Performance Improvement Quarterly, 8(2) pp 95 113

Promoting HPT Innovation: A Return to Our Natural Science Roots

Carl Binder
Precision Teaching and Management Systems, Inc

ABSTRACT

The core innovations represented by the field of Human Performance Technology (HPT) trace their origins, by way of Programmed Instruction, to the field of Behavior Analysis, a natural science methodology for the study of behavior developed by B F Skinner This methodology, like all experimental natural science, rests on a foundation of functional analysis and standard units of measurement Functional analysis is basic experimental method, whereby the investigator or practitioner keeps all but one variable constant, changes the variable in question (an "intervention"), and measures the effect on other variables Behavior Analysis, like HPT, emphasizes prediction and control of individual behavior rather than determination of average effects across groups of individuals In order for HPT to support and encourage greater and more effective innovation, it must re-emphasize reliance on standard units of measurement and functional analysis and promote policies and procedures that increase variation of interventions The combination of encouraging variation and selecting interventions by means of functional analysis and objective measurement will ensure steady, reliable progress in Human Performance Technology

Introduction

Human Performance Technology (HPT) theorists and practitioners claim their work is research-based, grounded in empirical science, and focused on *results* Yet a review of NSPI publications over the last few years reveals that fewer than 5% of the tables or displays in articles or chapters contain measures of performance, comparisons of measured results, or measures of change in behavior or accomplishments (Lindsley, 1994) And only 4 out of 60 contributors to the *Handbook of Performance Technology* (Stolovitch & Keeps, 1992) shared samples of performance data What should we make of this embarrassing fact? Can we say for sure that we're consistently discovering and implementing performance interventions that produce measured results in the performance of individuals and organizations they serve? How can we tell what works and what doesn't? How can we select from the cafeteria of options, approaches, and alleged innovations that seem to roll by like waves?

As a field, how far have we gotten beyond so-called "level-one evaluation" (Kirkpatrick, 1976)—assessment of whether or not people *like* what we're doing? How can we be sure that the field *as a whole* is advancing toward ever-more effective performance solutions? Are the many approaches and interventions that our publications describe merely passing fads, trends in thinking and practice that arise, peak, and are replaced by others, without regard to measured effectiveness? Or do they truly represent innovation in measurably effective instruction and management technology?

This paper reviews the natural science origins of Human Performance Technology, describes how the experimental methodology of Behavior Analysis that gave rise to HPT can continue to ensure innovation and progress based on measured results, and offers some suggestions for promoting innovation in the field

HPT Roots in Behavior Analysis

Rosenberg, Coscarelli, and Hutchison (1992) reviewed the evolution of Human Performance Technology, emphasizing its foundation in Instructional Systems Design (ISD) and, even more fundamentally, in behavioral psychology Technically, *behavioral psychology* is a popular derivative of Behavior Analysis, a natural science approach to the study of behavior invented by B F Skinner (Bjork, 1993) that gave rise to Programmed Instruction and Instructional Systems Design (ISD), and which, in turn, led to Human Performance Technology

Despite this history, some current writers in the field refer to behavioral psychology or "behaviorism" as though it were an ancient mythology, an anachronism, a limited view of the universe with naive assumptions and primitive methodologies They contrast the behavioristic foundation of our field with current-day cognitive science, and more recently, with constructivism (Ertmer and Newby, 1993) These more recent disciplines, they argue, are more sophisticated, relevant, and effective than old-fashioned "behaviorism," because they are better able to deal with complexity

Much of the apparent rejection of "behaviorism" by current-day HPT professionals is based on a fundamental misunderstanding of its origins, principles, and methodologies In order to explain how the natural science approach represented by Behavior Analysis can continue to support solid innovation in HPT, it will be necessary to clarify this misunderstanding

Widespread Misunderstanding of Behavior Analysis

What, we might ask, is the "behaviorism" to which current-day critics refer? Is it the simplistic and mechanistic stimulus-response theory advocated by philosophers and experimental psychologists such as John B Watson and Ivan Pavlov during the early part of this century? Or is it the natural science of behavior, based on B F Skinner's single-subject research paradigms—a scientific methodology that led to unprecedented discoveries of order and regularity in the relationships between behavior and the variables of which it is a function (Bjork, 1993, Johnston and Pennypacker, 1980, Sidman, 1960, Skinner, 1938)?

Unfortunately, it is a simplistic stimulus-response account of behavior, which many undergraduate textbooks and popular articles inaccurately equate with Skinner's work, that colors the understanding of current-day critics Prompted most dramatically by an inaccurate and misleading representation of Behavior Analysis by Noam Chomksy in his infamous (and some might say academically irresponsible) review of Skinner's book *Verbal Behavior* (Chomksy, 1967, MacCorquodale, 1970), a mechanistic rendition of the science spread across academe and into the general literate public, largely without reference to primary sources This misrepresentation of the science continued to multiply through several generations of graduate students and professors, whose misrepresentations of Skinner's work suggest that they either did not read or did not comprehend scholarly articles or books by Skinner himself or by any of those who followed him in the field of Behavior Analysis It is *this* "behaviorism" to which most critics refer today, often unwittingly accepting rendition after simplistic rendition, rather than referring to the primary texts or to any of the numerous contemporary research journals in Behavior Analysis (e g , *Journal of Experimental Analysis of Behavior, Journal of Verbal Behavior, Journal of Applied Behavior Analysis, Journal of Organizational Behavior Management, The Behavior Analyst*) As a result, little of the rich methodological and conceptual contribution of this science has spread beyond a community of practicing behavioral researchers and application specialists

