

# Beyond a Narrow Conception of Usability Testing

PATRICIA SULLIVAN

*Abstract*—Construed narrowly, “usability testing of documentation” can be limited to validating the usability of a nearly completed draft. This paper explores ways that technical communication can take a broader view of usability and then situate new approaches and studies. Appropriately situated, this broader interpretation can give technical communicators ways to talk with others interested in usability and can build a platform for an understanding of usability research that looks beyond “testing drafts for usability” to the tough issues driving usability research.

“Usability testing of documentation” can sound simple if we adopt the narrow definition Weiss critiques—“draft tested in a usability lab.” [1] It can even sound simple when we say that usability investigates whether users can successfully use computer products. But the study of users and their use of products is not straightforward. Consider some of the decisions faced by a particular study—how to measure *using*, what to consider *successful*, or where in the development cycle to conduct the test. These decisions are directed by the goals of the study. However, not every kind of investigative measure yields the same kind of information; success at getting started does not imply later ease of use, and testing of a working system in a natural environment has radically different requirements than a test of mockups for a new text design. The complexities of studying usability naturally push us toward a broader view.

In this article, I argue that we technical communicators need to interpret the charge of usability more broadly than a narrow interpretation of usability testing allows, turning to the whole range of usability research for a conception that takes advantage of the exciting investigations that are being carried out in various places. But there are consequences to broadening our view. The broad view of usability includes so much work that we need to build a framework for the studies. Thus, I argue that we handle the problems by situating a particular investigation in light of its research group, its research questions, its methods, the issues it settles or raises, and its major constraints. The situating will help us fashion a coherent landscape of usability, a landscape that will help us to assess more cannily what we know about usability and what we need to know. To that end, I examine the groups, the methods, and the research questions associated with usability research to begin to depict that landscape of usability research from the viewpoint of technical communication. In essence, I ask: What is it that we are doing in relation to others who study usability, and what might we need to be doing?

Patricia Sullivan, who received the Ph.D. from Carnegie Mellon University, directs the technical writing program at Purdue University. She has won the National Council of Teachers of English award for the best article in teaching scientific and technical communication.

## A RATIONALE FOR BROADENING USABILITY TESTING TO USABILITY RESEARCH

Weiss ably critiqued usability testing of documentation conceived in its narrowest form and called for usability to follow the lead of engineering and insert usability into the product development cycle at an earlier phase. [1] He called for developing the usability engineering of documentation. Most of us concerned with usability share Weiss' goal of inserting usability concerns into the entire development process. But the question immediately raised is how do technical communicators impact the process earlier? Some of the usability problems in documentation are really system or interface problems. How do we connect with the other processes in ways that will improve the general usability of computer products? I argue here that the first step is to take a broader view of usability, to look at the larger body of usability research that has been undertaken for other reasons, and to situate our work in relation to the broader goals of improving usability. I think it makes sense for technical communicators to be involved in the larger quest for usability because we have a close connection to end users, because we have an audience theory, and because we have a growing body of information about users in action. But I also think interdisciplinary work needs to begin with a cautious study of the “other.”

Thus, the focus of this article is on exploring and situating the varieties of usability research. I begin with a situated definition for usability research that also articulates an analytic tool for building situated definitions. I then consider several of the components that filter the results in more detail: the groups conducting the research, the methods they typically employ, and the questions they pose. I then apply this discussion to the task of framing the issues and opportunities facing usability research.

## A GENERAL APPROACH TO SITUATING USABILITY RESEARCH

If we take a broader view of usability, then we must deal with a flood of information, not all of which is ultimately relevant. Usability research, broadly construed, includes the work of people who design systems, test them, develop educational materials, and study users. The subject matter of usability research can be usefully described as the studies of users learning to use and using various computer products, in the design and evaluation phases of development, for the pur-

poses of learning more about how people use machines and of improving the machines and educational materials. How do we make sense out of it all? We remember our interest is in helping users to learn, and we approach situating new material within that general stance of advocating users and their learning.

Two steps can be used to situate those studies *vis-à-vis* the landscape of usability research, an analysis of the context surrounding a particular study and an analysis of the structure of the study itself. The reason for dividing this “situating” into a two-step process arises out of the multidisciplinary nature of the usability research project. When we are working inside a discipline, everyone shares values and goals, a common educational platform, and a common world view. We cannot assume that commonality in multidisciplinary work. Thus, the step of articulating the context of a study helps us to better understand why a study is as it is.

