an impression of magnitude of change, use:
        + Lowest: Short-wavelength (blue) colors.
        + Highest: Long-wavelength (red) colors.
    - Actions:
      * To convey an impression of action, use:
        + Required: Long-wavelength (red) colors.
        + Not required: Short-wavelength (blue) colors.
    - Order:
      * To convey an impression of order with color, use:
        + Low end of a continuum: Short-wavelength (blue) colors.
        + High end of a continuum: Long-wavelength (red) colors.
      * When displaying an array of ordered colors, position:
        + Short-wavelength colors at the left side or at the bottom.
        + Long-wavelength colors at the right side or at the top.
      * To convey an impression of order with value or lightness, use the lightness order of a color (darkest to lightest or vice versa).
    - Neutrality:
      * To convey an impression of neutrality, use black, gray, and white.

Colors yield different physical impressions. Bright, saturated colors convey a feeling of large and close. Dark, saturated colors mean heavy, far, and impression of depth. De- saturated, light colors indicate a light weight and height. Desaturated dark colors mean smaller. Long-wavelength (red) colors are associated with high rate of change, action required, and the high end of a continuum. Short-wavelength (blue) colors are associ- ated with low rate of change, no actions required, and the low end of a continuum. Neutrality is best indicated by black, gray, or white.

#### Choosing Colors for Web Pages

* + - Purpose:
      * Color must always have a meaningful purpose.
    - Palette:
      * Use the browser 216-color palette.
    - Presentation:
      * Minimize the number of presented colors.
      * Always consider color in context.
      * Use similar or the same color schemes throughout.
        + For foregrounds: Use black or strong colors for text and headings.
      * For backgrounds: Use weaker contrasting colors such as off-white or light gray.
      * Use a uniform color in large areas.
  + The smaller the element, the more contrast is required between it and its back- ground.
  + Larger images should use:
* Flat, Web-safe colors.
* Fewer colors than small images.
  + Select colors that can be easily reproduced in black and white.
* Links:
  + Use default colors for links.
* Make unselected/unvisited links blue.
* Make selected/visited links purple.
  + Do not display non-link text in link colors.
* Testing:
  + Test all colors.

**Purpose.** Color should always be used for a meaningful purpose. Acceptable uses include highlighting or calling attention to information, grouping related infor- mation, designating selected links, giving a site an identity, or communicating re- alism for images. Color without a purpose is gratuitous and visually distracting.

**Palette.** Use colors that succeed on a variety of platforms or browsers. There are 216 col- ors (out of the standard 256 colors) that will always look the same on any monitor at any resolution. This is called the *browser-safe* palette, and it is illustrated on several Web sites. A search on “color palette” or a similar term will lead to these Web sites.

**Presentation.** Colors should be carefully chosen to aid users in understanding con- tent, to keep a page well balanced, and to keep it graphically simple. Always *min- imize* the number of presented colors. Too much color makes it difficult for people to notice something that might be really important, and makes it difficult to com- prehend how color is being used to aid in understanding the screen’s content. Christmas trees should be reserved for December 25; the “Las Vegas effect” (as the gratuitous use of color is also sometimes called) should be confined to a region in southern Nevada. Minimizing the use of color will also have page download benefits. Fewer colors means faster downloads.

Always consider color in *context,* never in isolation. One background color in- teracts with other background colors. Foreground text colors and graphics inter- act with background colors. When choosing colors, consider these interactions and use colors that work well together. Use *similar* or the same color schemes throughout a Web site. This will give the site an identity and help the user main- tain a sense of place.

*Foreground* colors should be as different as possible from *background* colors. A contrasting foreground will maximize text legibility. A background color should organize a group of elements into a unified whole, isolating them from the re- mainder of the screen. Use background colors that do not compete with the fore- ground. A background must be subtle and subservient to the data, text, or symbols on top of it. With today’s high-resolution monitors, the most recom- mended foreground text color is black presented on a light-colored background of low intensity, either off-white or light gray. Most other pastel colors, especially spectrum center colors, have also been found to be acceptable as backgrounds

(Pastoor, 1990). Use dark backgrounds only when establishing contrast between an area of the screen and the main screen body. High intensity colors as back- grounds (such as red, magenta and bright green) must be avoided as they can be visually fatiguing when viewed for a period of time. Short wavelength, cool col- ors (blue and black) have been found to be preferred for dark backgrounds (Lalo- mia and Happ, 1987). Cool, dark colors visually recede, providing good contrast to the advancing lighter, foreground colors. Blue is especially good because of the eye’s lack of sensitivity to it in the retina’s central area and increased sensitivity to it in the periphery.