> ***Is it coincidental that the volume of data-based research in HPT has waned along with the influence of Behavior Analysis?***

This historical turn of events has had a profound effect on current-day understanding and application of core HPT principles The underlying analysis and evaluation methodology of HPT has drifted away from its data-based, scientifically secure origins This development may also have had a decelerating effect on the pace of empirically validated innovation in our field, reflected by the lack of objective performance measures in NSPI publications

HPT's Natural Science Foundation in Measurement

If asked to identify B F Skinner's most important contributions, the majority of professionals would likely cite one or more of the *findings* associated with his study of reinforcement schedules (Skinner, 1938, Ferster and Skinner, 1957), stimulus discrimination (Skinner, 1933), or perhaps programmed instruction (Skinner, 1968) However, in Skinner's own view, his most important contributions were use of *response rate* as the basic measure of behavior strength, and invention of the *cumulative response recorder* which monitors moment-to-moment changes in response rate (Evans, 1968, Skinner, 1938) In other words, it was his *measurement technology* that Skinner considered most important, and upon which he founded a natural science of behavior unlike any that preceded or followed it (Bjork, 1993, p 93) Beyond the measurement tools themselves, it was the analytical methodology for the evaluation of variables, known as *functional behavior analysis*, that was Skinner's greatest legacy (Sidman, 1960, Johnston and Pennypacker, 1980)

Ingredients of Functional Behavior Analysis

For the uninitiated, it will be worthwhile to review the essentials of functional behavior analysis in order to understand its fundamental contribution to our field Whether we choose to design performance interventions based on behavioristic, cognitive, or constructivist assumptions, the *method* of functional behavior analysis remains an essential foundation for HPT in natural science

In its simplest terms, we might say that functional analysis is based on three methodological premises

> **First premise** The goal of any science or technology of behavior is the prediction and control of behavior

While the unvarnished directness of this statement got Skinner into lots of trouble (e g , with publication of *Beyond Freedom and Dignity*, 1971), one might ask what other purpose we could possibly pursue Either we seek methods to improve education and training, therapy, management, and other activities intended to influence the way people behave, or not If we *are* concerned about changing or improving how people behave, then let us be blunt we seek to discover and apply laws of nature that govern behavior, to determine which specific interventions are most likely to affect behavior, and to assess their relative impact This is, in essence, prediction and control (The politically correct term might be *influence*) Such an orientation contrasts with an approach that selects programs or theories based on personal preference, consensus opinion, or other decision criteria not grounded in measured results

In passing, it might be worthwhile to mention that current-day constructivists may be unable to accept this first premise, to the extent that they adhere to radical subjectivism and therefore deny the very possibility of scientific laws regarding behavior The constructivist view seems to question the very essence of HPT, which is ostensibly aimed at developing reliable, cost-effective methods for producing or enhancing desired learning and performance outcomes, and which therefore must rest on the possibility of predicting the effects of interventions

> **Second premise** When assessing the effects of variables on behavior, it is best to observe and analyze the behavior of *individuals* rather than basing conclusions on *average* results across groups

Skinner's cumulative response recorder, still a standard tool in many basic research laboratories, monitors and produces graphic records of moment-to-moment patterns in *individual* response rates of target behaviors or accomplishments Skinner's overall approach was to understand, replicate, and refine interventions capable of reliably changing or maintaining patterns of *individual* behavior in order to discover *general* laws or rules that hold for most, if not all, individual organisms under specified conditions (Johnston and Pennypacker, 1980, pp 255 ff) *The method of functional behavior analysis is based on repeated demonstrations of effectiveness across many individuals rather than average effectiveness for a group*

Focus on the individual was an important characteristic of systematic instructional technology from the beginning (e g, Markle, 1964) *Averaged* group response measures may mask individual differences A particular curriculum or management intervention may produce an *average* increase in performance across a large group, but we cannot predict on the basis of such data that it will be effective in every *individual* case On the other hand, if we can identify variables powerful enough to affect the behavior of many or most *individuals*, and if we can repeatedly demonstrate such results, then we will have developed a basis for implementing robust, *generally effective* interventions As a practical example, a user-tested computer interface is a *general* performance intervention designed on the basis of many *individual* observations and tests Such an individual orientation is part of the legacy given by Behavior Analysis to HPT

Skinner's focus on the individual established an important precedent for Gilbert's (1978) emphasis on observing and replicating the conditions that support the exemplary accomplishments of *individual* performers Focus on individual learning and performance was also a key assumption in Mager's (1988) formulation of Criterion-Referenced Instruction, which enables *individuals* to achieve measurable objectives, by sometimes divergent paths, at their own pace

There is, by definition, no set of observations or procedures that cannot be described using the basic temporal sequence of functional analysis: what comes before the behavior in question, the behavior itself (whether covert or overt), and what comes after the behavior.