The first step in situating a study is to build an interpretation of the context that the study grows out of. We begin to situate studies when we place the study:

- In a time
- In a place
- In an investigative group
- In a project/event sequence
- In a society
- In a set of research goals

Posing these *place* questions about a study clarifies it in several ways. It makes clear what components of the study were (or should have been) focal to the effort. It also clarifies why particular groups were studied and why certain methods were selected. Further, it allows others to interpret and apply that study’s findings to situations that are analogous, at least roughly. If we consider the IBM Watson work on the minimal manual in context, for example, we see that the minimal manual was a part of a larger research effort to study why computer programs were difficult to learn and how various components of the system might aid the learner. This same group developed an interface that had “training wheels” (or advanced features disabled) as another response to the learning problems people experienced.

The second step in situating a study is a more conventional analysis of the structure of a usability study. It assumes that decisions summarized in table 1 are made in such a way that a study is done:

- With a group of target users
- Using a product
- In a setting
- To perform a task
- For a reason
- With a result

Placing a usability study within both of these frameworks—the contextual setting and the structural setting—helps us understand it more reliably.

TABLE 1  
QUESTIONS USEFUL FOR ANALYZING THE STRUCTURE  
OF A USABILITY STUDY

What is the focus of the study?	user task product
What product is focal?	interface CBT online documentation paper documentation coordinated use
What constrains the study?	cost facilities time/timeliness availability of materials/subjects appropriateness of methods
Where in the product cycle is the study placed?	beginning of development midcycle near the end pre beginning of new cycle
What do we know about the users?	classification of users beginning to learn later in learning (in/not in) environment of use
What methods are used?	product development model (e.g., iterative lab tests) cognitive model (e.g., protocols) cultural model (e.g., field test) combination
What users are the results put to?	product development diagnosis feedback validation critique theory-building users methodology systems learning writing

The process is often not as transparent as the two-step process outlined above. Because this research is interdisciplinary and because of the current conventions for reporting/publishing research, information that is important to situating the study must often be inferred. For example, in a published study, most journals do not provide a biography of the authors. But the groups’ identity reveals considerable information about their approach to a study, the methods they choose, and their research questions. Thus, it becomes important to examine components that filter the results, namely, information about the identity of the researchers, information about the methods, and information about the research questions.

#### *Filter 1: Identities of Groups Conducting Usability Research*

One major filter for a study is the group that conducts it. Several overlapping groups study usability, approaching it with different questions in mind. Psychologists and engineers in human-computer interaction typically use experiments and case studies to study the usability of interfaces and systems; sociologists and anthropologists use ethnography and field

methods when they study the computing of organizations; marketers typically use interviews and surveys to study consumer preferences; document designers, educational psychologists, and writers in technical communication use various exploratory and text-based methods to study the usability of educational materials.

*Human-Computer Interaction*—One group involved with usability research is human-computer interaction, an interdisciplinary linking of ergonomics engineers, human factors psychologists, and cognitive scientists. The focus of their research has been on the design of interfaces, that space where humans and computers interact. Human-computer interaction, because it is interdisciplinary, does not project a single attitude; its strictest cognitive psychologists aim to build theories of users or of learning and its strictest engineers aim to build systems that solve problems they notice. The strain of differing goals can be seen at times. In *Interfacing Thought*, a recent book edited by Carroll [3], cognitive psychologists gathered together work they saw as practical theory of users and systems. But the review of the book in *SIGCHI Bulletin* charged the text with being too theoretical and exclusively aimed at psychologists:

If the authors wish to bridge the gap between theory and application, the best approach is to demonstrate application of the theory. . . Examples of developed systems are needed; no such examples are included in this book. [4]

The review goes on to say that the book should be read by cognitive scientists, as working designers would not know how to deal with the articles. The only article deemed accessible is one by Whiteside and Wixon, an article that called for the study of human-computer interaction in the broadest context possible and was critical of laboratory studies, citing an example of an installation study that worked in the lab but was inhibited in real life by the boxes that piled up in the person's office.

The uneasiness is not restricted to psychologist-engineer clashes, as psychologists disagree on the relationship of theory/practice as well. In 1985 and 1986, for example, there was a heated exchange in *Human-Computer Interaction* between Carroll and Campbell [5] and Newell and Card [6] over an article Newell and Card [7] wrote about building a psychological science in human-computer interaction. Carroll and Campbell criticized the article for calling researchers to make a "hard" science rather than sticking to the top priority of human-computer interaction, namely, "the articulation of an alternative explanatory program, one that takes as its starting point the need to understand the real problems involved in providing better computer tools for people to use." Newell and Card replied that identifying underlying mechanisms is basic to human-computer interaction. The tensions exposed in this exchange serve to remind us that human-computer interaction researchers are not a monolith pursuing science. They are united as they pursue the work of improving computer systems; they focus their work on interfaces; and they value "IT WORKS FOR PEOPLE." But they differ on their pursuit of theory.

