

Characteristic-Based, Task-Based, and Results-Based: Three Value Systems for Assessing Professionally Produced Technical Communication Products

Saul Carliner
City University of Hong Kong

Technical communicators have developed different methodologies for evaluating the effectiveness of their work (whether the information can be used by the intended audience), such as editing, usability testing, and determining the value-added. But, as vastly differing assessments of the same professionally produced technical communication products suggest, at least three broad value systems underlie the assessment practices: characteristic-based (assessing against a set of criteria), task-based (assessing users' observed ability to perform tasks), and results-based (assessing the contribution to the publisher, usually in financial terms). The systems do not overlap with one another; rather, they embody different values about what makes technical communication effective. The most complete form of assessment may involve multiple assessment approaches and triangulated results.

Situation One: An author submitted a paper on generic writing to a refereed publication. For readers unfamiliar with it, generic writing is an approach to writing content so that it avoids being unnecessarily specific and, as a result, minimizes the need for revision (Holden). For example, rather than preceding a list with "these two items," which would have to be revised to "three" if the list were expanded, the lead-in would merely say "these items." A reviewer, then working as a professor, panned the article, saying that the author's company should be embarrassed to have its name attached to the idea. Nearly 20 years later, the reviewer, now working as a corporate consultant, preached generic writing as she tried to implement single sourcing and content management systems in her organization.

A concept that the reviewer dismissed in 1984 proved to have practical utility in 2002.

Situation Two: A team of judges in a Society for Technical Communication competition argued over the award-worthiness of an entry. The size of the type font differed in different sections of the publication because the author of the entry reprinted material from another source. Rather than going through the expense of re-formatting it, the author used the material as it was. One judge found the inconsistency jarring, but another judge applauded the author's resourcefulness. Both judges used appropriate criteria to assess the entry but the value systems clashed: one judge valued consistent appearance more highly; the other valued resourcefulness.

Situation Three: The director of marketing communications for a major corporation was horrified by an advertisement that one of the product managers placed in a trade magazine. The advertisement was text-heavy, and the camera-ready copy came from an inkjet printer. In his mind, it represented the antithesis of good advertising design. Before the director of marketing could call the product manager on the carpet for circumventing corporate rules intended to avoid such visual disasters, the ad won a reader's poll sponsored by the magazine. Write-in comments lauded the non-slick appearance and its wealth of information, from which they could make trustworthy purchasing decisions. The criteria for excellence among communications professionals were antithetical to the criteria for excellence among the intended audience.

Who is right in each of these instances? What seems appropriate in one person's professional judgment is not in another's. These differences of opinion are not arbitrary; they are rooted in the differing values of those making the assessment. In the first case, the values in conflict pertain to the utility of the content. An idea that seemed to have little utility in the academy has much in industry. In the second case, the values in conflict are editorial standards versus business efficiency. In the last case, the value in conflict is what determines credibility. To the designers, credibility comes from aesthetics; to readers, those aesthetics cause suspicion. Instead, content offers credibility. These conflicts represent the value systems encoded in the assessments that technical communicators use to assess professionally-produced work.

This article explores three dominant approaches to assessments for professionally produced technical communication products: characteristic-based, task-based, and results-based. After clarifying terminology regarding assessments, this article first explores the role of beliefs and values in the assessment of professionally produced work. Second, it presents overviews of the approaches. It defines each, explains how the approach emerged and is used in professional practice, names some of the underlying beliefs, and identifies issues with the approach. Next, this article suggests areas of conflict among the three approaches and, last, suggests implications for practice and research.

Quality: Productivity or Effectiveness?

Faced with contradictory assessments of technical communication products, one of the most vexing questions professionals confront in the field is "What is good work?" In the literature, the term quality is often used to label the concept of good work (Bandes; Spilka). But the term "quality" causes confusion, because we have no operationalized definition of it within the field of technical communication. Definitions range from "meets requirements" (Bandes) to effectively servicing sponsors (Fredrickson; Robinson and Robinson).

Quality issues can be divided into two categories. One category is productivity, which refers to the quantity of output that technical communicators produce in a given time period, such as the number of published pages and screens per day. The other is effectiveness, which means "Does the communication product do what it's supposed to do?" (Walmer).

Productivity (how much) and effectiveness (how good) are actually separate issues. In fact, communicating more effectively may involve producing less output with the same resources (Walmer). Assessment of communication products exclusively focuses on effectiveness, so assessing effectiveness is the sole focus of this article.

