Technical Report 941

# Setting Performance Standards: A Review of Related Literatures and Identification of Future Research Needs

**Philip Bobko and Adrienne Colella**
Rutgers University

December 1991




**United States Army Research Institute for the Behavioral and Social Sciences**



# U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

**A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel**

**EDGAR M. JOHNSON**
**Technical Director**

**MICHAEL D. SHALER**
**COL, AR**
**Commanding**

Research accomplished under contract
for the Department of the Army

Rutgers University

Technical review by

Michael G. Rumsey
Clinton B. Walker



**NOTICES**

**DISTRIBUTION**: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

**FINAL DISPOSITION**: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

**NOTE**: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

| | |
|---|---|
| . REPORT SECURITY CLASSIFICATION<br>nclassified | 1b. RESTRICTIVE MARKINGS<br>-- |
| . SECURITY CLASSIFICATION AUTHORITY<br>- | 3. DISTRIBUTION / AVAILABILITY OF REPORT<br>Approved for public release; distribution is unlimited. |
| . DECLASSIFICATION / DOWNGRADING SCHEDULE<br>- | |
| PERFORMING ORGANIZATION REPORT NUMBER(S)<br>- | 5. MONITORING ORGANIZATION REPORT NUMBER(S)<br>ARI Technical Report 941 |
| . NAME OF PERFORMING ORGANIZATION<br>utgers University | 6b. OFFICE SYMBOL (If applicable)<br>-- |
| 7a. NAME OF MONITORING ORGANIZATION<br>U.S. Army Research Institute | |
| . ADDRESS (City, State, and ZIP Code)<br>chool of Business<br>epartment of Management<br>anice H. Levin Bldg., New Brunswick, NJ 08903 | 7b. ADDRESS (City, State, and ZIP Code)<br>5001 Eisenhower Avenue<br>Alexandria, VA 22333-5600 |
| a. NAME OF FUNDING / SPONSORING ORGANIZATION<br>U.S. Army Research Institute for the Behavioral and Social Sciences | 8b. OFFICE SYMBOL (If applicable)<br>PERI-RS |
| 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER<br>DAAL03-86-D-0001,<br>Delivery Order No. 2331-02 | |
| c. ADDRESS (City, State, and ZIP Code)<br>5001 Eisenhower Avenue<br>Alexandria, VA 22333-5600 | 10. SOURCE OF FUNDING NUMBERS |

| PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO. |
|---|---|---|---|
| 162785A | 791 | 2211 | C7 |

1. TITLE (Include Security Classification)
Setting Performance Standards: A Review of Related Literatures and Identification of Future Research Needs

2. PERSONAL AUTHOR(S)
Bobko, Philip; and Colella, Adrienne

| 3a. TYPE OF REPORT | 13b. TIME COVERED | 14. DATE OF REPORT (Year, Month, Day) | 15. PAGE COUNT |
|---|---|---|---|
| Final | FROM 91/01 TO 91/11 | 1991, December | |

6. SUPPLEMENTARY NOTATION
Contracting Officer's Representative, Michael G. Rumsey

| 7. COSATI CODES | | | 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | Standard setting, Performance standards, Minimum standards, Criteria, |
| | | | |
| | | | |

9. ABSTRACT (Continue on reverse if necessary and identify by block number)

Because performance standards are used for evaluation by many organizations, this paper considers aspects of the standard-setting process that influence affective and behavioral reactions of job incumbents. Literatures used to specify future research needs and directions in the use of performance standards examined educational standard setting, goal setting, feedback and framing, performance appraisal, utility analysis, and job satisfaction. This report outlines research suggestions (and specific hypotheses when possible) concerning the ease of internalization of external performance standards, the definitional content of standards, the difficulty levels of performance standards, and the communication of standards. Implications within the organizational context of the U.S. Army are also considered.

| | |
|---|---|
| 0. DISTRIBUTION / AVAILABILITY OF ABSTRACT<br>☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS | 21. ABSTRACT SECURITY CLASSIFICATION<br>Unclassified |
| 2a. NAME OF RESPONSIBLE INDIVIDUAL<br>Michael G. Rumsey | 22b. TELEPHONE (Include Area Code)<br>(703) 274-8275 / 22c. OFFICE SYMBOL<br>PERI-RS |

D Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

**Technical Report 941**

# Setting Performance Standards: A Review of Related Literatures and Identification of Future Research Needs

**Philip Bobko and Adrienne Colella**
Rutgers University

**Selection and Classification Technical Area**
**Michael G. Rumsey, Chief**

**Manpower and Personnel Research Laboratory**
**Zita M. Simutis, Director**

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

**December 1991**

**Army Project Number**
**2Q162785A791**

**Manpower, Personnel, and Training**

## FOREWORD

In 1980, the Assistant Secretary of Defense directed all services to pursue a long-range systematic program to validate the Armed Services Vocational Aptitude Battery (ASVAB) and to reevaluate enlistment standards and on-the-job performance. As a result, the Army has sponsored two major research efforts. The first effort, known as Project A, investigated the validity of the ASVAB and several new predictor measures for a sample of diverse military occupational specialties (MOS). The second effort, the Army's Synthetic Validity Project (SYNVAL), considered both synthetic validation techniques for determining MOS specific selection composites and evaluation of methods for setting standards on these composites.

In theory, adoption of selection standards should be associated with assessment of performance standards. This report provides a preliminary theoretical framework for analyzing the role of performance standards and their relationship to the affective and behavioral reactions of job incumbents. This framework will provide guidance in generating hypotheses about performance standards and help guide future research in this domain. Further development of the ideas in this report will lead to increased understanding of performance standards in the scientific community and to production of more accurate and well-defined procedures for standard setting in the U.S. Army.

EDGAR M. JOHNSON
Technical Director

ACKNOWLEDGMENTS

This work was supported by the U.S. Army Research Institute for Behavioral and Social Sciences (Michael G. Rumsey, POC) under the auspices of the U.S. Army Research Office Scientific Services Program administered by Battelle.

The authors wish to thank Jane Arabian, Angelo DeNisi, Bea Farr, Avi Kluger, George Lawton, Cynthia Lee, Dave McGrath, Michael Rumsey, and Clint Walker for their helpful comments on earlier versions of this work. We also wish to thank Kitty Collins for her cheerful disposition and efficiency in managing our contract.

SETTING PERFORMANCE STANDARDS: A REVIEW OF RELATED LITERATURES AND IDENTIFICATION OF FUTURE RESEARCH NEEDS

## EXECUTIVE SUMMARY


### Requirement:

This report provides basic information critical for determining the directions for future research and the selection and evaluation of standard-setting methodologies for the U.S. Army.

### Procedure:

While the literature directly addressing performance standards is scant, there is a variety of related scientific literature that is useful for thinking about setting performance standards and job incumbents' reactions to those standards. For this project, the authors conducted a literature review of these domains, including literatures in education, goal setting, performance appraisal and feedback, utility analysis, and job satisfaction. During this literature review, they developed an initial model of aspects of the standard setting process that influence affective and behavioral reactions of job incumbents. This model attempted to compare and synthesize findings from the related literatures listed above. The authors also identified research needs for basic researchers in the field and considered how these needs might be stated within U.S. Army contexts.

### Findings:

From the model and literature review, the authors identified a variety of research suggestions and presented a set of thirteen global hypotheses that require basic empirical research in the future. They categorized all research needs into four groupings: the definitional content of performance standards, the difficulty levels of standards, the ease of internalization of external performance standards, and the communication of standards. While the specific hypotheses are too numerous to mention here, the basic tenet of the model was that external performance standards may be internalized by job incumbents in many ways and that the process needs to be managed with care so that incumbent motivation and affective reactions are as intended. The authors conclude that standard setting procedures should incorporate both absolute levels and positive changes in performance, that continuous utility functions can help identify multiple levels of standards, and that a consistency between organizational and

individual values will increase the likelihood that external standards are internalized.

Utilization of Findings:

This paper presents a series of research needs required for an understanding of how individuals react to setting organizational performance standards. Results of such basic research should allow U.S. Army decision makers to better manage setting and communication of performance standards so that they are internalized by job incumbents, increase incumbent satisfaction, and increase incumbent motivation and performance.

SETTING PERFORMANCE STANDARDS: A REVIEW OF RELATED LITERATURES AND IDENTIFICATION OF FUTURE RESEARCH NEEDS

CONTENTS

Page

## LIST OF FIGURES

# SETTING PERFORMANCE STANDARDS: A REVIEW OF RELATED LITERATURES AND IDENTIFICATION OF FUTURE RESEARCH NEEDS

## INTRODUCTION

The intent of this review is to motivate research in a crucial, but neglected area: how individuals react to the setting of performance standards. The notion of "standard" (e.g., as in "selection standard" or "performance standard") is ubiquitous in its use within both private and public sector organizations. However, research on standard setting has almost exclusively focused on selection standards (e.g., the educational literature noted below). Indeed, Cascio, Alexander, and Barrett (1988) provide an excellent review of selection standards and the setting of cut-off scores within a legal context. On the other hand, the textbook approach to human resources management makes it clear that a focus on performance (e.g., a good job analysis) is necessary before any informed decisions can be made about selection, training, or performance appraisal systems. For example, requirements about performance standards should directly inform optimal selection standards. However, there is little (if any) research directly related to how organizations set performance standards, whether the purposes of the standard setting are met, or whether inappropriate performance standards can lead to dysfunctional outcomes (Murphy & Cleveland, 1991). While the literature here is scant, it is assumed that related literatures (e.g., *performance appraisal*, *goal setting*, utility analysis, job satisfaction) can help delineate future research needs in our understanding of the purposes and procedures for setting performance standards.

### What is a Performance Standard?

Dictionaries present many definitions of the word "standard." For example, the Webster's definition most directly related to the setting of performance criteria is

> 3: something established by authority, custom, or general consent as a model or example: CRITERION (Webster's Seventh New Collegiate Dictionary, 1965, p. 853)

It is clear from this definition, and common organizational usage, that there are three features that define standards. First, standards imply evaluation, as noted by the definitional use of the term "criterion." Second, standards are criteria that are established externally and imposed on an individual's work task. In this sense, they do not necessarily specify individual or organizational intentions of desired performance levels. As noted later, "standards" are therefore to be distinguished from "goals" in the sense that goals are performance benchmarks that are intended to be adopted internally by the

individual and function to direct behavior. Finally, standards, as established entities, are usually considered to remain somewhat stable over time and individuals, therefore further distinguishing them from goals or objectives.

Some textbooks equate performance standards with "minimum standards" of performance (cf. Carroll & Schneier, 1982, p. 131). However, this view is myopic since standards can be set for a wide range of performance levels. For example, a cursory review of a local newspaper for one week revealed that performance standards could also be used to define "standards for excellence" (in two articles about standards for excellence in athletes and maintenance of high standards in a public school system). Thus, our perspective is one with a performance continuum (from very poor performance to outstanding performance), where standards can be used to define any point along that continuum. For example, as illustrated by Bobko and Wise (1987), one could define a continuum of performance standards as

(a) a "minimum standard" -- below which a person was demoted or fired. Presumably, such a person would have negative utility to the organization.
(b) a "training warning standard" -- below which some action was required for remediation (e.g., retraining).
(c) an "acceptable standard" -- requiring no action on the part of the organization.
(d) an "excellence standard" -- above which a positive reward would be provided.

