Keeping Your

To control errors in human-resource information,

Controlling errors in information—data contamination—is the critical link between generating data and applying them to meaningful human-resource-management decisions. Understanding the various sources of contamination is necessary for controlling them. The goal is to control the incidence of error in judgment, so that you can make predictions with perfect accuracy.

The three main contaminants

Realistically, total control and perhaps even understanding are unlikely. There are two kinds of error: random and systematic. The unpredictability of random error makes it quite difficult to explain, let alone control. But you can control systematic error to some degree, because you can have a reasonably good grasp of the factors that tend to be responsible for it—inconsistency, instability, and subjectivity.

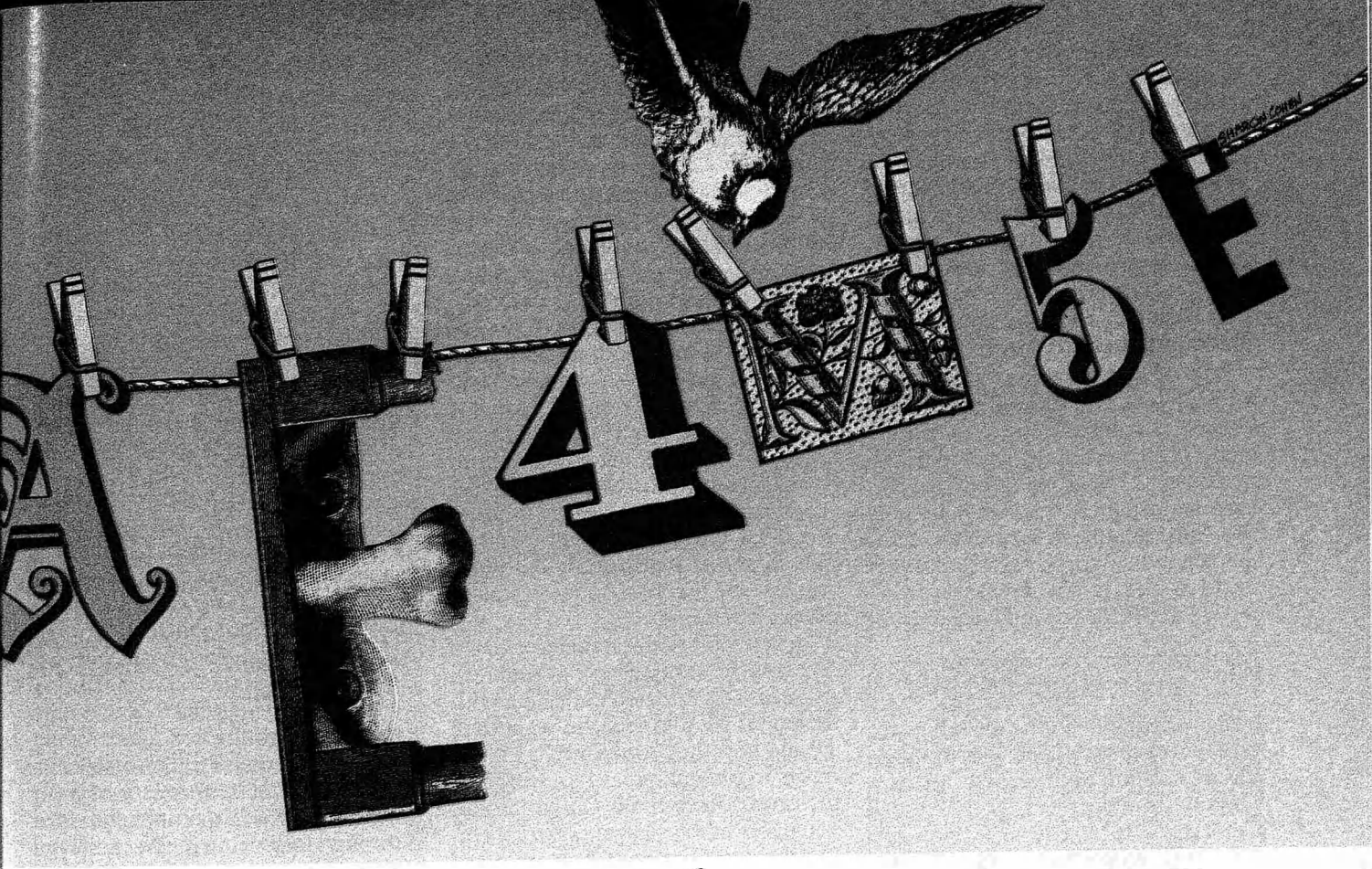
Briefly, inconsistency refers to the manner, method, or approach you employ when making evaluation decisions. It includes such factors as job relevance, standards of performance, standardization procedures and policy, and the data-gathering process itself.