Value Systems and Assessment

As the opening scenarios of this article demonstrate, values play a role in assessing technical communication products, whether that role is tacit or overt. In some cases, the assessor is not schooled in different approaches to assessment, so the assessor follows his or her instincts about what makes an effective technical communication product. In other cases, the assessor consciously uses a particular assessment approach, such as an assessment against a given set of criteria or a usability test. Encoded in those assessment approaches are beliefs about what makes effective technical communication products, as well as which assessment method is most appropriate for evaluating the effectiveness of a technical communication product. These beliefs emerge from cultural values (Lincoln and Guba) that favor certain characteristics, view others with disfavor, and ignore others. The differing assessments of single communication products described in the opening of this article result from people applying different cultural values to assess the same product.

One can view systems of assessment as a means of codifying value systems. For instance, if assessment systems are based on empirical data, both the observation of the data and its interpretation occur through the filter of the researcher's and analyzer's value systems (Denzin and Lincoln; Carpsecker and Apple; Lincoln and Guba). The concept of user-centered design, which guides the design of many technical communication products and much software, actually represents specific characteristics of a value system that utilizes a meta-analysis of several empirical usability tests.

At times, however, two value systems come into conflict. Because practicing technical communicators need not subscribe exclusively to a single value system, they can simultaneously value user-centered design and the concept of “value-add” (which states that technical communicators add value to the corporate performance). These systems conflict, however, when an organization interested mainly in return on investment chooses to postpone changes that would enhance the usability of a document or software—or, in value-add terms, the organization chooses to cut development expenses by reducing up-front investment in order to increase overall return later. Similarly, technical communicators might find the value systems used to assess professional communication products at odds with those of the people with whom we collaborate. For example, although user-centered design seems to be popular among technical communicators, graphic designers take a more skeptical view of it (Nattress; Cloninger). Naming and understanding the value systems behind the assessment approaches can aid technical communicators as they develop assessment methods used in their organizations and as they consider the differing features of varying assessment approaches.

The Three Broad Value Systems

Three broad approaches have emerged in the assessment of professionally produced technical communication products. *Characteristic-based* assessment focuses on characteristics of technical communication products, like the readability of type and the use of language. *Task-based* assessment focuses on the ability of users to perform tasks, like the ability to install hardware after reading an installation procedure. *Results-based* assessment focuses on business results generated by the communication product, such as reducing technical support costs.

The following sections describe each of these types of assessment. After defining each type and describing its emergence and underlying values, I explore the methods typically used to assess technical communication products within that type of assessment, and describe the limitations of that type of assessment.

Characteristic-Based Assessment

Characteristic-based assessment evaluates the effectiveness of a technical communication product against one or more stated criteria. The presence of certain criteria and absence of others is associated with effectiveness. The criteria range from the specific and objective (such as the type font and size for body text) to the very broad and subjective (“are graphics attractive and of high quality?”).

Characteristic-based assessments are in wide use in technical communication. Specific assessments include heuristic evaluation, an evaluation methodology emerging from the field of usability testing, in which someone assesses the potential ease with which a user might go

through a technical communication product and identify potential usability issues; the blind review process, such as the one used for professional journals; technical reviews of draft communication products; and competitions, such as the STC Technical Publications Competition (for printed communication products) and the STC Online Communication Competition.

The Emergence of Characteristic-Based Assessments

Of all the forms of characteristic-based assessment, many practicing professionals regard editing as the earliest, (Corbin, Moell, and Boyd). Admittedly, many others view editing as “comma chasing”—that is, looking for grammatical, stylistic, and spelling errors—and others consider it to be a publishing activity rather than a form of assessment. But editors view themselves as a means of assuring the quality of content, predicting a publication’s effectiveness, and then enhancing that effectiveness.

One major development in editing that permitted its use in assessment was the concept of “levels of edit,” a system intended to provide developmental feedback to an author in stages, so that the meaning and presentation of content are firmly established before editors focus on mechanical issues (Van Buren and Buehler). Unfortunately, editing assesses technical communication products in their formative stages; professional technical communicators seek assessment of published work (summative evaluation).

In industry, much of the conversation on effectiveness has focused on what to assess. In an attempt to assess published work, some communicators have attempted to quantify effectiveness, giving a communication product a score that indicates effectiveness. Some efforts have focused on a single measure, like a reading on the Flesch index, which measures the reading level of content. Other efforts focus on identifying characteristics associated with quality and measurement of their presence (Hosier, Rubens, Krull, and Velotta) and try to develop an index that represents a combination of measures.