We also note that performance standards are value judgments. That is, these established standards are externally set, defined, and operationalized by supervisors, or management, or unions, or peer advisory groups, etc. Wherever they come from, there is no "ultimate criterion" against which to judge what the "right" standard is. The educational literature most clearly points out the subjective nature of standards and the judgments underlying them (Glass, 1978). As Cronbach (1949) states when discussing the setting of any cut-off scores, "The choice ... cannot be made scientifically. It is based on personal, social, and economic values, combined with practical considerations." (p.424). In fact, this subjectivity is what leads us to conclude that research is needed to systematize our knowledge about how performance standards are set, how they should be set, their impact on organizational performance, and their impact on affective reactions to the workplace.

Finally, it is useful to draw a distinction between job performance standards and the selection standards mentioned earlier. Selection standards refer to cut-off scores on some selection criteria (paper and pencil tests, interviews, etc.). The purpose of these standards is to select individuals who are expected to perform well on the job. Our use of the term "performance standard" refers to standards concerning any aspect of an individual's performance once he or she has entered the organization. Thus, performance standards can be applied to

numerous domains of performance (e.g., promptness, technical knowledge, and specific task competency). These performance standards can serve many purposes relating to personnel decisions such as training, firing, and promotion.

## A Framework for Thinking About Performance Standards

In line with the above definition of standards, Carroll and Schneier (1982) state that performance standards are "criteria against which to evaluate performance." Most research on standards (primarily in the educational literature) has focused on methods of setting standards so that they accurately describe specified performance levels. Also, as noted above, this research has mainly demonstrated that despite careful methodology, resulting standards are value judgments and that different techniques will lead to different standards reflecting the same performance level. A recent example of this is the U.S. Army's synthetic validity (SYNVAL) project (see Wise, Peterson, Hoffman, Campbell, & Arabian, 1991) where two different standard setting procedures were used: a task-based approach and a behavioral incident approach. The task-based approach led to a "marginal" standard in which 29% of current soldiers failed to meet the cut-off, whereas the behavioral incident approach led to a "marginal" standard that only 6% of current soldiers failed to meet. Thus, given a multiplicity of methodologically sound procedures for developing standards, the question still remains as to how an organization should choose a standard-setting procedure.

A major issue that may help clarify the above question but has previously been ignored in the standard setting literature is how individuals react to standards. Generally, organizational concern has been with the evaluative potential of performance standards, e.g., "does a marginal standard really reflect marginal performance?" However, when performance standards are communicated and applied, those to whom the standards apply will most likely have some reaction to the standard, especially since standards are used to make personnel decisions. For example, considering the above SYNVAL results, soldiers would be likely to have different reactions to a standard in which only 70% of them would be at least "marginal," compared to the other system in which 94% would be at least "marginal."

The purpose of this paper is, therefore, to encourage thinking about how performance standards may influence reactions of those to whom the standards apply. Figure 1 is a schematic representation of the main points to be made in this paper. We argue that different aspects of setting performance standards, such as the difficulty and communication of standards, will influence whether or not these performance standards are internalized. In other words, the way in which standards are set can impact on whether or not individuals accept, or at least attend to, organizational performance standards. Once external standards are internalized, then two types of reactions can occur. First, individuals may have an affective reaction toward the

organization, job, or the individuals applying the standards. For example, standards viewed as too difficult or as being set unfairly may lead people to be less satisfied with their job. Affective reactions would then be related to turnover and retention (Mobley, Griffeth, Hand, & Meglino, 1979). Second, individuals may have a motivational reaction to the standard. For example, standards viewed as too difficult may lead employees to just give up and be demotivated. On the other hand, it may be possible to set standards which will actually increase motivation. The motivating potential of standards should be related to employee performance within the environmental constraints which make this possible. Thus, this paper draws a link from the setting of performance standards to their impact on organizational and individual effectiveness by discussing how performance standards may influence employees' motivational and affective reactions.

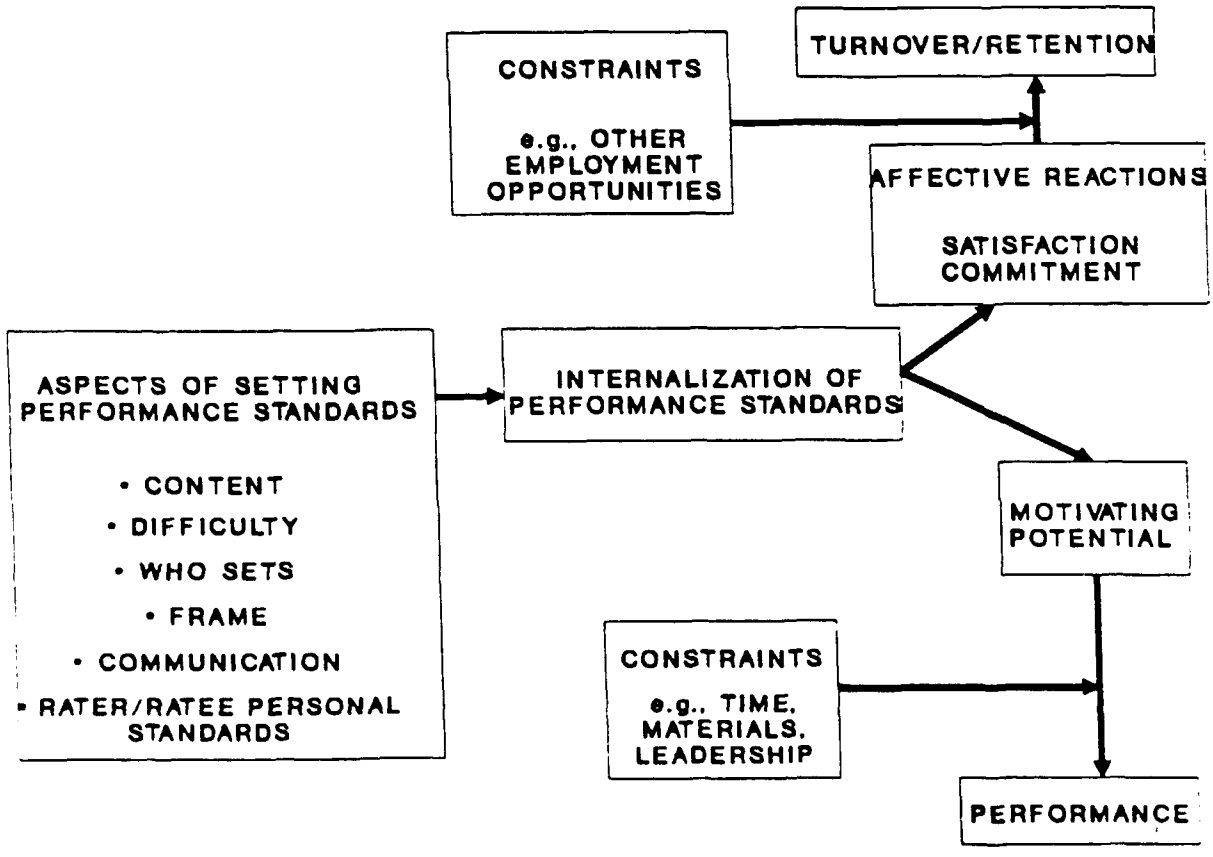


Figure 1. Underlying model of employee reactions to performance standards.

The issue of employee reactions to performance standards has gone essentially unresearched. Thus, we review other lines of research which have implications for the influence of standards on employee reactions to those standards. Specific literatures reviewed are: goal setting, performance appraisal, feedback, framing, utility analysis, and job satisfaction. The end result is a section which puts forth specific hypotheses (when possible) and outlines future research directions. These research directions are related to:

- the difficulty of the standards being set.
- the content of the standards being expressed: How specific should the standards be? Should performance feedback be given in positive or negative frames (half full vs. half empty)?
- the suitability of absolute standards versus standards based on percentage change.
- the decision about who has input into the standard setting procedures.
- the communication of standards.

Finally, we briefly present the implications of our hypotheses for the application of standard setting procedures within the Army context.

## A REVIEW OF RELATED LITERATURES

### Standard Setting in Education

While there is little research on performance standards in the organizational literature, educational researchers have suggested and studied procedures for setting standards -- usually in the arenas of pupil diagnosis, certification, or program evaluation (Shepard, 1980). Reviews of this literature may be found in Berk (1986), Hambleton and Eignor (1979), Jaeger (1976), Pulakos, Wise, Arabian, Heon, and Delaplane (1989), and Shepard (1980). The two major approaches to developing standards involve (i) judgments about test content or (ii) judgments about mastery vs. non-mastery groups. Note that both approaches are judgmental in nature.

Angoff (1971), Nedelsky (1954), and Ebel (1972) have all proposed standard setting methodologies which require judgments about item content. For example, the Angoff procedure asks judges to consider a minimally qualified individual and estimate the probability that such a person would answer each item correctly. The sum of such probabilities, across all items in a test, provides an estimate of where the standard (cut-off score) should be set. The Nedelsky method is similar, but asks judges to arrive at these probability estimates by first casting out options (on multiple choice items) which minimally competent individuals would clearly know are wrong. In the Ebel method, items are additionally stratified by relevance and difficulty (such categorizations are also made by the judges). Several reviews (Berk, 1986; Hambleton & Eignor, 1979) have concluded that the Angoff method

is favored (due to its simplicity and technical adequacy) over other item-based, or mastery-group based, procedures.

Zieky and Livingston (1977) offered two standard setting procedures based on judgments about groups of individuals, rather than judgments about items. In the "contrasting groups" method, judges identify individuals who are clearly either "masters" or "non-masters" and a cutoff score on the test in question is set such that the two groups are maximally discriminated. Zieky and Livingston's "borderline group method" is similar, but only operates on individuals identified as borderline. The cut-score is then set at the median of this group.

Studies which have empirically compared standard setting methods in education have found that different standard setting methods can yield substantially different results (Andrew & Hecht, 1976; Brennan & Lockwood, 1980; Halpin & Halpin, 1983, Koffler, 1980; Skakun & Kling, 1980). This is not surprising, given that standards are judgments of value, and given the variety of frames of reference in the standard setting methods, potential differences in the types of judges used, and differences in types of stimuli used (items, individuals who are masters, individuals who are borderline, etc.).

Other methods for setting standards have been proposed which involve the use of external criterion measures and a decision-theoretic evaluation. Shepard (1980) calls these methods "empirical methods for adjusting standards". For example, assume there is some dichotomous, external criterion (e.g., "successful" vs. "unsuccessful"). Then cut-points on a selection test can be evaluated from the number (and cost) of both false positives and false negatives when using the selection cutoff to predict group membership on the criterion. Of course the problem with this technique is apparent: it begs the fundamental question by assuming that there <u>is</u> a (dichotomous) standard of performance available. Thus, such methods are not useful for setting <u>performance</u> standards, where the very assumption is the central issue.

Pulakos et al. (1989) note two further application problems in using the educational literature for setting job <u>performance</u> standards. First, examinee-based methods (e.g., those of Zieky and Livingston) may be susceptible to range restriction when judges attempt to identify the full range of performance. That is, "unacceptable" performers may exist only in minimal numbers in the reality of some organizations (due to the effects of good training procedures or attrition of poor performers), while exceptionally high performing individuals may have been promoted out of the job. Second, Pulakos et al. (1989) note "item-based methods do not ... hold much promise for use in conjunction with continuously scored performance measures (e.g., rating scales)" (p.25). That is, estimating the probability that a "master" will get a <u>rating scale</u> item "correct" makes little sense. To solve this dilemma for behaviorally anchored rating scales (BARS), Pulakos et al. suggest that future researchers might ask judges to sort critical incidents into two categories ("below minimum competency" and "above minimum competency") and identify a scale value on the BARS which represents a minimum

performance standard.