Third premise The domain of behavior and the variables that might affect it can be divided into three parts, based on temporal sequence

- *Antecedent events* The events and conditions that precede behavior
- *Behaviors* The overt actions or covert thoughts and feelings we seek to analyze, predict or control, and
- *Subsequent events* The events or conditions that follow target behaviors

This categorization of behavioral and environmental events, based on temporal sequence, is a very basic and generally applicable approach—independent of one's theoretical framework, whether the behaviors or their environments are simple or complex, whether the behaviors in question are overt or covert, or whether the situation we are analyzing or managing is isolated in the laboratory or part of a highly complex world There is nothing theoretical or biased about this observational and analytic "chunking" strategy, since behavior and performance do, in fact, occur in temporal relationship with the environment

Once having specified events and behaviors in time, functional behavior analysis seeks to identify those antecedents and/or subsequent events that have reliable effects on the form or frequencies of behaviors—and which can therefore be described as *functionally related to* the desired behavior change, with precisely the classical scientific or mathematical meaning of *functional relationship* (e g , Y as a function of X on a graph)

Objective Descriptions Operational and Functional

As a requirement for performing functional analysis, Behavior Analysis draws an important distinction between *operational* description and *functional* description of behaviors and environmental events as they occur in time

An *operational description* specifies the events or conditions one is observing or evaluating, clearly and completely enough so that other scientists or practitioners can recognize and/or replicate the situation by referring to the written or verbal description Operational description includes the "operations" performed in order to affect behavior change, and also descriptions of target behaviors

Operational description is a basic requirement for any science or technology, a key differentiator from art or craft Art or craft can survive in the form of peculiar or unique instances of creativity and innovation, which may or may not be replicable by others Science and technology, which aim for general solutions or laws, cannot survive unless they use descriptions of events, conditions, and procedures able to be repeated and verified in the future and by others Empirical validation and effective communication of the effects of innovative procedures cannot occur without a solid foundation in operational description

By arranging events and measuring the effects of interventions, Behavior Analysis moves from operational to *functional* description—specification of behaviors and events with reference to what they *do* to one another

Lindsley (1964) clarified the distinction between operational and functional description by using what he called *IS* (operational) and the *DOES* (functional) terminology for expressing relationships among behavioral and environmental variables

IS (operational terms)

Antecedent Event
↓
Behavior
↓
Subsequent Event

Operationally speaking, the only certain relationship among the events is temporal some events or conditions occur *before* behaviors, while other events or changes in conditions occur *after* behaviors (Those interested in "cognitive processes" should keep in mind that behaviors can include covert thoughts or feelings, and that behaviors can be antecedents for other behaviors) By changing one of the variables and simultaneously measuring for effects on the others it is possible to determine whether there are any functional relationships among these events—whether they have any reliable *effects* on one another Such measurement and evaluation of effect leads to *functional* description, expressed by Lindsley in the *DOES* terminology

DOES (functional terms)

Discriminative Stimulus
↓
Response
↓
Consequence

> ***What separates the "performance-based" orientation of HPT from other approaches to performance improvement is the assumption that it is possible to discover regularity in the relationships between behavior and the factors that influence it, and to use that regularity to help produce desired performance outcomes.***

The functional *(DOES)* terminology specifies what the events *do*, or how they *function*, with respect to one another If, for example, we change or introduce an antecedent event (e g , by providing an instruction to perform a task in a different way, or supplying a job aid) and a different behavior occurs reliably, then we can describe the antecedent event as a *discriminative stimulus* and the behavior as a *response* This term indicates a cause-and-effect (functional) relationship between the antecedent and the behavior *Stimulus* is thus a functional term, and can only be used if we have determined that the event or condition to which it refers has an effect on (or function with respect to) the behavior *Response* is also a functional term that we use only after determining that the probability of behavior changes in relation to previously verified stimuli or consequences (Using this technical terminology, we might say "That instruction just isn't a *stimulus* for Bob," in the event we cannot yet demonstrate a functional relationship) Similarly, if we change what happens *after* a behavior, and the rate of behavior increases, then we can say that the event is actually a functional consequence (or *reinforcer*) insofar as it increases the strength of behavior it follows (Again, not all subsequent events function as *reinforcers* And there are individual differences - "different strokes for different folks ")

For example, compensation or incentive programs may or may not function as reinforcers with respect to the job behaviors they are intended to increase Whether or not they do is an empirical question, and can only be answered by "running the experiment" – varying incentive arrangements and measuring their effects on behavior

In order to apply functional description, one must use objective, standard measurement procedures and instruments to monitor what happens This might be as simple as counting standard units of behavior or accomplishment without instrumentation (e g , from accounting records, self-tallying, or test scores), or as complex as using an automated monitoring and recording environment (e g , built into computerized workflow automation software or a laboratory apparatus) But in each case, we change conditions and objectively measure the effects to determine the functions of behaviors and environmental events

Strictly speaking, it is also necessary to repeat the "experiment" more than once, measure the effects, and determine that there is a reliable relationship that we can predict and use to influence or control what happens The principle of *replication* in behavior analysis replaces the principle of "average effect" as a means of demonstrating generality in more statistically oriented social science methodologies (Sidman, 1960) (This principle foreshadowed the evaluation-revision cycle of Instructional Systems Development, whereby one "replicates" a particular intervention and refines or modifies it until it produces the desired results with most or all individuals)

