

MF-1

Manual for Developing  
Job Performance Measures

Robert K. Branson  
Principal Investigator

Gail T. Rayner  
Project Director

Penelope Fry  
Research Assistant

Center for Educational Technology  
Florida State University  
Tallahassee, Florida

In fulfillment of Contract # 78-DF-AX-0211 with the Florida Police  
Officer's and Training Commission.

May 31, 1981

U.S. Department of Justice  
National Institute of Justice

This document has been reproduced exactly as received from the  
person or organization originating it. Points of view or opinions stated  
in this document are those of the authors and do not necessarily  
represent the official position or policies of the National Institute of  
Justice.

Permission to reproduce this copyrighted material has been  
granted by

~~Public Domain~~  
~~LEAA, U.S. Dept. of Justice~~  
to the National Criminal Justice Reference Service (NCJRS).

Further reproduction outside of the NCJRS system requires permis-  
sion of the copyright owner.

Table of Contents

NCJRS

FEB 12 1982

	PAGE
Introduction.....	1
Chapter 1 - Job Performance Measures.....	2
Chapter 2 - Job Performance Measure Specification Sheets.....	53
Chapter 3 - Reliability and Validity.....	57
Chapter 4 - Legal Considerations in Test Development.....	61
References.....	68
Recommended Reading List.....	69
Appendix 1 - Sample Job Performance Items.....	70

82/334

## Introduction

Recent court rulings have indicated that employers must be able to demonstrate that their hiring, training, testing, and promotion practices are based on actual job related criteria. No longer can an employer require a job applicant or incumbent to demonstrate specific characteristics without some type of empirical data to support those requirements. The results of the Job and Task Analysis of Florida Law Enforcement Officers survey provide the necessary data base for developing tests that comply with current rulings and legislation. However, the data base must be translated into valid measures. Chapter I of The Job Performance Measures Manual provides a methodology for using task data to develop job related tests. Chapter II describes a format for developing the measures. Chapter III discusses reliability and validity of tests. Chapter IV describes the legal guidelines on tests and testing. The Appendix contains four sample Florida Law Enforcement related test items.

## Chapter I Job Performance Measures

### Overview

Job Performance Measures -- called JPMs for short -- are probably the most useful evaluation tools available to a law enforcement agency or organization. Their flexibility and utility is such that they can be adapted for use in a variety of law enforcement activities, such as the basis for screening applicants during recruiting; further screening in the selection for training; for interim and end-of-training tests during the training; and for the performance appraisal or evaluation on-the-job.

JPMs are the building blocks of a realistic and reliable testing method that will probably stand up to court scrutiny. JPMs are written at task level for tasks that have been identified by an empirically based job-task analysis. They are constructed to measure performance of tasks, under conditions and standards as close as can be achieved to those that actually exist on the job. As such, JPMs are fundamental to the development and control of any performance appraisal or training program. They provide compatibility and continuity in the various personnel management programs of an organization.

With JPMs you can achieve the best possible compromise between the requirements of direct performance measurement and the constraints that exist within your agency. Because of this attribute and the flexibility and utility described above, JPMs are a cost effective investment for any organization.

Before we get into the actual development of JPMs, you should be familiar with some of the important terminology of this manual. Here are some definitions and discussion of what the terms mean in relation to JPMs.

### Job

The duties and tasks performed by single workers constitute their job. If identical duties and tasks are performed by several individuals, they all hold the same job. The job is an organization or agency's basic unit for the personnel actions of selection, classification, training, and assignment.

### Duty

A duty is one of the major subdivisions of work performed by one individual. A job is made up of one or more duties.

The following are some of the characteristics of duties:

1. A duty is one of the job incumbent's main functions. It sometimes may be a particular job incumbent's total job.
2. A duty is a grouping of closely related tasks.
3. Duty requirements often are the basis for initial assignment to a job, for determining the qualifications required to perform in the job, or for determining requirements for post-assignment training.

Duties can usually be defined by asking a supervisor what he thinks are the 5 or 6 most critical factors of a job or what he would demand of a person being considered for a job. In the case of a clerk/typist, for example, the supervisor may say, "(1) Ability to operate equipment, (2) ability to route correspondence, (3) ability to maintain files, and (4) ability

to prepare correspondence." The job of clerk/typist, even after exhaustive analysis, probably will be found to consist of four major duties, i.e., OPERATING equipment, MAINTAINING files, ROUTING correspondence, and PREPARING correspondence.

### Task

Job analysis actually is accomplished at the task level. As you will recall, duties are actually clusters of tasks, the performance of which constitute the duties. Job analysis goes much deeper into job activity description at the task level than it does with the more general duty statements. A task is the lowest level of behavior in a job that describes the performance of a meaningful function in the job under consideration. Examination of the job at the task level allows the job to be described in sufficient detail to serve as the basis for a complete instructional system.

Task statements must be constructed carefully to assure that the final analysis yields usable job performance data. The following are characteristics of tasks and task statements:

1. A task statement is a statement of a highly specific action.

The statement has a verb and object.

### EXAMPLE

1. "Perform emergency repairs on vehicle" is not sufficiently specific to be a good task statement. To one individual, such a statement might mean performing such actions as "replace wiper blades" and "replace burned-out head lamp". Another person might think it means "replace water pump" or "change flat tire".

2. Also "inspect and repair exhaust system" is not sufficiently specific. However, "inspect exhaust system," might be one task and "repair exhaust system," might be another task.

2. A task has a definite beginning and end.

#### EXAMPLE

Such action phrases as "have knowledge of" or "take responsibility for" are not time-ratable and therefore should not be included in a task statement.

3. Tasks are performed in relatively short periods of time, i.e., seconds, minutes, or hours, but rarely, if ever, days, weeks, months, or years. Although no definite time limit can be set, the longer the period of time between the beginning and completion of the activity, the greater the probability that the activity is a generality or goal rather than a task.
4. Tasks must be observable in that by observing the performance of the job holder or the results of his efforts a definite determination can be made that the task has been performed.

#### EXAMPLE

"Understand legal principles" is not observable. Neither the process nor the results can be observed. (However, certain actions that require an understanding of legal principles can be observed.)

5. A task must be measurable; that is, in the real world, a technically proficient individual can observe the performance of the task or the product produced by the task and be able to conclude that the task has or has not been properly performed.

#### EXAMPLES

1. "Know how to" or "be able to" are not measurable. Neither are they observable.
2. "Assure success of operation" is too general to be measurable.
6. Each task is independent of other actions. Each task statement must describe a finite and independent part of the job. Tasks are not components of a procedure. In the eyes of a job holder, a task is performed for its own sake in the job situation. A task is either performed or not performed by any one job holder. The job holder is never responsible for only part of a task. If he or she is responsible for only a part of a work activity that would otherwise be defined as a task, the part for which he or she is responsible is the task.

TABLE 1.1

Samples of Good Task Statements

FUNCTION	SATISFACTORY TASK STATEMENT
Sorting items of mail into pigeon holes.	Sort mail.
Fixing carburetors.	Adjust carburetor Replace carburetor Rebuild carburetor
Deciding where to begin troubleshooting (specific electronic item).	Select troubleshooting strategy for specific item of equipment.
Establishing the objectives for a course.	Specify course objectives.

### Element

An element is the smallest "package" of behavior that has practical meaning. By "has practical meaning," we mean that further subdivision of the element is unnecessary since the element is the lowest part of a task that is useful in designing training and job performance measures. To be useful as a basis for developing instruction, step-by-step direction and guidance is required as to how the task is performed. The work activities that make up this step-by-step direction and guidance are the elements that make up the task.

The elements that make up each task must be determined for two reasons. First, since many of the tasks will be selected for training, the trainer must have sufficient details of the tasks to provide a solid basis for training. If individuals are going to be trained to do a task or be provided with job aids to help them perform a task, those who develop the training or job performance aids must know exactly how the task is done. (Note: Job Performance Aids (JPAs) are manuals, checklists, or any other devices--often attached to equipment--that assist individuals in performing certain operations.)

The second reason for determining the elements that make up a task is that some task statements look alike even though the tasks are quite different. Some task statements may have the same verb and object and only appear different when the elements are added. As an example, note that in Table 1.2, while the task statement is the same, the actual task is quite different for the different job levels. The elements that make up the task give a special "flavor" to the task at each job level.

TABLE 1.2  
Relationship Between Task, Job Level, and Elements

TASK	JOB LEVEL	ELEMENTS
Prepares Reports	Very low	Fills out logs Count units of materials Compute indices
Prepares Reports	Intermediate	Combine totals Integrate information Obtain concurrence
Prepares Reports	Upper	Check accuracy Release reports Interpret reports

### Predictive Validity

This is the quality of accurately predicting an individual's performance on the job from his or her performance on a test. Your JPM for a particular law enforcement task will have good predictive validity if someone who scores high on or passes your JPM can also perform the task out there on the street. Perfect predictive validity means that with no exception, anyone who passes your JPM can perform the actual task, and anyone who fails your JPM can't perform the task on the job.

There are four possible outcomes for a job task based JPM.

1. The person passes the JPM and also can perform the task on the job. This is a correct classification.
2. The person fails the JPM and also cannot perform the task on the job. This is a correct classification.

3. The person passes the JPM but cannot perform the task on the job. This, of course is a classification error.

4. The person fails the JPM but can perform the task on the job. This is also a classification error.

The intent of all testing should be to minimize both of the types of classification errors.