Finally, Shepard (1980) identified two additional methodological issues when standards are being set. First, wherever a standard is set on a particular continuum, "individuals on either side of the standard will be virtually indistinguishable from one another" (p.448). This implies that each cut-point should have an associated standard error routinely computed as part of the standard setting process (Bobko & Wise, 1987). Second, Shepard (1980) distinguishes between two types of standards: an "individual diagnosis" standard which is focused on improvement, and an "individual certification" standard which "renders a final judgment of quality" (p. 449). This distinction is consistent with our above notion that standards can be used for either motivational or evaluative purposes.

## Goal Setting

Empirical findings in the goal setting literature are relevant to the extent that performance standards have motivational value. The basic premise underlying goal setting theory is that an individual's personal goals are the immediate (although, not only) regulator of his or her actions (Locke & Latham, 1990). Over the past 25 years, an abundance of research has been conducted in both laboratory and field settings which tests and develops the conditions, mediating mechanisms, and limits of goal setting theory. Numerous reviews and meta-analyses have been conducted on these empirical studies, the most recent and comprehensive being that of Locke and Latham (1990). In order to avoid repetition, this section of the paper simply points out those findings which appear to be most robust and consistent across studies. The major focus of this section is to draw the connection between performance standards and one's personal goals and then to illustrate how standards, through their influence on personal goals, can affect the motivation to perform and, consequently, performance.

Locke and Latham (1990) provide an in-depth discussion of the differences and similarities between terms such as intention, norm, purpose, level of aspiration, goal, and standard. Our current review is concerned with issues involved in setting external standards (i.e., the performance levels specified by the organization, to individuals, which are necessary to achieve a designated level of performance). On the other hand, goal setting theory primarily focuses on the personal goals held by an individual (i.e., the internal goal which an individual is striving to achieve).

There are two general findings in the goal setting literature which have been extensively studied and consistently supported (Locke & Latham, 1990). First, subjects who work towards difficult, specific goals perform better than those who work toward easy, vague, or no goals. Second, performance tends to increase linearly with goal difficulty. Goal setting has been proposed and found in many cases to have positive effects on performance through the following motivational

mechanisms: behavioral direction (e.g., Organ, 1977; Rothkopf & Billington, 1975), effort (e.g., Bandura & Cervone, 1983, 1986; Bryan & Locke, 1967a, 1967b; Campion & Lord, 1982), persistence (e.g., Bavelas & Lee, 1978), and task specific strategies (Earley & Perry, 1987; Terborg, 1976).

In order for external standards to have a positive influence on motivation through the goal setting process, externally defined standards must be translated by the individual into personal goals which are specific and difficult. This situation is similar to one in which subjects are assigned goals which then are either translated into personal goals or rejected.

The literature most relevant to the issue of the translation of assigned goals into personal goals is that on commitment to assigned goals. Both Locke and Latham (1990) and Hollenbeck and Klein (1987) have reviewed this literature and developed summaries of the conditions under which individuals will be committed to assigned goals. These are presented in Figure 2. It should be noted that many of these conditions have been empirically examined and supported (e.g., rewards [Latham, Mitchell, & Dossett, 1978; Pritchard and Curtis, 1973], peer influence [Earley & Kanfer, 1985; Matsui, Kakuyama, & Onglatco, 1987], and conflict [Erez, Earley, & Hulin, 1985; Latham, Erez, & Locke, 1988]) while others have not received much empirical attention (e.g., ego involvement and goal intensity).

One important dimension on which external standards and assigned goals differ is that some organizations use external standards to define minimally acceptable performance (although individuals may try to perform beyond the "standard"). On the other hand, assigned goals, as usually presented in goal setting theory, reflect the level of performance that external sources intend individuals to try to achieve. This distinction raises the question, "When will an individual come to interpret a performance standard as if it was a personal goal?" If some of the conditions in Figure 2 are present when a standard is set (e.g., the standard is set by a respected authority figure, it is tied to rewards and/or punishment, and it is public), will individuals adopt the standard as their own maximal performance goal to the extent that they will not be motivated to go beyond that particular level of performance? This issue deserves empirical investigation.

Thus, if the organization decides to use standards as a way of increasing performance (rather than insuring a minimal level of performance), then care must be taken that performance standards are sufficiently difficult and specific and that conditions exist which will lead to commitment. On the other hand, if the organization sets standards without considering their possible translation into individual performance goals, individuals may adopt minimal level standards as their personal goals, and consequently both motivation and subsequent performance may be lower than had more difficult personal goals been set.

* Goals are assigned by authority figures who are ...

  - seen as legitimate
  - physically present
  - supportive
  - trustworthy
  - provide a rationale for goals
  - exert reasonable pressure
  - are knowledgeable and likeable

* Assigned goals

  - imply rewards and punishments
  - convey positive self-efficacy information
  - foster a sense of achievement
  - imply opportunities for self-improvement
  - are challenging
  - are public
  - have high instrumentality and valence
  - are made for ego involving tasks
  - do not conflict with other goals
  - have a high expectancy for success
  - are intense

Figure 2. Factors increasing commitment to assigned goals.

Related to the goal setting literature is the management by objectives (MBO) literature, which concerns the application of goal setting theory on an organization-wide basis. In theory, MBO programs begin at the top of an organization, whereby top management sets organizational goals and objectives. These objectives are then filtered (cascaded) downward in a performance evaluation system where managers meet with their subordinates to develop individual objectives and plans. These objectives should follow from the goals and objectives specified at higher levels. Although the idea of MBO (Drucker, 1954) was postulated long before goal setting research, the positive effects of MBO can be mainly attributed to goal setting theory (cf. Tosi, Rizzo, & Carroll, 1990).

One feature of MBO systems relevant to the setting of performance standards, and not clarified in the empirical research on goal setting theory, is its focus on employee participation and commitment by top management. By definition, MBO requires employee participation in the setting of objectives (Griffin, 1990). Furthermore, top management's participation and commitment to the MBO program has been found to be a crucial determinant to the program's success in increasing motivation and performance (Carroll & Tosi, 1973; Rodgers & Hunter, 1986). This

implies that participation in standard setting procedures by both top management and those to whom the standards apply should increase the overall motivating potential of the resulting standards.

Performance Appraisal

Although the procedures, processes, and methods of performance appraisal have been the subject of much empirical and theoretical study, little of this work has centered on the role of external performance standards in the performance appraisal process. While there have been many recent reviews of the performance appraisal literature (DeNisi, Cafferty, & Meglino, 1984; Feldman, 1981; Ilgen & Feldman, 1983; Landy & Farr, 1980; Murphy & Cleveland, 1991; Smith, 1976; Wexley & Klimoski, 1984), only the Murphy and Cleveland (1991) review has examined performance standards closely. Indeed, these authors concede that the topic of performance standards has not been given sufficient attention in the performance appraisal literature.

There are two roles that external performance standards play in the performance appraisal process. The first is the manner in which external standards influence how raters rate people. The second concerns the impact of standards on feedback and how people react to feedback. This section of the review focuses on the first issue. The impact of standards on the feedback process is reviewed in the next section.

Murphy and Cleveland (1991) describe the performance appraisal/judgment process as one in which raters compare their perceptions of a ratee's behavior to their own, internal standards, resulting in an evaluation. Again, as in the goal setting literature, a distinction needs to be made between internal (on the part of raters) and external performance standards. External performance standards can influence ratings by influencing any of the following components of the performance appraisal process: 1) a rater's perception of ratee behavior; 2) the rater's internal performance standards; and 3) the evaluation process.

The existence of standards can influence perceptions of ratee behavior by focusing raters' attention on certain aspects of performance over others. Standards can define which behaviors are important, unacceptable, desirable, etc., and thus define which behaviors have evaluative implications. Research has shown that behaviors which have strong evaluative implications are likely to receive attention (Murphy, 1982; Wegner & Vallecher, 1977). Therefore, explicit performance standards may have a beneficial impact on the rating process by drawing raters' attention to those aspects of performance explicitly defined as important by the organization. On the other hand, explicit standards may lead raters to ignore aspects of performance, which although important to effectiveness, are not specified in standards (Murphy & Cleveland, 1991).

Standards may also influence raters' perceptions of performance by influencing how performance information is cognitively encoded, retained, and retrieved. Performance appraisal research has consistently pointed out that the manner in which raters cognitively categorize and encode information has an important impact on how that information is subsequently used in making performance judgments (Murphy & Cleveland, 1991). DeNisi and Feldman, along with their colleagues (DeNisi, Cafferty, & Meglino, 1984; DeNisi & Williams, 1988; Feldman, 1981; Ilgen & Feldman, 1983), have developed models of the performance appraisal process which borrow heavily upon the social cognition literature concerning person perception.

One main feature of this work has been to examine how people categorize information about others' performance (Borman, 1987; Nathan & Lord, 1983). Wyer and Srull (1986) have suggested that category accessibility is an important factor in determining how individuals categorize information. It may be that externally set standards help in defining the performance categories held by individuals and make certain categories more accessible through priming. However, little research has been done on the content of raters' performance categories (Murphy & Cleveland, 1991). For example, we really don't know what type of information is contained in one's "good worker" category or how such categories develop over time. Thus the impact of external performance standards on the content and development of raters' cognitive performance categories is in prime need of investigation.

The second way in which standards can influence performance judgments is by influencing raters' internal performance standards, used as a comparison for perceived behaviors. Clearly, this issue relates to how external standards influence raters' cognitive categories of performance. Murphy and Cleveland (1991) have also hypothesized that the degree to which external performance standards get translated into internal performance standards depends on the rater's identification with and commitment to the goals of the organization.

Finally, performance standards may affect ratings by influencing the evaluation process. This may occur because providing explicit performance standards can reduce the amount of bias in ratings due to extraneous factors (e.g., gender, handicap status). It has been argued that when raters are given specific performance related information (in this case performance standards), the amount of variance in their ratings due to bias or rating error will be reduced (Czajka & DeNisi, 1988; Terborg, 1977). Two laboratory studies which directly examined the impact of specific performance standards on raters' judgments had differing results. Huber, Neale, and Northcraft (1987) found that performance standards did not impact on managers' judgments of performance, compensation, training, and promotion. In this case, the performance level of employees (described on paper) had the most impact on ratings. On the other hand, Czajka and DeNisi (1988) found that providing subjects with specific performance standards reduced the bias in their performance ratings of emotionally disabled ratees (described

using videotape). Clearly this issue is in need of further research.

Feedback

However performance standards are used, an evaluation of performance (relative to the set standard) needs to be fed back to the individual worker and relevant decision makers. Research has demonstrated that, in general, providing people with feedback about their performance will have a positive effect on their future performance (Guzzo, Jette, & Katzell, 1985; Ilgen, Fisher, & Taylor, 1979; Kopelman, 1986).

However, before feedback can have an effect on performance, people must perceive, accept, and respond to feedback (Ilgen et al., 1979). External performance standards can have profound effects on any of these processes (Taylor, Fisher, & Ilgen, 1983). There is empirical research which has examined the attributes of feedback and their effects on employee reactions and behavior (Guzzo et al., 1985; Ilgen et al., 1979). However, little empirical work has directly examined how performance standards enter into the feedback process.

Taylor et al. (1983) have developed a model to explain how individuals react to feedback. Their model is based on control theory (Carver & Scheier, 1981; Powers, 1973) which states that reactions to feedback result from a comparison process in which perceived feedback is compared to internal performance standards. Based on this model, external performance standards *might* play two important roles in the feedback loop. The first (already mentioned in the previous section) deals with how individuals translate external performance standards into internal performance standards. The second issue is the effect that external performance standards have on the comparison process and resulting response.

