

Performance & Performance Standards

An Opinion

Fred Nickols, CPT

5/15/2010



Performance is commonly equated with what people do; that is, performance and behavior are seen as one and the same. In this “opinion piece,” I intend to define what I mean by performance, distinguishing it from behavior, and to discuss what I see as some of the costs of not making such a distinction. I also intend to suggest some initial guidelines for establishing valid performance standards – guidelines that recognize performance as something more than just behavior.

Performance & Performance Standards: An Opinion

Introduction and Overview

Performance standards play a key role in determining the success or failure of training programs and performance improvement efforts because they provide the “yardsticks” against which performance is measured. When performance standards are valid, such efforts have a better than even chance of being successful; when performance standards are invalid, the odds weigh heavily against success. In short, the validity of performance standards is the *sine qua non* of performance measurement, and performance measurement is in turn the *sine qua non* of successful performance improvement programs.

A fundamental issue regarding the validity of performance standards is the way in which performance is defined. I am of the opinion that the way in which performance is ordinarily defined makes the task of establishing valid performance standards a misleading and an unnecessarily difficult one.

Performance is commonly equated with what people do; that is, performance and behavior are seen as one and the same. In this “opinion piece,” I intend to define what I mean by performance, distinguishing it from behavior, and to discuss what I see as some of the costs of not making such a distinction. I also intend to suggest some initial guidelines for establishing valid performance standards – guidelines that recognize performance as something more than just behavior.

Performance Defined

Performance, as I view it, is defined by the outcomes of behavior. Behavior is individual activity, whereas the outcomes of behavior are the ways in which the behaving individual’s environment is somehow different as a result of his or her behavior. The concept of outcomes is similar (if not identical) to the concept of “accomplishment” described by Thomas F. Gilbert in *Levels and Structure of Performance Analysis* (1974), and to the concept of “achievements” described by Gilbert Ryle in *The Concept of Mind* (1949).

Performance, then, is the achievement of some condition that reflects one or more outcomes of the behavior of one or more individuals. (A definition of performance that is not behavior-dependent is particularly useful in integrating and understanding the many different kinds and levels of performance that occur within organizations.)

To illustrate these points, let’s analyze the following behavioral objective, taken from Robert F. Mager’s *Preparing Instructional Objectives* (1962, p.59):

"Given a properly functioning audiometer of any model, the student must be able to make the adjustments and control settings necessary prior to the conduct of a standard test."

The behavior described in this objective may be conveniently termed as “adjusting” behavior, whereas the outcomes of this adjusting behavior are reflected in the state of the audiometer, namely, controls that are set and adjusted.

The performance implied in this objective is also reflected by the audiometer, that is, when the student completes his or her activity, the audiometer is to be in a condition suitable for conducting a standard hearing test. The measure of performance for this objective is the extent to which the audiometer is or is not in that condition upon completion of student activity. From the presumably verifiable condition of the audiometer, it is reasonable to infer that the proper “adjusting” behavior did or did not occur. (It is also possible to deduce, from an improperly adjusted audiometer, where the adjusting process went awry and to make some reasonably useful inferences about how and why it went awry. However, the analysis of performance is not the subject of this paper, so comments about performance diagnostics will have to wait.)

The preceding analysis of a behavioral objective illustrates my central point: Performance is (or ought to be) defined more by the outcomes of behavior than by the behavior itself. It follows, therefore, that the development of performance standards should be based on an examination of the outcomes of behavior instead of the behavior that is thought to lead to those outcomes. As Gilbert (1974, p.13) wrote, “If you

Performance & Performance Standards: An Opinion

think about it, then, it is only the accomplishments of performance that we value – never the behaviors that produce them.”

The failure to distinguish behavior from performance is a misleading and a costly one. There are at least two ways in which behaviorally-oriented attempts to establish performance standards are misled:

1. So-called “behavioral” descriptions are mistakenly thought to reflect performance, and
2. Performance standards are imposed on activity instead of the outcomes of activity.

A discussion of both these false trails to performance standards is next, following by a discussion of some of the costs of taking these false trails.

False Trails to Performance Standards

Statements of behavior are abstractions; for example, the behavior of people called computer programmers is often referred to as “programming,” obviously an abstraction, and a generalization. Regardless of the degree of specificity used in describing programmer behavior (whether done by an observer, by the programmer, or by a panel of experts), the resultant descriptions are still abstractions – descriptions of behavioral events are simply not the same as the events they describe. More important, because behavioral statements are the direct products of the conceptual-verbal processes of the person formulating them, there is some question as to just what a supposedly behavioral statement actually describes – behavior, or the conceptual-verbal structure of the person formulating the description.