*Sociology*—Another group, though peripheral to the day-to-day study of usability, is made up of sociologists and anthropologists. These researchers have contributed studies of people working with computers that point out dimensions of usability that are not seen in a laboratory study. Their studies also serve as methodological guides for naturalistic studies of usability. Two areas are typically studied, education and office environments. Turkle, in a typical sociological study of education, pointed to ways that interacting with computers challenged the child's distinction between physical and social. [8] Suchman, in a study of an office environment, tackled the question of the nature of the interaction that people had with the expert help system for a copy machine. [9] She found that communication breakdowns may be traced, in part, to the process used in design to logically represent solution paths, and therefore she argues that more attention needs to be paid to the differences between people and computers.

A problem with classical fieldwork is that it is founded on interviews and notes taken by the researcher (after the observation). It assumes that the people are competent in their grasp of their own culture, and that the researcher participates until she/he understands the culture enough to go back and tell us about it. This assumption has trouble when the study deals with the introduction of new elements into a culture. Then the "natives" are not expert informants. Further, the interactions with the machine and the documentation are so rapid and complex that it is difficult for the researcher to capture the problems in notes after the fact. Thus, classical methodology is in the process of being adapted to the needs of studying usability. Successful adaptation is needed if field studies are to make a robust contribution to usability.

*Marketing*—Marketers, another group sometimes involved with usability research, have goals developed out of their responsibility in a company to promote profitable products. To this end, the marketers are constantly working to build markets and satisfy customers. Usability can be a part of their research. When marketers are interested in usability, that interest is product specific and is customer satisfaction driven. They tend to run focus groups about current and potential products, analyze the customer response cards for a product, and conduct telephone surveys. A product that works well and is an innovation will not be seen as successful if the customers are dissatisfied.

*Technical Communication*—A final group is technical communicators. Writers, document designers, and educators focus on the educational materials that surround a system and on how users employ those aids to help them learn a system. This collection of groups is responsible for producing usable documentation, online help, and computer-based training, and meets together in technical communication organizations. Like those in human-computer interaction, technical communicators are educated in a number of disciplines, including English, education, graphics, and technical writing. Most have no training in empirical research and find the empirical investigation of usability a challenge. At the same time,

technical communicators have sophisticated methods for analyzing audiences and an empathy for users, giving them a unique contribution to the development of computer products.

Technical communication's approach to research can be traced to competing ties to humanities (and qualitative research) and to education (and both to the grading of texts and to the Gagne approach to instructional design). On one hand, their close ties to humanities promote the valuing of qualitative research. Case studies are the most frequently reported research in the technical communication venue. Further, a number of researchers have begun to explore field studies of how people actually use systems and documentation. Gould and Doheny-Farina, for example, argue for the importance of field studies to the development of usable documentation by pointing to information that writers can gain in the field. [10] But this development is quite new, and to date, adaptation of field methods to the study of usability has been a major topic. [11] On the other hand, technical communication's close ties to education have also influenced the types of research pursued. Finding an efficient way to "grade the text" has been a persistent goal, an idea founded on the notion that a text holds the meaning. Writers have looked for a textual measure of quality, and of readability, and of comprehensibility, and of usability. Warren, for instance, articulated a case for readability formulas as a measure of usability. [12] Further, military research into the structuring of manuals has adapted Gagne's approach to instruction. [13] In the final analysis, the tension between exploration and grading has encouraged the variety of approaches to the study of usability in technical communication. They hold equally compelling beliefs that the natural experience can be a test of usability and that grading the text would offer the best (most efficient) test of usability.

In general, few usability researchers think about the missions of usability researchers outside their fields. Usability of documentation can illustrate the differing views. Baecker and Buxton, in a text that features the importance of usability to the field of human-computer interaction, reflect the marginal position of documentation to that field as they relegate documentation to a miscellaneous chapter near the end of the book. [14] Even their forward-looking text sees writing as "tacked on." Writers, on the other hand, generally make human-computer interaction just as marginal to their effort. As both Nelson [15] and McDowell [16] point out, human factors is not a recognized part of the education of technical communicators. Thus, we can expect that few technical communicators read what human-computer interaction researchers are saying about users and usability and then think about how that work may apply to the writing of documentation.

#### *Filter 2: Methods Employed in Usability Research*

Methods provide another filter on a study. Because many of the researchers in technical communication and in human-computer interaction view research from a problem-solving rather than a theory-building framework, they treat methods as guidelines rather than as safeguards in the process of experimentation and they adapt methods freely. The "freeness"

springs in part from the interdisciplinary nature of the field, in part from the exploratory nature of the task of improving usability, and most of all, to the newness of it all. There hasn't been time for a particular method to gain dominance over the others by being legislated through the education of all those people who do usability research.