Due to the amount of role conflict and ambiguity often reported by employees, especially new employees (French & Caplan, 1973; Katz & Kahn, 1978), the importance of setting standards which will be internalized is clear. The literature on organizational entry (Wanous, 1980; Wanous & Colella, 1989) emphasizes the importance of providing organizational newcomers with specific, honest, and clear cut information. Whether this information (including performance standards) becomes internalized depends to a large part upon the socialization process (Van Maanen & Schein, 1979) and the degree to which standards are valued by individuals (Taylor et al., 1983). Thus, three important factors to consider when setting standards are (i) the clarity of the standard, (ii) how adequately information about the standard is disseminated, and (iii) how important the standard is likely to be to employees.

Once external performance standards have been internalized, individuals must compare their feedback to these standards. There are two attributes of standards which may lead to frequent comparisons: 1)

when employees are unsure about their ability to meet standards (Taylor et al., 1983; Weiss, Ilgen, & Sharbaugh, 1982), and 2) when standards are valued (Ashford & Cummings, 1982; Taylor et al., 1983). Therefore, in order to increase the frequency with which employees compare feedback to performance standards, the internalized standards should be both difficult and rewarding. [Although beyond the scope of this review, C. Walker (personal communication, July 23, 1991) has noted that the concept "difficulty" can be defined along many dimensions. For example, difficulty can be increased by reducing time to task completion, extending cognitive ability requirements, increasing endurance requirements, or increasing the need for dependence on co-worker performance. Future research is needed to investigate whether these dimensions moderate the effects of goal difficulty.]

Once a comparison is made, individuals may react in several ways to feedback (see Taylor et al., 1983, for a review). Taylor et al. distinguish between cognitive, behavioral, and affective reactions, listed in Figure 3. Individual reactions are a function of many factors (e.g., peer influence, individual differences), but the focus here is on how external performance standards may influence reactions.

<u>Cognitive</u>
* Assessment of Feedback Accuracy
* Evaluation of Feedback Source Credibility
* Evaluation of System Fairness
* Adjustment of Expectancies for Standard Attainment
* Adjustment of Internal Behavioral Standards

<u>Behavioral</u>
* Changing Direction of Behavior
* Altering Effort
* Persisting versus Quitting
* Responding Against the Feedback System

<u>Affective</u>
* General Job Satisfaction
* Satisfaction with Appraisal System

Figure 3. Reactions to feedback

One attribute of external standards likely to influence feedback reactions is the level of difficulty of the performance standard. From a control theory perspective (and congruent with goal setting theory), increases in motivation (i.e., direction, intensity, or persistence of behavior) will likely result when feedback indicates that people have performed <u>below</u> standards, thus supporting the notion that standards should be difficult (Campion & Lord, 1982; Taylor et al., 1983). If

external standards are easily attained, and no reward is provided for exceeding standards, then people may actually lower their internal standards and motivation (Janz, 1982; Taylor et al., 1983). On the other hand, if standards remain too difficult, subjects may lower their expectancy of success to the point where they just give up (Carver & Scheier, 1981) or lower their own internal standards (Campion & Lord, 1982). Extremely high standards may also cause individuals to perceive the feedback system as unfair and not credible (Dornbusch & Scott, 1975). Further, since a primary influence on affective reactions to feedback is the sign of feedback (Ilgen et al., 1983; Taylor et al., 1983), unrealistically difficult standards, to the extent that they result in negative feedback, will lead to negative affect.

Another influential attribute of performance standards might be the specific content of the standards. Research on feedback has demonstrated that specific, descriptive feedback, compared to evaluative outcome feedback, results in more accurate evaluations of expectancy for success (Taylor et al., 1983), leads to perceptions of source credibility and system fairness (Dornbusch & Scott, 1975; Jacobs, Jacobs, Feldman, & Cavior, 1973), and increases performance by allowing people to make accurate attributions about past performance (Carver & Scheier, 1981). Standards which are clear, descriptive, and specific, and consequently allow for feedback along these dimensions, should produce more desirable responses.

Finally, the issue of who sets performance standards may be important in determining reactions to feedback. Typically, research on standard setting has focused on how judges or subject matter experts set standards. Allowing those to whom the standards will be applied to participate in the standard setting process may be beneficial for several reasons. First, employees will be more likely to perceive standards as fair (Taylor et al., 1983). Second, employees may be more committed to reaching the performance standards and more likely to internalize them (Erez & Kanfer, 1983). Third, reactions against the system may be less likely (Dornbusch & Scott, 1975).

The framing of performance information. The framing of performance information is also critical to the feedback process. Decision making research has demonstrated that individuals process information differently depending upon whether the information is framed in terms of positive outcomes or negative outcomes (e.g., are performance standards set in terms of "success rates" or "failure rates"?). For example, according to Kahneman and Tversky's (1979) "prospect theory", judges are risk-seeking when confronted with information in terms of potential loss, but risk-averse when faced with information in terms of potential gain. The impact of frame occurs even when the possible outcomes of the negatively and positively framed problems are mathematically identical. This has been illustrated by Tversky and Kahneman's (1981) "epidemic problem" where, in response to an imminent epidemic, subjects were found to select a risky option when the outcome was expressed in terms of lives lost, but a more certain option when the identical outcome was expressed in terms of lives saved.

Bazerman (1983) has suggested that the framing of information is one of the crucial factors which affect policymaker judgments. Indeed, Bobko, Shetzer, and Russell (1991) demonstrated that positive and negative frames clearly affected judgments about the perceived worth of work performance. Further, the existence of framing effects has been noted at the group level of behavior (Schurr, 1987). Thus, since standards are value judgments, policymaker judgments about levels for standards are predicted to be affected by whether or not the standards are stated in positive or negative ways. Further, even if there is agreement about the level of a standard, it may be that individuals receiving performance feedback in relation to existing standards will react differently depending upon how the feedback is framed (e.g., "you achieved 80% of the standard" versus "you failed to achieve 20% of the standard").

## Related Issues in Utility Analysis and Performance Effectiveness

The appropriate level of difficulty for performance standards can also be informed by a review of utility analysis -- which has recently focused on judgmental procedures for assessing the effectiveness (or overall worth) of particular levels of individual performance. There has been a resurgence of interest in determining the utility of selection instruments and organizational interventions (Bobko, Karren, & Kerkar, 1987; Boudreau, 1991; Vance & Colella, 1990b). Computations of utility are based upon decision theoretic equations developed by Cronbach and Gleser (1965). In computing utility, the standard deviation of performance worth, SDy, is a focal parameter. Knowledge about estimates of SDy are critical here because they usually involve the subjective estimation of the worth of particular work performance levels (cf. Bobko et al., 1987, for a review of SDy estimation).

For example, a number of studies (Burke & Frederick, 1984; Karren & Bobko, 1983; Mathieu & Tannenbaum, 1985; Weekly, Frank, O'Connor, & Peters, 1985) have reported <u>negative</u> estimates of worth (for individual performance at 15th percentiles) even though the judgment task had explicit demand characteristics to give positive values. Subsequently, Sadacca, Campbell, DiFazio, Schultz, and White (1990) explicitly gave subjects both positive and negative scale anchors in order to obtain better psychometric properties for judgments of performance worth. The point here is that negative and positive effectiveness ratings imply the possibility of a "zero-point" in the utility curve. Such a point would be a natural estimate for a "minimum" performance standard when standards are used as evaluative devices.

Further, Pritchard, Jones, Roth, Stuebing, and Ekeberg (1988) have extended the notion of negative performance utility estimates by explicitly mapping the entire range of performance effectiveness. That is, they generate "contingency curves" where the horizontal, X-axis is an increasing scale of performance on a criterion and the Y-axis is a

measure (in ratio scale) of the estimated effectiveness of that performance to the organization, ranging from -100 to +100. It should be noted that Pritchard et al.'s curves also cross the "zero-point" of effectiveness, so that some minimum standard might be identified. Additionally, the effectiveness curves they generate are not usually linear: some ranges of performance show large changes in effectiveness (large slope), while other performance ranges are associated with "flat" effectiveness curves. Thus, the contingency curves often look like step functions (see Figure 4, to be discussed later). It may be that the points at which the steps change slope are good candidates for different levels of performance standards.

It has also been suggested that different organizational environments can effect beliefs and values about the perceived effectiveness of work behaviors (e.g., utility values may change across hostile environments, benign environments, after the introduction of new technologies, after a union vote; see Bobko et al., 1987). Again, because they are value-driven, standards may also be affected by these environmental factors. Indeed, it is common folklore that organizations will "lower their standards" to meet labor market conditions (see Carlson, 1967, for an empirical verification).

The metric in which utility analyses are reported has also received some theoretical attention. Typically, utility is assessed in dollar terms, although recent reviews (Bobko et al. 1987; Vance & Colella, 1990b) have concluded that other metrics, specific to organizational needs, are more useful to decision makers. For example, both Eaton, Wing, and Mitchell (1985) and Sadacca et al. (1990) developed metrics in terms of hypothetical individuals (e.g., perhaps 5 superior people are worth the same as 8 average, 50th percentile, individuals). On the other hand, the work of Pritchard et al. (1988) finesses the metric problem by assessing utility in terms of an abstract effectiveness scale ranging from -100 to +100. It should be clear that performance utility (and, hence, how we conceive of performance standards) can be approached in many formats, such as an abstract value, a subjective rating along some criterion dimension, measured output (quantity of production, production value), or hypothetical individuals (i.e., some sort of prototypic individual, ranging from minimally qualified to outstanding).

## Job Satisfaction

There has been a plethora of research concerning employees' affective reactions to jobs. We focus on job satisfaction, because it is clearly the most researched affective reaction and generally correlates highly with organizational withdrawal behaviors. For example, meta-analyses have found that job satisfaction has a mean correlation of -.29 with absenteeism (Scott & Taylor, 1985) and -.23 with turnover (Carsten & Spector, 1987). Also, those aspects of standard setting most likely to influence job satisfaction should also have similar effects on other affective reactions such as

organizational commitment. Several models exist which attempt to explain job satisfaction (Hulin, Roznowski, & Hachiya, 1985; March & Simon, 1958; Smith, Kendall, & Hulin, 1969; Thibaut & Kelley, 1959). A review of this literature is clearly beyond the scope of this paper (see Hulin, 1991, for the most recent review). For our purposes, only those areas of the literature which appear relevant to the issue of setting performance standards will be reviewed.

Discrepancy theories have been a popular approach to explaining or predicting job satisfaction. In general these theories (e.g., Lawler, 1973; Locke, 1976; Lofquist & Dawis, 1969; Porter & Steers, 1973; and Smith et al., 1969) state that it is the discrepancy between one's current perception of job characteristics or experiences and some internal comparison standard which, in part, explains job satisfaction. The presumed dimensions of comparison vary across researchers. For example, Porter and Steers (1973) focused on the discrepancy between what one perceives to have and what one expects to get, while Locke (1976) focused on the discrepancy between what one perceives to have and what one wants to have. Empirical research has generally supported discrepancy theories of job satisfaction (Michalos, 1986; Wanous, Poland, Premack, & Davis, 1991).

Probably the most critical dimension of performance standards in influencing satisfaction is the individual's probability of successfully meeting or exceeding existing standards. Research has demonstrated that there is a strong relationship between successful task performance and job satisfaction (see Locke & Latham, 1990, for a review). Thus, in order to facilitate satisfaction, standards should be set at a difficulty level which makes success probable.