Statements of outcomes are also abstractions; however, they describe the products of activity. The outcomes of programming activity include function hierarchies, coding sheets, flowcharts, and other tangible products. Standards defining the quality of tangible products are much more easily developed and agreed upon than are standards for abstract descriptions of behavior. Performance standards relating to the outcomes of behavior are also more valid for assessing performance than standards relating to behavior. This is because the outcomes of behavior define the performance wanted, whereas behavioral statements reflect little more than an abstract description of the presumed causes of that performance (which brings us to the problem of imposing standards on activity).

There are three generally accepted classifications of performance standards: quality (accuracy), quantity (amount), and time (speed). A point seemingly overlooked in many attempts to establish performance standards is that quality, quantity, and time are classification categories for standards, not sources of standards. Standards should be derived from the outcomes of activity and they may then be classified in accordance with the three aforementioned categories. Instead, standards are all too often generated from these abstract categories and then arbitrarily imposed on activity. In the most extreme case imaginable, it is possible to have a supposedly measurable performance objective that not only does not describe performance, but that also has measures of performance that are arbitrary and totally lacking in validity. As a case in point, consider the following incident from the data systems area of a very large telecommunications company:

A computer programmer was transferred out of the programming unit because of a performance problem. She “took too long” from the start of her programming assignments to the point where she was ready to compile and edit her programs.

Later, it was discovered that this same programmer had been taking practically no time at all to compile, edit, and de-bug her programs. This owed to the care she exercised in their initial development.

Performance & Performance Standards: An Opinion

Consequently, the total elapsed time from her initial programming assignments to the point where her programs were up and running was no more than that for the other programmers in her group. More important, because her initial programs contained fewer errors than those of other programmers, the dollar-costs of her program were significantly lower than those of the other programmers.

It seems the other programmers were hurrying through the initial program development phase to meet the compilation deadline and then using inordinate amounts of very expensive computer time in de-bugging their programs.

All in all, the programmer who was transferred out of the unit was creating the best programs in terms of both quality and cost. Sadly, she fell victim to a performance standard that was arbitrary, invalid, and imposed on activity, a standard that had not been derived from an analysis of the outcomes of programming activity. (But, cheer up; we managed to get her reinstated.

The important point is to recognize that the true measures of performance are found in the outcomes of behavior, in the ways in which the performer's environment is somehow different as a result of his or her behavior. When behavior and performance are considered equivalents, attempts to measure performance wrongly focus on behavior and the behaving individuals instead of the outcomes of behavior. As we have just seen, this practice is very misleading. As we are about to see, it is very costly.

The Costs of Following False Trails

A focus on behavior, especially a focus on the behaving individuals, often leads to the use of normative measures of performance. Norm-referenced measures are largely irrelevant to the task of measuring performance because they measure differences between individuals and the task of performance measurement is to measure the consequences of interaction between individuals and their environment.

As William Powers put it in *Behavior: The Control of Perception* (1973, p.12), "It is unfortunate but true that measures and predictions obtainable only through averaging the performance of many persons are applied to individuals, so that a person's life may be seriously affected by his performance on a test that is valid only for predicting behavior en masse."

It is also unfortunate but true that the money spent on norm-referenced attempts to measure performance is by and large wasted because that which is being measured in such attempts is not performance.

To concentrate on behavior as the source or application point for performance standards is also to zero in on activity instead of results. Measures of activity (normative or otherwise) are, of necessity, measures of efficiency and not measures of effectiveness.

Given the misleading nature of many activity-oriented standards, it is likely that measurement based on such standards leads to false conclusions regarding efficiency and to the neglect of measures of effectiveness. I think it is impossible even to begin to estimate the costs of what appear to be frequent sacrifices of the measurement of effectiveness to what are certainly misleading and possible invalid measurements of efficiency.

Equating behavior with performance leads also to unnecessary conflict between the individual and the organization. Human behavior is but the means to various ends, and it must usually satisfy the ends of at least two separate parties: the behaving individual and those interested in his or her performance. If performance is described only in terms of behavior, then there is no alternative but for the management of performance to rest squarely on a strategy of direct control and manipulation of human behavior. The consequence of such an approach is predictable, inevitable, and costly: Conflict. On the other hand, as the percentage of performance that is described in terms of outcomes is increased, the requirement to manage

Performance & Performance Standards: An Opinion

performance through the control and manipulation of individual behavior is decreased, as is the potential for conflict between the individual and the organization.