Methodologies for Performing Characteristic-Based Assessment

In characteristic-based assessments, the criteria serve as a guide, and assessors use the guide to evaluate the extent to which they believe the communication product is effective. Specific methodologies for performing characteristic-based assessment vary, depending on the type of assessment. For example, when reviewing competition entries, reviewers often take a combination of the checklist approach (does the entry have the characteristics identified in the competition guidelines?) and the critical approach.

One challenge for this approach is ensuring consistently applied standards. Some types of assessment, such as heuristic evaluations, are performed by a single person. When using formative evaluation on draft communication products, the same person ideally assesses each

draft. The single perspective is believed to provide continuity throughout the process. In addition, limiting the assessment to a single person significantly reduces the time and cost of performing the assessment. Style guides, editors, and heuristic reviews are intended to minimize variation of quality because, in theory, assessors all look for the same thing based on common criteria.

In other instances, several people assess the same communication product in an effort to triangulate the results. Organizations sponsor assessments with several reviewers for many reasons. Among them is the concern that the review criteria are not sufficiently specific or that the results are public enough that the organization does not want to rely on a single assessment. Multiple reviewers are commonly used in blind reviews and competitions. Both traditionally request three independent reviews; a team leader then consolidates the reviews and makes a final assessment of the communication product.

Values Embodied by Characteristic-Based Approaches

The lists of criteria on which these assessments are based codify particular value systems. Furthermore, these characteristics represent a norm of what technical communicators believe to be an effective technical communication product.

Consider the codifying of value systems. For example, in addition to the rules of grammar (which technical communicators must follow to ensure that readers can accurately understand their work), editorial criteria codify values about language, usage, punctuation, and the appearance of text. These are called style guidelines or, as Rook calls them, “pseudo-rules” (274). While rules of grammar affect how readers interpret a sentence (for example, transposing a verb and its subject could lead to serious misunderstandings), style guidelines represent prevailing opinion.

Communicators can violate style guidelines without risking misunderstanding, but the professional image of the corporate author may be affected by these stylistic variations. Organizations often make decisions about matters of style, sometimes democratically, sometimes autocratically. They codify these decisions in an editorial style guide, which represents their value system. Because many issues have been “coded” in style guides, technical communicators may accept these codes as rules, rather than as value systems.

Editorial style guidelines work well with concrete issues, but have difficulty specifying more abstract criteria. For example, how does an editor codify the characteristics that render text clear? One possibility is defining all new terms in text (an oft-stated criterion), although doing so does not guarantee clarity of the entire text. Similarly, some guidelines try to encompass international audiences; but simply avoiding culturally-based references, such as social security numbers and six-character postal codes, does not guarantee that a message will not cause confusion or anger among users in another culture.

Competition guidelines also reveal underlying value systems. Guidelines for the STC Online Communication Competition do so at two levels. One level addresses standards common to all communication products; the other pertains to features common to specific genres of communication products and represents the values held by people who specialize in developing materials in that genre. For example, one of the criteria in the Tutorials category is “Does the entry clearly define the learning objective?” because one belief of instructional designers is that every tutorial begins with an instructional objective (Mager, “Measuring”; Deutsch; Wedman and Tesmer). But the type of objective that might be acceptable to a technical communicator might not meet the stringent values of instructional designers, most of whom believe that objectives must be stated in observable and measurable terms (Mager, “Preparing”; Dick and Carey). The variety of definitions of “instructional objective” may lead to inconsistent assessments of the same product.

Limitations of the Approach

Three limitations underlie the characteristic-based approach. One is untrained assessors. The second is resolving conflicts. The third is a failure among practicing professionals to recognize the limits of characteristic-based assessments.

Training assessors or reviewers to conduct characteristic-based assessment poses a challenge. Because they typically assess a communication product against a given norm, assessors need to be familiar with those norms. But, at times, they are not. This may happen because the assessor did not take time to become familiar with the guidelines, because the coordinator of the assessment failed to provide the criteria, or because the coordinator and the assessors assumed that the criteria were universally understood and “obvious.”

For example, in a technical review, a subject matter expert is supposed to review a technical communication product to assure its completeness and accuracy. Unsure of their responsibilities, technical reviewers may perform what appear to be edits, even though they may be unacquainted with the editing style guide, and overlook the technical content. Such reviews reflect an incomplete socialization of the assessors; because they do not fully understand their role and how to perform it, these subject matter experts assess technical communication products as they assume the task should be performed. Some attempt to copyedit the document, focusing on typographical and grammatical errors. Some make sure that all of the material in product specifications is covered in the manual. Often, technical reviewers do not consider how consumers will use the product, and whether the instructions accurately tell consumers how to do so. Conflicting understandings of the review process and of the appropriate criteria for assessment are serious problems in characteristic-based assessment.