Another dimension of performance standards which can affect job satisfaction concerns the particular rewards which are linked to standard attainment. According to the Hulin et al. (1985) model of job satisfaction, the outcomes (e.g., salary, fringe benefits, status, working conditions, intrinsic outcomes) received from meeting a standard should be favorably compared to the inputs (e.g., skills, time, effort, training) involved in performing to meet the standard. The specification of particular outcomes is beyond the scope of this paper, and furthermore, this is an issue which is heavily influenced by specific organizational goals. However, one important intrinsic reward, a sense of achievement, needs to be mentioned since it directly bears on the issue of standard difficulty. In the preceding paragraph it was postulated that standards should be easy enough so that success is likely. However, research has shown that meeting or exceeding difficult and challenging standards will lead to a sense of achievement not found when meeting easy standards (Locke & Latham, 1990). Thus, standard difficulty should be set at a level which, when met, will be intrinsically rewarding, as well.

A third, related aspect of standards which can influence satisfaction is the degree to which success in meeting the standard can be attributed to one's personal ability. That is, satisfaction with

performance has been found to be greater when individuals can attribute success to personal qualifications or abilities rather than to luck, chance, or the easiness of the task (Weiner, 1986). Thus, standards which reflect aspects of job performance which are under the control of the individual (e.g., the number of calls made in the case of a phone salesperson) should have more positive effects on satisfaction than those aspects of performance which are influenced by extraneous factors (e.g., the number of actual sales).

Fourth, the perceived rationale for a given standard can influence satisfaction. Phillips and Freedman (1988) found that students who were given a reasonable explanation for an assigned goal experienced more satisfaction than those who were given no reason. Perhaps participation in the standard setting process is one mechanism through which employees can come to understand the rationale behind given performance standards.

Finally, it has been demonstrated that people consider their past performance when judging their current performance (Simon, 1988; Vance & Colella, 1990a) and that satisfaction may be enhanced when feedback is presented in terms of changes (i.e., improvements) in performance. The implication for standard setting is that standards which are defined in terms of performance improvements may facilitate satisfaction.

## IDENTIFICATION OF FUTURE RESEARCH NEEDS

All organizations set and communicate performance standards, implicitly or explicitly. While there is scant organizational research on how this is best accomplished, we have reviewed related literatures which focus on the appraisal, feedback, and valuation of performance. Based on these literatures, we identify four domains where both basic and applied research on performance standards are clearly needed. Such research is necessary if we are to better employ performance standards to both evaluate and motivate employees. Specific hypotheses are also stated when possible. The four general topic areas are:

1) The internalization of external performance standards
2) The definitional content of performance standards
3) The communication of performance standards
4) The difficulty of performance standards

### Translating External Standards Into Internalized Standards

The goal setting, performance appraisal, and feedback literatures reviewed above all point to the importance of the translation of external performance standards into internalized performance standards. This translation seems so fundamental to the efficacy of setting

performance standards, yet there is no direct literature which verifies the linkage. We hypothesize that:

> Hypothesis 1: **Individual reactions (both motivational and affective) to standards depends on the degree to which external performance standards are internalized by both those who apply the standards and those to whom the standards apply.**

Our review turned up several hypotheses about conditions under which internalization is most likely to occur. First, Murphy and Cleveland (1991) made the following hypothesis:

> Hypothesis 2: **The degree to which external performance standards are internalized depends on the degree to which the employee identifies with, and is committed to, the goals of the organization.**

Thus, one objective in setting standards is to develop standards which are likely to build upon an individual's organizational commitment (and/or reason for joining the organization in the first place). The goal setting literature has discussed this issue in great detail, and the factors thought to influence external goal commitment were presented in Figure 2. Thus, we posit:

> Hypothesis 2a: **Standards which are communicated and constructed in accord with the factors presented in Figure 2 are more likely to be internalized than those which do not. For example, standards that are clear, do not conflict with other standards, are tied to rewards, and foster a sense of achievement will result in greater likelihood of internalization by individual workers.**

It should also be noted that there are a variety of factors related to goal <u>difficulty</u> which influence goal commitment, such as the degree of challenge in the goal or expectancies for success. These issues are discussed further in a following section.

One other factor which may influence commitment to (and hence, internalization of) externally set performance standards involves the degree to which individuals to whom the standards apply participate in the standard setting process. Goal setting researchers are divided on the importance of participative goal setting (see Locke & Latham, 1990). However, researchers in other areas tend to agree on the importance of participation in standard setting (e.g., Dornbush & Scott, 1975; Drucker, 1954; Taylor et al., 1983). Thus, the following hypothesis deserves further attention within the framework of standard setting:

> Hypothesis 3: **Standards which are participatively set will be more likely to be internalized than those which are not.**

Definitional Content

When referring to the definitional content of performance standards, we are referring to those aspects of performance which define various levels of standards. For example, a minimal performance standard can be defined as: 1) being of marginal utility to the organization; 2) making 20 widgets per day; or 3) not being absent for more than five days a year. Another way of considering the definitional content is to consider which aspects of performance should be subject to standard setting procedures.

From an evaluation standpoint, the issue of what should be rated during performance appraisal has received attention in the existing literature. This literature focuses on those aspects of performance which are valuable parts of individual effectiveness and are amenable to tracking and rating. Prescriptive accounts of standard setting point out that the content of standards should be derived from a careful job analysis. Carlyle and Ellison (1984) provide the following list of attributes, which if truly considered when setting standards for performance evaluation, should result in a high quality appraisal system:

1. Standards should be concrete and specific.
2. Standards should be practical to measure.
3. Standards should be meaningful.
4. Standards should be realistic and based on a sound rationale.
5. Standards should be consistent across similar jobs.

On the other hand, those who approach performance appraisal from a more cognitive, descriptive perspective (Borman, 1987; DeNisi et al., 1984; Feldman, 1981) would argue that raters have their own personal theories of what aspects of performance are worthy of defining standards, and it is the rater's personal standards which drive how they rate individuals. One's personal theory of performance determines what behaviors are attended to, how these behaviors are recalled, and how these behaviors are interpreted and judged. Thus, we make the following hypothesis:

> Hypothesis 4: **To the degree that external performance standards match the internalized standards of raters, rater consistency will increase.**

From a motivational perspective, the definitional content of standards is important to the extent that it influences the type of feedback that individuals receive and the manner in which feedback information is framed. In this case, the focus is on the internalized standards of those to whom the standards apply. Based on the feedback literature cited earlier, we make the following hypothesis:

> Hypothesis 5: **To the degree that external performance standards match the internalized standards of ratees, feedback regarding those standards will more likely be**

**accepted and perceived as appropriate.**

These hypotheses suggest it is important that externally set performance standards are similar to those adopted by both raters and ratees. For example, suppose that for the job of university professor, external standards indicate that professors are to be judged upon the amount of their research activity. However, if a rater (e.g., department chair) feels that collegiality is of utmost importance to the professor job, then the rater will attend to information which he or she perceives to be most related to collegiality, and perhaps ignore or not remember information related to research activity. This problem is compounded by considering the ratee. For example, in the above case, if the professor feels that teaching effectiveness is the most important factor in job performance, feedback regarding collegiality and research activity may be discounted as unimportant or reinterpreted to coincide with the ratee's internal performance standards.

Regarding specificity of performance standards, both feedback literatures (e.g., Taylor et al., 1983) and goal setting models (e.g., Locke & Latham, 1990) point out that specific feedback and goals lead to better motivational effects than vague feedback and goals. Thus, to the extent that standards allow for concrete and specific feedback and goals, we hypothesize:

> Hypothesis 6: **Performance standards which are defined in specific, behavioral terms will lead to greater motivation, and consequently greater performance, than those which are vaguely defined.**

A third aspect to the content of performance standards concerns how the definition of the standard is framed. Based on the literature concerning framing (e.g., Kahneman & Tversky, 1979), it is likely that individuals will react differently to standards framed in a positive or negative format. According to Kahneman and Tversky's (1979) prospect theory, individuals receiving negatively framed information will be more risk-seeking than those receiving positively framed information. This suggests that standards defined with negative frames (e.g., misses 2 to 4 days of work per month) will lead to changes in behavior more readily than those framed in positive terms (e.g., attends work 16 to 18 days per month).

On the other hand, the feedback literature (e.g., Taylor et al., 1983) suggests that individuals will be less likely to accept negative feedback than positive feedback, and will only respond with increased motivation to negative feedback if their expectancies for success on future performance is relatively high. This body of literature would suggest that negatively framed standards, leading to negatively framed feedback, may have less positive effects (feedback being discounted or ignored) than positively framed standards. While these conflicting literatures make it difficult to develop specific, directional hypotheses, the issue of "how to frame the definition of standards" clearly deserves study.

Finally, the valence (or value) of standards is crucial to their motivational potential. As specified in Figure 2, goal setting researchers have posited that external goal valence (associated rewards, punishments, sense of achievement, opportunities for self-improvement) is a factor which determines whether individuals become committed to assigned goals. Expectancy theory (Vroom, 1964) also argues that the valence of performance outcomes is a major influence on motivation. Therefore, we posit:

> Hypothesis 7: **Performance standards which are related to outcomes valued by those to whom the standards apply will lead to higher motivation and performance than standards which are related to performance outcomes not valued.**

Based on our review of the job satisfaction literature, the importance of attaching valued outcomes to achievement of performance standards was reiterated, and it was further specified that outcomes associated with standards should be perceived as fair reward for the employee inputs. Thus we hypothesize:

> Hypothesis 8: **Performance standards which result in rewards which are perceived as fair given the inputs necessary to successful performance will lead to higher satisfaction than standards resulting in outcomes which are not perceived as fair reward.**

Our review of the job satisfaction literature also suggested that people will be more satisfied when success can be attributed to personal qualifications or abilities rather than external factors, such as luck or leadership. This suggests that performance standards be defined by aspects of job performance which primarily reflect individual effort, skill, and ability. Thus we posit:

> Hypothesis 9: **Performance standards which are defined by aspects of performance heavily influenced by individual effort will lead to more satisfaction than standards which are defined by aspects of performance heavily influenced by external factors.**

## Communicating Performance Standards

There is also scant research on how performance standards get communicated to employees. Above, several aspects of the communication process were hypothesized to be important, such as the clarity of the standard's definition. The first line of research in this area should be descriptive in its attempt to identify the channels which organizations use to communicate performance standards.

One area of research which may be useful for examining how, and when, standards get communicated is that on the socialization of

newcomers. For example, Feldman's (1976) stage model of socialization suggests that norms and standards are most likely to be communicated to newcomers during the second stage of socialization, accommodation. This stage is the point at which the newcomer's role begins to be made clear. Further, descriptive accounts of socialization may inform us about the manner in which standards get communicated. For example, Van Maanen's (1976) description of the socialization of police officers points to the importance of the recruit's Field Training Officer in both stating and demonstrating performance standards. Although the socialization literature is relevant, it has not specifically addressed the issue of the communication of performance standards, precluding us from making specific hypotheses about this particular process.

Murphy & Cleveland (1991) draw upon research concerning social norms to hypothesize how standards get communicated. They point out the importance of direct communication, observing the application of standards to others, identification with reference groups, and direct experience as means by which standards get communicated. Again, while existing research has not explicitly focused on the communication of organizational performance standards, all of these processes deserve future study.

Finally, our review of the job satisfaction literature pointed out that providing a rationale for external standards may facilitate satisfaction. This aspect of standard setting should be considered when communicating performance standards. Therefore we posit:

> Hypothesis 10: **Communicating performance standards along with a rationale will result in greater satisfaction than use of equally appropriate, but unjustified, performance standards.**

Setting the Difficulty of Standards

Performance standards are often presented in terms of the minimal level of accepted performance (Carroll & Schneier, 1982). However, as argued earlier, standards may be set for any level of performance and thus, for any level of difficulty. From a motivational standpoint, the research reviewed above suggests that minimal standards will be less effective than more difficult standards. Even the research on the effects of positive feedback (an outcome likely to result if standards are easy) suggests that if standards are easily attained and no reward is provided for meeting standards, people may actually become less motivated (Janz, 1982; Taylor et al., 1983).