*Methods Employed in Usability Research*—To survey the methods, first take the narrow lens of usability testing. Weiss argues that usability test methods are monolithic:

Draft manuals are crash-tested by appropriate subjects, who are interviewed before and after, and observed during. The statistical results are even bounced off significance tests. [1]

But articles that articulate the lab methodology Weiss critiques point to the variety possible within that lab orientation. Dumas, for example, described three of a number of methods: designer-user dialogues, a number of users working together, and a user working alone. [17] All sorts of studies are possible in a well-equipped lab.

Methods used are numerous. Conservatively, they include the following:

- Direct questioning
  - Surveys
  - Interviews
  - Comprehension tests
- Observation
  - Informal observation
  - Laboratory observation
  - User protocols
  - Reading protocols
  - Keystroke records
- More traditional evaluation methods
  - Computer text analysis
  - Editorial review
  - Technical review

These methods are used alone or put in combination with others to study the question of interest. In practice, most laboratory and field studies blend together direct questioning and observation methods into one study.

The direct questioning methods of surveys and interviews tell us what users think and feel; comprehension tests tell us what they learned. For example, surveys and interviews are used alone very early in the product cycle for needs analysis and very late in the product cycle for customer acceptance; they are also linked to most observations as a way to get information about response to the product; they are also a mainstay in field studies, as researchers rely on interviewing informants from the target culture. These methods are most helpful in finding out what users expect to happen before they use a product, how they respond to using a product, what they thought was happening during a session, their attitudes, their memory, their learning, and their judgments. Surveys, interviews, and comprehension tests are not particularly helpful in finding out what moves users make while they use a product.

The observation methods are the main route to answering

TABLE 2  
VARIOUS EVALUATION METHODS' STRENGTHS FOR YIELDING INFORMATION ABOUT DOCUMENTATION QUALITY

Method	Place- ment (of text on page)	Accuracy (of information)	Level of Explanation (whether explains what users need)	Style (appropriateness)	Satisfaction (user response)
<b>Direct Questioning</b>					
Surveys			may		good
Interviews		may	may		good
Comprehension tests [21]		may	good		
<b>Observation</b>					
Informal observation [22, 23]	good	may	good		may
Lab observation [24, 25]	good	good	may		
User protocols [26]	good	may	good		may
Reading protocols [27]			good	good	may
Keystroke records		good			
<b>More Traditional Evaluation</b>					
Computer text analysis				good	
Editorial review [28]	may	may	may	good	
Technical review [29]		good			

questions about what users actually do when they try to use a product. Observation ranges from an informal method, where the researcher watches a user and takes notes, perhaps interrupting and asking questions, to a very formal method, where a video camera records and time stamps a user at work while a program records the keystrokes used. Users may work naturally, or they may be asked to talk aloud as they use the system (user protocol) or read the text (reader protocol). These methods of observation help answer questions related to what happens when a user tries to use a product. They can capture the time it takes, the actions taken, and the success the user had. If the user talks aloud, they can also tap the strategies and reasons for action. But they do not necessarily answer questions about the user's response to the product.

The traditional evaluation methods involve a computer or an expert evaluating the product for how well it meets preset criteria. Consider evaluation methods for a manual as an example. A computer program such as Writer's Workbench can give valuable feedback about the readability level of a prose passage and can pinpoint a number of stylistic problems. Editorial experts can find style problems more efficiently, and more reliably, than users can. Technical experts can locate problems with the technical content of a manual more efficiently, and more reliably, than users can. Although traditional evaluation handles many problems well, it does not necessarily expose usability problems. It uses some ideal text as the arbiter for the actual text. Only when that ideal text is usable does this type of arbitration work well.

The methods can be usefully grouped according to the philosophical models they adhere to. At least three models can be usefully employed in this effort: a product development model, a cognitive model, and a cultural model.

*Product Development Model With an Example From Testing of Documentation*—The product development model adheres to a pragmatic, engineering model. It focuses on deliv-

ering usability information important to the various stages of product development in a timely manner. Because of the time constraints, it normally employs laboratory rather than field studies. Thus, it selects research methods on the basis of how well they deliver the necessary information. Dieli [18] and Lasselle [19] have identified a number of products that could be tested at each point in the development cycle—from others' products and prototypes in the design phase, to storyboards, to documentation modules, and to full product tests. These lists suggest a range of places to test that are new to usability testers who focus on testing completed, or near-completed, products.

Documentation testing can serve as an example in discussing methods aimed at improving product development. We can group the product tests according to those aimed primarily at yielding information about documentation quality and those aimed primarily at providing information about the usability of a product. Table 2 (adapted from Sullivan, [20]) depicts the methods that are more useful in delivering feedback about quality and includes references to appropriate studies. Table 3 depicts the methods that are more useful in studying the usability of documentation and includes references to appropriate examples. These two groupings point out that usability is a part of typical evaluation (as in Table 2) and is also separated from some of the typical evaluation methods (Table 3). When users and their actions are the focus of the product studies, some of the traditional methods become less attractive.