The second limitation emerges from the first: even if the review criteria are explicitly stated, assessors may disagree about the absence or presence of features in particular communication products, or give different weights to the same criterion—that is, one assessor might say grammar represents 50 percent of effectiveness, another might say 30 percent.

The last limitation of characteristic-based assessments is a failure among practicing professionals to recognize its limits. For example, some practicing technical communicators seek to “prove” the effectiveness of all technical communication products by developing a single list of criteria that can be assessed. This belief guides efforts in industry to develop effectiveness metrics for technical communicators, such as the U-Metric (Carliner, “Demonstrating”). But the presence of certain characteristics does not guarantee the effectiveness of communication products. For example, although the number of index entries per page is closely correlated with effectiveness, providing a specific number of index entries does not guarantee effectiveness of a communication product. Characteristic-based assessment does not take into account the specific context of communication. Task-based assessment, discussed next, differs in that it assesses communication products’ effectiveness for their particular situations.

Task-Based Assessment

Task-based assessment evaluates the effectiveness of a technical communication product by observing whether or not users can perform the activities that the communication product is supposed to describe. For example, in task-based assessment, an installation guide is effective if users can successfully install the equipment or software described in it. Usually, task-based assessment identifies tasks in observable and measurable terms, and sets targets for completing the task. These targets usually include a time limit and an acceptable number of errors, but might also include the conditions of performance, such as with or without the use of product information, and satisfaction levels, such as might be measured on a Likert scale or some similar quantitative scale (Dumas and Redish). Tasks and targets are specific to the project, and are usually established either in the early phases of the project (Dick and Carey) or while preparing the assessment (Dumas and Redish).

The Emergence of Task-Based Assessment

An early example of task-based assessment is used in technical training: the criterion-referenced test designed to assess the extent to which learners mastered the content. Unlike the character-based criteria discussed above, the criteria in these tests are tasks defined at the beginning of the instructional design process. They differ from course to course, based on subject matter, the entry skills of the learners, and the level of skill expected by the sponsor of the training

program (Mager, “Measuring”). Criterion-referenced tests date back at least to the 1940s, when the instructional design process was first developed (Deutsch).

A more recent form of task-based assessment, usability testing, is more widely used with technical communication products like user’s guides, references, and online help. In a usability test, people who represent intended users are instructed to perform designated tasks under the guidance of the technical communication product. Observations of their behaviors are recorded and used to assess the effectiveness of the technical communication product and to guide efforts to improve it.

Values Embodied by Task-Based Approaches

Behaviorist values underlie task-based assessment, especially the criteria-referenced assessment, which derives, in part, from the behaviorism of B. F. Skinner (Mager, “Preparing”). Skinner’s belief that observable behavior matters most motivated instructional designers to develop a form of assessment based on observation of behavior. For criterion-referenced assessment to be effective, people must begin the design process by defining observable and measurable objectives for the communication product (Rossett). The assessment must derive from those objectives—and only those objectives. In response to this value that the only valid assessments emerge from objectives, nearly 100 percent of instructional designers develop objectives for their training courses (Wedman and Tesmer).

Methodologies for Performing Task-Based Assessment

Specific assessment techniques vary, depending on the type of communication product, but all rely on users demonstrating that they can perform the intended tasks in an observable and measurable way. In criterion-referenced testing, learners must respond to test questions or perform a task that an instructor observes. Test questions emerge directly from the criteria—that is, each objective forms the basis of one or more test questions (Horton). For example, if an objective states, “Learners must be able to match customer needs with the appropriate model of a PC,” then learners might participate in a test in which they are provided with customers’ needs, and must correctly match that with the appropriate PC model. Although criterion-referenced tests can involve an observation, they usually involve a test because tests are easier and quicker to grade.

Usability tests work in a similar way, but generally involve an observation and test the communication product, not the user. If problems arise in the test, they are presumed to derive from the communication product, not the person using it. Test subjects are chosen to meet the demographics of the intended users as closely as possible. They receive information about tasks to perform. As they do so, users typically consult the communication product for guidance,

and independent observers record their actions. Observers typically track the time needed to perform the tasks, the number and type of issues that arise as users perform the tasks, and users' satisfaction with the communication materials and the products they support. After compiling the results, the usability test team makes the development team aware of the issues that arose during the review (Barnum; Dumas and Redish; Duin).