Of course, the feedback and goal setting literatures also noted that if their respective interventions are to be effective, then individuals must maintain high levels of expectancy for success, self-efficacy, and a sense of accomplishment. This is particularly true when valence for the standard is not extremely high (Hollenbeck & Klein, 1987). Thus,

Hypothesis 11: **Individuals working under standards which are realistically difficult will be more motivated, and consequently perform better, than those working under easy standards.**

The research on job attitudes makes the same prediction about the optimal level of standard difficulty as the motivation literature: i.e., standards should be realistic but difficult enough to provide a sense of achievement. Therefore we also hypothesize:

Hypothesis 11a: **Individuals working under and meeting standards which are realistically difficult will be more satisfied than those working under easy standards.**

These hypotheses point out the importance of doing research which examines the conditions under which standards will be seen as difficult, yet achievable and motivating. So-called "optimal level theories of motivation" (cf. Arkes & Garske, 1982; Dember & Earl, 1957) should provide assistance in determining appropriate difficulty levels. These theories posit an "inverted-U" relationship between the difficulty of the task and its motivating potential. The optimal level of arousal (often called the "pacer") is determined by locating the maximum value of the inverted-U function.

It should also be clear that, for evaluation purposes, there may be many levels of standards. For example, in the introduction, we proposed a performance standard continuum of: (a) a "minimum" standard, (b) a "training warning" standard, (c) an "acceptable" standard, and (d) an "excellence" standard. In a military context, Wise et al. (1991) adopted a related continuum of "unacceptable", "marginal", "acceptable", and "outstanding". Their scale was defined by the degree of future training a soldier required, as well as the value (negative to positive) of performance to the organization. We hypothesize that:

Hypothesis 12: **Organizations which adopt a continuum of performance standards (such as listed above), rather than a single common standard across all individuals in a particular job, will increase individual motivation and subsequent overall organizational effectiveness.**

In turn, organizations which set and emphasize only minimum performance standards may actually be <u>demotivating</u> its employees. Setting low external goals can reduce an individual's intrinsic motivation to perform (Deci, 1975). We hypothesize that:

Hypothesis 12a: **To the extent that only minimum performance standards are emphasized by organizations, (i) these minimum standards will be internalized, (ii) they will serve as relatively easy individual goals, and (iii) intrinsic motivation to perform at greater levels will be reduced.**

Furthermore, we note that minimum standards are often adopted

because a judgment is made that performance below the standard would be deleterious to the organization. For example, a company which produces bolts for high speed aircraft can not afford to have "too many" defective products. However, the notion of "too many" bears closer examination. That is, if one defective product causes a plane to crash, then even one defect is too many. Thus, the real "minimal standard" is no defective products. About the only thing that minimum standards do is indicate an acceptable tolerance for error (e.g., our standard is no more than 2% defective products), but the real standard is no defects. This logic applies to just about any situation (e.g., graduating at least 90% of high school students when one really wants to graduate 100%). Of course, the magnitude of this tolerance is a value judgment: if less than 2% is acceptable, why not less than 3%, and so on? Glass (1978) calls this the "counting backwards from 100%" approach (p. 244). Thus, the setting of minimal standards can often be considered as placing an arbitrary standard error around the real goal of "no errors".

The particular level of acceptable error is often scientifically indefensible. This has led some researchers (Austin & Bobko, 1985; Glass, 1978) to conclude that standards and goals should be cast in a framework of change. That is, one can tell whether production (or any other performance behavior) has gone up, remained the same, or gone down. Psychometrically, this focus on change has the advantage of having a true zero-point (i.e., zero change). Motivationally, a focus on change places standards for behavior closer to the optimal, "pacer" level of intentional behavior. Thus, we posit the following:

> Hypothesis 12b: **A focus on the evaluation of changes in performance will enhance individual motivation and increase subsequent individual and organizational performance, relative to the adoption of minimal standards of performance.**

Also, as noted in our review of the job satisfaction literature, it has been suggested that satisfaction may be enhanced when feedback is presented in terms of changes (i.e., improvements) in performance. Thus, we hypothesize:

> Hypothesis 12c: **A focus on the evaluation of changes in performance will enhance individual satisfaction, relative to the adoption of inflexible standards of performance.**

[There are problems with a focus on change, such as measuring change reliably and deciding how much change is required. Further, absolute behavior can not be completely ignored in some situations. However, we believe the focus on standards of change has not been sufficiently exploited by organizational researchers.]

In sum, it is clear that we are suggesting that organizations take a hard look at the notion of minimum standards in the workplace. While there may be good reasons for adopting and communicating minimum performance standards, it is critical for organizations to engage in

answering the self-reflective question, "Why did we do this in the first place?". We have suggested that there might be dysfunctional consequences of setting only minimum performance standards. Further, the process of "thinking back" to the original reasons for any standard is a valuable aspect of organizational analysis and strategic planning (Bobko, 1985; Thompson & Strickland, 1990). As Thompson and Strickland state, "Policy helps ... translate the corporate philosophy into how things are done, how people are treated, and what the corporate beliefs and attitudes mean in terms of everyday activities." (p.242). In fact, it has been argued that the questions arising from such corporate self-reflection can only be answered by an appeal to values and ethical criteria (Keeley, 1983) -- and we have then come back full circle to the notion of performance standards as being value statements.

Identifying Performance Levels from Utility Analysis. The focus of this review has been on employee reactions to performance standards, and not on the psychometrics of where to set cut-points. Nonetheless, is has become clear that the difficulty level of standards affects employee reactions. We believe that recent psychological literature, which tries to identify the function relating performance to perceived value, can help inform the appropriate levels of performance standards.

As noted above, Pritchard et al. (1988) have developed functions which map task outcomes (e.g., failure rates for a manufactured product) against judgments of organizational effectiveness (see Figure 4 for a typical pattern). These judgments are presumed to be on a ratio

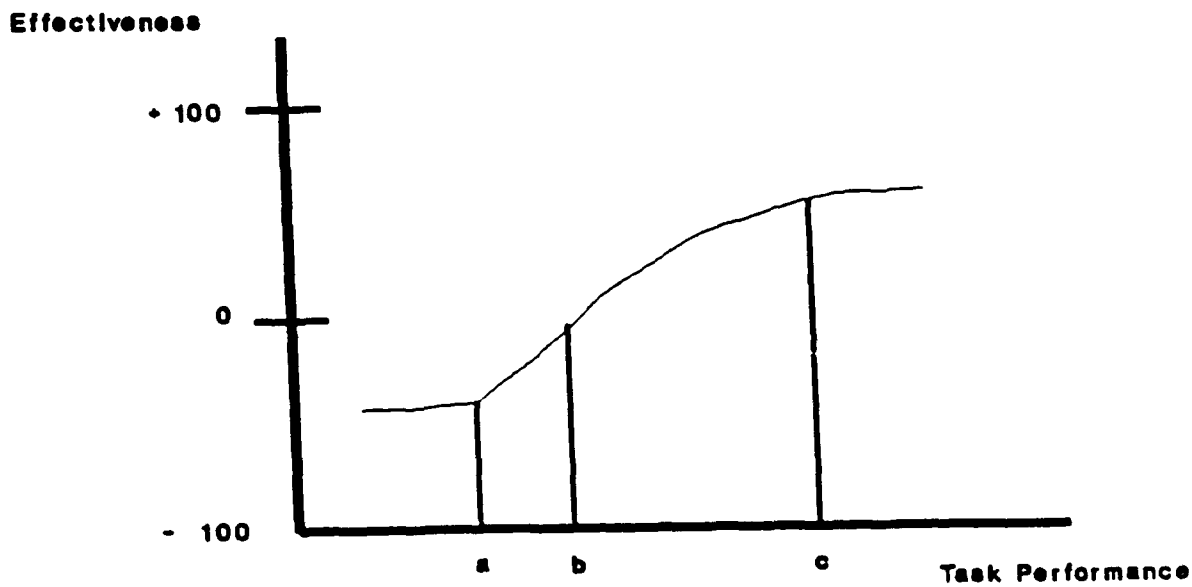


Figure 4. Typical function relating utility/effectiveness to task performance.

scale, such that a zero-point of effectiveness can be identified. The shape of these functions is not linear. In fact, the effectiveness curves tend to be S-shaped. Pritchard et al. generated such functions for jobs in the U.S. Air Force. It is interesting to note that almost identically shaped functions were independently developed within the context of the U.S. Army's classification system (cf. Sadacca et al., 1990).

We propose that the non-linearity of the utility/effectiveness function can be exploited. For example, point "c" on the task output dimension (horizontal axis) would be a likely candidate for a "standard of performance excellence". That is, from an organizational point of view, increasing performance beyond this already high level shows little marginal gain in organizational effectiveness (and an individual's motivation might best be directed to other components of the job once performance reaches level "c" on this component).

Further, there are two candidates for "minimum performance" standards in Figure 4; i.e., points "a" and "b". If the effectiveness scale is truly ratio in nature, then point "b" should be set as the minimum standard, since performance below this standard yields negative returns to the organization. However, the psychological literature demonstrates that the construction of truly ratio scales is not easy. Therefore, point "a" identifies another candidate for the minimum performance standard -- below this point all performance is equally poor, while above this point motivated effort may lead to gains in organizational effectiveness. In either case, we hypothesize that:

> Hypothesis 13: **Organizations which set multiple performance standards defined by natural breaks in utility/effectiveness functions will show increased organizational effectiveness, relative to other organizations.**

Of course, standards which are set using such functions should be inspected for their reasonableness, specificity, and acceptability, as well as any other characteristics noted in the prior hypotheses.

## CONCLUSIONS AND THE ARMY CONTEXT

### The Army Synthetic Validity Project

In the late 1980's the US Army sponsored a project to investigate the transportability of selection and classification equations from known Military Occupational Specialties (MOS) to new, or unstudied, MOS. In order to help generate and investigate the utility of synthetically generated selection devices, a substantial portion of this project was devoted to how the Army might set performance standards within MOS (in order that selection cutoffs could be related to success rates on performance). This "standard setting" aspect of the synthetic validity (SYNVAL) project is one of the largest empirical

investigations about performance standard setting in the field of industrial psychology. An extensive summary of the results may be found in Wise et al. (1991).

The SYNVAL project, after empirical and theoretical consideration of many alternatives, settled on the investigation of two procedures for setting standards: a task-based approach and an incident-based approach. The task-based approach required raters to identify the percentage of task steps (percent GO scores) that would be required to attain particular standards of job performance (unacceptable, marginal, acceptable, outstanding). The incident-based approach required raters to rate behaviorally-based samples of job performance as either unacceptable, marginal, acceptable, or outstanding. The percentage of current Army soldiers who fell into each of these categories using each of these two methods could then be calculated using known empirical distributions of incumbent performance and incumbent behavioral incident ratings.

As noted by Wise et al. (1991), "The most significant conclusion of the standard setting research was that the different methods that we developed and evaluated led to different results" (p. viii). In particular, standards set using the task-based method were stricter than the standards using the incident-based method, particularly for the lower-level cutoffs (e.g., marginal performance).