#### EXAMPLES

1. There are forms and reports required by just about every aspect of a law enforcement job whether it's one in the Highway Patrol, Municipal Police Department, Marine Patrol, or the Game and Fresh Water Fish Commission. A JPM for the task of completing a form or report is almost always identical to the task on the job. This means that the JPM you develop for such a task will have very high predictive validity. That is, if an officer or recruit can correctly complete a citation form, investigation report, or a radio log in your job performance measure, he or she can do it correctly on the job as well.
2. Another example of a job-related requirement for law enforcement officers is to give chase when in pursuit of a suspect, sometimes at high speeds in a vehicle, aircraft, or boat. While high speed chase is not technically defined as a task, it is a skill for which a JPM can be designed. For instance, high speed vehicle pursuit often requires that an individual drive at a high speed under a variety of road and weather conditions. Your JPM could be:

Action: Drive a vehicle while in hot pursuit

#### Conditions:

- a. on wet driving surfaces
- b. on gravel or sand
- c. around fixed obstacles
- d. at speeds of 45 mph or more
- e. around sharp corners

#### Standards:

- a. without losing vehicle control
- b. without injuring passengers
- c. without endangering pedestrians, other vehicles, or property
- d. without damage to the vehicle
- e. without losing the suspect

The predictive validity of this JPM should be reasonably high. If a law enforcement officer or trainee passed this JPM, he or she should be able to safely and effectively drive a vehicle in hot pursuit under actual conditions on the job. If the person failed the test, he or she may not be a good candidate for road duty if the ability to give high speed chase is a requirement.

Whenever you can measure a law enforcement task as it is actually performed on the job, you can determine the predictive validity of the related JPM. In the example above, your primary objective for your JPM is to achieve the highest possible predictive validity but keep testing factors, such as safety, time, and cost within acceptable limits for your agency.

Paper and pencil tests can often have high predictive validity as job performance measures, and they are usually very economical in terms of cost, time, safety, and resources required. Two types of paper and pencil tests are particularly suitable as JPMs.

The first type is designed to measure actual paper and pencil tasks, such as a patrol supervisor's task of reviewing a patrolman's accident report prior to its authentication. This is a real world paper and pencil task, and your JPM for it would, therefore, be a paper and pencil test. Any task in a law enforcement job which requires paper and pencil should have a corresponding paper and pencil test as the JPM in order to have the greatest predictive validity.

The second type of paper and pencil test is appropriate for tasks in which mental discrimination is the most important aspect of the task. Such a task might be decide types of evidence to search for at the crime scene. Your JPM would provide the trainee or officer with adequate information

about physical characteristics of the crime scene, the crime committed, and any witness testimony. Six types of evidence might be offered as possibilities with one type or combination of several types clearly being the most appropriate in light of the conditions in the background information. Such a JPM would be a medium high fidelity test of an individual's ability to make a correct decision on what types of evidence to look for at a crime scene--yet it is presented in a multiple choice paper and pencil format.

There is a difference between this kind of a multiple choice performance test and an ordinary multiple choice test. A multiple choice performance test usually provides much more background information related to each question, and it also offers a varying number of answer options. How many options are offered depends on the number of options available in the real world. The option of doing nothing, or in this case deciding not to look for any evidence, is offered in the multiple choice performance test but rarely included in the standard multiple choice test.

#### Physical Fidelity

This is the quality of closely approximating the actions, conditions, cues, and standards of a task through a Job Performance Measure. Your JPM for a law enforcement task will have high physical fidelity if the JPM performance is very close to the actual task performance. In the first example of the previous section, the JPM for completing a form or report has the highest possible physical fidelity because the JPM performance is identical to the task performance. The second example of a JPM for high speed chase has lower but still reasonably high physical fidelity. With these two examples, however, predictive validity is more important than physical

fidelity. If a paper and pencil test or some other testing approach with much lower physical fidelity can be shown to have high predictive validity and is preferable in terms of cost, time, and safety, use it. We say this because high physical fidelity is not as critical as high predictive validity in assessing the quality of a JPM. Although high physical fidelity is usually closely related to high predictive validity. However, whenever predictive validity cannot be used, you should work toward high physical fidelity in your JPM.

#### EXAMPLE

Any law enforcement officer must be able to return fire with and disable an armed assailant according to a specific procedure and with a high degree of accuracy. This skill is a vital one for the protection of innocent citizens as well as for the safety of the individual officer. To measure whether the JPM has high predictive validity would require that an officer actually exchange fire with another individual in order to duplicate the real world task. Obviously, such a test is out of the question because of safety factors. (Also, few individuals will volunteer to be shot as part of a performance test.)

In the example above, predictive validity cannot be used to determine the quality of the JPM for the task. Since there's no way to test the task under real world conditions, there's no way to test predictive validity either. In such a case, the quality of the JPM is often based on the degree of physical fidelity between the JPM and the task. Your JPM for this task will have to be based on performance on a firing range. There will be some physical fidelity, but not a high degree because the conditions, cues, and standards of the task cannot be duplicated.

Figure 1.1 is a graphic picture of some of the possible degrees of fidelity between your JPMs and tasks.



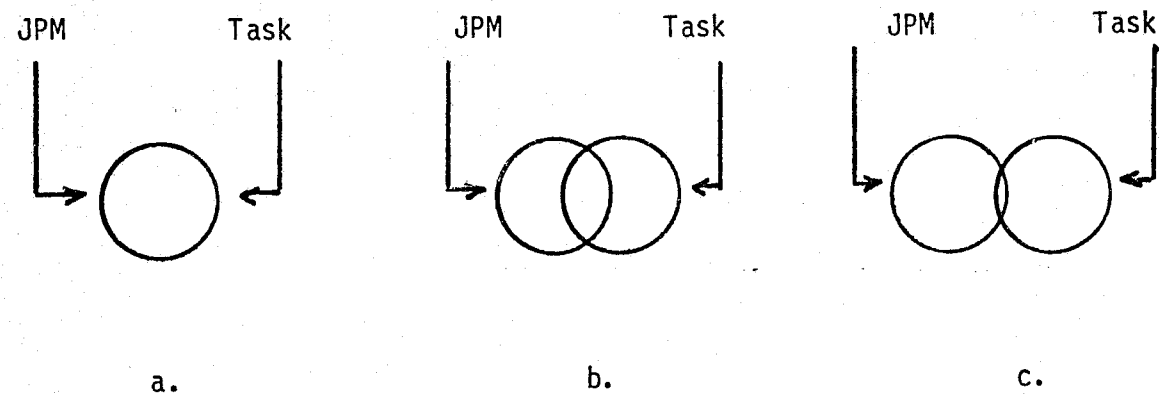


Figure 1.1. Degrees of Physical Fidelity Between JPMs and Tasks

The three examples in this figure could represent the three examples used earlier. In the first one, the JPM for completing reports or forms was identical to the actual task. This is like "a" in the figure. In the second example, the JPM for high speed chase driving has reasonably high physical fidelity but its performance is not identical to the task--this is like "b" in the figure. In the last example, the JPM for returning weapons fire will have to be considerably different from the actual task. The physical fidelity, of necessity, will be lower. This is like "c" in the figure.

#### Simulation

Simulation, in terms of JPMs, is any change from reality or any imitation of reality. Whenever a performance test can't be given under real world conditions, some form of simulation must be used.

#### EXAMPLE

In the previous example the task required the law enforcement officer to return fire and disable an armed assailant. We already pointed out that real world conditions can't be

duplicated because the officer would be in jeopardy and we can't put the officer under fire and endanger his or her life just to test the performance. Also, we can't provide an "attacker" who will be willing to get shot. Since we won't endanger the officer's life by putting the officer under fire, we can't provide a high fidelity initiating cue. There is no way to supply the same element of danger; therefore, the JPM will be more like "c" in the previous figure. However, we can simulate the initiating cue and have the officer go through some or most of the steps. The initiating cue may be an order to fire at a pop-up target on a weapons course that will fall if hit with a disabling shot.

In this example, the stationary and pop-up targets and any other imitations of the real world task that you design are simulation.

#### Unitary Tasks

A unitary task is any task that is always performed in exactly the same way with exactly the same inputs. Construction of a job performance measure of the task is always the same.

#### EXAMPLE

"Disassemble a .38 Police Special" and "assemble a .38 Police Special" are unitary tasks. The input, the weapon, is always the same and the task of assembling or disassembling is always performed in the same way.

#### Multiple Tasks

A multiple task is one that has a number of possible inputs. There are two types of multiple tasks based on whether or not the performance varies according to the input. One type of multiple task is always performed in the same way no matter what variations occur in the input.

#### EXAMPLE

Multiplication of three-digit numbers by three-digit numbers is a multiple task. Even though there are almost one million pos-



sible inputs or combinations for the two sets of three digit numbers, the multiplication is always performed following the same mathematical procedure.

In this example it would be impossible for you to test all possible three-digit number inputs just to make sure an individual knows how to multiply all of them. Therefore, your JPM must measure a representative sample of the task inputs.

The other type of multiple task is one in which the inputs vary, and the task can be performed differently depending on the input. This means that the input is a cue that initiates an appropriate response.