Instability refers to the effects of time on rating or judgment. Those effects include not only the literal movement from one period in time to another (six months, two days, or one hour) but also the events that occur during that movement.

Last, and perhaps most recognizable in judgment-making, is the error caused by subjectivity. Unlike instability and inconsistency, subjectivity is the source of error with which most managers can identify, understand, and even explain. It emanates from a basic principle of human behavior: everything is in the eyes of the beholder. With many different beholders with many different perspectives, the pres-

Cohen is senior vice-president of product development at Wilson Learning Corporation, 7500 Flying Cloud Drive, Eden Prairie, MN 55344. *This article is the second in a three-part series on managing human-resource data; Part I addressed the development of useful and accurate data bases.*

The author wishes to acknowledge the influence of his many years at Assessment Designs, International (a Wilson Learning division where he was a founder), and thank its staff for its invaluable contribution to the foundation for both the thoughts and the spirit of this article.



Data Clean

By Stephen L. Cohen

you first have to understand what causes them.

ence of subjectivity is probably a given, unless you can control it. The various forms of subjectivity are well researched, but their underlying sources are relatively ill-defined, given the attention paid to them.

Subjectivity accounts for a considerable amount of systematic error—that is, the error is not all random. Many people, for example, make what are called constant errors—they display certain tendencies or predispositions that lead to predictable errors in judgment. For example, the tendency to evaluate all people, regardless of their abilities or qualifications, as high, low, or average is a constant error. That common tendency defies the bell-shaped, normal distribution of performance that more accurately describes the actual world. Likewise, the tendency to generalize from one characteristic of an individual to all other characteristics can be a constant error, as is the tendency to contrast people with oneself.

But the most dangerous and pervasive error of subjectivity is a person's inability to differentiate observations from inferences—the tendency to allow subjective perceptions to interfere with fact, to jump to conclusions. For example, a man is running into an office building. What is the first thing you think? He's late for an appointment; someone is chasing him; or he's got to go to the washroom. Often, whatever you think becomes the observation when, in fact, it is merely an inference taken from the actual observation— running into the office building. Nothing more, nothing less. To assume cause is common in making observations, but obviously detrimental to accurately evaluating people. The problem is at least threefold:

- you have to observe accurately;
- you have to observe long enough;
- you have to observe frequently enough.

Only then can you begin to draw the appropriate conclusions.

Now, think about the potential damage an unsubstantiated inference might create when you make selection, promotion, or career-planning decisions. Remember the articulate, well groomed sales rep you hired, only to discover that he lacked critical analytical, planning, and follow-up skills? Or that top-notch engineer you promoted to project manager, only to find out she didn't have the necessary leadership skills to manage a group effectively? Why did you make those decisions in the first place? Probably

because you inferred from one set of characteristics that your candidates would possess another set.

The issue of subjectivity goes beyond inference, and is related to the types of rating errors people often make in evaluating others. There are several ways to control the influence of those errors in the decision-making process.

Errors in judgment

The three sources of error—inconsistency, instability, and subjectivity—contribute to errors in judgment. If the goal of management is to make good decisions about people, then management's penultimate goal should be to remove the contaminants in the judgment process. You can help decision-makers to recognize and then reduce the incidence and magnitude of inaccurate people-related decisions. But nothing is perfect—when dealing with human behavior, you must assume less-than-perfect predictions.

In many cases, factors outside the control of the manager (and perhaps of the employee) interfere with the perfection of the system. Motivational variations, personal and home issues, introspection and insight, and variations in maturity and growth, not to mention economic, technological, and political changes, all can contribute to a less-than-perfect formula. The role of managers, then, should be to make the most of what they have available to them—namely, what they can actually see of any employee's on-the-job performance. Better observation, documentation, and evaluation skills will improve the data base and subsequent people decisions.