Functional analysis is the essence of what we claim to do in HPT to ensure that our interventions are effective We use objective measurement coupled with experimental or evaluation designs to identify what procedures have desirable, reliable effects on behavior, and thereby on production of target accomplishments (Gilbert, 1978) While this methodology is fundamental, and implicit in our claims for empirically validated methods, the fact that our publications generally lack reports of performance outcomes suggests that it is not widely practiced by those espousing HPT

Why HPT Has Abandoned Skinner's Legacy

Simplistic renditions of the "three-term contingency" (Skinner, 1953) or functional relationships among discriminative stimuli, responses, and consequences have contributed to misunderstandings about "behaviorism " In fact, functional analysis does *not* suppose behavior to be a collection of simple, mechanistic stimulus-response linkages Rather, it reveals a dynamic *field* of changing probabilities in which different elements shift in their relative prominence and frequency to form *a continuous fabric or stream of interaction between individuals and environments* (In fact, Behavior Analysts' study of complex behavior-environment interactions closely resembles Ecologists' study of multi-dimensional, organism-environment ecosystems) The three terms of Skinner's model support an analytical methodology that enables scientists or technologists to identify the effects of various elements in the stream Even in simple laboratory situations, behavior-environment interactions are not mechanical or binary They are not like the on/off switching of digital computers Rather, they are *probabilistic*, more appropriately modeled in the field of artificial intelligence by neural networks than by sets of decision rules

I believe that the over-simplification of Skinner's three-term contingency has resulted in a general ignorance about the power and generality of functional behavior analysis This has led to a gradual degradation in the extent to which HPT has relied on data-based functional behavior analysis as a scientific methodology, or "innovation engine"

One source of this over-simplification has been the experimental conditions under which some basic research scientists analyze behavior Often, laboratory behavior analysts have chosen lower organisms, easy-to-repeat responses, and simple stimulus dimensions in order to isolate, manipulate and thereby measure the effects of specific variables In using such simple conditions for basic research, they have applied the same rationale as when physicists manipulate sub-atomic particles in accelerators or cyclotrons to understand, predict, and sometimes control events occurring in the complex universe Experimental science of all types studies simple situations as elements of more complex "real-world" conditions, in order to isolate and measure the effects of specific variables Behavior Analysis is no less complex in its implications than experimental physics, biology, or chemistry Surely we would not accuse modern-day formulations of chemistry or physics of being "too simplistic" merely because experimental scientists in those fields work with relatively simplified conditions prior to extrapolating to more complex situations In fact, just as mechanical or electrical engineers apply simple principles of physics with incredible complexity, so performance engineers attempt to apply laws of behavior in complex situations The simplicity of basic laboratory research conditions should not be misconstrued to allow only simplistic applications in the real world

The most fundamental purpose for measurement is to decide whether and how much a given intervention affects the performance of a given individual. The self-correcting character of HPT depends on measurement in this form.

Gilbert's Extrapolation From the Three-Term Contingency

Human Performance Technology, as formulated by Gilbert (1978) and others who came from the tradition of Behavior Analysis, represents a very successful extrapolation from basic science to the complexities of everyday life, just as the design of airplanes combines application of many relatively simple principles of physics Gilbert's (1978, p 85) Behavior Engineering Model, which divided possible behavior influences into six categories, mirrored Skinner's three-term contingency

- *Information* (corresponding to *discriminative stimuli*) divided into *data* in the environment and *knowledge* in the individual,
- *Instrumentation* (corresponding to *responses*) divided into *instruments* in the environment and *response capacity* in the individual, and
- *Motivation* (corresponding to *consequences*) divided into *incentives* in the environment and subjective preferences or *motives* in the individual

Gilbert created a matrix with Skinner's three-term temporal sequence on one dimension and the environment/individual distinction on the other Whether or not one adopts this particular categorization of the variables affecting performance, the underlying scientific methodology of identifying variables and measuring for possible effects of changes in those variables provides a foundation for systematic, data-based decision-making about what is needed and what "works" to improve performance When HPT practitioners conduct front-end analyses, use try-outs and pilot tests, or continuously improve their interventions based on measured results, they follow basic principles derived from functional behavior analysis This approach also provides a strong foundation for continued innovation

Functional Analysis Still Works!