When choosing foreground and background colors, ensure that contrasting combinations are selected. See Tables 12-3 and 12-4 for color selection guidance. Also, always test all selected foreground and background colors. What may look good in theory may not always look as good in reality.

Use a *uniform* color in large screen areas. Large areas of the same color down- load faster. They compress well and are an efficient use of the GIF format. The *smaller* the element, the more contrast is required between it and its background. To reduce image file sizes for *larger* images, use flat Web-safe colors. Also use fewer colors with smaller images. Finally, select colors that can be easily *repro- duced* in black and white on a screen or printer. This ensures that those who use a monochrome display or print in black and white will have a faithful reproduction of the intended screen.

MAXIM *Content* is always more important than color.

**Links.** Unselected/unvisited links must be distinguishable from selected/visited links. Stick with the default colors of *blue* for links already followed and *purple* for links not yet ventured down. While the choice of blue as a text color was poor because of its degraded reading ability, it is well learned, its use is recommended because it is now very familiar. Using nonstandard link colors can lead to severe problems. It is difficult to remember what color means what, increasing link se- lection errors. It can also lead to confusion with normal underlined text in a doc- ument. Never display general screen text in the link colors of blue and purple. This will create confusion between linked and non-linked text. It is acceptable, however, to use these colors in text that is large and decorative in nature and act- ing as a graphic and not plain text.

**Testing.** The possibility always exists that identical colors may appear differently on different monitors and platforms. Color choices should be tested on a variety of displays.

#### Uses of Color to Avoid

* + - Relying exclusively on color.
    - Too many colors at one time.
    - Highly saturated, spectrally extreme colors together:
      * Red and blue, yellow and purple.
    - Low-brightness colors for extended viewing or older viewers.
* Colors of equal brightness.
* Colors lacking contrast:
  + For example, yellow and white; black and brown; reds, blues, and browns against a light background.
* Fully saturated colors for text or other frequently read screen components.
* Pure blue for text, thin lines, and small shapes.
* Colors in small areas.
* Color for fine details.
* Non-opponent colors.
* Red and green in the periphery of large-scale displays.
* Adjacent colors that only differ in the amount of blue they possess.
* Single-color distinctions for color-deficient users.
* Using colors in unexpected ways.
* Using color to improve legibility of densely packed text.

The proper use of color in screen design also suggests some things to avoid.

**Relying exclusively on color.** Consider the needs of color-blind viewers and the ef- fects of ambient lighting on color perception. Do not underestimate the value and role of other techniques such as spatial formatting and component locations in good screen design.

**Too many colors at one time.** Using too many colors at one time can eliminate the ben- efits of color. Response times are increased, erroneous associations made, the han- dling of information is interfered with, and confusion is created. The objective is a screen that communicates; a colorful screen is not the objective. Use just enough colors to create effective communication. Again, consider the value of other tech- niques such as spatial formatting and consistent component locations in good design.

**Highly saturated, spectrally extreme colors together.** Spectrally extreme combina- tions can create eye focus problems, vibrations, illusions of shadows, and afterim- ages. In addition to red/blue and yellow/purple, other combinations that might cause problems are yellow/blue, green/blue, and red/green.

**Low-brightness colors for extended viewing or older viewers.** The eye adapts to color during extended viewing. The eye’s capability also diminishes with age as the amount of light passing through the lens decreases. All colors will look less bright, and colors that are dim to begin with may not be legible. Brighter colors are needed to prevent reading problems.

**Colors of equal brightness.** Colors of equal brightness cannot be easily distinguished.

A brightness difference must exist between adjacent colors.

**Colors lacking contrast.** Colors lacking contrast also cannot be easily distinguished. Similar foreground and background colors often do not have sufficient contrast with each other.