## SYNVAL Results as Value Judgments

The SYNVAL results are certainly consistent with our review above, in the sense that different procedures result in different answers (why this occurs is considered in the next section). This is what one would expect of value judgments -- they are affected by the context in which they are derived. Second, it is obvious that if one wants to set more difficult standards (with all the resultant properties noted above), then the Army should choose a task-based standard setting procedure. Of course, doing so would imply that a greater percentage of current Army incumbents are "unacceptable" and that, once again, is a value judgment which needs to be considered by the organization as a whole. However, we believe that the interesting psychological research implied by SYNVAL's finding concerns why the differences occurred. This issue is now considered.

## Maximal Versus Typical Performance

Based on the development of the two standard setting methods (and the instructional sets they evoked), Wise et al. (1991) conjectured that task-based standards were based on perceptions about maximal, "can-do" soldier performance while incident-based standards were based on perceptions about a soldier's typical, "will-do" levels of performance. Further, Sackett, Zedeck, and Fogli (1988) have investigated the commonly used terms of maximal and typical performance

and found relatively low correlations between these two performance measurements. One implication is straightforward: when setting performance standards, the Army (and probably most other organizations) need to consider whether these standards should be set on maximal or typical performance (or some combination of the two). Otherwise, confusion between the two measures could be a source of rater/ratee disparity (noted earlier), lack of goal clarity (noted earlier), or create unintended difficulty levels of performance standards.

It may be that standards based on typical performance are adequate under peacetime conditions. On the other hand, maximal performance standards might be more critical (i) during hostilities, (ii) on tasks for which there are no back-up, "fail-safe" remedies or, (iii) on tasks which are rated as having a high cost of error. Future research should systematically delineate the organizational and environmental conditions under which the Army might want to use maximal versus typical performance standards.

Further, it is theoretically interesting to suggest that the difference between a person's typical performance and his/her maximal performance can be considered as an index of a soldier's motivation to perform. As such, a research hypothesis would be that the difference between maximal and typical performance will, as a motivational index, be related to other constructs in the nomological net, such as commitment to the organization (Army). Indeed, the SYNVAL results noted in the introduction of our review indicate that this difference is substantial in the Army context: the cutoff for "marginal" performance using the incident-based method is at the 6th percentile of current Army performance; the cutoff for "marginal" performance using the task-based method is at the 29th percentile of performance (see Wise et al., 1991, Table 5.23). Thus, there may be lots of opportunity for increased improvement in Army performance by focusing on maximal performance standards.

A Qualitative Comment

During an interview with Army subject matter experts, we asked the question, "When a soldier thinks about 'performance standards', what does that soldier think about?" The answer was, "The soldier probably thinks of a list of tasks suitable to his/her MOS and assumes that all tasks should be capable of being performed." This is a form of the "counting from 100%" approach considered earlier in this paper. That is, the standard is performance capability on all tasks, with some allowable slippage. below that standard. However, SYNVAL's findings for the task-based standard setting approach would indicate that the organizational reality is far from this "counting back from 100%" standard. (Indeed, the instructions used in that portion of the project directed subject matter experts to consider less than 100% performance as being acceptable.) As already noted, the task-based results indicate that "marginal" performance is at the 29th percentile of incumbent distribution. Further, this same study indicates that only 20% of the

current soldiers could be labeled "outstanding". Clearly, the conjecture about a typical soldier's statement and the organization's reality are not congruent.

Individual Performance and Satisfaction

It has been our working assumption that performance standards are set and communicated by almost all organizations, including the Army. We have drawn upon literatures related to motivation, performance, and job affect in order to focus attention on a new issue: i.e., how people react to performance standards. The reader is referred back to our list of hypotheses for future research and new considerations in the setting of performance standards. Our paper proposes that standards can be used to motivate and help retain employees in addition to being a means for assessing performance. Specifically, we hypothesize that the Army can increase motivation, and consequently, individual performance by setting performance standards which are 1) congruent between soldiers' internal standards, what the Army says, and what the Army does; 2) specific; 3) tied to valued rewards; 4) accepted by the soldier; and 5) set at challenging levels of difficulty. Furthermore, standards set in this manner should also lead to increased job satisfaction and more positive affect toward the job, leading to increased levels of soldier retention. All of these effects are further hypothesized to occur through the mediating mechanism of the internalization of externally set performance standards, suggesting future research in the area of how soldiers interpret and accept standards set by the Army. We believe that the research areas outlined in this paper will be fruitful. While several of these relationships are assumed to exist by some management textbook writers (see most chapters on setting business policy and subsequent performance objectives), there is little direct empirical evidence examining how employees react to performance standards. The hypotheses herein should help guide this necessary empirical research.

## REFERENCES


Andrew, B., & Hecht, J. (1976). A preliminary investigation of two procedures for setting examination standards. Educational and Psychological Measurement, 36, 45-50.

Angoff, W. (1971). Scales, norms, and equivalent scores. In R. Thorndike (Ed.), Educational measurement (pp. 508-600). Washington, D.C.: American Council on Education.

Arkes, H., & Garske, J. (1982). Psychological theories of motivation (2nd ed.). Monterey, CA: Brooks/Cole.

Ashford, S., & Cummings, L. (1983). Feedback as an individual resource: Personal strategies of creating information. Organizational Behavior and Human Performance, 32, 370-398.

Austin, J., & Bobko, P. (1985). Goal setting theory: Unexplored areas and future research needs. Journal of Occupational Psychology, 58, 289-308.

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. Journal of Personality and Social Psychology, 45, 1017-1028.

Bandura, A., & Cervone, D. (1986). Differential engagement of self-reactive influences in cognitive motivations. Organizational Behavior and Human Decision Processes, 38, 92-113.

Bavelas, J., & Lee, E.S. (1978). Effect of goal level on performance: A trade-off of quantity and quality. Canadian Journal of Psychology, 32, 219-240.

Bazerman, M. (1983). Negotiator judgment: A critical look at the rationality assumption. American Behavioral Scientist, 27, 211-228.

Berk, R. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. Review of Educational Research, 56, 137-172.

Bobko, P. (1985). Removing assumptions of bipolarity: Towards variation and circularity. Academy of Management Review, 10, 99-108.

Bobko, P., Karren, R., & Kerkar, S. (1987). Systematic research needs for understanding supervisory-based estimates of SDy in utility analysis. Organizational Behavior and Human Decision Processes, 40, 69-95.

Bobko, P., Shetzer, L., & Russell, C. (1991). Estimating the standard deviation of professors' worth: The effects of frame and presentation order in utility analysis. Journal of Occupational Psychology, 64, 179-188.

Bobko, P., & Wise, L. (1987, September). Towards a methodological understanding of standard setting: Are we satisfied with our satisfaction benchmarks? Paper presented at the Measurement Research Symposium, Bell Communications Research Corporation, Cherry Hill, NJ.

Borman, W. (1987). Personal constructs, performance schemata, and "folk theories' of subordinate effectiveness: Exploration in an Army Officer sample. Organizational Behavior and Human Decision Processes, 40, 307-322.

Boudreau, J. (1991). Utility analysis for decisions in human resource management. In M. Dunnette & L. Hough (Eds.), Handbook of Industrial & Organizational Psychology (2nd ed., pp. 621-745). Palo Alto, CA: Consulting Psychologists Press.

Brennan, R., & Lockwood, R. (1980). A comparison of the Nedelsky and Angoff cutting score procedures using generalizability theory. Applied Psychological Measurement, 42, 219-240.

Bryan, J., & Locke, E. (1967a). Goal setting as a means of increasing motivation. Journal of Applied Psychology, 51, 274-277.

Bryan, J., & Locke, E. (1967b). Parkinson's law as a goal-setting phenomenon. Organizational Behavior and Human Performance, 2, 258-275.

Burke, M. & Frederick, J. (1984). Two modified procedures for estimating standard deviations in utility analysis. Journal of Applied Psychology, 69, 482-489.

Campion, M., & Lord, R. (1982). A control systems conceptualization of the goal-setting and changing process. Organizational Behavior and Human Performance, 30, 265-287.

Carlyle, J., & Ellison, T. (1984). Developing performance standards. In J. Bernardin & R. Beatty, Performance appraisal: Assessing human behavior at work. Boston: Kent.

Carlson, R. (1967). Selection interview decisions: The effect of interview experience, relative quota situation, and applicant sample on interviewer decisions. Personnel Psychology, 20, 259-280.

Carroll, S., & Schneier, C. (1982). Performance appraisal and review systems. Glenview, IL: Scott, Foresman and Company.

Carroll, S. & Tosi, H. (1973). Management by objectives: Implications and research. New York: MacMillan.

Carsten, J., & Spector, P. (1987). Unemployment, job satisfaction, and employee turnover: A meta-analytic test of the Muchinsky model. Journal of Applied Psychology, 72, 374-381.

Carver, C., & Scheier, M. (1981). Attention and self-regulation: A control theory approach to human behavior. New York: Springer-Verlag.

Cascio, W., Alexander, R., & Barrett, G. (1988) Setting cutoff scores: Legal, psychometric, and professional issues and guidelines. Personnel Psychology, 41, 1-24.

Cronbach, L. (1949). Essentials of psychological testing (3rd ed.). New York: Harper & Row.

Cronbach, L., & Gleser, G. (1965). Psychological tests and personnel decisions. Urbana, IL: University of Illinois Press.

Czajka, J., & DeNisi, A. (1988). Effects of emotional disability and clear performance standards on performance ratings. Academy of Management Journal, 31, 394-404.

Deci, E. (1975). Intrinsic motivation. New York: Plenum.

Dember, W., & Earl, R. (1957). Analysis of exploratory, manipulatory, and curiosity behaviors. Psychological Review, 64, 91-96.

DeNisi, A., & Williams, K. (1988). Cognitive approaches to performance appraisal. In G. Ferris and K. Rowland (Eds.), Research in personnel and human resource management, Vol. 6. Greenwich, CT: JAI Press.

DeNisi, A., Cafferty, T., & Meglino, B. (1984). A cognitive view of the performance appraisal process: A model and research propositions. Organizational Behavior and Human Performance, 33, 360-396.

Dornbusch, S., & Scott, W. (1975). Evaluation and the exercise of authority. San Francisco: Jossey-Bass.

Drucker, P. (1954). The practice of management. New York: Harper.

Early, P., & Kanfer, R. (1985). The influence of component participation and role models on goal acceptance, goal satisfaction, and performance. Organizational Behavior and Human Decision Processes, 36, 378-390.

Earley, P., & Perry, B. (1987). Work plan availability and performance: An assessment of task strategy priming on subsequent task completion. Organizational Behavior and Human Decision Processes, 39, 279-302.

Eaton, N., Wing, H., & Mitchell, K. (1985). Alternative methods of estimating the dollar value of performance. Personnel Psychology, 38, 27-40.

Ebel, R. (1972). Essentials of educational measurement. Englewood Cliffs, NJ: Prentice-Hall.

Erez, M., Earley, C., & Hulin, C. (1985). The impact of participation on goal acceptance and performance: A two-step model. Academy of Management Journal, 28, 50-66.

Erez, M., & Kanfer, F. (1983). The role of goal acceptance in goal setting and task performance. Academy of Management Review, 8, 454-463.

Feldman, D. (1976). A contingency theory of socialization. Administrative Science Quarterly, 21, 433-452.

Feldman, J. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66, 127-148.

French, J., & Caplan, R. (1973). Organizational stress and individual strain. In A.J. Morrow, (Ed.), The failure of success. New York: Amacon.

Glass, G. (1978). Standards and criteria. Journal of Educational Measurement, 15, 237-261.

Griffin, R. (1990). Management (3rd ed.). Boston: Houghton Mifflin.

Guzzo, R., Jette, R., & Katzell, R. (1985). The effects of psychologically based intervention programs on worker productivity. Personnel Psychology, 38, 275-293.