#### EXAMPLE

One of the sensitive tasks of a law enforcement officer is to apprehend a suspect. The procedure for performing this task depends heavily on the input cues. For example:

Is the suspect armed?  
Is the suspect under the influence of drugs or alcohol?  
Is the suspect in a vehicle or on foot?  
Is the suspect alone?  
Is the suspect male or female?  
Is the officer alone?  
Is the officer in a vehicle or on foot?

Answers to questions like these will determine how the task will be performed.

As with the first type of multiple task, your JPM for this task cannot measure the total task. Instead, you must measure a sample of the possible variations. However, your JPM must do more than measure the adequacy of the performance. You must also measure the appropriateness of the performance--whether the performance on the JPM was the correct response to the particular cue or cues.

#### Constructing Job Performance Measures

The JPMs you are going to develop will look much like task statements from the Task Summary Sheets (TSSs) as in Figure 1.2.

Like tasks, JPMs are statements of actions to be taken. Your JPMs will also state test conditions, cues for particular responses, performance standards, and the elements of the total JPM performance. Later, we will talk more about these parts of a JPM. Right now, we're going to talk about constraints.

There are eight major steps for developing and validating JPMs. You should go through these steps for each task. That is, take one task, determine any testing constraints, make the decisions in each of the other steps and develop the JPM. Then take your next task and do the same. Your first decisions are only tentative; after you work through a group of tasks you'll probably want to go back and make changes. After you have developed JPMs and scoring procedures for tasks, you will validate the JPMs for predictive validity and physical fidelity.

#### Procedures

Since you want your JPMs to be high in predictive validity and physical fidelity, your goal is to make your JPMs identical to the tasks they are intended to measure. Ideally, a JPM would consist of observing a law enforcement officer while performing the task on the job and noting whether he or she met the job standards. However, as you probably know, too many things can interfere, so your first step in JPM development is to:

This page for Figure 1.2

This page for Figure 1.2

### 1. Determine testing constraints.

Constraints are factors which limit your capability to observe and test performance--such as time, manpower, costs, and facilities or equipment. These constraints are all interrelated--time availability, manpower availability, equipment availability, and costs (financial availability) often are all different aspects of the same problem. You must analyze a task to determine just what constraints apply to that particular task. The more constraints you identify, however, the more of a compromise your JPM must become between your testing situation and the real world task, and compromise in a JPM usually means lower physical fidelity. Here are the practical constraints that can force a change from a high fidelity JPM to a lower one.

- Time. Availability of time is a common constraint. Often it's not practical to test a task as it is stated--you just do not have enough time available to you.

#### EXAMPLE

The law enforcement task of Conduct intelligence gathering operations is one that can occur over an extended period of time, often ranging from several hours to several weeks or months. In most situations the task would take much too long to test, so your JPM must be modified to permit testing in less time.

Inventory equipment is another example of a task that may require too much time for testing. Your JPM for this task would most likely test an officer's ability to inventory selected items rather than a whole room full of equipment.

In most performance testing situations, you must place time limits on test administration; this in turn limits the amount of time that can be spent on each of your JPMs. Therefore, if performing some of the tasks

requires more time than you have available for testing, your JPMs will have to sample elements of those tasks.

- Manpower. Availability of manpower necessary for performance testing is another common constraint.

#### EXAMPLE

If you want to test performance of the task Confront or monitor groups (such as demonstrators, rioters, or crowds) you're going to need some people to play the parts of the demonstrators. You're probably going to need several law enforcement trainees, or officers too, since in most situations a single officer does not confront a crowd but is a member of a crowd control or riot control team. If all these individuals are not available, you will have insufficient manpower for testing the task under normal job conditions.

In the example above, it would be preferable for you to arrange simultaneous JPMs to test a group of trainees or officers in a crowd control or riot control situation. However, each individual would have to be scored separately on performance of the task. In this way you make efficient use of the personnel to be tested but require more personnel to support the test.

- Cost. Cost is almost always a constraint when you develop JPMs. You must keep the cost of test administration within the limits of your agency's testing budget.

#### EXAMPLE

Consider the task Locate downed aircraft, off-road vehicle accidents, or swamped or capsized vessels. Under most circumstances, it would be entirely too costly (and unreasonable) to deliberately crash an aircraft or vehicle or sink a boat just to test an individual on performance of the task. There are other more practical means of testing such a task such as simulating a crashed aircraft or vehicle. Occasionally an

agency may have a confiscated vehicle or vessel which could be used to simulate a crash site or a capsize situation on a permanent and reusable basis.

- Facilities or equipment. Sometimes sufficient equipment or facilities are not available for your test administration. This is often true for sophisticated equipment and highly specialized facilities.

#### EXAMPLE

Non-availability of specialized equipment can be a serious constraint for law enforcement tasks such as Estimate speed of moving vehicles (via radar device) or Select or set up electronic surveillance devices. A JPM with physical fidelity cannot be developed unless the equipment will be available for testing use. That is, a law enforcement officer probably could demonstrate how to estimate a vehicle's speed with only a mock-up of a radar device, but it wouldn't be as accurate a performance test--or have the physical fidelity--of using the actual radar equipment.

A serious facilities constraint can exist for a task such as Inspect Marinas, launching ramps, or terminal facilities. It's doubtful that training facilities have marina or terminal facilities in place for training and testing use.

If you have many tasks whose testing would tax your facilities or equipment beyond manageable limits, try to select those tasks for testing which will sample the officers tasks while causing the least inconvenience. Other tasks may be simulated in some way. For tasks such as the inspection task above, you may have to simulate important features of the specialized facilities or arrange for actual on-site testing.

- Other constraints. Some of the other constraints you may have to deal with are:

logistical constraints

legal constraints

ethical constraints

safety constraints

inconvenience factors

Remember that constraints are related in most cases. The practical constraint in the example of the electronic surveillance devices was given as equipment availability, but it could also have been given as a cost constraint. Considering the limitations in equipment, personnel, time, space, safety requirements and other factors that you may face in your agency, complete physical fidelity may not always be practical or even desirable. If such is the case, the JPMs you develop must be the best possible trade-off with reality. If you have no constraints, however, you can use the actual job task as your JPM without any tradeoffs.

#### 2. Determine JPM conditions.

JPM conditions are based on actual task conditions. By conditions we mean those on-the-job conditions that can greatly affect the performance of a particular law enforcement task. The job conditions should be listed on the TSS. Such conditions may always be the same or they may vary at different times. Conditions that remain the same are factors that can be controlled both on the job and in your JPM. Conditions that vary are those factors that cannot be controlled, such as weather, the amount of natural light, temperature, etc. All of these conditions should be listed in your JPM.

## EXAMPLE

Action: Testify at trials, hearings, or grand juries.

Conditions: In courtrooms, chambers, or attorney offices  
In uniform or off-duty attire  
Using personal notes

In this example the task conditions are fairly constant; that is, the task will always be performed under the same basic conditions. In such cases your JPM conditions probably can match the actual job conditions. For other job tasks, however, you must sample the conditions for any one JPM.

## EXAMPLE

The law enforcement task is: Track humans or animals. The constant condition is:

Using a map of the area

The variable conditions are:

All types of weather  
Across all terrains found in Florida  
At any time of day or night  
Alone or as part of search party  
Using a lensatic compass (not always necessary)

In developing a JPM for this task, you must list all the conditions of the task that will make a difference in performance of the JPM. So, in addition to the fixed condition of using a map of the area, you must list:

- All possible types of weather--dry, damp, rain, snow clear, cloudy, no wind, light wind, high wind, cold, below freezing, ice, temperate, hot, etc.
- All types of terrain--flatlands, cities, forests, towns, villages, hills, swamps, salt marsh, palmetto scrub, etc.
- Any time--daytime, nighttime with stars, no stars, moon, no moon, etc.

If your testing constraints preclude using all the conditions, you will have to use a sample of the conditions. You should include those conditions that you think will best reflect an ability to perform the task under all the possible conditions. Also you must choose conditions that are realistically available. There are few cold days in Florida in the summer and not all areas of the state have all types of terrain mentioned.

3. Determine JPM cues

A cue is the state of events or the signal in a job environment that determines when a job incumbent performs a particular job task according to a particular procedure. The job task initiating cue should be described on the TSS. Often your JPM cues will have to be different from the actual job cues even though that lowers the JPM fidelity.

## EXAMPLE

A law enforcement officer may have to perform mouth-to-mouth resuscitation in the task Resuscitate persons (such as drowning, electrocution, or suffocation victims). The cue for the performance is that the person is not breathing. Realistically, you are not going to be able to reproduce this exact cue for a JPM. While changing a cue does reduce the fidelity of the JPM, the best you can probably do is to use an artificial cue. One common way to do this is for the test administrator to verbally cue the performance in this way: "The injured person (dummy) you're treating at this accident scene is unconscious. You find a weak pulse; then you notice the person has stopped breathing. Take immediate action."

Some cues are more important than others--or critical--because they signal a particular response which will determine the success or realistic task performance. Critical cues in your JPMs should be as close to the job as you can make them within your agency's testing constraints.

## EXAMPLE

The task of Perform emergency repairs on vehicle or vessel might be one that is pretty simple if an individual knows which repair to make. The critical cue of the task, then, is the particular mechanical signal or condition that indicates which repair to make. If the JPM cue for this task is a test administrator statement such as, "Assume your vehicle requires such-and-such repair," your JPM will have very low fidelity and probably very low predictive validity, too. In this case, it is the law enforcement officer's ability to diagnose the problem or respond to a critical cue in the vehicle or vessels condition--that determines whether he or she can successfully perform the task.