Confusing behavior with performance greatly reduces the value of feedback as a technique for maintaining and improving performance. The concept of feedback requires the monitoring of outputs, not activity, and behavior is activity. Technically speaking, feedback is information about an actual condition with regard to some reference condition. (See William Powers' book, cited above, for a fascinating description of human behavior as the means whereby individuals control their perceptions of input stimuli as opposed to being controlled by them.)

If an individual is truly to receive feedback, then that individual must be able to recognize the reference condition, interpret information about actual conditions relevant to the reference condition, and act to reduce the difference between the two. Otherwise, data from his or her environment is not feedback; it is just so much "noise."

The reference condition for a given performance is defined by the standards for that performance and, as we have seen from the example involving the audiometer, standards should be based on the outcomes of behavior. True feedback is possible only when the standards for performance are based on the outcomes of behavior, otherwise, feedback is nothing more than a device to ensure obedience and accountability for performance is lost in a struggle for the control of behavior.

One implication of the relationship between feedback and performance standards is obvious but frequently overlooked, namely, that performers must be able to recognize quality in the outcomes of their behavior. (Robert M. Pirsig provides some interesting insights into the basic nature of quality in his bestseller, *Zen and The Art of Motorcycle Maintenance*.) If, for example, the standards for good computer programs are specified so the programmers can discriminate between good programs and bad, they are in a position to monitor and correct their own program development activities. In 1969, the *Programmed Instruction (PI) Writer's Course* at the Navy's Instructor Training School in San Diego, California was redesigned based on just such an approach. The trainees were taught to tell the difference between good PI and bad, and then provided with the opportunity to write good PI. The premise was that the trainees would be capable of providing their own feedback. Indeed, they were. On the other hand, if programmers are held responsible only for adherence to a set of abstractly defined developmental procedures, then all that can be monitored is activity, not output, and feedback will consist only of vague indications that the programmers aren't doing it right. The net result is the proliferation of myths and mythunderstandings about the "right way" of doing things.

Behavioral descriptions of performance have their origins in the field of training development. There, too, behavioral descriptions of performance incur high costs. There is usually more than one way to achieve a given outcome or result. Consequently, when "one right way" is prescribed, through a rigid behavioral or procedural description of performance, the individual performer's flexibility and adaptability are severely reduced, if not completely eliminated. More important, overly precise procedural descriptions of performance prevent the drawing of an accurate picture of the individual performer's true capabilities, that is, it is entirely possible that the individual can produce the desired results, but not through the specified activities. In fact, it is even possible that many training programs are created only because unduly confining procedural models of performance are specified and performers are then restricted to these models. The performers so restricted are of course subsequently identified as "trainees." The costs of such unnecessary training programs are probably astronomical.

To summarize thus far, when behavior is equated with performance, attempts to establish performance standards can be misled. Being misled, they incur high costs in the form of inappropriate measurement methodologies, false measures of efficiency, neglected measures of effectiveness, unnecessary conflict between the individual and the organization, vague and confusing performance feedback that really isn't feedback at all, and irrelevant and unnecessary training programs.

It is obviously easier to deal with issues of performance when performance standards are valid than when they are invalid. A likely question at this point is, "How does one go about establishing valid performance standards?" To answer that question would be to focus on activity, the very act I have been

Performance & Performance Standards: An Opinion

criticizing. A question more in keeping with the ideas presented thus far is, “What are the criteria for distinguishing between valid and invalid performance standards?” I choose to answer this latter question because I think it is more important to begin the task of “setting standards for standards” than to describe procedures for establishing standards, especially in light of the fact that the qualities of good performance standards have not yet been specified.

Standards for Standards

Performance standards for both formative and summative evaluation should reflect the quality of the outcomes of behavior that define the performance in question.

The preceding statement really contains two points: (1) performance standards should pertain to the outcomes of behavior instead of behavior, and (2) performance standards should pertain to the quality of those outcomes. In the case of computer programmers, this means that performance standards should relate not to the programmers’ behavior but to the function hierarchies, flowcharts, coding sheets, and other products they produce. Moreover, the performance standards for these products should reflect what is known about the quality of such products. (As an aside, I have found subject-matter experts or SMEs to be of more assistance in specifying standards for work products than in specifying skills or subject matter.)