Why have we taken this apparent digression into the methodology of behavior science? Simply stated, the science underlying the origins of HPT has gotten a bad rap Misrepresented by simplistic renditions, it has appeared to the larger community of professionals and the literate public as a crude and simple-minded approach that attempts to describe the behavior of people as though they were rats or pigeons responding under the influence of colored lights and food pellets, caged in boxes! This misunderstanding of the science thoroughly ignores the enormous range of *human* behavior analysis research and application, the increasing sophistication of quantitative behavior analysis and *behavioral economics*, and the growing links of functional behavior analysis to behavioral biology (Malott, Whaley & Malott, 1993)

In the wake of this misunderstanding, philosophical approaches represented by cognitive science and now constructivism have come to fill the perceived gap in a "behavioral" account supposedly created by the complexity of human cognitive behavior and the "real world" environment Is it coincidental that the volume of data-based research in HPT has waned along with the influence of Behavior Analysis? Whether or not there is a relationship, it is important that HPT not abandon the powerful methodology from which it arose

Consider that there is, by definition, no set of observations or procedures that cannot be described using the basic temporal sequence of functional analysis what comes before the behavior in question, the behavior itself (whether covert or overt), and what comes after the behavior Whether simple or complex, behavior occurs in a stream of environmental events and individual responses, overt or covert Any approach that claims scientific validity will need to take temporal sequence, and functional relationships, into account To the extent that any analysis of performance seeks to identify the effects of one variable upon others, it will be functional analysis – like experimental science in physics, chemistry, or biology

Without absolute, standard units of measurement, the field of HPT is unlikely to produce reliable, scientifically solid innovation.

Human Performance Technology, if it seeks to understand and optimally arrange the factors that influence behavior in the workplace or elsewhere, needs to remain firmly rooted in operational description, direct measurement of results, and functional analysis Whether it be called "Behavior Analysis" or not, this natural science foundation is what makes HPT potentially so powerful it offers the benefits of self-correction and continuous improvement based on data

Cognitive science has traditionally taken Behavior Analysis to task by claiming that "simple stimulus-response" cannot account for the complexities of human problem-solving, conceptualization, and other advanced behavior (Ertmer and Newby, 1993) As a field, it claims to solve this problem by advancing models of mental processes and other hypothetical constructs supposed to exist in the mind or in the brain of the performer Its research methodology is based on hypothesis testing Researchers make predictions using mental models or other hypothetical constructs, and then test those hypotheses in statistical designs The research is intended to confirm or invalidate the hypothetical constructs But even when HPT seeks to apply cognitive constructs, it still must focus on and measure the behaviors and accomplishments it aims to influence In cognitive science, the variables being manipulated tend to be one form of antecedent or other either complex visual or auditory inputs, complex real world situations, or internal self-cueing and other forms of covert behavior that prompt additional behavior None of these behavioral or environmental elements fall outside the scope of the three-term contingency or functional analysis The primary contribution of cognitive science to HPT, then, may be that it has focused our attention on more complex stimulus configurations and more complex, and often covert, behaviors Methodologically, HPT still demands functional analysis, and probably can't afford to rely on hypothesis testing in applied settings

Constructivism—in its extreme version a form of radical philosophical subjectivity—is based on the view that each individual "constructs" his or her own reality, and that any attempt to objectively specify either desired learning outcomes or the elements of a complex environment is, by definition, impossible, since everyone's reality is different At some point this position rejects the possibility of applying scientific method to human affairs But in its less extreme form, this view merely claims that people learn best in complex, real-world environments, and that they learn in highly individualized and unpredictable ways and with highly individualized outcomes (Ertmer and Newby, 1993) Again, nothing about the environment, subjective experience, or overt behavior of persons is beyond the scope of the three-term contingency as a descriptive framework, unless one takes the extreme constructivist position—in which case, there is no basis for any form of technology that reliably defines or produces outcomes, let alone a science The primary contribution of constructivism to our field, then, may be that it has led us to prepare people for more complex environments by shifting attention to problem solving and other adaptive repertoires, and it has reminded us of a broader set of individual differences

The point is that what separates the "performance based" orientation of HPT from other approaches to performance improvement is the assumption that it is possible to discover regularity in the relationships between behavior and the factors that influence it, and to use that regularity to help produce desired performance outcomes In that context, the principles of functional behavior analysis represent an underlying discovery method, implicit in the ISD model, which recommends refining interventions through repeated loops of Analysis-Design-Development-Implementation-Evaluation-Revision, until they produce optimal results (Rosenberg, Coscarelli, and Hutchison, 1992)

Measurement and Innovation

Without direct, standard measurement of outcomes, it is not possible to objectively evaluate or compare interventions In other words, the most fundamental purpose of measurement in HPT is to determine the "functions" of various interventions intended to affect human performance In fact, the progress of natural science over the centuries has occurred largely because of progress in measurement technology (Johnston and Pennypacker, 1980) Without objective evaluations and comparisons of effects, HPT as a field cannot support its claims to be based on scientific research methods or to produce measurably superior results

Reasons People Measure

In the practice of HPT, there are three possible reasons for measuring the effects of what we do

- **Validation** to prove that some general method or program "works" (often associated with publications or academic theory-testing),

- **Accountability** to meet administrative criteria that hold practitioners and managers accountable for the results of their interventions, and
- **Decision-making** to support individualized or group decisions about what's working and what to try next

Any good system for making individualized decisions about effectiveness supports the other two purposes, as well If enough individuals respond positively to a given intervention, that intervention might be “publishable” and considered validated If we take the trouble to collect individualized decision-making data (e g , progress toward criterion performance), those data also support organizational accountability But ultimately, the most fundamental purpose for measurement is to *decide* whether and how much a given intervention affects the performance of a given individual The self-correcting character of HPT depends on measurement in this form And innovation will be best served by measurement systems designed to serve this purpose

If our field were to move more aggressively toward standard units of measurement, then it would strengthen the foundation for innovation, continuous progress, and improved efficacy.