Limitations of the Approach

Task-based assessment is affected by four limitations. The first pertains to the methods. On the one hand, usability tests begin with defined scenarios and goals, suggesting that usability testing has its roots in behaviorism. But usability testing differs from criterion-referenced tests because the tasks (criteria) are not specified before development begins. Rather, most are defined after the software and communication product being tested are already developed. In addition, usability testing methodologies have their roots in ethnography and other qualitative approaches.

The appropriate use of test results is the second limitation. Because usability tests often gather detailed quantitative statistics on user performance, some people advocate using the data to develop inferential statistics about user performance in general (Hughes). But because usability tests involve purposeful sampling rather than random sampling, and because these tests use small samples, questions remain about whether these results can be used to reach generalizable conclusions about user performance.

The third limitation pertains to the timing of usability tests. Most usability tests are performed as formative evaluation, assessment of communication products that are at a draft stage, rather than summative evaluation, assessment of published work. The purpose of formative evaluation is to ensure that a technical communication product is usable and to identify areas for improvement. Formative evaluations cannot be used to make assessments about user performance with the polished publication. Only summative evaluations can be used to assess the usability of the published product (Dick and Carey).

The last limitation pertains to the high costs of usability testing, which limits its use. Despite wide discussion of usability testing within the field of technical communication, it does not appear to be widely performed by technical communication departments. For example, in my study of management practices in larger technical communication departments (ones with 20 or more workers), I found that these departments test fewer than 10 percent of their work products (Carliner, "Preliminary").

The high cost of usability testing (estimates range from \$5,000 to \$50,000) may lead organizations to test only a limited number of communication products. To accommodate the principle of usability

without the cost, some organizations have tried to develop a list of generalizable characteristics against which reviewers can assess the effectiveness of a communication product without the expense of a test. This heuristic evaluation is close to being characteristic-based assessment and may defeat the purposes of task-based assessment. Because of these limitations of task-based assessment, several organizations have begun to explore results-based assessment.

Results-Based Assessment

Results-based assessment evaluates the effectiveness of a technical communication product by its tangible contribution to the performance of the organization. Under ideal circumstances, the contribution can be calculated in financial terms.

The Emergence of Results-Based Assessments

Discussions of effectiveness in other branches of communications, such as marketing communication and training, have shifted to business measures. These discussions focus on the financial effect that publishing content has on the business that underwrote the expense of the project. For example, people in advertising measure the effectiveness of communication products like advertising, catalogs, and brochures by the sales that they generate. Tracking codes published on communication products help communicators tie results to a particular communication product.

People in the field of instructional design attempt to measure the “return on investment” of a training program. For example, if workers receive training on safety procedures, trainers first try to determine whether that training led to a reduction in the number of accidents, then try to calculate the cost savings resulting from that reduction. If the reduction in accident-related costs exceeds the cost of designing, developing and delivering the training, then the company experiences a positive return on investment (P. Phillips; Horton; Kirkpatrick; J. Phillips).

Focusing on business measures can encompass other types of effectiveness. After a course of safety training, if the number of accidents goes down, learners have likely mastered the skills taught in the training course.

Some technical communicators have attempted to transfer these results-based approaches to our field. The best-known effort is a 1995 study commissioned by the Society for Technical Communication, which explored ways that technical communication adds value to the businesses that sponsor their work (Redish). The study suggested a number of ways that technical communicators can demonstrate the financial impact of their work, typically by showing that effective documentation reduced the amount of re-work required by users or that it decreased the number of calls to a support line.

Methodologies for Performing Results-Based Assessment

Results-based assessment relies on designing methods for measuring those results after a product is distributed to users. Actual measures and techniques vary, depending on the particular document involved and the business problem it is intended to address. Reva Daniel's case study of the reduction in re-work resulting from the redesign of a form and cover letter by the U. S. Veteran's Administration provides an example of this approach. In this case, she calculated the improvement in response rate and the increase in the number of correctly completed forms users sent to the Veteran's Administration on the first try. If the Veteran's Administration calculated costs associated with re-sending letters and the time needed to correct improperly completed forms, the agency could calculate the results provided by the technical communication product.

