
Can You Measure That? Using Promotability Ratings as a High Risk/High Potential Assessment Tool

Presented at the IPMA Assessment Council Conference

By

Jennifer French and Kristine Smith
Darany and Associates
1250 San Pablo
Redlands, CA 92373
(909) 798-4475

June 10, 1999

Use of Promotability Evaluations

- P Internal promotional processes
 - P Typically used for public safety positions
 - P Similar in concept to oral examination
-

Why is it high risk?

- P Rater subjectivity (Typical rater errors)
 - P Multiple panels
 - P Raters' relationship with candidates
 - P Obtaining useful, standardized samples
 - P Questionable quality
 - P Increased potential for appeals
-

Why is it high potential?

-
- P Management wants input into selection decisions
 - P Opportunity to capture actual performance over a longer time period
 - P Opportunity to address behaviors tied to organizational values (organizational citizenship)
-

Project Background

Structured Chain of Command Evaluation for Police Captain

P Had been used in the past (Unhappy with results)

P New Police Chief - Organizational changes

P Focus groups evaluating promotional processes

Project Approach

Goal: Objectify critical performance behaviors valued in the new organization

P Literature review

P Link to new organizational values: Community policing model

P Implement as a component of a three part test battery

Critical Quality Improvement Areas

P Developing useable indicators tied to defined organizational achievements and behaviors valued by the organization.

P Rating panel training

Rating Dimensions

P Work involvement

P Personal Impact

P Leadership

Key Component: Structured Prompt

Goal: Standardize information used in rating; obtain most useful information for rating

P Described process

P Identified and defined dimensions

P Enabled raters to evaluate behaviors without over dependence on historical records

P Provided candidates an opportunity to provide input

Key Component: Structured Rating Process

Goal: Minimize subjectivity and rater error

P Use rating panels consisting of three raters

P Emphasis on purpose of the process

P Oral examination training model

P Additional types of error addressed:

< “Ownership” effect

< Exposure/contact effect

P Discussion of feedback process

Training Session Content

- P Overview of entire exam process
 - P Overview of evaluation process
 - P Discussion of information sources
 - < Relevance
 - < Balance
 - P Discussion of rating dimensions
 - < Practice categorizing “evidence” and information
 - P Rating scales and rating process
 - P Rater tendencies
 - P Frame of reference: examples, practice
 - P Defining ratable observations: specificity, relevance, recency
 - P Providing feedback
-

Results

P Client department pleased with results

< “This is the best we’ve seen”

P Obtained acceptable range of scores

< Converted score range = 60 - 100

P Interrater reliability estimate = .86

< Intra-class correlation method

STRUCTURED CHAIN OF COMMAND EVALUATION PROCESS PROJECT SUMMARY

INTRODUCTION

While organizations strive to develop the most effective predictors of performance, they also struggle with the aspects of performance that the assessment process does not capture. This is particularly true in the area of law enforcement promotional examinations. Though specific job knowledge and a variety of skills and abilities are effectively captured by traditional assessment processes, there are also some critical aspects of performance that may be missed. These might include commitment to changing philosophies such as community policing and actual behaviors demonstrating leadership and effectiveness on the job.

Additionally, in such cases, the individuals in the candidate group are known to each other and have had the opportunity to observe one another's performance. Thus, outcome on the selection procedure is inevitably compared to perceptions of job performance held by both candidates and supervisors. Inconsistencies between the two leads to the belief that to get promoted a person only has to perform well on the day of the examination and that performance over time does not matter as much as it should. This belief not only damages the credibility of the assessment process, but also impacts the morale of the organization.

It was these concerns that prompted the development of a structured chain of command evaluation (SCOCE) process for police promotional classifications. With the active cooperation and participation of a large city police department, a process was developed and implemented which effectively captured achievements and behaviors in areas valued by the organization.

BACKGROUND

The Police Department had adopted community policing both as a philosophy and as an organizational strategy. As a result, systemic changes were introduced throughout the organization to support this commitment. In the area of promotional processes, focus groups were established to evaluate process effectiveness and make recommendations for improvements. The implementation of some kind of promotability evaluation was strongly recommended. Over the years, the organization had used a variety of written and behavioral examination processes, including promotability evaluations. The latter had not been used in the more recent promotional processes because of the typical problems of score inflation, leniency and restriction of range. Though not pleased with the results of previous promotability evaluations, the department was reluctant to discard the process.

INSTRUMENT DEVELOPMENT

Key aspects of instrument development included incorporating focus group input into the job analysis process and conducting a literature review to integrate relevant research findings. The focus groups provided examples of critical behaviors leading to success or failure on the job with emphasis on the role of community policing. Literature in the area of performance evaluations was also utilized to identify those aspects of assessment that are unique to supervisory evaluations. For example, potential rating errors such as contact effect and ownership effect were identified and incorporated in rater training materials.

The final instrument addressed three performance areas through the use of behaviorally anchored rating dimensions. These dimensions included work involvement, personal impact, and leadership. Each dimension area was defined and included benchmark descriptions based on actual behaviors that were provided in the focus groups and likely to be observable on the job. Dimension benchmarks reflected performance expectations at the level of the position being applied for rather than the position currently held.

ADMINISTRATION AND EVALUATION

Candidates were provided with definitions of the dimension areas and given an opportunity to prepare a written statement of their achievements in each of the areas. This document was included along with historical file information to the chain of command panel. This panel consisted of the next three levels of the command structure. Adjustments in panel composition were made in instances in which the supervisory relationship had been in place for less than six months. The panels convened to review the candidates' submissions, file information and to share individually observed examples of achievements and behaviors within the dimension areas. Ratings were made independently, however, ratings were shared and discussed to minimize rating errors or inconsistencies.

Prior to convening, all individuals serving as raters participated in a rater training session. These sessions included discussion of the process, rating criteria, various information sources including issues such as relevance, balance, and recency, rater tendencies (errors), and frame of reference which included classifying and rating sample achievements and behaviors. Raters were also made aware of the candidate feedback process and their role in that process. It was considered essential that the raters understood that they were likely to be called upon to meet with the candidates and describe the basis of their ratings.

RESULTS

The initial administration of the process occurred without incident. The department was especially pleased with the outcome and there were no candidate challenges or appeals of the SCOCE process. The instrument produced a wide range of scores (converted scores ranged from 60-100) and acceptable reliability (inter-rater reliability estimate = .86).

As with any process, notes for potential improvements were made. The primary issue to be addressed in subsequent administrations involves the development of a system of checks and balances to both ensure rating criteria are uniformly applied and to increase candidates' comfort level with the objectivity of the process.

Darany and Associates, 1999