Most people commit rating errors in one way or another and to different degrees. The first step in dealing with the errors is to understand what they are and how they contribute to various errors in judgment, and to recognize the ones you commit. Only then can you discover the ways to control them. See the box for a brief description of the common rating errors; figure 1 shows how those errors relate to the three sources of judgment error.

Another dimension important to controlling the sources of error are the skills of the decision-maker in making judgments about people. The skills are

- the accurate, empirical **observation** of performance;
- the **documentation** (written or mental records) of those observations so that they are available on demand;
- the **evaluation** of those recorded observations in as objective a manner as possible, which is critical to effectively making sense out of the available information.

Figure 2 shows the relationship between those error-control skills and the typical rating errors. Applying those skills reduces the likelihood of rating errors, which in turn controls the judgment errors related to making people decisions.

Figure 3 combines the first two figures to demonstrate a three-dimensional relationship.

To control errors requires more than simply demonstrating their relationship to their sources and type. But managers can use some methods and techniques to help control the contamination of data.

Common Rating Errors

Accuracy of recall: the tendency to forget the relevant details of a situation based on time passage and intervening events.

Halo effect: the tendency to either positively or negatively generalize from one characteristic of an individual to other characteristics of that same individual.

Contrast effect: the tendency to compare people to each other rather than to some pre-established standard.

Stereotyping: the tendency to classify or evaluate employees in a certain way because of apparent membership in a particular category of people (such as religion or sex).

Differences in standards: the tendency to evaluate consistently all people on the basis of a pre-defined set of experiences and expectations.

Fixed impression: the tendency to allow a momentary observation to affect indefinitely an overall evaluation of someone.

The effects of time: the tendency to allow the time proximity of observation to influence an evaluation.

Projection: the tendency to allow one's own characteristics or values to influence one's ratings.

Inference: the tendency to confuse true observations with internal thoughts about those observations, so that the observation of behavior is assumed rather than actually having been seen or heard.

Controlling data contamination

Total control of the sources of error in judgment, both systematic and random, is nearly impossible. We try to account for the true causes of performance, knowing full well that the dynamics of life provide a continually moving target that is difficult to hit. Even in the physical sciences, total truth eludes us. While the course of nature may be relatively predictable, the instruments used to assess those conditions may not always be operating at full capacity and hence create errors. Imagine then the difficulty of reliably controlling the instrumentation in evaluating others, when part of the accuracy of the instrument includes those doing the rating. Nonetheless, managers can use certain techniques to minimize judgment error and enhance the chance for reliable evaluations.

Enhancing Stability

Because the major causes of instability are the effects of the passage of time, the best way to control for that is through improving the frequency of observation over the time period. By doing so, you can at least understand and manage the factors that would contribute to instability, such as maturation and history. Understanding the effects of the situation at different times will shed light on the actual causes and account for truer variance around those events. Similarly, allowing for not only frequent but also qualitatively different observations under varying conditions will enhance the reliability of the ultimate decision-making process.

That alone is not enough. The passage of time brings with it the likelihood that you will forget what has occurred, so you also need a system for recalling events. Recording or documenting the observation in a meaningful and relatively unobtrusive manner will capture the essence of the events at the time they occur.

Ensuring Consistency

Inconsistency refers to a variance in approach or methodology of evaluation. It is usually associated with the criteria against which you compare a person and the instrumentation you use in the measurement process. The greater the variance in both criteria and measurement devices, the greater the likelihood of inconsistency in the way

you make decisions. You can achieve consistency, then, by improving definitions of performance criteria, and creating more effective standardization and more valid instrumentation.

■ **Define performance criteria.** Understanding the duties and responsibilities of a job and the knowledge, skills, and abilities required to carry them out effectively is critical if you want to achieve consistency in the evaluation process. Such information gives observations the proper focus. It signals to the decision-maker just what to look for and how important it is to overall job performance.

Once available, the information allows for relevant and meaningful comparisons between candidates for the same positions. Furthermore, it focuses attention on the most important performance indicators, allowing a better job-to-person match (though enough can't be said about the value of simply understanding the specific job parameters). Indeed, you could argue that the foundation for a good personnel-evaluation system is a thorough job analysis. Without such an analysis, observations may be off-target, documentation erroneous, and evaluation irrelevant. A complete analysis of the job tasks, responsibilities, and skills required for effective performance is necessary.