Values Embodied by the Approach

Results-based assessment is rooted in the market economy. Assessing the effectiveness of technical communication by its return assumes that organizations will only continue to invest in technical communication if the return on that investment is favorable to that of other investment options (Bassi and Youngblood).

In one sense, results-based assessment is related to task-based assessment. According to this value system, if users can perform the designated tasks, then a tangible benefit accrues to the organization that published the technical communication product (Robinson and Robinson; Stolovich and Keeps). Conversely, performing the tasks only has value if the publisher of the content receives a tangible benefit.

The most significant belief underlying this type of assessment, however, is that a tangible value can be assigned to technical communication products and services, especially those used internally or provided as part of a product package. This belief is held by the founders of the fledgling (and, by some accounts, floundering) intellectual capital movement, who almost exclusively represent the financial community. They believe that quantifying the value of an organization's intellectual assets is an ethical issue, because balance sheets represent only tangible assets. Intellectual assets are intangible ones, but in a knowledge economy, may represent the primary value of an organization (Edvisson and Malone; Stewart). The values embodied in this approach have the potential to increase (or decrease) the importance that organizations afford to all communication efforts, including technical communication.

Limitations of the Approach

Three significant issues limit results-based assessment. First, no one knows how to calculate the exact financial value of technical communication products. Most technical communication products

are intended to make users self-sufficient so they do not need assistance from technical support. When users do not request these services, organizations do not incur a charge, and balance sheets cannot represent expenses that were never incurred (Redish). Nor do organizations typically charge for help and other product information, so one cannot assess financial value by the revenue generated from sales of this content. Until more precise methods are devised, this approach will remain somewhat limited in its usefulness.

The second limitation is the inadequate amount of study in this area leading to a paucity of generalizable insights. Mead noted that most efforts to assess effectiveness of technical communication products have been one-time or theoretical efforts, and do not have the power of sustained study behind them. No sustained studies have been published since that time (Carliner et al.).

The third limitation of this approach is largely a theoretical one. Perhaps because the methodologies for calculating financial returns are not yet honed, many larger technical communication departments do not use them (Carliner, "Preliminary"). Indeed, many technical communicators are not even aware of them. Rather, the typical technical communicator may remain perplexed about results-based assessment because few of the ideas are diffused widely into practice.

Conflicts among the Approaches

The three different types of assessment represent different value systems that overlap, but conflicts exist among the three types and even sometimes within a type. For example, conflicts arise between the task-based and results-based value systems. Some advocates of task-based assessment (especially usability) assume the primacy of usability over all other issues, while some advocates of results-based assessment assume the primacy of results. Each claims that it has the other value system at heart. For example, proponents of task-based assessment say that financial benefits accrue if users can easily perform the designated tasks (Donoghue). But the ability to easily perform tasks does not mean that the tasks are worth performing in the first place. Similarly, proponents of results-based assessment say that the presence of financial results only occurs if users can perform the intended tasks (J. Phillips). When users have no alternate source of content, they might find a way to achieve results regardless of the usability of the content.

Clashes exist, too, within characteristic-based assessments, placing people in one profession against those in another. Typically, technical communicators advocate a usability-oriented approach to the assessment of functional content, especially online. But the graphic design community believes that aesthetics plays a key role in the effectiveness of websites and has issues with characteristic-based assessments that pay little or no attention to the aesthetic value of websites (Horn; Cloninger).

But more fundamental than the conflicts among these value systems are their uniquenesses. Each of the three types of assessment represents a different way of considering a technical communication product. To some extent, each looks at something that the other two do not. These uniquenesses suggest that the true value of acknowledging the multiple perspectives offered by the different types of assessment lies in applications of practice and research.

Applications of these Concepts in Practice and Research

Although many applications exist for the acknowledgment of different types of assessment and their value systems, three immediately spring to mind. The first is that, rather than advocating only one type of assessment, an assessment program might encompass several different approaches (characteristic-based, task-based, and results-based). This idea is consistent with the concept of *triangulation* in qualitative research in which a researcher attempts to gain multiple perspectives on a given situation in order to develop a rich understanding of it (Lincoln and Guba). The Kirkpatrick methodology that is widely used in the field of training relies on several perspectives to assess the effectiveness of training courses. These perspectives include learner satisfaction, the extent to which learners “learned” the material after taking a course (as measured by a criterion-referenced test), the extent to which learners have changed their behavior after taking the course, and the impact that the course has had on the organization that funded it. Carliner suggested an application of this methodology to technical communication (“Demonstrating”). Multiple perspectives for assessing the effectiveness of technical communication products can provide explanations for contradictory evidence about effectiveness.