Halpin, G., & Halpin, G. (1983, August). Reliability and validity of 10 different standard setting procedures. Paper presented at the annual meeting of the American Psychological Association, Anaheim, CA.

Hambleton, R., & Eignor, D. (1979). Competency test development, validation, and standard setting. In R. Jaeger & C. Tittle (Eds.), Minimum competency testing. Berkeley, CA: McCutchan.

Hollenbeck, J., & Klein, H. (1987). Goal commitment and the goal setting process: Problems, prospects, and proposals for future research. Journal of Applied Psychology, 72, 212-220.

Huber, V., Neale, M., & Northcraft, G. (1987). Judgment by heuristics: Effects of ratee and rater characteristics and performance standards on performance-related judgments. Organizational Behavior and Human Decision Processes, 40, 149-169.

Hulin, C. (1991). Adaptation, persistence, and commitment in organizations. In M. Dunnette & L. Hough (Eds.), Handbook of industrial and organizational psychology, (2nd ed., pp. 445-505). Palo Alto, CA: Consulting Psychologists Press.

Hulin, C., Roznowski, M., & Hachiya, D. (1985). Alternative opportunities and withdrawal decisions: Empirical and theoretical discrepancies and an integration. Psychological Bulletin, 97, 233-250.

Ilgen, D., & Feldman, J. (1983). Performance appraisal: A process focus. In L. Cummings & B. Staw (Eds.), Research in organizational behavior, Vol. 5. Greenwich, CT: JAI Press.

Ilgen, D., Fisher, C., & Taylor, M. (1979). Consequences of individual feedback on behavior in organizations. Journal of Applied Psychology, 64, 349-371.

Jacobs, M., Jacobs, A., Feldman, G., & Cavior, N. (1973). Feedback II-the credibility gap: Delivery of positive and negative emotional and behavior feedback in groups. Journal of Consulting and Clinical Psychology, 41, 215-223.

Jaeger, R. (1976). Measurement consequences of selected standard setting models. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.

Janz, T. (1982). Manipulating subjective expectancy through feedback: A laboratory study of the expectancy-performance relationship. Journal of Applied Psychology, 67, 480-485.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Karren, R., & Bobko, P. (1983, August). Conducting utility analysis: Some methodological concerns. Paper presented at the meeting of the Academy of Management, Dallas, TX.

Katz, D., & Kahn, R. (1978). The social psychology of organizations (2nd ed.). New York: Wiley.

Keeley, M. (1983). Values in organizational theory and management education. Academy of Management Review, 8, 376-386.

Koffler, S. (1980). A comparison of approaches for setting proficiency standards. Journal of Educational Measurement, 17, 167-178.

Kopelman, R. (1986). Objective feedback. In E. Locke (Ed.), Generalizing from laboratory to field settings. Lexington, MA: Lexington Books.

Landy, F., & Farr, J. (1980). Performance ratings. Psychological Bulletin, 87, 72-107.

Latham, G., Erez, M., & Locke, E. (1988). Resolving scientific disputes by the joint design of crucial experiments by the antagonists: Application to the Erez-Latham dispute regarding participation in goal setting. Journal of Applied Psychology (Monograph), 73, 753-772.

Latham, G., Mitchell, T., & Dossett, D. (1978). Importance of participative goal setting and anticipated rewards on goal difficulty and job performance. Journal of Applied Psychology, 63, 163-171.

Lawler, E. (1973). Motivation in work organizations. Monterey, CA: Brooks/Cole.

Locke, E. (1976). The nature and causes of job satisfaction. In M. Dunnette (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand-McNally.

Locke, E., & Latham, G. (1990). A theory of goal setting & task feedback. Englewood Cliffs, NJ: Prentice Hall.

Lofquist, L., & Dawis, R. (1969). Adjustment to work. New York: Appleton-Century-Crofts.

March, J., & Simon, H. (1958). Organizations. New York: Wiley.

Mathieu, J., & Tannenbaum, S. (1985, March). Supervisors' estimates of the dollar value of job performance: Some qualitative and quantitative findings. Paper presented at the meeting of the Southeastern Psychological Association, Atlanta.

Matsui, T., Kakuyama, T., & Onglatco, M. (1987). Effects of goals and feedback on performance in groups. Journal of Applied Psychology, 72, 407-415.

Michalos, A. (1986). Job satisfaction, marital satisfaction, and the quality of life: A review and preview. In F. Andrews (Ed.), Research on the quality of life. Ann Arbor: University of Michigan Institute for Social Research.

Mobley, W., Griffeth, R., Hand, H., & Meglino, B. (1979). Review and conceptual analysis of the employee turnover process. Psychological Bulletin, 86, 493-522.

Murphy, K. (1982). Difficulties in the statistical control of halo. Journal of Applied Psychology, 67, 161-164.

Murphy, K., & Cleveland, J. (1991). Performance appraisal: An organizational perspective. Boston, MA: Allyn and Bacon.

Nathan, B., & Lord, R. (1983). Cognitive categorization and dimensional schemata: A process approach to the study of halo in performance ratings. Journal of Applied Psychology, 68, 102-114.

Nedelsky, L. (1954). Absolute grading standards for objective tests. Educational and Psychological Measurement, 14, 3-19.

Organ, D. (1977). Intentional vs. arousal effects in goal setting. Organizational Behavior and Human Performance, 18, 378-89.

Phillips, J., & Freedman, S. (1988). The task-related competency and compliance aspects of goal setting: A clarification. Organizational Behavior and Human Decision Processes, 41, 34-49.

Porter, L., & Steers, R. (1973). Organizational, work, and personal factors in employee turnover and absenteeism. Psychological Bulletin, 80, 151-176.

Powers, W. (1973). Behavior: The control of perception. Chicago: Aldine.

Pritchard, R., & Curtis, M. (1973). The influence of goal setting and financial incentives on task performance. Organizational Behavior and Human Performance, 10, 175-183.

Pritchard, R., Jones, S., Roth, P., Stuebing, K., & Ekeberg, S. (1988). Effects of group feedback, goal setting, and incentives on organizational productivity. Journal of Applied Psychology, 73, 337-358.

Pulakos, E., Wise, L., Arabian, J., Heon, S., & Delaplane, S. (1989). A review of procedures for setting job performance standards. (Tech. Report No. 840). US Army Research Institute, Alexandria, VA (AD A210 717)

Rodgers, R., & Hunter, J. (1986). The impact of management by objectives on organizational productivity. Unpublished manuscript, University of Texas, Austin, TX.

Rothkopf, E., & Billington, M. (1975). A two-factor model of the effect of goal descriptive-directions on learning from text. Journal of Educational Psychology, 67, 692-704.

Sackett, P., Zedeck, S., & Fogli, L. (1988). Relations between measures of typical and maximum job performance. Journal of Applied Psychology, 73, 482-486.

Sadacca, R., Campbell, J., DiFazio, A., Schultz, S., & White, L. (1990). Scaling performance utility to enhance selection/classification decisions. Personnel Psychology, 43, 367-378.

Schurr, P. (1987). Effects of gain and loss decision frames on risky purchase negotiations. Journal of Applied Psychology, 72, 351-358.

Scott, K., & Taylor, G. (1985). An examination of conflicting findings on the relationship between job satisfaction and absenteeism: A meta-analysis. Academy of Management Journal, 28, 599-612.

Shepard, L. (1980). Standard setting issues and methods. Applied Psychological Measurement, 4, 447-467.

Simon, K. (1988). Effects of self-comparison, social comparison, and depression on goal setting and self-evaluative reactions. Unpublished manuscript, Psychology Department, Stanford University.

Skakun, E., & Kling, S. (1980). Comparability of methods of standard setting. Journal of Educational Measurement, 17, 229-235.

Smith, P. (1976). Behaviors, results, and organizational effectiveness. In M. Dunnette (Ed.), Handbook of industrial organizational psychology. Chicago, IL: Rand-McNally.

Smith, P., Kendall, L., & Hulin, C. (1969). The measurement of satisfaction in work and retirement. Chicago: Rand-McNally.

Taylor, M., Fisher, C., & Ilgen, D. (1983). Individuals' reactions to performance feedback in organizations: A control theory perspective. In K. Rowland & G. Ferris (Eds.), Research in personnel and human resources management. Greenwich, CT: JAI Press.

Terborg, J. (1976). The motivational components of goal setting. Journal of Applied Psychology, 61, 613-621.

Terborg, J. (1977). Women in management: A research review. Journal of Applied Psychology, 62, 647-664.

Thibaut, J., & Kelley, H. (1959). The social psychology of groups. New York: Wiley.

Thompson, A., & Strickland, A. (1990). Strategic management: Concepts and cases (5th ed.). Boston, MA: Irwin.

Tosi, H., Rizzo, R., & Carroll, S. (1990). Managing organizational behavior (2nd ed.). New York: Harper & Row.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. Science, 211, 453-458.

Van Maanen, J. (1976). Rookie cops and rookie managers. Wharton Magazine, 1, 49-55.

Van Maanen, J., & Schein, E. (1979). Toward a theory of organizational socialization. In B. Staw (Ed.), Research in organizational behavior. Greenwich, CT: JAI Press.

Vance, R., & Colella, A. (1990a). Effects of two types of feedback on goal acceptance and personal goals. Journal of Applied Psychology, 75, 68-76.

Vance, R., & Colella, A. (1990b). The utility of utility analysis: A case for objective metrics. Human Performance, 3, 123-139.

Vroom, V. (1964). Work and motivation. New York: John Wiley.

Wanous, J. (1980). Organizational entry: Recruitment, selection, and socialization of newcomers. Reading, MA: Addison-Wesley.

Wanous, J., & Colella, A. (1989). Organizational entry research: Current status and future directions. In K. Rowland & G. Ferris (Eds.), Research in personnel and human resources management. Greenwich, CT: JAI Press.

Wanous, J., Poland, T., Premack, S., & Davis, S. (1991). The effects of met expectations on newcomer attitudes and behaviors: A review and meta-analysis. Unpublished manuscript, WPS91-41. The Ohio State University, College of Business.

Webster's seventh new collegiate dictionary. (1965). Springfield, MA: G. & C. Merriam Company.

Weekly, J., Frank, B., O'Connor, E., & Peters, L. (1985). A comparison of three methods of estimating the standard deviation of performance in dollars. Journal of Applied Psychology, 70, 122-126.

Wegner, D., & Vallecher, R. (1977). Implicit psychology. New York: Oxford University Press.

Weiner, B. (1986). An attributional theory of motivation and emotion. New York: Springer-Verlag.

Weiss, H., Ilgen, D., & Sharbaugh, M. (1982). Effects of job stress on information search behaviors of organizational members. Journal of Applied Psychology, 67, 60-66.

Wexley, K., & Klimoski, R. (1984). Performance appraisal: An update. In K. Rowland and G. Ferris (Eds.), Research in personnel and human resource management, Vol. 2. Greenwich, CT: JAI Press.

Wise, L., Peterson, N., Hoffman, R., Campbell, J., & Arabian, J. (February, 1991). Army synthetic validity project: Report of phase III results. Tech. Report 922. Alexandria, VA: US Army Research Institute. (AD A235 635)

Wyer, R., Jr., & Srull, T. (1986). Category accessibility: Some theoretical and empirical issues concerning the processing of social stimulus information. In E. Higgins, C. Herman, & M. Zanna (Eds.), Social cognition: The Ontario symposium on personality and social psychology. Hillsdale, NJ: Erlbaum.

Zieky, M., & Livingston, S. (1977). Manual for setting standards on the Basic Skills Assessment Tests. Princeton, NJ: Educational Testing Service.