■ **Standardize the evaluation process.** Similar in concept to job relevance is standardization. But the issue of relevance concerns a consistent definition of job criteria across employees, while standardization refers more to consistency across the process by which you evaluate employees. That includes the administration of not only the evaluations but also the people making the evaluations.

Some examples of standardization in evaluation include adhering to similar policies and procedures (that is, using a systematic approach); using the same evaluation forms or formats; applying the same principles or decision points across all people; providing similar training for people responsible for evaluations; and using the same criteria for performance ranges (including forced distributions, if appropriate). In brief, the process of standardization helps to ensure a consistent and fair approach to evaluations, one that is relatively the same for similar positions. It controls the extraneous error-variance of different evaluators, who may use different standards and procedures to evaluate different people. The absence of standardization results in inconsistent treatment of people and invites personal bias to intrude in the evaluation process.

■ **Validate instrumentation.** Another way to ensure consistency in evaluation is to develop valid and reliable measurement instruments. Such instruments range from paper-and-pencil tests, to rating scales for evaluating on-the-job performance, to simulations of real-world environments.

Underlying the use of the measurement device is the assumption that it does not contribute per se to uncontrollable or unsystematic variance in the evaluation process. But, an instrument that doesn't measure what it purports to measure will yield results that are inconsistent with what you intend. Often, however, instruments do ignore relevant variables or facilitate individual bias by design, consequently

Figure 1—Relationship of rating errors to sources of variance

	Source of Error Variance		
Type of Rating Error	Instability	Inconsistency	Subjectivity
Accuracy of recall	X		
Halo effect	X	X	X
Contrast effect		X	
Stereotyping		X	X
Difference in standards		X	X
Fixed impression	X		
Time effects	X		X
Projection			X
Inference			X

Figure 2—Relationship between rating errors and skills

	Error Control Skills		
Type of Rating Error	Observing	Documenting	Evaluating
Accuracy of recall		X	
Halo effect	X	X	X
Contrast effect	X	X	X
Stereotyping	X		X
Difference in standards	X		X
Fixed impression	X	X	
Time effects	X	X	
Projection	X		
Inference	X		

providing inconsistent evaluations—they permit individual evaluators, evaluatees, and situations to influence their use. For instance, performance-appraisal forms that leave the definition of listed factors open to interpretation invite bias. Selection tests for the same job that vary in length and content invite inconsistency in results and, therefore, usage.

Even an instrument's calibration can affect accuracy. Clearly defined scaling with meaningful intervals, anchored by readily discernible differences, is more likely to yield truer results. Research exploring different types of performance-appraisal rating scales has demonstrated that you can control certain constant rating errors by using different types of scales. The more you define calibrations behaviorally, for example, the less likelihood of the halo effect (see the sidebar) in your results.

Similarly, research concerning selection has demonstrated that the more relevant and performance-based the evaluation instrument, the greater the likelihood of predicting on-the-job performance. That explains why simulation-related assessments often are better predictors of performance than traditional paper-and-pencil, multiple-choice tests, particularly when the predicted performance involves skill and process abilities rather than job knowledge or technical information.

The less ambiguity in test items, the less room for error in interpretation by both test taker and test interpreter. To the extent that the instrumentation used to evaluate different people on the same dimensions is uniform, greater consistency and less error-variance will result. Likewise, if an instrument is consistent within itself (that is, if it doesn't fluctuate in its ability to measure reliably what it is supposed to measure), unexplained errors in evaluation are less likely to occur.