Technical communication research can also profit from an awareness of these approaches and their value systems. One fruitful area to research may be a comparison of different assessment methodologies. Ideally, researchers would not undertake this comparison to determine which methodology is most effective; that effort is likely to be futile. Rather, researchers might undertake this comparison to identify the type of issues that different types of assessments provide and the situations in which the assessments contradict one another. Practicing technical communicators who have limited resources for assessment could use the insights learned to determine which type of evaluation would provide them with the type of data they seek about their communication product, within their budget and time constraints. The results of this research could provide guidance for designing many types of assessment.

Teaching technical communication could also benefit from understanding these approaches and their value systems. In practice-

oriented classes, students could learn about the different means of assessment and use each of them to assess their own work. For example, students could apply usability testing (in a simple format), editing, and heuristic evaluations, and attempt to identify the types of financial results that might result from publishing a technical communication product.

In theory-oriented classes, especially at the graduate level, students could learn about the different value systems underlying the assessments, and use them to develop an understanding of the value systems at work in the workplace, explore the contradictions within our professional culture, and consider how value systems between academics and practicing professionals differ.

In Closing

The acknowledgment that different value systems underlie the assessment of professionally produced technical communication products is a call to each of us to articulate our own value systems. We need to identify which values influence assessments of technical communication quality and acknowledge those in our teaching. By doing so, we can become open to other value systems. Indeed, articulating our value systems helps us to see better the complexities of technical communication and to expand our understanding of the factors influencing the effectiveness of technical communication products.

Acknowledgment

I would like to thank Margaret Hundleby for encouraging me to submit this article, and Marjorie Rush Hovde for her patience and encouragement in shaping it. I would like to thank all of the reviewers for their feedback; it has significantly strengthened this article.

Works Cited

- Bandes, Hanna. "Defining and Controlling Documentation Quality—Part I." *TC* 33 (1986): 6-9.
- Barnum, Carol. *Usability Testing and Research*. New York: Allyn and Bacon, 2001.
- Bassi, Laurie, and Curt Youngblood. "Human Capital: The Measures that Matter." American Society for Training and Development International Conference and Exhibition. New Orleans. 3 June 2002.
- Carliner, Saul. "Preliminary Research Report of What Do We Manage: A Survey of the Management Portfolios of Larger Technical Communication Departments." STC Research Grant. 2002.
- . "Demonstrating the Effectiveness and Value of Technical Communication Products and Services: A Four-Level Process." *TC* 44 (1997): 252-65.
- Carliner, Saul, et al. "What We Can Learn about Productivity and Effectiveness Metrics from Software Design and Development, Training, and Marketing Communications." In progress.
- Carpsecker, P. F., and M. Apple. "Critical Qualitative Research and Theory, Methodology, and Practice." *Handbook of Qualitative Research in Education*. Ed. M. D. LeCompte, M. D. Millroy, and J. Preissle. San Diego: Academic Press, 1992. 507-53.
- Cloninger, Curt. "Usability Experts Are from Mars, Graphic Designers Are from Venus." *A List Apart*. June 9, 2002. <<http://www.alistapart.com/stories/marsvenus/index>>.
- Corbin, Michelle, Pat Moell, and Mike Boyd. "Technical Editing as Quality Assurance: Adding Value to Content." *TC* 49 (2002): 286-300.
- Daniel, Reva. "Revising Letters to Veterans." *TC* 42 (1995): 69-75.
- Denzin, N. K., and Y. S. Lincoln. "Introduction: Entering the Field of Qualitative Research." *Handbook of Qualitative Research*. Ed. N. K. Denzin and Y. S. Lincoln. Newbury Park: Sage, 1994. 1-17.
- Deutsch, W. "Teaching Machines, Programming, Computers, and Instructional Technology: the Roots of Performance Technology." *Performance and Instruction* 31.1 (1992): 14-20.
- Dick, Walter, and Lou Carey. *The Systematic Design of Instruction*. 3rd ed. New York: Harper, 1990.
- Donoghue, Karen. *Built for Use: Driving Profitability through the User Experience*. New York: McGraw-Hill, 2002.
- Duin, Ann Hill. "Test Drive: Evaluating the Usability of Documents." *Techniques for Technical Communicators*. Ed. Carol Barnum and Saul Carliner. New York: Prentice-Hall, 1993. 305-36.
- Dumas, Joseph S., and Janice C. Redish. *A Practical Guide to Usability Testing*. London: Intellect, 1999.