I recognize that it is neither possible nor perhaps desirable to utilize summative or end-result performance standards as the sole basis for a performance measurement scheme. However, I do not believe that the requirement for formative or en route performance measurement should be used as an excuse or pretext for shifting the focus of performance measurement from outcomes to activity. Here’s why.

In systems terms, “process” refers to activity of the system, to the interactions between the system’s processors and the inputs to the system. These interactions produce the incremental state changes that inputs undergo in the course of being transformed into outputs. The ordinary practice of system analysis calls for (1) identifying the inputs and outputs, and (2) identifying the functions that cause the transformation of inputs into outputs. In contrast, the form of analysis to which I here refer is one of identifying the incremental changes or “deltas” that inputs undergo during the transformation process and then and only then identifying the alternative functions that might bring about these desired transformations. I call this particular process analysis technique “Delta Analysis.”

For example, in the Navy’s *Programmed Instruction Writer’s Course*, the trainees were taught not only the difference between “good PI and bad,” but also the difference between acceptable and unacceptable analyses, objectives, tests, terminal frames, teaching frames, editorial comments, field test data, and program introductions and administration.

In any event, formative evaluation, like summative evaluation, should be based on standards for outcomes, namely, the sub-products that constitute the developmental stages of the end product.

Almost all written materials go through a number of stages before becoming a finished product. Quality standards for each stage (e.g., first-draft, revised versions, and final copy) can and should be established.

If the first-draft of an instructional program is to be “as lean as possible,” then some measures of “leanness” must be established and some activity devoted to the achievement and measurement of that “leanness.” Presumably, editors and writers alike would have knowledge of these criteria and the long-standing issue of “How lean is lean?” could be put to rest.

The “leanness” of a first-draft instructional program seems to be a function of the writer’s analytical and writing capability, that is, the way he or she views and describes the performance. Accordingly, it might be better to establish criteria for “leanness” that would achieve consistency between an individual’s analysis and his or her first-draft program than to attempt to achieve consistency between the way individual program writers view or describe performance.

Performance & Performance Standards: An Opinion

The criteria for “lean programs” used in the Navy’s *Programmed Instruction Writers’ Course* were as follows: (1) the number of terminal frames is one per objective, (2) the minimum number of teaching frames is one per terminal frame, and (3) the maximum number of teaching frames shall not exceed the number of steps identified in the writer’s analysis of the objective.

The important point, whether conducting formative or summative evaluation, is to stay focused on the outcomes of activity and to use standards that are primarily focused on the quality of those outcomes.

Performance standards should be perceived as valid, realistic, and open to influence by those who are subject to them. The importance of the way in which performance standards are perceived cannot be overemphasized. If performance standards are seen as unrealistic or invalid, then people simply refuse to acknowledge as valid evaluations based on those standards. More important, when performance standards are seen as invalid, the time and energies of the performers will frequently go into finding ways to beat the system instead of finding ways to improve performance.

If the standards cannot be influenced by those who are subject to them, it is unlikely that the standards can be changed to reflect more valid and realistic concerns because the best sources of data – the performers – have no reason to speak out. Most important, the control process breaks down. Because the standards are not seen as valid, they do not serve as a reference condition. The performers do not “buy in” to the stated standards and operate instead in accordance with their own.

Performance standards should be stated in ways that facilitate their use by those who are subject to them. This requirement is important for two reasons:

1. When performers cannot use the standards to which they are subject to monitor and manage their own performance, they tend to discredit or discount the standards.
2. If feedback is to be immediate so as to shape performance instead of simply being an after-the-fact accounting, the standards and any feedback in relation to them must be usable by and useful to the performer.

It might seem desirable, for instance, to limit computer programmers to some average dollar-cost of de-bugging time on the computer. At first glance, it might then seem that performance standards would be based on cost and feedback would take the form of information about the costs incurred by each programmer. However, as such costs are a function of time, it might make more sense to provide the programmers with time-based standards. (Time can be measured and monitored by the programmers, whereas the costs of computer time vary with factors unrelated to the programmer’s use of time.) De-bugging time, however, is directly proportional to the numbers and kinds of errors in the initial program. Consequently, it might make the most sense to provide the programmers with standards for initial programs that are expressed in terms of the kinds and numbers of errors that will be tolerated.

Performance standards for one area of performance should be consistent with the standards for related areas. Performance occurs within and between various levels in an organization; for example, the level of the individual, the unit, and the entire organization. The performance standards for each unit must be consistent vertically and laterally. *Vertical consistency* means that standards between levels should add up. *Lateral consistency* means that the standards within a level should fit with one another. Examples of both follow.