4. Determine if the JPM will measure a task product, process, or both.

Task performance can be measured in several ways. You can measure a task product which is the final outcome of performing a task because you can easily observe and inspect the task product.

## EXAMPLE

Prepare daily log of law enforcement activities.

Sketch accident or crime scene.

In both of these examples there is a concrete, observable product from performing the tasks--a log and a sketch--and the products signal the completion of the tasks. Your JPM for a task should measure the product of task performance when:

- The product can be observed and inspected.
- The process by which the product was produced can not easily be observed.
- The process is not as important as the product.

Task performance can also be measured in terms of the process involved. The process is the series of actions or procedures taken by an individual

in reaching the desired goal or end of the task. The completion of the process signals the end of the task, but the process itself often leaves no record. Your JPM for a task should measure the process of task performance when:

- Performance of the task does not leave a product that can be easily observed or measured.
- Failure to use the correct process in task performance could result in injury to the performer or others, damage to equipment, property, etc., or violation of the law.

## EXAMPLE

Advise persons of rights (per Miranda).

Interrogate suspects.

Administer first aid to injured persons.

In the first two examples above, the failure to use the correct warning or interrogation process can result in violation of individual legal rights. In the third example, failure to use proper first aid procedures can result in further injury or death to the person being treated. In all three examples, there is not an easily observed or measured product. Although a signed statement acknowledging Miranda advisement is obtained by the law enforcement officer, it can be thrown out by a judge if the process used in obtaining it is proven to be illegal. The same is true, of course, for statements or confessions obtained by illegal interrogation procedures. It can also be difficult to measure whether an individual has responded to first aid even though correct procedures were used. Therefore, your JPMs for tasks like these should evaluate process.

Sometimes you must measure both product and process as task performance output. Often your JPM will evaluate a process which results in a product. JPM evaluation of the product as well as the process will provide important feedback on errors in the process which may affect the product. In certain cases, processes in a task may be critical if they insure individual safety or prevent equipment damage.

## EXAMPLE

Conduct field, frisk, or pat down search.

In this example, a law enforcement trainee or officer could correctly perform the different search procedures yet fail to find a concealed weapon on the suspect. Evaluation of the product--in this case the finding of the concealed weapon--and evaluation of the process would insure that legal restrictions on the search were observed and provide feedback on performance of the process.

## EXAMPLE

Transport persons (such as injured, deceased, or lost persons, mental patients, prisoners, or suspects).

The law enforcement officer charged with this task may get the persons from point A to point B as required, but violate laws and safety regulations in the process. In such cases, product evaluation by itself is not always adequate.

5. Determine JPM standards.

JPM standards are based on task standards. By standards we mean the acceptable quality of task performance in the real-world job environment. Task standards are the statements of how well the tasks must be performed, if and when performed. Performance measures must have these same specific standards that are applied to all people taking the test so that each person is rated in the same way as all the others.

JPM standards will generally evaluate products in terms of accuracy, tolerances, completeness, format, clarity, number of errors, and quantity. Processes will normally be evaluated in terms of sequence, completeness, accuracy, and speed of performance. Most of your job task standards should have already been established and documented on the Task Summary Sheets. Such documentation can provide a starting point for you to derive JPM standards that reflect actual law enforcement requirements. You can partially

derive standards for some tasks from the characteristics of the equipment required to perform the tasks. Other task standards are implied such as "accurate and complete", "submitted on time", and "correct solution".

JPM standards usually include time limits and required procedures; some describe the product required and criteria for evaluating that product. Time limits simply mean that the law enforcement officer must perform correctly within a certain time. The time limits may be set by the task.

## EXAMPLE

In the task Confront or monitor groups, use of teargas may be an element of performance tested by a JPM. After the pin is removed from a teargas grenade or canister and the handle is released, there is a fixed number of seconds before the grenade or canister releases the gas. The standard for tossing it must be well within that time limit.

Some standards, however, have time limits that are not job related. These are standards for tasks where there can be a delay between the time the need for task performance becomes evident and the time the actual job performance must begin.

## EXAMPLE

The task, organize special operations is not one that would normally have to be performed within a specific time limit, such as five hours, two days, one week, etc. The time limit on-the-job would depend on the nature of the target of the special operations. The rescue of a hostage might require that the task be accomplished very quickly. If the target were an established drug smuggling ring, however, the task might be performed within a deliberately longer time period.

You, as a law enforcement officer trainer or evaluator, most likely can not afford to test one such task over an extended period of time. Therefore, you have to impose a time limit standard for testing purposes. In such



situations you have to be very careful that the time limit you impose is reasonable in terms of actual job requirements. Unrealistically high standards of any kind can be very costly to you in your performance testing.

#### EXAMPLE

In the task Plot boat position using charts or equipment, the actual job task standard may predetermine the JPM standard. Probably only a fraction of an inch in error can be allowed in plotting such a position since any deviation may be magnified in terms of miles on the water. Therefore, to allow the position plot to be useful in locating a vessel on the water, it must be accurate to the specified degree on the chart.

In this task the job standard would dictate the performance measure standard, including whatever margin of error that can be allowed.

Standards for tasks that call for a procedure require that certain steps be performed, perhaps in a specific order. If a law enforcement trainee or officer omits such a step, or performs the step out of order (when order is important in the procedure), then that person does not meet the standard. The idea behind this type of standard is that there is no such thing for some skills as performing them half right or half wrong. As you are well aware, this is true of many law enforcement tasks. An officer either knows or does not know when and how to give a suspect the Miranda advisement. If order is not critical to the performance procedure, then your JPM can allow an individual to vary the order of the steps.

Some of your JPM standards may include requirements related to the final products from task performance. For certain tasks the quality of the final product may be more important than the procedures followed in producing the product. However, if an individual follows a certain procedure,

the chances of the final product meeting the standard will be much greater. Therefore, your JPM standard may measure both the product and process.

#### EXAMPLE

For the task Fingerprint persons, the final product is most important. The prints must be clear, complete, and correctly identified. Although there could possibly be variation in the method of taking the prints, law enforcement practical experience dictates a "best way" to fingerprint a person in order to obtain the best results. Your JPM might measure the procedure as well as the product for this task.

Some performance standards are highly qualitative--or subjective--in nature, such as those for court testimony, briefings, staff studies or tactical plans. It is difficult to break down elements of effectiveness for an oral briefing that can be objectively scored on a right-wrong basis. Several different tactical plans might be good; some may be better than others; one may be the best of all. Evaluation of these tactical plans and the oral briefing, however, is judgmental and qualitative. You must be careful when you evaluate such task performances to be sure that you do not fall into the trap of measuring only the most easily scored parts of the tasks. You must develop quantitatively measurable standards in your JPM for the other qualitative or judgmental portions of the tasks. You must also include the more difficult to measure--or qualitative standards in your JPM as in the following example.

#### EXAMPLE

- The room is clean.
- The equipment is correctly stowed in the proper lockers.
- The locker doors are closed and locked.
- Authorized persons have keys to the lockers.

# 6. Determine if all or part of the task will be tested.

Your decision to test all or part of the performance of a law enforcement task depends mainly on whether the task is unitary or multiple. For unitary tasks--those with only one officially recognized procedure--and the same inputs for all parts of the task are usually tested.

## EXAMPLE

Fingerprint persons is a unitary task. The input (a person to be fingerprinted) is always the same, as is the procedure. The task would be performed in roughly the following manner:

1. Enter the appropriate individual identification information on a standard 8x8 FBI fingerprint card.
2. Place the card in the card holder.
3. Apply ink to the glass or metal inking plate and spread it with a roller to the proper consistency.
4. Position the inking pad.
5. Check the person's fingertips to make sure they are clean and dry.
6. Position the subject.
7. Starting with the fourth finger of the person's left hand; ink the finger pad by rolling it from side-to-side, left-to-right on the inking pad.
8. Make a print of the finger in the pre-labeled space for the fourth finger, left hand, by rolling the finger pad in the same side-to-side motion on the card.
9. Follow the same procedure for the remaining fingers of both hands in the following sequence: left hand--third finger, second finger, first finger, thumb; right hand--thumb, first finger, second finger, third finger, fourth finger.
10. Take palm prints, if required.

11. Check to make sure that fingerprints are full, not partial; deltas and cores are legible; hands and fingers are not reversed or printed in wrong location.
12. Refingerprint the person, if necessary to get good prints.
13. Have the person clean his/her hands with special cleaner, towels.
14. Have the person sign the card.
15. Make any necessary special notations on the card.
16. Place fingerprint card in person's file.
17. Clean the inking plate.

In this example, all of the task procedure is tested because it is unitary, and your JPM is basically identical to the task.

Multiple tasks are those that have a variety of inputs. One kind of multiple task may be performed in the same way no matter what the different inputs are:

## EXAMPLE

Multiplying three-digit numbers by three-digit numbers may be required for calculating vehicle speed in the task Estimate vehicle speed using physical evidence, mathematical formulas, or graphs. There are almost one million possible combinations of 3-digit numbers multiplied by three-digit numbers. Obviously, you can only measure a representative sample of these task inputs with your JPM.