*The Cognitive Model With an Example From Developing User Models*—The cognitive model aims to model learning and use. Most of the work done within this framework is done by psychologists from the human-computer interaction camp who work in research labs or universities. Though this work employs products, it normally focuses on a finished product chosen for its representativeness or a concocted product that has the features of interest.

TABLE 3  
VARIOUS EVALUATION METHODS' STRENGTHS IN YIELDING INFORMATION ABOUT DOCUMENTATION USABILITY

Method	Situated Use (in workplace)	Strategy for Use (how plan to use)	Action Record (how actually use)	Success (tasks correctly done)	Satisfaction (user's response)
<b>Direct Questioning</b>					
Surveys [30, 31]	may			may	good
Interviews [32, 33]	good	may		may	good
Comprehension tests		may		may	
<b>Observation</b>					
Informal observations		good		good	may
Lab observations [34, 35, 36]		may	best	good	
User protocols [37, 38]		best	good	good	may
Reading protocols [37]		may	good	good	
Keystroke records[39]			good	good	
<b>More Traditional Evaluation</b>					
Computer text analysis [37, 12]					
Editorial review [37]		may		may	
Technical review				may	

The milieu of the cognitive work is the laboratory. Though surveys and interviews may be part of the work, classical experiments or protocols are the mainstays of the research. Informal observation, editorial review, technical review, and field studies are not often employed.

Singley and Anderson's model of learning a word processor serves as a good example of this work. [40] They are modeling the learning of a word processor as a cognitive activity, with a goal of building a model for transfer of skills that can explain user actions at a keystroke level. Although they use skilled office workers as subjects, they downplay the milieu of word processing, the particulars of the word processor itself, and the particulars of the instruction. They are not interested in the success of a particular manual or user satisfaction or program strengths. Singley and Anderson are trying to build an explanation of how people move to a new word processor, making the focus clearly cognitive.

*The Cultural Model With an Example From Describing Social Use*—The cultural model takes a sociological approach to the task of studying usability, and its methods aim to study usability in the context of normal use. Thus, it prefers field studies to laboratory studies. Because this is a relatively new approach for both technical communicators and human-computer interactionists, there are relatively few studies and no set methods to pinpoint. In fact, Campbell, Mack and Roemer [41] call for increasing attention to both field work and to improving its methodology.

Work at Xerox PARC serves as a good example of a blend of social and cultural views with the cognitive view of developing usable systems. Xerox PARC has been interested in collaborative computing systems and in the relationship between computer information systems and the larger culture. Brown and Newman present a philosophy of design that advances the notion that computing systems can affect the social

milieu as they articulate a blending of cognitive and social concerns. [42] Their article is the most visionary statement related to field studies of computing, as it gives a reason for entering the field that fits with a philosophy of computing system design. In the field, researchers can learn about the social complexities of use and gain strategies for (1) improving communication with users and (2) designing systems that support the natural means of social feedback and interaction. Malone, for example, reports on a series of interviews that probe how professionals and clerical workers organize the information on their desks. [43] He finds that reminding themselves of information is as important as pointers to finding information, and that workers view classification of information as their most difficult task. Malone's work fed into the development of a hypertext program (NOTECARDS) that Xerox PARC later built.

#### *Filter 3: Research Questions That Can Be Posed in Usability Research*

Research questions give us another filter on the studies, particularly if we see these questions as means of assessing whether a study focuses on one type of question or cuts a broader swath. By grouping the typical research questions posed about usability into questions that focus on the product, questions that focus on the user, and questions that focus on the milieu, we can quickly see whether a study goes deeply into one type of research question.

Because no one group focuses on usability exclusively, developing a comprehensive approach to studying it during product development, testing, and natural use, it certainly makes a difference to a particular study which group is conducting the study. If we were to ask each of these groups to study users learning to use a word processor, the resulting studies would differ in foci, methods, goals, findings, and philosophy. Thus, studies that grow out of collaboration of people from different groups have interesting possibilities for creat-

ing new and complex approaches to the study of usability, as sure as they have the possibility of becoming a battleground for philosophical disputes.

They may aim to study the product for any of the following purposes:

- To diagnose needs
- To give feedback to developers on strengths/weaknesses
- To validate approaches/systems
- To develop methods of study

They may aim to study the users for any of the following purposes:

- To elicit response
- To locate initial learning strategies
- To model learning
- To study using
- To develop methods of study

They may aim to study the milieu for any of the following purposes:

- To describe context for using
- To describe learning in context
- To diagnose context-dependent product problems
- To develop methods of study

Obviously there are parallels between the methods and the questions. Product methods are usually linked to product questions, cognitive model methods to user questions, and cultural model methods to milieu questions. Obviously too, the bulk of the literature on usability testing focuses on product issues, while the literature on user interfaces focuses both on product issues and user issues. Holes in published literature seem especially prominent in the area of the milieu, though work on feedback during product development is also needed, as are integrated studies of product, users, and milieu, and studies that develop methods. In short, almost everything needs to be done.