**Fully saturated colors for frequently read screen components.** Fully saturated col- ors excessively stimulate the eye, possibly causing visual confusion and fatigue.

**Pure blue for text, thin lines, and small shapes.** Because of its physical make-up, the eye has difficulty creating a clear and legible image of small blue shapes. They will look fuzzy.

**Colors in small areas.** Distortions in color, lightness, and saturation may occur for small areas of color.

**Colors for fine details.** Black, gray, and white will provide much better resolution.

Reserve other colors for large areas or attracting attention.

**Non-opponent colors.** Non-opponent colors, red/yellow or green/blue, produce poorer images. Opponent colors, red/green or yellow/blue are good combinations for simple displays.

**Red and green in the periphery of large-scale displays.** The edges of the retina are not particularly sensitive to red and green. Avoid these colors in the periphery, especially for small symbols and shapes. Yellow and blue are much better periph- eral colors.

**Adjacent colors only differing in the amount of blue they possess.** Because of the eye’s difficulty in dealing with blue, differences in color based upon varying amounts of blue in the color’s mixture will not be noticed.

**Single color distinctions for color-deficient users.** For those people with color- viewing deficiencies, distinguishing certain colors may be difficult or impossible. Always provide a redundant coding scheme.

**Using colors in unexpected ways.** Colors have become associated with certain meanings. Red, for example, is always associated with stop or danger. To display a critical or error message in green would violate an ingrained association and cause confusion. Always use colors in the way people expect them to be used.

**Using color to improve legibility of densely packed text.** Space lines between para- graphs of text or after about every five lines of data will work much better.

**Too many colors at one time (again).** Never overuse color (again, this is important). Too many colors at one time may make a screen confusing or unpleasant to look at. Use only enough color to fulfill the system’s objectives.

**Table 14.4** Continued

**TEST: ENSURE THAT:**

Backgrounds and color. Backgrounds are compatible with foregrounds.

Backgrounds are not distracting. Monitor color rendering differences do not negatively affect site usability.

Graphics and icons. Graphics are meaningful and efficient. Graphics are not distracting to the user.

Page breaks. Actual page breaks are obvious.

False signals for page breaks do not exist.

Page printing. The page prints completely and does not bleed off the page.

Accessibility. Compatibility with accessibility utilities.

Compatibility with accessibility guidelines.

##### Test Participants

* Assemble the proper people to participate in the test.

Assembling the proper people to participate in the test is critical to its success. Al- ways recruit participants with the proper qualifications, those currently performing the job where the product will ultimately be used. While the “boss” may profess to be knowledgeable, he or she is too far removed from the grind of daily activities and sel- dom knows exactly what is going on. Also, recruit participants covering the spectrum of user characteristics, including age, gender, and experience, in order to allow general conclusions based on the test results. Recruit strangers, not friends. A stranger is less likely to withhold comments that might possibly hurt the tester’s feelings. There will also be less embarrassment if problems are uncovered.

Nielsen (2000) recommends that the optimum number of users for a usability test is five. He states that research indicates about 85 percent of a system’s usability problems will be identified with five participants, and the law of diminishing returns sets in at this point. It is much better, he reasons, to conduct more types of tests with fewer par- ticipants, than to conduct fewer tests with more participants. All in all, this is a very rea- sonable recommendation. One caveat—if a system like a Web site has several highly distinct groups of users (children, parents, and senior citizens, for example) then their interface behaviors may differ. In this case, each different group must be tested. He rec- ommends using 3 to 4 users from each category if one is testing two groups of users, and three users from each category if one is testing three or more user groups. Also, do not forget to include users with disabilities.

Always advise selected test participants of what to expect in the test. They will ap- proach the test with less apprehension. Finally, never refer to test participants as “sub-

jects.” While many of us, including myself, have been conditioned to use this term, and have used it for very many years, the correct term today is “evaluator” or “participant.”

##### Test Conduct and Data Collection

To collect usable data, the test should begin only after the proper preparation. Then, the data must be properly and accurately recorded. Finally, the test must be concluded and followed up properly. Following are guidelines for conducting a usability test. Many are from Schrier (1992).