The second kind of multiple task is one that will be performed differently depending on the different inputs. In this sense, the input is a cue because it initiates a particular task response.

11. Check to make sure that fingerprints are full, not partial; deltas and cores are legible; hands and fingers are not reversed or printed in wrong location.
12. Refingerprint the person, if necessary to get good prints.
13. Have the person clean his/her hands with special cleaner, towels.
14. Have the person sign the card.
15. Make any necessary special notations on the card.
16. Place fingerprint card in person's file.
17. Clean the inking plate.

In this example, all of the task procedure is tested because it is unitary, and your JPM is basically identical to the task.

Multiple tasks are those that have a variety of inputs. One kind of multiple task may be performed in the same way no matter what the different inputs are:

#### EXAMPLE

Multiplying three-digit numbers by three-digit numbers may be required for calculating vehicle speed in the task Estimate vehicle speed using physical evidence, mathematical formulas, or graphs. There are almost one million possible combinations of 3-digit numbers multiplied by three-digit numbers. Obviously, you can only measure a representative sample of these task inputs with your JPM.

The second kind of multiple task is one that will be performed differently depending on the different inputs. In this sense, the input is a cue because it initiates a particular task response.

#### EXAMPLE

The inputs for the task Extinguish vehicle or vessel fires are obviously varied. The type of fire can vary (electrical, oil or gas, wood or other combustible material) and the location can vary (vehicle or vessel), also the type of fire fighting equipment may vary. Each type of fire requires different fire-fighting procedures. The factor of vehicle or vessel can also make a difference in task performance.

You may not have sufficient resources, time, personnel, etc., to measure performance for each of the possible inputs. Therefore, you have to select the most appropriate one(s), such as testing a deputy sheriff on extinguishing an oil or electrical type fire in a vehicle, or testing a Marine Patrol officer on extinguishing the same type(s) of fires on a vessel. If you can only test for one type of fire at a time, you can develop alternate forms of your JPM, including each of the possible inputs. When you are constructing alternate forms of your JPM, look for the total number of inputs that will cause all possible procedures to be followed for a task. This specific number of inputs will be your pool for developing alternate JPMs.

#### 7. Determine simulation requirements.

Once you have determined your testing constraints, and your JPM conditions, cues, and standards, you have essentially determined your simulation requirements. Your JPMs will require some kind of simulation, or function as simulation, any time they are changed from reality. If the JPM is identical to the real world task performance, then the JPM requires no simulation. Besides the JPM simulation we have already discussed, such as

simulating real world cues with artificial cues, you may have other simulation requirements. You may wish to simulate:

- Part of a system. You may have to test performance related to a particular piece of equipment and you may decide to create a mock-up of that equipment. The mock-up could be low to high fidelity and non-operational to fully operational according to your testing needs.
- Operation of a system. You may wish to simulate system operation using a "talk-through" technique which allows identification of major decisions, required actions and alternatives, and information gaps. You may also have the capability of using a complex simulator or a computer to simulate operational activities under numerous conditions.
- The environment in which the system will operate. You may have a requirement to determine under what conditions people will have difficulty operating the system. You can do this by having both experienced law enforcement officers and novices operate the system using simulated inputs. Then you can measure the problems encountered.

Once you determine that you need simulation, decide on what type. For some JPMs, the resemblance will have to be very close to the real world. In other JPMs, you will find high fidelity simulation is not necessary for making correct classification decisions.

In short, look at your list of actions, standards, conditions, cues, and constraints. What can you realistically do in your JPM? What should you simulate? If you can test the whole task as is, do it. If not, develop

a part-task test with simulation as necessary that will predict job performance. You will validate your JPMs and have a chance to revise them later.

#### 8. Develop sampling plan.

So far in your JPM construction process, you have identified task-related conditions, cues, standards, and constraints; you have made a decision on testing whole or parts of each task; and you have specified what outputs (process or product) you will measure. For all those factors listed above that are related to a unitary task (a task that presents only one option), you should use that factor. When the task you are testing involves multiple possible actions, conditions, or cues, you must decide if all of the factors are equally likely to occur and if there is the same criticality involved with the occurrence of each one. If they are equally likely to occur and equally critical, sample equally from them until each is included in one alternate form of the JPM. If the likelihood of occurrence or the criticality is unequal for some factors, then your sampling plan should reflect that.

Several of the above examples use the term "alternate form of the JPM." These alternate forms are equal but different versions of the JPM. Constructing alternate forms makes it possible for you to include all of the important variables without making any single test too long to administer. However, sometimes, it is not practical to construct alternate forms.

#### 9. Document JPMs.

Once you have completed the previous steps, your JPM will be half complete. All that you have left to do is document your decisions and

develop a scoring procedure for the JPM. Documenting JPMs is quite similar to documenting tasks. As you have already seen, JPMs have conditions, cues, and standards just as tasks do. JPMs also have procedural steps that often are similar to task elements. We will use a Job Performance Measure Specification Sheet in the examples in this section and in Appendix 1. Figure 1.2 is an example of a JPM documented on such a form. By attaching your JPMs to the TSS you will have a complete package of information to use as a basis for fine-tuning your evaluation program. See Chapter IV for more discussion on JPM documentation.

#### 10. Develop Scoring Procedures.

There are two basic scoring procedures you can develop to provide a record of a job incumbent's or trainee's performance. Your approach will depend upon the nature of the task you are testing, i.e., is the output of task performance a product or process or both? To develop scoring procedures you must determine the most appropriate procedures for recording the adequacy of the product, process, or combination of both. From a scoring point of view, the best JPM is one that allows an individual to produce a record of his or her tested performance, such as the holes in a range target, a clear set of fingerprints, or an accurate accident diagram. These outputs can be objectively scored by an evaluator and they also have high physical fidelity with the actual job tasks. (Hits on a practice target are a record of task performance with only somewhat less fidelity than gunshot wounds in a suspect.) Some products of task performance, however, may require evaluation by a qualified supervisor before the performance can be judged.

Some tasks do not provide an observable record of performance. These most often involve a process output--the performance of a series of procedural steps. Knowledgeable supervisors or evaluators will have to either observe and judge the performance being tested or inspect a final product from the process in order to determine the adequacy of process performance. Here are some guidelines for scoring product and process adequacy, selecting appropriate rating scales, and detecting common scoring errors. We will discuss some techniques in each area which should help you develop and document scoring procedures for your JPMs.

#### Product Rating Method

You have already seen that a product output of task performance is usually a tangible object. This makes product rating easier to describe than process rating. To develop a scoring--or rating--procedure for a task performance product, you need to identify all scorable characteristics as accurately and specifically as you can. Scorable characteristics are those that distinguish a satisfactory product from a product that is not satisfactory. One of the most common ways to measure a product is to compare it with an example of the desired product, such as comparing fingerprints taken for a JPM with a set of perfect prints. Another way to measure a product is by some kind of instrument. For this kind of scoring you must identify characteristics to be measured and provide specific measurement instructions for the scorer. On the basis of your JPM standard, you must then decide what is a qualifying score. To insure that your standard does provide clear guidance, test it, and then revise it until does.

### Process Rating Methods

Process output of task performance is often harder to measure and score. A good procedure for process measurement and scoring will provide your scorer with explicit directions on what an individual being tested should do at each stage of the JPM. This usually takes the form of a checklist with detailed step-by-step descriptions of the process by which the JPM should be performed. Breaking task performance into several observable elements through the use of a checklist greatly reduces measurement error and provides an easily-used basis for scoring. The example in Figure 1.3 shows a portion of a checklist for rating instrument flying proficiency. The scorer who uses the checklist will indicate whether each step is completed in a satisfactory or unsatisfactory manner.

Reliability is usually high in checklist rating because of the nature of the decisions required and specificity of the items listed. Usually, the more specific your items are on the checklist (such as in Figure 1.3) and the more items you list, the higher your measurement reliability will be. A general rating category such as in the example below is just about worthless because it is not specific and is inadequate as a single rating factor.

#### EXAMPLE

"Follows pre-flight safety procedure," is too general; "makes the visual inspection of aircraft," "uses pre-flight checklist," and "starts engine after starting signal" all provide a much better opportunity to make reliable ratings of performance because they are very specific and greater in number.

### Select Rating Scales

#### Numerical scales

A numerical rating scale divides performance into a fixed number of points, such as between one and five with one indicating low performance

-CHECKLIST-	
INSTRUCTIONS: If the performance is satisfactory place a + sign in the space provided. If the performance is unsatisfactory, place a - sign in the space.	
1. Maintains constant heading . . . . (within 50° of course)	<input type="checkbox"/>
2. Maintains constant altitude . . . . (within 50 feet)	<input type="checkbox"/>
3. Can make a timed turn (gyros . . . caged) (within 10° of new heading)	<input type="checkbox"/>
4. Can make a steep turn (within . . . 50 feet of altitude)	<input type="checkbox"/>

Figure 1.3. Example of Checklist of Rating Proficiency in Instrument Flying

and five indicating high performance. The number of points on the scale depends primarily on the ability of your scores to differentiate between degrees of performance. Most people are able to make at least five such distinctions, but few trained scorers can reliably make more than nine distinctions. As a result, most numerical scales have five to nine points.

### Descriptive scales

This rating scale uses descriptive words and phrases to indicate levels of performance ability. The example in Figure 1.4 is a descriptive scale for rating navigational ability, with five levels of ability described. The different degrees of performance can be varied to suit whatever performance is being tested.