#### ISSUES AND OPPORTUNITIES THAT EMERGE IN THE EXPLORATION OF USABILITY RESEARCH

It is from knowing where we are that we make sense out of what is happening around us. It is from knowing where others are that we determine a course we can take to reach them. So, too, is the broader landscape of usability. We need to explore it to place our own work on it, to place the work of others on it, and to build the connecting roads among the places on the map. The exploration is an analytic one of constructing interpretations for studies and placing those studies on this emerging map.

As technical communicators, we are always going to be more interested in issues related to the development of usable educational materials and interfaces. The challenge for us is to figure out how to incorporate the growing knowledge of users into the development of manuals and interfaces. How do we connect research findings to improved documentation and computer products? How do we connect the findings to

usability in context? How do we connect the findings to communication theory and writing practice?

#### *Issue 1: Can Usability Research Develop Methods for Enhancing Product Development?*

The answer can be *yes*. This answer is tied, in part, to the effort to test products early in the cycle. Many researchers, in technical communication and elsewhere, are arguing for earlier and more exploratory testing. We know from published research that product research is developing methods rapidly. But so few of the product tests are published. Because so few are published, as a group we have fewer methods for enhancing product development than we as individuals actually employ. To fully answer this question we must publish the designs of the product tests aimed at yielding feedback early in product development.

A second means for enhancing product development can come from applying the results of validation research. If evaluations of existing products are written to point to generic learning and use problems and to suggest appropriate long-term research, as well as to handle the current evaluation, then the cumulative effect may well be a change in development approach.

#### *Issue 2: Can Usability Research Develop an Adequate Model for the Study of Usability in its Naturally Occurring Contexts?*

I do not think we can definitively answer this question yet. An adequate answer involves developing a way to select appropriate sites and also involves developing adaptation of classical ethnographic methods to the milieu of computing. But, there is tremendous interest in field study. Gould and Doheny-Farina articulate some of the opportunities for insight that field studies offer to technical communicators. [9] In human-computer interaction, Whiteside and Wixon [44] call for field research like the studies reported by Good [45] to underpin a process of making the users' concerns central. More and more people believe that we need to know about "natural use" to know about "usability" and to build more usable systems. But the natural study of usability is in another dimension from the laboratory test of usability. It may be difficult for groups to have enough patience to wait for the results from tests that take many months to complete.

#### *Issue 3: Can Usability Research Present Writing Theory With Compelling Arguments for Integrating Usability More Centrally Into the Writing Development Process?*

The answer to this question is a resounding *yes*. When Pakin and Associates [46] reported that 80 percent of all calls to software support help lines concern routine and repetitive problems, they voiced an accepted critique of software and documentation, a critique that has given rise to usability testing of products. Although the narrowly defined usability test critiqued by Weiss is the industry's response to information like Pakin's, it is not necessarily true that even the narrow usability test has no impact. Subjecting a completed text to a test of usability may not have a profound impact on the current text. But the writer can learn about users. If the learning about

users becomes a habit that shapes writing, a cumulative study that informs future writing, then every usability test can contribute to that writer's knowledge of users. It becomes a way of building a concrete theory of audience.

Such a concrete theory of audience can lead writers to change the ways in which they think about manuals in the early stages of writing. They may read more about user behavior before starting to work on a project to a new group of users. They may, for example, ask for more testing of competitors' packages before shaping their own documentation. They may run exploratory user tests on modules that take experimental approaches, or on the basic language that is central to the book, or on the outline of the text. These exploratory tests will serve as litmus paper for their ideas, and any of these changes to the concrete thoughts that writers have about users will change the ways in which they develop manuals.

### SUMMARY

A growing number of psychologists, engineers, and technical communicators want to make the user more integral to the whole development process. Whiteside and Wixon [44] call for field research to underpin a process of making the user's concerns central; Norman [47] calls for the interface design process to adopt a "user-centered system design" and calls for traditional lab work on learning; Wright [48] argues for a user-centered development of documentation. It is through reaching toward the broader meaning of usability, from the vantage of what we have been learning from usability testing that we will find new ways to confront the hard issues associated with making computer systems and documentation more usable.