- Edvisson, Leif, and Michael S. Malone. *Intellectual Capital: Realizing Your Company's True Value by Finding Its Hidden Brainpower*. New York: Harper, 1997.
- Fredrickson, Lola. "Quality in Technical Communication: A Definition for the 1990s." *TC* 39 (1992): 394-99.
- Holden, Norman. "Generic Writing." IBM Technical Report 07.793. Rochester: IBM Systems Products Division, 1984.
- Horn, Robert. "Information Design: The Emergence of a New Profession." *Information Design*. Ed. Robert Jacobson. Cambridge: MIT Press, 1999.
- Horton, William. *Evaluating E-Learning*. Alexandria, VA: ASTD Press, 2001.
- Hosier, William J., et al. "Basing Documentation Quality Standards on Research." *Proceedings of the 39th Annual Conference of the Society for Technical Communication*. Arlington, Virginia: STC, 1992: 428-31.
- Hughes, Michael. *Usability Testing for Knowledge Sharing*. Saddle River: Prentice Hall. In Press.
- Kirkpatrick, Donald L. *Evaluating Training Programs: The Four Levels*. 2nd edition. San Francisco: Berrett-Koehler, 1998.
- Lincoln, Yvonna S., and Egon G. Guba. *Naturalistic Inquiry*. Newbury Park: Sage, 1985.
- Mager, Robert. *Measuring Instructional Results*. 3rd ed. Atlanta, GA: Center for Effective Performance, 1997.
- Mager, Robert. *Preparing Instructional Objectives: A Critical Tool in the Development of Effective Instruction*. 3rd ed. Atlanta: Center for Effective Performance, 1997.
- Mead, Jay. "Measuring the Value Added by Technical Documentation: A Review of Research and Practice." *TC* 45(1998): 353-79.
- Nattress, Paul. "The Indie Life: Talking with Louis Rosenfeld." *Boxes and Arrows*. 5 Sept. 2002. <http://www.boxesandarrows.com/archives/the_indie_life_talking_with_louis_rosenfeld.php>.
- Phillips, Patricia. *The Bottom Line on ROI: Basics, Benefits, and Barriers to Measuring Training and Performance Improvement*. Atlanta: Center for Effective Performance, 2002.
- Phillips, Jack. *Handbook of Training Evaluation and Measurement Methods*. Gulf Publishing Company, 1997.
- Redish, Janice. "Adding Value as a Technical Communicator." *TC* 42 (1995): 26-39.
- Robinson, Dana, and James Robinson. *Training for Impact*. San Francisco: Jossey-Bass, 1989.
- Rook, Fern. "Remembering the Details: Matters of Grammar and Style." *Techniques for Technical Communicators*. Ed. Carol Barnum and Saul Carliner. New York: Macmillan, 1993. 274-90.
- Rossett, Allison. *Fast Things First: A Handbook for Human Performance Analysis*. San Francisco: Jossey-Bass, 1998.

- Society for Technical Communication. *Guidelines for the International Online Communication and Technical Publications Competitions*. Arlington: Society for Technical Communication, 2002.
- Spilka, Rachel. "The Issue of Quality in Professional Documentation: How Can Academia Make More of a Difference." *TCQ* 9(2000): 207-220.
- Stewart, Thomas A. *Intellectual Capital: The New Wealth of Organizations*. New York: Doubleday, 1997.
- Stolovich, Harold, and Erica Keeps, eds. *Handbook of Human Performance Technology: Improving Individual and Organizational Performance Worldwide*. 2nd ed. San Francisco: Jossey-Bass, 1999.
- Van Buren, Robert, and Mary Fran Buehler. *The Levels of Edit*. Pasadena: Jet Propulsion Laboratory, 1980.
- Walmer, Daphne. Personal conversation. 8 Feb. 1999.
- Wedman, J., and M. Tessmer. "Instructional Designer's Decisions and Priorities: A Survey of Design Practice." *Performance Improvement Quarterly* 6.2 (1993): 43-57.

Saul Carliner is a visiting assistant professor at the City University of Hong Kong. His research interests include the emerging genres of online communication, the information design process, and economic and business models underlying information design. His books include *Designing E-Learning* (ASTD), *An Overview of Online Learning* (HRD Press), and *Techniques for Technical Communicators* (Prentice-Hall). He has received two Best of Show awards in the Frank R. Smith Outstanding Article Competition, and is a fellow and past international president of the Society for Technical Communication.