Encouraging Objectivity

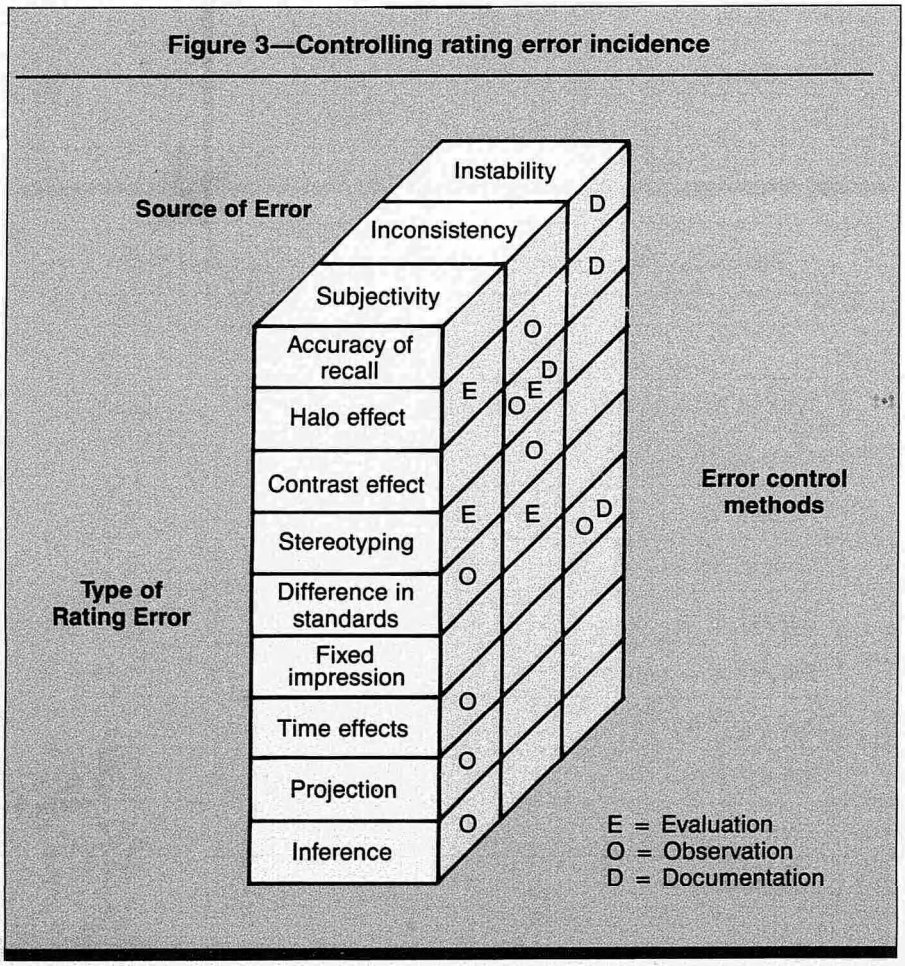
Of the three general sources of error variance, the easiest one with which to identify is subjectivity. Indeed, you probably have been on the other side of the evaluation process, and perhaps have experienced the unfairness of pure subjectivity—that should allow you better to understand its underlying causes.

To simplify the causes of subjectivity would be an oversight, because they require a relatively complex understanding of the principles of human behavior. But you can confidently say that evaluation errors of subjectivity arise from a complex interaction of past experiences, individual thought-processes, unique preferences, personality styles, and situational influences. And you also must recognize that the vast differences among people yield both constant evaluation errors from one person to the next and varied errors specific to individuals or groups. Several factors affect the likelihood of those errors intruding into the evaluation process. Some are familiar to most people, and some are easier to fall into than others.

The great complexity of human behavior as it relates to evaluation seems to be a barrier to objectivity in the rating process. Many techniques can help enhance objectivity. Simple awareness of your own susceptibility to certain rating errors can be important. But awareness coupled with specific procedures and formalized training appears to add a larger dimension of objectivity to the admittedly subjective evaluation process.

For example, simply separating observations from inferences will help to reduce stereotyping. Classifying and categorizing observations by distinct performance factors will significantly diminish the likelihood of the halo effect. And documenting or formally recording those observations can improve the accuracy of your recall. Add rating training to those techniques, and you can effectively reduce subjectivity and increase accuracy in your evaluations of job performance.

Figure 3—Controlling rating error incidence



Conclusion

You can reduce rating errors in many ways. There are well-established methods for ensuring consistency, enhancing stability, and encouraging objectivity. The thread that ties those factors together is the data base from which you draw them. In order to mitigate judgment error, managers must learn how first to generate the type of objective information that lends itself to more accurate evaluations. Human-resource management systems are readily available to help collect those data; you will see them next month in Part III of this series.