### REFERENCES

- Weiss, E., "Usability: Stereotypes and Traps," in *Text, Context, and Hypertext*, E. Barrett (ed.), Cambridge, MA: The MIT Press, 1988.
- Sullivan, P., "User Protocols: Tools for Building an Understanding of Users," *IPCC 1988 Conference Record*, pp. 259-263.
- Carroll, J. M. (ed.), *Interfacing Thought*, Cambridge, MA: MIT Press, 1987.
- Gray, W. D., and Atwood, M. E., "Interfacing Thought," *SIGCHI Bulletin* 20, 2 (1988), pp. 88-91.
- Carroll, J. M., and Campbell, R. L., "Softening Up Hard Science: Reply to Newell and Card," *Human-Computer Interaction* 2, 227-249.
- Newell, A., and Card, S., "Straightening Out Softening Up: Response to Carroll and Campbell," *Human-Computer Interaction* 2 (1986).
- Newell, A., and Card, S., "The Prospects for Psychological Science in Human-Computer Interaction," *Human-Computer Interaction* 1 (1985), pp. 209-242.
- Turkle, S., *The Second Self*. New York: Simon and Schuster, 1984.
- Suchman, L. A., *Plans and Situated Actions: The Problem of Human-Machine Communication*, Cambridge: Cambridge University Press, 1987.
- Gould, E., and Doheny-Farina, S., "Studying Usability in the Field: Qualitative Research for Technical Communicators," in *Effective Documentation: What We Have Learned From Research*, S. Doheny-Farina (ed.), Cambridge, MA: The MIT Press, 1988.
- Sullivan, P., "The Potential Contribution of Longitudinal Research to Usability Research," *IPCC Conference Record*, 1989.
- Warren, T. L., "Readers and Microcomputers: Approaches to Increased Usability," *ITCC Proceedings* 35 (1988), pp. RET120-RET123.
- Gagne, R. M., "Military Training and Principles of Learning," Presidential Address, Division of Military Psychology, American Psychological Association, September 1961, reported in *Figuring Things Out: A Trainer's Guide to Needs and Tasks Analysis*, R. Zemke and T. Kramlinger (ed.), Reading, MA: Addison-Wesley, 1982.
- Baecker, R. M., and Buxton, W. A. S., *Readings in Human-Computer Interaction: A Multidisciplinary Approach*, Los Altos, CA: Morgan Kaufman, 1986.
- Nelson, R. J., "Beyond the Basic Technical Writing Course: Status Report, Part 1," *The Technical Writing Teacher* 13 (1986), 140-146.
- McDowell, E. E., "An Investigation of Graduate Courses in Technical Communication at Universities in the United States," *ITCC Proceedings* 36 (1989), pp. ET28-ET30.
- Dumas, J. S., "Stimulating Change Through Usability Testing," *SIGCHI Bulletin* (in press).
- Dieli, M., "Integrating Usability Evaluation Into the Computer Documentation Development Cycle," paper presented at the ACM SIGDOC Conference, Ann Arbor, MI, October 1988.
- Lasselle, J., "Usability Testing in the Product Development Cycle," paper presented at the College Composition and Communication Conference, Seattle, March 1989.
- Sullivan, P., "Incorporating User Feedback into Technical Manuals," paper presented at the College Composition and Communication Conference, Minneapolis, March 1985.
- Goswami, D., et al., "Evaluating the Document," in *Writing in the Professions*, Washington, DC: American Institutes for Research, 1981.
- Atlas, M. A., "The User Edit: Making Manuals Easier to Use," *IEEE Transactions on Professional Communication* 24, 1 (1981), pp. 28-29.
- Krauss, D. A., "Usability Testing for Lotus Documentation," *ITCC Proceedings* 35 (1988), pp. RET165-167.
- Vanderlin, G., Cocklin, T., and McKita, M., "Testing and Developing Minimalist Tutorials: A Case History," *ITCC Proceedings* 35 (1988), pp. RET196-RET199.
- Soderston, C., "The Usability Edit: A New Level," *Technical Communication* 32, 1 (1985), pp. 16-18.
- Sullivan, P., and Seiden, P., "Educating Online Catalog Users: The Protocol Assessment of Needs," *Library HiTech* 4, 13 (1985), pp. 29-36.
- Bond, S. J., Hayes, J. R., Janik, C. J., and Swaney, J. H., *C-MU TOPS-20*, vol. 2, Carnegie Mellon University, Pittsburgh.
- Buehler, M. F., "Defining Terms in Technical Editing: The Levels of Edit as a Model," *Technical Communication* 28, 4 (1981), pp. 10-14.
- Dye, K. L., "When is a Document Accurate and Complete?," *IPCC 1988 Conference Record*, 1988, pp. 269-272.
- Ramey, J., "A Self-Reporting Methodology for Rapid Data Analysis in Usability Testing," *IPCC 1988 Conference Record*, pp. 147-150.
- Wagner, C. B., "Quality Control Methods for IBM Computer Manuals," *Journal of Technical Writing and Communication* 10 (1980), pp. 92-102.
- Compton, J. C., "Results of a Telephone Survey of Technical Documentation Users," *IEEE Transactions on Professional Communication* 29, 4 (1986), pp. 93-99.
- Bethke, F. K., "Measuring the Usability of Software Manuals," *Technical Communication* 30, 3 (1983), 13-16.
- Vanderlin, G., Cocklin, T., and McKita, M., "Designing Tutorials That Help Users Learn Through Exploration," *IPCC 1988 Conference Record*, pp. 295-299.
- Schell, D. A., "Testing Online and Print User Documentation," *IEEE Transactions on Professional Communication* 29, 4 (1986), pp. 87-92.
- Chaboya, H. H., and Bolden, M. A., "Usability Testing and You: Improving Your Product for Your Customer," *ITCC Proceedings* 35 (1988), pp. RET96-RET98.
- Dieli, M. A., "Designing Successful Documentation: An Investigation of Document Evaluation Methods," doctoral dissertation, 1986, Carnegie Mellon University.
- Mack, R. L., Lewis, C. H., and Carroll, J. M., "Learning Word Processors: Problems and Prospects," *ACM Transactions on Office Information Systems* 1, 3 (1983), 254-271.
- Neal, A. S., and Simons, R. M., "Playback: A Method For Evaluating the Usability of Software and Its Documentation," *IBM Systems Journal*, 23 (1984), pp. 84-96.
- Singley, M. K., and Anderson, J. R., "A Keystroke Analysis of Learning and Transfer in Text Editing," *Human-Computer Interaction* 3 3 (1987-1988), pp. 223-274.
- Campbell, R. L., Mack, R. L., and Roemer, J. M., "Extending the Scope of Field Research in HCI," *SIGCHI Bulletin* 20, 4 (1989), pp. 30-32.
- Brown, J. S., and Newman, S. E., "Issues in Cognitive and Social Ergonomics: From Bauhaus to Our House," *Human-Computer Interaction* 1 (1986), pp. 359-391.
- Malone, T. W., "How Do People Organize Their Desks? Implications