The Importance of Standard, Objective Units of Measurement

Natural science deals with *standard* units of measurement (meters, grams, minutes, liters, counts of objects, etc), as does business accounting Neither accounting nor scientific discovery could proceed very far with rating scales or percentage correct measures in the absence of absolute, standard measures Fortunately, HPT has been influenced by both business and the natural science of behavior, so one might expect practitioners of HPT to use objective measures for evaluating interventions Nonetheless, most of the data we see in HPT publications and presentations lacks standard units of measurement Rating scales (which are essentially refined opinion) and percentage correct calculations (which “cancel out” the absolute counts on which they are based) do *not* allow us to objectively quantify or evaluate the behaviors or accomplishments we claim to produce or improve (Johnston and Pennypacker, 1980) Without absolute, standard units of measurement, the field of HPT is unlikely to produce reliable, scientifically solid innovation

Gilbert's table of “performance requirements” (Gilbert, 1978, p 45) presents a reasonable list of standard measures for HPT The following list represents a slight modification of Gilbert's original, focused on ensuring that each type of measure is something that we can *count*, and thereby use to assess change over time

Quality (counting by type or category)

- correct vs incorrect answers or acceptable vs unacceptable units
- different classes or categories, defined by objective criteria
- unique or innovative accomplishments or behaviors, using criteria
- timeliness (counting those completed within a specified time limit)

Quantity (Counting by amount produced)

- number
- volume
- market value (in units of currency)

Cost (counting by dollars or time spent)

- labor
- materials and environment
- management

The field of performance management, a sub-set of HPT, has been perhaps most aggressive in applying such objective measures in organizations (Daniels, 1989) Journals such as the *Journal of Organizational Behavior Management* are filled with articles containing such performance measures If HPT publications and practitioners more frequently and consistently reported results using one or more of these standard measures, while also providing clear operational descriptions of interventions, we would be far better able to evaluate and compare the effects of specific interventions We would also be in a much stronger position, as a field, to continuously improve our technology based on progressive refinement of interventions

This is why it is disconcerting to see so few NSPI publications or articles containing objective measures of results If we are research-based, what is the research? Without performance data, effectiveness is a matter of opinion—even if that opinion is formalized in a five-point rating scale (telling us merely what people *think* or *feel* works, or what they like or dislike most) On the other hand, if our field were to move more aggressively toward standard units of measurement, then it would strengthen the foundation for innovation, continuous progress, and improved efficacy Binder (1988), Lindsley (1994), Geis and Smith (1992), Smith and Geis (1992), and others have made specific methodological recommendations for measuring performance, *some* of which emphasize objective, standard units of measurement

Recommendations for Supporting Innovation in HPT

The previous sections of this article summarize key elements of the scientific methodology from which HPT evolved as a research-based approach to improving performance If natural science, with its focus on objective verification, is to continue as a model for HPT, then these core methodological elements must continue to drive innovation in the field

My first recommendation for supporting effective innovation in HPT, then, is that practitioners increase the frequency with which they gather and report results in the form of standard measures of behavior and accomplishment Given the previously cited lack of reported results in HPT publications, I have always wondered whether practitioners were actually gathering such information, but not sharing it for reasons such as client confidentiality. While this is a credible explanation in some cases, it is hard to believe that this problem can fully account for the lack of reported data. Depending on the source of the problem, we must either do a better job of convincing our clients to collect and use objective measures of performance, or we need to develop standard methods for reporting such information while maintaining confidentiality. In either case, we need to work harder to “put our money where our mouths are”—to gather and report more objective measures of performance.

My second recommendation is to more rigorously apply the scientific methodology of functional analysis (embodied in the evaluation/revision cycle of the ISD model). Adhering to these guidelines would enable the field to rest on a much stronger foundation for comparing the effects of different interventions and identifying the variables that reliably produce performance improvement in individuals and groups. Again, there is a question about how often HPT practitioners actually pilot test and evaluate interventions, based on objective measures (not merely rating scales or other “Level 1” assessments). If we are neither pilot testing interventions and revising them until they are measurably effective, nor gathering and using objective measures of behavior, then our field is, frankly, hypocritical. If, as a field, we are attempting to conduct effective evaluation-revision cycles, but running into resistance, then it behooves us to focus our attention on this problem and to support a common effort toward more frequent and objective evaluation and revision on the basis of measured results.

> ***NSPI and those concerned with increasing the discovery of effective methods, procedures, tools, and programs should make rewards and recognition contingent upon demonstration of replicable, objective results.***

The remaining recommendations for increasing effective innovation in our field depend on application of several basic principles of behavioral and cultural evolution.

A Model for Innovation Variation and Selection

Skinner (1986), Johnson and Layng (1992), and others have emphasized that the same general principle of “selection by consequences” applies to biological evolution, individual learning, and evolution of culture. The evolutionary dynamic in each domain is the same: variation of alternatives and selection by consequences. In evolution, the variation is genetic and selection is by “survival of the fittest.” For the individual, various biological, physical, cultural and educational factors prompt new behaviors, and learning occurs when those behaviors make contact with their consequences. Cultures, too, evolve through variation of cultural processes and selection of those that work best to maintain the well-being or survival of the group.