- *Vertical consistency* refers to the interface between the standards for different levels of the organization; for example, the interface that exists between the standards for a work team and the standards for its individual members. In one organization, a major repair operation consisted of six discrete phases, all performed by different people. One of the standards for the entire operation was that 95 percent of all reported troubles were to be corrected within two hours of the time the trouble was reported. The time standards for the individual phases added up to two and one-half hours. It was possible for each individual to meet his or her

Performance & Performance Standards: An Opinion

time standard but for the operation as a whole to not meet its standard. In short, the standards didn't add up.

- *Lateral consistency* refers to the interface between standards for two areas of performance within a level; for example, the interface between standards for computer system analysts and computer programmers during the design and development of a computerized information system. In one company, this interface presented a problem. The programmers alternately accused the analysts of providing too little or too much detail and the analysts accused the programmers of demanding too much detail or of “getting into the analysts’ bailiwick.”

The product involved at this lateral interface was known as the Program Specifications Document (PSD). It served as the end product of the analysts’ activity and as the main input to the programmers’ activity. Unfortunately, there were no standards for the PSD. In other words, the interface between the analysts and the programmers was undefined.

In this case, the remedy turned out to be remarkably simple. The level of detail required by a beginning programmer was established as the maximum level of detail the analysts could provide, and the level of detail required by an experienced programmer was defined as the minimum level of detail the analysts could provide. Agreement was also reached to work out concrete illustrative examples during a series of upcoming projects. As long as the PSD fell in this mutually agreeable “zone of detail,” the programming group would have the capability of writing the required programs and neither group would complain of territorial encroachment.

Summary and Conclusions

A major focal point of this article has been the distinction between behavior and performance, namely, that behavior is individual human activity and that performance is defined by the outcomes of behavior. I have argued that performance standards should be based on the outcomes of behavior instead of behavior. Basing performance standards on the outcomes of behavior offers several advantages:

- The focus of the performance measurement effort is taken off the performer and places where it belongs – on performance.
- The selection of measurement methodologies is likely to be appropriate to the task at hand.
- The issue of effectiveness is unlikely to be obscured by mistaken notions about efficiency.
- The potential for conflict between the individual and the organization is decreased.
- The value of feedback as an integral element in controlling performance is maintained.
- The solutions to problems of performance, including training programs, are more accurately determined.

Basing performance standards on the outcomes of behavior is also advantageous because it draws attention to many performance-related issues that are not highlighted by purely behavioral descriptions of performance.

For example, behavior serves the individual’s needs as well as those of the organization. It is necessary, then, to examine the outcomes sought by the individual as well as those sought by the organization and to examine how these two sets of outcomes relate to one another.

Human behavior in organizations has many outcomes that are in no way task or work-related, so it is important to examine the social as well as the technical aspect of performance.

Perceptions and expectations regarding performance are held by many people in organizations and these might relate to behavior, to its outcomes, or to both. It is therefore essential to identify and reconcile the perceptions and expectations of all those having a stake in a given performance.

Performance & Performance Standards: An Opinion

Finally, many organizational outcomes are the result of more than one level of performance. Accordingly, it is important to determine how the outcomes of the behavior of many individuals interact and combine so as to produce unit and organizational outcomes. These are all issues that must be resolved in the course of defining desired performance and establishing standards for that performance.

At least, that's my opinion.

References

1. Gilbert, Thomas F., *Levels and Structure of Performance Analysis*. Praxis Corporation. Technical Series – No. 1. The Praxis Corporation, Morristown, NJ: 1974.
2. Mager, Robert F., *Preparing Instructional Objectives*. Fearon Publishers, Palo Alto, CA: 1962.
3. Powers, William T., *Behavior: The Control of Perception*. Aldine Publishing Company, Chicago, IL: 1973.

Author's Note:

This article originally appeared in the "Speaking Out" section of the *NSPI Journal*, Vol 16, Issue 1, pp.14-17 (1977). It was titled "Concerning Performance and Performance Standards: An Opinion." After dusting it off, and doing a little light editing, it seems as sound as when I first wrote it. At least that, like the rest of this article, is *my opinion*. I am quite proud of the fact that the definition of performance I presented in this article is the first definition of performance cited by Stolovitch and Keeps (1999) in Chapter One of the *Handbook of Human Performance Technology*.

Contact the Author

Fred Nickols can be reached by e-mail at fred@nickols.us. Many other articles of his can be found by going to www.nickols.us and clicking on the link to articles.