- for the Design of Office Information Systems," *ACM Transactions on Office Information Systems* 1, 1 (1983), pp. 99-112.
44. Whiteside, J., and Wixon, D., "Discussion: Improving Human-Computer Interaction—A Quest for a Cognitive Science," in *Interfacing Thought*, J. M. Carroll (ed.), Cambridge, MA: The MIT Press, 1987.
45. Good, M. (ed.), "Seven Experiences With Contextual Field Research," *SIGCHI Bulletin* 20, 4 (1989), pp. 25-32.
46. Pakin, S. and Associates, "Using Troubleshooting Information to Reduce Support Calls," *Journal of Documentation Project Management* (1987), p. 12.
47. Norman, D. A., "Cognitive Engineering," in *User-Centered System Design: New Perspectives on Human-Computer Interaction*, D. A. Norman and S. W. Draper (ed.) Hillsdale, NJ: Lawrence Erlbaum, 1986.
48. Wright, P., "Manual Dexterity: A User-Oriented Approach to Creating Computer Documentation," *CHI'83 Proceedings*, 1983, pp. 11-18.

---

**The All-Purpose, Handy, Dandy Editor's Time-Saving Response Form**

- \_\_\_\_\_ 1. Thank you for your recent communication.
- \_\_\_\_\_ 2. You want me to do what?
- \_\_\_\_\_ 3. Surely you jest.
- \_\_\_\_\_ 4. I should prefer to leave the wording as I sent it. Please pretend that the words are carved in stone.
- \_\_\_\_\_ 5. Yes, it does appear that the reviewer(s) did not read the paper you thought you sent.
- \_\_\_\_\_ 6. Last I heard, this journal had an editor.
- \_\_\_\_\_ 7. Your request, while obviously eminently reasonable to you or you would not have made it, does reveal a certain flawed understanding of the constraints of journal publishing. These constraints become apparent through editorial experience, which is not recommended.
- \_\_\_\_\_ 8. Your kind and thoughtful words are very much appreciated.
- \_\_\_\_\_ 9. When in doubt, read the instructions.
- \_\_\_\_\_ 10. Mother told me there would be days like this.
- \_\_\_\_\_ 11. You are right, I was wrong.
- \_\_\_\_\_ 12. Please deduct the costs from the editor's salary, starting after my retirement party.
- \_\_\_\_\_ 13. Luke 10:5
- \_\_\_\_\_ 14. I Corinthians 13:1
- \_\_\_\_\_ 15. Hmm.
- \_\_\_\_\_ 16. Trust me.
- \_\_\_\_\_ 17. Oops.
- \_\_\_\_\_ 18. Call me.
- \_\_\_\_\_ 19. My day could have been brightened in many ways. I wish you had picked one.
- \_\_\_\_\_ 20. It takes a little longer than you think.

*Dean F. Martin  
from CHEMTECH,  
reprinted with permission of  
the American Chemical Society*