INSTRUCTIONS: Place a check mark in the scale above the word that most accurately describes the officer being rated.

ACCEPTABLE	FAIR	GOOD	EXCELLENT	SUPERIOR

FIGURE 1.4. Example of descriptive scale for rating navigational ability

For example, suppose a Marine Patrol operations officer wants to evaluate the navigational ability of Marine Patrol officers. He feels that all of them satisfy performance requirements, but he wants to know to what degree each is better than satisfactory. By using a descriptive scale, the operations officer gives his scorers a frame of reference. Here the lowest rating possible is labeled "acceptable."

The major disadvantage in using descriptive scales is a semantic one. An "excellent navigator" does not mean the same thing to all scorers. Another disadvantage is that it is hard to select phrases which describe

degrees of performance that are equally spaced. When the scale shown in the example is used, most people feel that there is less distance between "excellent" and "superior" than between "fair" and "good."

### Graphic scale

The graphic scale is a combination of the numerical and descriptive scales. Numerical values are set above a continuous horizontal line and various adjectives or descriptive phrases are set below the line in matching positions relative to each other. (The line represents the range of the performance ability or trait being measured.) In using the graphic scale, the scorer must consider not only the numerical range of the scale but also the phrases that describe the various positions on the scale.

Three typical forms of the graphic scale are described below. In the first Example, the scorer is given instructions for judging the trait of "industry" and told to mark the scale after considering "energy and application to duties, day in and day out." These instructions help reduce errors and improve objectivity and reliability. They also encourage scorers to consider the same things about each person.

### EXAMPLE

Industry: Consider energy and application to duties day in and day out.

	1	2	3	4	5
	Lazy	Indifferent	Diligent	Energetic	Untiring



The second example shows a graphic scale in which certain types of behavior are described for each point on the scale. With most scales, the scorer must not only observe, but also must evaluate the observation in order to form a rating. People can observe more accurately than they can evaluate what they have observed. Whenever ratings can be based on observations alone, reliability is greatly improved. In this example, the scorer is required only to record, not evaluate, the actions of the person being rated. Hence, this type of graphic scale incorporates much objectivity. In preparing this type of scale, make sure that the behavior described for each point is actually an improvement over the point just below it, and that distances between the points appear to the scorer to be about equal.

#### EXAMPLE

Cooperation: Demonstration  
of willingness to work with  
others.

1	2	3	4	5
Creates friction	Indifferent to others	Gets along with most people	A harmonious team worker	Actively promotes harmony in working with others

The scale in the third example is similar to that in the second, except descriptive phrases are not provided for all points. Many times scorers feel that the rating should fall somewhere between two points; such a rating is possible with this form of the graphic scale. The fuller descriptions of the last example increase the likelihood that observed behavior can be pinpointed on the scale. Generally, more detailed descriptions contribute to better rating results.

#### EXAMPLE

Initiative: Action taken  
on own responsibility

1	2	3	4	5
Slow to act even when a decision is much needed. Waits for others. Lets opportunities pass. Does not volunteer. Reticent.		Takes needed action without delay. Volunteers for some tasks. Undertakes all routine jobs without supervision. Dependable.		Anticipates needs. Works ahead and prepares for possibilities. Actively seeks opportunities. Eager.

#### Establish Cut-off Scores

Very often you are going to find that it is impractical to insist on perfect test scores. In these cases, you will have to decide upon a cut-off point (a score below which performance is considered failing or "no-go"). If you are assessing complex skills with your JPM or greatly varied types of process or product, you will have a greater danger of misclassifying the individual's performance).

There are no fixed rules or formulas for establishing the cut-off point for scoring, but you should consider these factors:

1. Immediate manpower needs. If your agency's needs are very high, you may have to lower your cut-off levels, especially if errors are less critical than no performance at all.
2. Feasibility of a score for an established "go". A target on a firing range may be placed in such a way that even

the best marksman may score only 50 per cent hits. If you set a 70 per cent cut-off score, no one will pass your JPM.

3. Consequences of inadequate or delayed task performance. If your JPM is testing task performance which must be accomplished in a specific manner or within a critical time period to avoid substantial risk or damage to persons or property, your cut-off score must be higher to ensure that those who receive a "go" can perform the task within the critical limits.

Establishing cut-off points is a complex matter, and you should make decisions on this matter only after you carefully consider the generally acceptable performance standards for your tasks and the consequences of only adequate or less than adequate performance. A good guideline is that if the consequences are severe for passing one incompetent officer, your cut-off point should be set high. The ratings for consequences of inadequate performance are on the TSSs.

In general, cut-off scores are useful when:

1. Absolute mastery of the task is not expected but a suitable level of performance can be specified.
2. Absolute mastery is possible but factors other than competence affect the score (such as careless errors, measurement errors, etc.).

#### Measurement Errors

Your JPMs must lead to decisions that are consistent and unbiased about individual performance if they are to be free of measurement error. Consistency implies that decisions made about an individual's ability to do his or her job will not vary over a period of time (assuming that the

individual remains the same) and that different scorers using the same JPM will make the same decisions in any given testing situation. An unbiased JPM is one which ensures that decisions are based only on the task in question and are not influenced by other unexpected or unaccountable factors.

Some factors that can lead to inconsistency and bias in your JPMs are:

1. a lack of clear standards for evaluating the particular product or process;
2. poorly written test items and/or testing directions;
3. untrained scorers/judges who may be biased or who do not understand the JPM;
4. the testing environment;
5. malfunctions of special tools or equipment used in the test; and
6. individual day-to-day differences in performance.

Rating errors can be classified into three broad groups:

1. Errors of standards
2. Errors of halo
3. Errors of logic

#### 1. Errors of standard

Some scorers tend to rate performers too high or too low because of differences in their personal standards. Standards that use physical measurement--inches, centimeters, ounces, grams--are fixed values. However, standards that involve judgment and abstract comparison may be as many and varied as the scorers themselves.

## EXAMPLE

Fried eggs are overcooked if there is a noticeable "lacing or bubbling" around the perimeter, which is a darker color than the remainder of the egg white.

Based on this standard, two scorers might have a great deal of difficulty agreeing on whether a particular egg is overcooked or not.

2. Errors of halo

A scorer sometimes allows his or her rating of the performance to be influenced by a personal impression of the person being tested. Such an impression is usually formed on the basis of personal observation or personal knowledge that is not relevant to the rating. If this impression or "halo" is allowed to influence a scorer's judgment, it will result in a shift of the true rating causing an error of halo. If the scorer is favorably impressed the shift will be toward the high end of the scale. Halo error can be either favorable or unfavorable, and it affects only certain persons rated. You may suspect error of halo in JPM situations; but you won't be able to positively identify it until many competent and experienced scorers rate a number of persons under identical conditions.

## EXAMPLE

Officer Jones and nine other scorers rated six person on communications skills. The criterion for consideration for a promotion was a rating of 5 or higher by all ten scorers. The ratings might occur as illustrated in Figure 1.5.

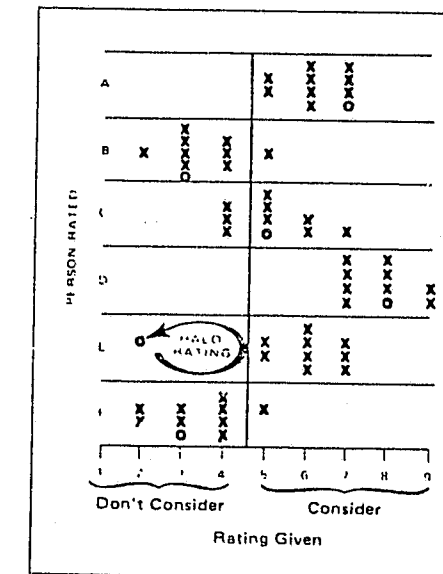


Figure 1.5. Error of Halo

(Jones' rating is indicated by 0, other raters by X.) All ten scorers agreed reasonably well on the communications skills of five of the six persons rated. All but Jones agreed on the communications skills of the sixth (E). For person E, 8 of the 9 scorers rated him 5 or above--only Jones rated him lower. Apparently, Jones allowed some unfavorable impression regarding E to heavily influence his rating; he made an error of halo. If two promotional slots were available, teacher D would clearly be considered; but, remembering Jones' halo rating, there would be some difficulty in deciding whether officer A or E should be considered for the second slot.

3. Errors of logic.

An error of logic may occur when two or more traits are being rated. This error is present when a scorer tends to give similar ratings to traits which do not necessarily go together. For example, some scorers may think that an industrious person is also efficient. Industrious persons may often be efficient, but not necessarily so.

"Error of logic" means that the traits--such as industry and efficiency--are related only in the mind of the scorer making the error. The relationship may not appear to be logical to someone else. As a matter of fact, the person who exhibits an error of this sort is probably not really aware of it.

#### EXAMPLE

In the illustration of error of logic in Figure 1.6, six scorers (A, B, C, D, E, and F) rated a certain officer on four traits (industry, promptness, efficiency, and courtesy) on a scale of 1 to 9. In three of the traits, the six scorers agree reasonably well; however, E gave a much higher rating on efficiency than did the other scorers. E also assigned the same rating to both efficiency and industry. It appears that E thinks industry and efficiency are much the same--an error of logic. There may be several reasons for this error: Scorer E may not have spent enough time observing the performer's efficiency, or he or she may not know the difference between efficiency and industry. In any event, scorer E and the other scorers have not rated the same thing when they rated efficiency. In effect, scorer E rated industry twice.