###### Usability Test Guidelines

* Before starting the test:
  + Explain that the objective is to test the software, not the participants.
  + Explain how the test materials and records will be used.
  + If a consent agreement is to be signed, explain all information on it.
  + If verbal protocols will be collected, let participants practice thinking aloud.
  + Ensure that all participants’ questions are answered and that participants are comfortable with all procedures.
* During the test:
  + Minimize the number of people who will interact with the participants.
  + If observers will be in the room, limit them to two or three.
  + Provide a checklist for recording:
    - Times to perform tasks.
    - Errors made in performing tasks.
    - Unexpected user actions.
    - System features used/not used.
    - Difficult/easy-to-use features.
    - System bugs or failures.
  + Record techniques and search patterns that participants employ when attempting to work through a difficulty.
  + If participants are thinking aloud, record assumptions and inferences being made.
  + Record the session with a tape recorder or video camera.
  + Do not interrupt participants unless absolutely necessary.
  + If participants need help, provide some response.
    - Provide encouragement or hints.
    - Give general hints before specific hints.
    - Record the number of hints given.
  + Watch carefully for signs of stress in participants:
    - Sitting for long times doing nothing.
    - Blaming themselves for problems.
    - Flipping through documentation without really reading it.
  + Provide short breaks when needed.
  + Maintain a positive attitude, no matter what happens.
* After the test:
  + Hold a final interview with participants; tell participants what has been learned in the test.
  + Provide a follow-up questionnaire that asks participants to evaluate the product or tasks performed.
  + If videotaping, use tapes only in proper ways.
* Respect participants’ privacy.
* Get written permission to use tapes.

Before Starting the Test

Most people participating in a test will approach it with some anxiety. Fears may exist that they themselves are being judged, or apprehension may be caused by a general fear of the unknown, a common human trait. These fears must be put to rest. Before the test begins, all participants must be told exactly what will happen in the test. Explain that the test objective is to evaluate the software, not the participants themselves. Also explain how the test materials and records will be used. If participants will be signing a consent agreement, review and explain all information on it before it is signed. If ver- bal protocols will be collected, that is, if the participants are going to be asked to think aloud, let participants practice this process. For most people this is not a common ex- perience, and it may require getting used to. Finally, do not start the test session until all participants’ questions are answered and people are comfortable with all of the test procedures. Providing this kind of information, and preparation, will enable partici- pants to relax faster at the start of the test.

During the Test

Minimize the number of people who will interact with participants. Many and strange voices must be avoided because they can be very distracting and disturbing. If observers will be in the room during the test, limit their number to two or three. Observers must *never* talk during the test.

For data recording, provide observers with a checklist reminding them what to record and for use in actually recording data. Useful information to collect includes the time needed to perform each of the test tasks, errors made, any unexpected actions taken by the participants, how often system features are used, those features that are not used, difficulties in using features, features that are particularly easy to use, and system bugs or system failures. When participants encounter a difficulty, record the techniques and search patterns they employ when attempting to work through the difficulty. If partic- ipants are thinking aloud, record the assumptions and inferences they make as they proceed. If practical, record the test with a tape recorder or video camera. This will per- mit more leisurely review of the test session later. Details missed during the session will be uncovered, and comparisons can be made between the approaches and activi- ties of the different participants. The entire design team will also be allowed to later re- view and evaluate the test results.

Never interrupt a test participant unless it is absolutely necessary. If, however, it is sensed that participants need help, provide a response of some kind, first through encouragement and then through hints. Provide general hints before specific hints, and record the number of hints given. Watch carefully for signs of stress. If it is detected, again give encouragement or provide hints, or provide a break. Signs of stress include a participant’s sitting for a long time doing nothing, blaming him- or herself for prob- lems, and flipping through documentation without really reading it. Provide short breaks when they are needed, and maintain a positive attitude no matter what happens (everything probably will). A tester with a negative attitude will influence the partici- pants in the same way, and the data collected will be contaminated.