This same two-stage process applies to any effort aimed at producing effective innovation, including the practice of experimental science. I suggest that attention to this model could help to drive innovation in HPT. In short, we need to encourage new and different types of innovations for improving performance, and then establish criteria and consequences based on effectiveness—the results or accomplishments produced by the interventions. As long as our selection criteria depend on measured effectiveness, the field will progress.

Promoting Variation in HPT

As a field, we can err in two ways with respect to encouraging innovative approaches. If we fail to encourage significant variation, sticking with narrow variations of existing methods, we may miss significant opportunities to improve our technology. Alternatively, as in education and many parts of the training industry, we could err by encouraging new and different procedures *for their own sake, independent of effectiveness*. The latter produces the kind of "faddishness" from which many of the education and training professions suffer. The goal is to encourage variations that are likely to produce superior measured results. The following methods might promote such effective variation.

Setting expectations that *effective* innovation will be rewarded. In industry forums and publications, if we emphasize the importance of measurably effective innovation, perhaps establishing professional award and recognition programs, we will likely see more efforts to generate new effective technologies among practitioners, their sponsors, academic professionals, and students. Clear expectations for effectiveness may discourage practitioners from valuing interventions simply because they are new.

Encouraging variations based on previously validated methods. This is the most conservative approach, but one which can produce steady innovation. Most of science operates by this principle. While scientists occasionally stumble on major breakthroughs, often by chance, most progress in science occurs by extending existing methods and approaches a little further each time. It should be the function of journals and professional meetings to support such gradual evolution of the field.

Challenging one another to address very difficult problems. Consider John Kennedy's challenge for NASA to reach the moon "by the end of the decade." This was a case of setting a very difficult goal which would, of necessity, require entire systems, approaches, tools, and methods that no one had yet created. A basic principle discovered in the laboratory about behavior is that when a given response fails to produce the desired outcome, then variation in responding increases (Skinner, 1938). When we encounter previously insoluble problems, initial attempts seldom produce the desired outcome. Therefore, setting very difficult goals or approaching very challenging problems is likely to prompt greater variation. Consider the rich scientific and technological by-products of the "space race." An example that is somewhat closer to home involves the Loebner Prize (Allen, 1994), a contest co-sponsored annually by a private donor and the Cambridge Center for Behavioral Studies In this contest, software developers submit programs intended to meet the criterion that a person interacting with them will not be able to tell if the printed responses they receive are from a software program or from a person at a keyboard on the other side of a wall. (This is known as the Turing Test for artificial intelligence, after the scientist who originally proposed it.) The point is that by proposing a very difficult challenge, this contest has generated a tremendous amount of innovative software development. NSPI or other professional organizations might consider arranging difficult performance improvement challenges as a means for promoting greater innovation. (Difficulty might be defined either by the type of problem, or by the desired magnitude of performance change.) Similarly, individual HPT practitioners might seek particularly difficult performance problems as a means of advancing their own innovation.

Combining principles from multiple disciplines. One of the strengths of HPT as a field (and of NSPI as an organization) is that its scope seems to be ever increasing, attracting bright people from training, management, human factors engineering, software development, communications, even stand-up comedy and magic! The benefit of increasing our community "repertoire" is that, as long as we hold firmly to objective criteria for effectiveness, it should be possible to generate a wider variety of potential solutions to learning and performance problems. Much as an individual musician or performance artist expands his or her capacity for improvisation by learning new skills, HPT as a field can benefit from its expanding scope. Examples of HPT innovation based on an integration of disciplines include electronic job aids (Front End Analysis: performance and training analysis software, 1994) which combine on-line computer technology with performance-based job aid methodology, and the Information Mapping® method (Horn, 1985, 1992;) which combines principles from a half dozen different fields of research into a systematic methodology for analyzing, organizing, and presenting information that produces measurable improvements in reading rate, comprehension, and task completion.

Rewarding effective innovation. The practice of natural science sets a model for rewarding effective innovation. Given objective measures of results and a functional analysis methodology aimed at reliably assessing and comparing results, effective innovation tends to "speak for itself." NSPI and those concerned with increasing the discovery of effective methods, procedures, tools, and programs should make rewards and recognition contingent upon demonstration of replicable, objective results. With clear, operational descriptions of interventions, and standard units of objective measurement, it should be possible to compare the effects of interventions aimed at achieving the same results, and to demonstrate which is best. For example, work in progress at a major semiconductor corporation is evaluating the effects of a literacy training program on the performance of standard job tasks that require quantitative and reading skills This program has previously been shown to produce criterion performance on standardized educational tests more than 20 times faster than average public school programs (Johnson and Layng, 1992) Using standard performance criteria and objective measures, it is possible to select those interventions that produce greatest effects Consistent with the accomplishment-based philosophy of NSPI and HPT, we should enshrine replicable, objective results as the highest possible accomplishment This is the bedrock of scientific progress

In addition to or instead of its annual awards for outstanding programs, methods, and publications, NSPI – the "home" organization of HPT – might consider awarding prizes and recognition for interventions, articles, or methods that demonstrate the greatest effectiveness in producing one or a number of standard, objectively measured performance outcomes

Conclusion

There will surely be readers who conclude that this author is merely a "behaviorist" seeking to re-impose a narrow view on an ever-widening field of endeavor Some may criticize my appeal to the principles of Behavior Analysis as anachronistic, in a period when these principles are being "replaced" by a new generation of cognitive and constructivist methodology To those readers I ask only this If you think that objective measurement and functional analysis no longer serve the purpose for which they were intended—the identification of variables capable of significantly influencing individual and group behavior—then what principles should we put in their place? If HPT cannot rely on the basic principles of experimental science, then what is to distinguish it from any other philosophical trend or fad? How, in short, can we argue that our overall approach is more likely to produce results than any other, if we neglect the principles and methodology of natural science?