Logical Error									
INDUSTRY							B, F	A, D, E	C
PROMPTNESS					A, C	D, B, E	F		
EFFICIENCY			A, B, D	C, F				E	
COURTESY		C, D, E	A, B, F						
	1	2	3	4	5	6	7	8	9
RATING GIVEN									

Figure 1.6 Error of Logic Illustration

#### 11. Validate and Revise Your JPMs.

A validated JPM is one that you test and find it to have high predictive validity or, where constraints restrict testing, verify that it has high fidelity. Many JPMs can be validated on the job because they test tasks that are performed regularly under a broad spectrum of real-world conditions. In such cases, JPM validation is similar to a job analysis. The validator observes the performer and thereby checks the JPM checklist or measure. The number of job incumbents you need to observe for validation is very small for unitary tasks, higher for multiple tasks.

There are also jobs that are rarely performed or performed only under certain circumstances, or where validation may be hazardous both to the job performer and the validator. In these cases, JPMs are "verified" rather than validated. Experienced performers should add any missing steps or conditions, delete any unnecessary ones, and verify the standards.

#### Prepare for the JPM Tryout

Follow the rules below.

- To conduct the draft tryout:
  - Draft general instructions for administering your JPM.
  - Ensure, through use of subject matter experts and job incumbents, that the items agree with the corresponding job tasks.
  - Complete all resource arrangements for trying out your draft JPMs.
  - Train your examiners and scorers. At least two (preferably three) scorers should be used.
  - Select examinees. Select several who are typical of the incumbents in the job.
- Prepare forms for recording the information discussed in item 3 below.

3. Select observers. Observers will watch the administration of the JPMs and record information that can be used to correct deficiencies in the items or in the assessment procedures themselves. (They may be used as scorers as well as observers to test reliability of the scorers).

#### Conduct the JPM Tryout

As a rule, you should conduct the tryout as if it were "for real" except for the following procedure. Conduct the tryout in steps. Measure five job holders and revise your JPMs to correct any difficulties that are found. Then, measure the next five job holders; they will be a check on the success of your JPM revisions. This procedure may uncover further difficulties, and it will demonstrate the success (or lack of success) of your JPM revisions. Continue this process of measurement and revision until all JPM deficiencies are corrected.

Here are some guidelines on how you can obtain information concerning problems with the testing situation itself.

1. Ask the trainee to repeat general and specific instructions in his or her own words to determine whether the instructions are clearly understandable. Note any significant deviations.
2. Record any questions asked by the examinees. From these, you can prepare written instructions to answer frequently asked questions.
3. Record any shortage of supplies or breakdown of equipment.
4. Note any ways in which the layout of equipment can be improved without hurting the validity of your JPMs.
5. Note any accidental injury to the examinee or damage to equipment.
6. Note the time required and any problems you experience in re-establishing a test station for the next examinee.
7. If a JPM is given in a series of stations, note any problems you experience in maintaining a smooth flow of examinees from test station to test station.

8. Note any testing conditions that may invalidate the JPM results.
9. When examinees make errors, question the examinees. Note if their wrong answers indicate a misunderstanding of the JPM.
10. Note any actions of a scorer which might give away the correct answers or confuse the examinee.

You should record this information in sufficient detail to provide a basis for correcting any deficiencies in your JPMs. The purpose of the tryout is to make your JPM highly reliable by eliminating as many causes of unreliability as possible. If a JPM item proves unsatisfactory even after successive revisions when given to each of at least 10 examinees, it should be replaced with a new one. Remember, the validator must verify JPM conditions, cues, and standards while he is validating the JPM task.

When the tryout is finished, you will have a corresponding job performance measure for each task on the list of tasks your agency has selected for instruction or evaluation. While JPMs are not the objectives of the instruction, they are used as an evaluative device for quality control of the instruction.

The tryout will uncover any problems with your JPMs. What is wrong usually can be determined from the inputs you get from the validators. After you correct the problems, you will have to try out the JPMs again. This cycle must be repeated until all your JPMs are validated or verified.

#### Outputs

When you have finished all of the steps in this manual you should have:

1. A validated or verified JPM for each task you have to train. Each JPM will include the required test performance and the test conditions, cues, and standards.

2. Instructions for administering your JPMs.
3. The rationale for trade-off decisions you made because of testing constraints.
4. A summary statement of implementation and the results of validating or verifying your JPMs.

With these JPMs you will be able to assess the quality of your agency's formal and on-the-job training programs, as well as individual officer ability to perform the tasks of a particular law enforcement job. You can also use JPMs as part of a selection or screening process for personnel management considerations such as promotion to a higher grade or assignment to a special law enforcement job. In any case, you'll discover that job performance measures are the best measurement and evaluation tools a trainer or personnel administrator can have.

## Chapter II JPM SPECIFICATION SHEETS

In order to design JPMs that are consistent with the tasks and consistent with each other, a plan is needed which can be successfully used over time by several people. One way to accomplish this is to prepare item specification sheets.

The JPM specification sheet helps give a definite structure to the development of each JPM and systematizes the process. The JPM specification sheet requires the JPM developers to analyze the task being tested, specify the characteristics of the action and the correct response, and describe qualities of alternate responses. With this kind of advance thinking, JPM developers can more easily, consistently, and coherently write questions, phrase correct answers, or describe correct procedure, and prepare scoring decision rules.

Although the format can vary according to the user's convenience, a correctly prepared JPM specification sheet needs several components:

- Some form of identification
- The job task
- The skill or element of the job task to be tested
- the initiating cues
- The question characteristics
- The response characteristics
- Scoring rules
- The test pool characteristics (if there are several possible alternate JPMs)

The identification should be sufficient for a new person to locate the JPM specification sheet from the curriculum or TSS or the curriculum TSS from the JPM specification sheet. The identification may be a number, title, or coded words, but it must relate clearly to the TSS for the task and any curriculum for which it is developed.

The job task should be derived from the PSTC job analysis. It may be identical to the job task or it may be modified as described in Chapter I.

The element or skill should be derived from the job task and should be written as it is specified in the Task Summary Sheet (TSS). There may be several elements or skills for each job task. If it is to be tested separately each element or skill will have its own question and answer characteristics. The simulation requirements are those that have been described in Chapter I.

The question characteristics describe what the instructions or directions will contain or specify.

The response characteristics describe what the correct response should be and often, but not always, what the incorrect response will be.

Figure 2.1 is an example of these five elements of a JPM specification sheet.

NUMBER: Task 098  
Equipment piece 217

TASK: Apprehend Suspects

ELEMENT  
OR SKILL: Apply Hand Irons

SIMULATION  
REQUIREMENTS: The "suspect" will be role played by an available person.  
The "suspect" should not put up resistance.

(Continued on next page)

#### QUESTION

CHARACTERISTICS: The person being evaluated will be told the following: The suspect has been searched and his weapon has been confiscated. The suspect is under investigation and is to be transported to the station for further questioning. He will make the trip in hand irons. You are to apply the hand irons.

#### RESPONSE

CHARACTERISTICS: The person will follow the steps listed in the TSS in that order

Figure 2.1 Partial JPM Specification Sheet

The sixth element is the item pool characteristics. The potential size of the JPM pool is different for different kinds of skills.

NUMBER: Task 098  
Equipment piece 217

TASK: Apprehend Suspects

ELEMENT  
OR SKILL: Apply hand irons

#### SIMULATION

REQUIREMENTS: The "suspect" will be role played by an available person.  
The "suspect" should not put up resistance.

#### QUESTION

CHARACTERISTICS: The person being evaluated will be told the following: The suspect has been searched and his weapon has been confiscated. The suspect is under investigation and is to be transported to the station for further questioning. He will make the trip in hand irons. You are to apply the hand irons.

#### RESPONSE

CHARACTERISTICS: The person will follow the steps listed in the TSS in that order.

#### ITEM POOL

CHARACTERISTICS: All of the JPMs will require the person to apply hand irons. In one the "suspect" will not resist; in one he will resist; in a third the "suspect" will wear a cast on one arm.

Figure 2.2 Partial JPM Specification Sheet



The six elements mentioned above are all needed prior to writing the first JPM. The remaining elements of the sheets are the JPM itself, the administrators' directions, the scoring rules and the JPM data.

The development of JPM specification sheets gives test-makers two ways of making new tests: by rearranging existing example JPMs from a JPM bank, or by using the specification sheets to generate altogether new JPMs.

### Chapter III RELIABILITY AND VALIDITY

Two final but basic essentials of test development are ensuring test reliability and validity. Reliability is a measure of how consistently each person scores on the test either at different times, on different forms, or within the same test. Briefly, the following is a list of the steps to establish adequate test reliability.