After the Test

At the test’s conclusion, hold a closing interview with the participants. During this inter- view, questions that occurred to the tester during the actual test can be asked; the partic- ipants can also ask questions, and the tester can tell the participants some of what has been learned. This will make the participants feel that their effort was worthwhile. Also provide follow-up questionnaires that ask participants to evaluate the product or tasks performed. Finally, if videotaping is performed, respect the participant’s privacy when the tape is later shown. If necessary, obtain the participant’s written permission to later use the tape.

#### Analyze, Modify, and Retest

* Compile the data from all test participants.
* List the problems the participants had.
* Sort the problems by priority and frequency.
* Develop solutions for the problems.
* Modify the prototype as necessary.
* Test the system again, and again.

**Data analysis.** Compile all the data collected from all test participants by listing the problems the participants had, sorting the problems by priority and frequency. See Table 14.1 for a problem-rating scheme. Make the results available to the en- tire design team for analysis, again to provide multiple insights into problem so- lutions. Then, develop solutions for the problems. Get expert advice if the solutions are not obvious.

**Prototype modification.** Prototypes must, of course, be modified based on the design recommendations made during testing, and the solutions decided upon.

**Test again.** The testing process continues in an iterative manner until all problems are satisfactorily solved and all criteria are met. After the prototyping is complete and all code written, a final system test must be performed to assure no software bugs exist and performance meets all specifications. The screens and interface must also be again tested to assure all established usability criteria are being met. The design steps and methods are identical to those for prototype testing.

#### Evaluate the Working System

* Collect information on actual system usage through:
  + Interviews and focus group discussions.
  + Surveys.
  + Support line.
  + Online suggestion box or trouble reporting.
  + Online bulletin board.
  + User newsletters and conferences.
  + User performance data logging.
* Respond to users who provide feedback.

Testing never stops with system implementation. The interface, like any part of a sys- tem, must be continually evaluated to ensure that it is achieving its design objectives. Problems detected can be corrected in system enhancements and new releases. This type of evaluation is necessary for a variety of reasons.

**Problems will have slipped through the cracks.** In spite of all the preimplementa- tion testing performed, problems will still exist. It is impossible to exercise all as- pects of a system in a testing environment as thoroughly as is done in actual system use. Also, actual use can capture the experiences of all users, not simply those used in the various testing phases.

**Initially impressive features may later be regarded as frustrating or completely ignored.** Initial impressions may change over time. Some parts of a system, which in the testing process were seen as neat or helpful, may, in everyday use, be found to have little or no value.

**Experienced users are more sensitive to time delays.** Response times that seemed adequate in the testing process may, as users become experienced, become a source of great irritation.

**Customizable interfaces may change.** Over a period of time customization may change the interface so much that original features are lost and tasks become more difficult to perform.

**The external environment may have changed.** The original system hardware used in the testing process may change over time. Added elements or features may affect the system’s usability in a negative way, or allow enhancement of the interface.

Many of the techniques used in requirements determination and prototype evaluation can also be applied in this evaluation phase.

**Interviews and focus group discussions.** Individual user interviews can identify specific areas of concern. These can be pursued to the level of detail needed. Focus groups can then be conducted to determine the generality of the identified problems.

**Surveys.** A survey or questionnaire is administered to users to obtain feedback on usability problems. These surveys can use e-mail, the Web, or traditional mail.

**Support line.** Information collected by the unit that helps customers with day-to- day problems can be analyzed (Customer Support, Technical Support, Help Desk, and so on).

**Online suggestion box or trouble reporting.** An online suggestion box can be im- plemented to solicit comments or enhancement suggestions.

**Online bulletin board.** An electronic bulletin board can be implemented to solicit comments, questions, or enhancement suggestions.

**User newsletters and conferences.** Improvements can be solicited from customer groups who publish newsletters or convene periodically to discuss software usage.

**User performance data logging.** The system software can capture users’ usage pat- terns and problems doing real tasks. When one uses log data, however, a user’s goals and needs can only be surmised.

**Respond to users who provide feedback.** When feedback is received, respond in an appreciative manner to the person providing it. A positive relationship will be es- tablished, and the person will be more likely to provide additional feedback. A failure to respond will discourage future suggestions and indicate to the user that his or her needs are not important. Finally, if a system change is made based upon a suggestion, inform the user or users making the suggestion of the change.