References

Allen, F (1994) Unreasonable facsimile Do we really want computers to be more like us? *Atlantic Monthly,* August, 20-23

Binder, C (1988) Measuring performance *CBT Directions*, October (Reprinted in *Data Training*, December, 1988)

Bjork, D W (1993) *B F Skinner A life* New York Basic Books

Chomsky, N (1967) A review of B F Skinner's *Verbal Behavior* In L A Jacobovitz and M S Miron (Eds), *Readings in the psychology of language* New York Prentice Hall, 142-171

Daniels, A C (1989) Performance Management Improving quality productivity through positive reinforcement Tucker, GA Performance Management Publications

Ertmer, P A and Newby, T J (1993) Behaviorism, cognitivism, constructivism Comparing critical features from a design perspective *Performance Improvement Quarterly,* 6(4), 50-72

Evans, R I (1968) *B F Skinner The man and his ideas* New York E P Dutton and Company, Inc

Ferster, C and Skinner, B F (1957) *Schedules of reinforcement* New York Appleton-Century-Crofts

Geis, G L , and Smith, M E (1992) The function of evaluation In H D Stolovitch and E J Keeps (Eds), *Handbook of Human Performance*

Technology San Francisco Jossey-Bass, 130-150
Gilbert, T F (1978) *Human competence Engineering worthy performance* New York McGraw-Hill
Horn, R E (1985) Results with structured writing using the Information Mapping® writing service standards In T M Duffy and R Waller (Eds), *Designing usable texts* Orlando, FL Academic Press
Horn, R E (1992) Developing Procedures, Policies, and Documentation Waltham, MA Information Mapping, Inc
Johnson, K R and Layng, T V J (1992) Breaking the structuralist barrier literacy and numeracy with fluency *American Psychologist 47,* 1475-1490
Johnston, J M and Pennypacker, H S (1980) *Strategies and tactics of human behavioral research* Hillsdale, NJ Lawrence Erlbaum Associates
Kirkpatrick, D (1976) *Technique for evaluating training programs* Baltimore ASTD Press
Lindsley, O R (1964) Direct measurement and prosthesis of retarded behavior *Journal of Education, 147*, 62-81
Lindsley, O R (1994) Performance is easy to monitor and hard to measure In R Kaufman, S Thiagarajan, and P MacGillis (Eds), *Handbook of Human Performance Systems* San Diego University Associates
MacCorquodale, K (1970) On Chomsky's review of Skinner's Verbal Behavior *Journal of the Experimental Analysis of Behavior, 13*, 83-99
Mager, R F (1988) *Making instruction work* Belmont, CA Lake
Malott, R W , Whaley, D & Malott, M (1993) *Elementary principles of behavior* Englewood Cliffs, N J Prentice Hall
Markle, S M (1964) *Good frames and bad* New York John Wiley and Sons
Rosenberg, M J , Coscarelli, W C , and Hutchison, C S (1992) The origins and evolution of the field In H D Stolovitch and E J Keeps (Eds) *Handbook of Human Performance Technology* San Francisco Jossey-Bass, 14-31
RWD Technologies (1994) Front End Analysis Performance and training analysis software (1994) Columbia, MD RWD Technologies
Sidman, M (1960) *Tactics of scientific research* New York Basic Books
Skinner, B F (1938) *The behavior of organisms An experimental analysis* New York Appleton-Century-Crofts
Skinner, B F (1986) What is wrong with daily life in the western world? *American Psychologist, 41*(5), 568-574
Skinner, B F (1933) The rate of establishment of a discrimination *Journal of General Psychology, 9,* 302-50
Skinner, B F (1953) *Science and human behavior* New York MacMillan
Skinner, B F (1968) *The technology of teaching* New York Appleton-Century-Crofts
Skinner, B F (1971) *Beyond freedom and dignity* New York Alfred A Knopf, Inc
Smith, M E and Geis, G L (1992) Planning an evaluation study In H D Stolovitch and E J Keeps (Eds), *Handbook of Human Performance Technology* San Francisco Jossey-Bass, 151-166
Stolovitch, H D , and Keeps, E J (Eds), (1992) *Handbook of Human Performance Technology* San Francisco Jossey-Bass

CARL BINDER is President of Precision Teaching and Management Systems, Inc , and Chairman of Product Knowledge Systems, Inc Areas of special interest include behavioral fluency and application of knowledge architectures to learning and reference systems *Mailing address* PT/MS, Inc , P O Box 95009, Nonantum, MA 02195 Email 73240 1134@Compuserve com