#### Establishing Adequate Test Reliability

1. Establish type of reliability coefficient to compute:
  - Test-retest method
  - Alternate forms
  - Internal consistency
  - Other.
2. Consider limitations of selected reliability coefficient in terms of test approach to be used:
  - Norm referenced tests (NRT) can use all methods. (NRTs compare the performance of test takers with each other.)
  - Criterion referenced tests (CRT) more appropriately use test-retest and alternate forms. (CRTs compare each test taker with some external criterion.)
3. Identify possible sources of error which will result in low reliability of test:
  - Familiarity with test form
  - Fatigue
  - Emotional strain
  - Physical conditions of testing site

- Health of test taker
  - Fluctuations of human memory
  - Amount of student's experience with the skill being measured
  - Specific knowledge that has been gained outside of the experience evaluated by the test
  - Poor or confusing directions
  - Incomplete or inaccurate test taking or scoring instructions
  - Bias in selected sample.
4. Administer JPM(s) in accordance with selected reliability coefficient.
  5. Compute reliability coefficient:
    - Test-retest: Compute the correlation coefficient between each person's scores on the two administrations of the test.
    - Alternate forms: Compute correlation coefficient between each person's score on the two JPMs testing the same task, element or skill.
- NOTE: It would be necessary to do a test/retest for reliability of each JPM (JPMs related to each task element or skill). JPMs on the same task would be tested on two successive days, preferably at the same time of day. The same JPMs would have to be tested both days for the same person. It is important that the people tested are from the target population and preferably have mastered the task. A coefficient can be calculated for each subtest or for each item to estimate reliability. The coefficient should be + .50 or above to say the test is sufficiently reliable.
6. Evaluate reliability of test.
  7. Revise if necessary.

### Establishing Test Validity

Validity is a measure of how the JPM matches what one intends to test. The following is a list of the steps to establish test validity.

1. Establish kind of validity to be measured. There are several kinds, some of which are:
  - Content:  
Evidence of the validity of a test by a content validity study should consist of data showing that the process and/or product called for in the test is representative of the important process and/or product called for on the job for which the person is being measured.
  - Concurrent:  
Evidence of the validity of a test by a concurrent validity study should consist of data showing that the content of the test correlates with another currently available criterion.
  - Criterion-related:  
Evidence of the validity of a test by a criterion-related validity study should consist of data demonstrating that the test is predictive of or significantly correlated with the important process or product of job performance.
  - Construct:  
Evidence of the validity of a test through a construct validity study should consist of data showing that the test measures the degree to which the person being measured has identifiable characteristics which have been determined to be important in performance in the job for which the person is being measured.
  - Predictive:  
Evidence of the validity of a test through a predictive validity study should consist of data showing that the test measures some identifiable characteristics that are predictive of success or failure on the job.
  - Face:  
Evidence of face validity of a test consists of an appearance of a close relationship of the test to the criterion (for instance, job sample tasks often have high "face" validity).

- Curricular:

Evidence of the validity of a test by a curricular validity study should consist of empirical data demonstrating that the test items are highly correlated with the curriculum objectives.

For a curriculum based test the most appropriate kinds are content and curricular validity, neither of which is statistical. A co-efficient is recommended as the statistical measure of validity.

2. Develop JPM content validation procedures:
  - Select representative sample of trainees or job incumbents.
  - Review JPM development procedures.
3. Establish jury of experts.
4. Develop procedure for assessing experts' opinions and perceptions of JPM.
5. Administer test to representative group.
6. Submit test results to jury of experts for analysis.
7. Revise as necessary, based on input from jury of experts.
8. Review methodology for assessing validity.

## Chapter IV

## Legal Considerations in Test Development

The following information consists primarily of excerpts from Uniform Guidelines on Employee Selection Procedures, (Miner & Miner, 1979), a publication incorporating the selection guidelines as published in the Federal Register (1978), and questions and answers on the guidelines prepared by a joint committee.

## Background

- A test must not be designed, intended, or used to discriminate.

One problem that confronted the Congress which adopted the Civil Rights Act of 1964 involved the effect of written preemployment tests on equal employment opportunity. The use of these test scores frequently denied employment to minorities in many cases without evidence that the tests were related to success on the job. Yet employers wished to continue to use such tests as practical tools to assist in the selection of qualified employees. Congress sought to strike a balance which would proscribe discrimination, but otherwise permit the use of tests in the selection of employees. Thus, in Title VII, Congress authorized the use of "any professionally developed ability test provided that such test, its administration or action upon the results is not designed, intended or used to discriminate...." (Section 703(h), 42U.S.C. 20003(2) (h).)

- A test must be professionally developed and must fairly measure or predict performance.

At first, some employers contended that, under this section, they could use any test which had been developed by a professional so long as they did not intend to exclude minorities, even if such exclusion was the consequence of the use of the test. In 1966, the Equal Employment Opportunity Commission (EEOC) adopted guidelines to advise employers and other users what the law and good industrial psychology practice required.... (U.S.L.W. 2137, 1966). The Department of Labor adopted the same approach in 1968 with respect to tests used by Federal Contractors under Executive Order 11246 in a more detailed regulation. The Government's view was that the employer's intent was irrelevant. If tests or other practices had an adverse impact on protected groups, they were unlawful unless they could

be justified. To justify a test which screened out a higher proportion of minorities, the employer would have to show that it fairly measured or predicted performance on the job. Otherwise, it would not be considered to be "professionally developed."

- All other selection procedures as well as tests are covered by the same guidelines.

In succeeding years, the EEOC and the Department of Labor provided more extensive guidance which elaborated upon these principles and expanded the guidelines to emphasize all selection procedures. In *Griggs v. Duke Power Co.* (1971), the Supreme Court announced the principle that employer practices which had an adverse impact on minorities and were not justified by business necessity, constituted illegal discrimination under Title VII. The elaboration of these principles by courts and agencies continued into the mid-1970's (*Albermarle Paper Co. v. Moody*, 1975) but differences between the EEOC and the other agencies (Justice, Labor, and Civil Service Commission) produced two different sets of guidelines by the end of 1976.

Since 1977, efforts were intensified to produce a unified government position. The following overview of the Uniform Guidelines discusses the result of that effort.

#### Synopsis

The following section provides an overview of the Uniform Guidelines based on the required selection procedures steps. The four steps are:

1. Overall evaluation. The selection procedure must be evaluated to measure whether any adverse impact exists.
2. Specific evaluations. The specific steps which cause the adverse impact must be isolated.
3. Elimination of cause. Any procedure found to cause adverse impact must be:
  - o Abandoned,
  - o Changed, or
  - o Proved to be job related.

4. Investigation of alternatives. Studies must be made to find suitable alternatives with less or no adverse impact. When more than one suitable alternative exists the one with the least adverse impact should be adopted.

- If a selection procedure has an adverse impact it must be justified on grounds of business necessity.

The fundamental principle underlying the guidelines is that employer policies or practices which have an adverse impact on employment opportunities of any race, sex, or ethnic group are illegal under Title VII and the Executive Order unless justified by business necessity ... (*Griggs v. Duke Power Co.*, 1971).... A selection procedure which has no adverse impact generally does not violate Title VII or the Executive Order.... (*Furnco v. Waters*, 1978). If adverse impact exists, it must be justified on grounds of business necessity. Normally, this means by validation which demonstrates the relation between the selection procedure and performance on the job.

- If the selection rates of any affected group is less than 80% of that of the dominant group (usually white males) the rates are considered evidence of adverse impact.

The guidelines adopt a "rule of thumb" as a practical means of determining adverse impact for use in enforcement proceedings. This rule is known as the "4/5ths" or "80 percent" rule. It is not a legal definition of discrimination, rather it is a practical device to keep the attention of enforcement agencies on serious discrepancies in hire or promotion rates or other employment decisions.

#### The "4/5ths Rule of Thumb"

A selection rate for any race, sex, or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact ....

The "4/5ths" rule of thumb" is calculated in the following manner:

- (1) Calculate the rate of selection for each group (divide the number of persons selected from a group by the number of applications from the group).

- (2) Observe which group has the highest selection rate.
- (3) Calculate the impact ratios, by comparing the selection rate for each group with that of the highest group (divide the selection rate for a group by the selection rate for the highest groups).
- (4) Observe whether the selection rate for any group is substantially less (i.e., usually less than 4/5ths or 80%) than the selection rate for the highest group. If it is, adverse impact is indicated in most circumstances.

For example:

Applicants	Hires	Selection rate	Percent hired
80 White.....	48	48/80	60%
40 Black.....	12	12/40	30%

A comparison of the black selection rate (30%) with the white selection rate (60%) shows that the black rate is 30/60, or one-half (or 50%) of the white rate. Since the one-half (50%) is less than 4/5ths (80%) adverse impact is usually indicated.

The determination of adverse impact is not purely arithmetic, however; and other factors may be relevant.

- When there is evidence of Adverse Impact the employer must modify or eliminate or validate the procedure.

#### Procedures for Adverse Impact

Once an employer has established that there is adverse impact, what steps are required by the guidelines? As previously noted, the employer can modify or eliminate the procedure which produces the adverse impact, thus taking the selection procedure from the coverage of these guidelines. If the employer does not do that, then it must justify the use of the procedure on the grounds of "business necessity" (McDonnell Douglas v. Green, 1973). This normally means that it must show a clear relation between performance on the selection procedure and performance on the job. In the language of industrial psychology, the employer must validate the selection procedure.

#### Validation

The majority of the text of the guidelines is devoted to outlining what is acceptable evidence of validity, which may be based on any of the three professionally accepted types of validation--criterion-related validation studies, content validity, or construct validity. Where criterion-related validation studies are conducted, they should include an investigation of the fairness of a selection procedure if technically feasible--that is, if there are large enough samples of women or members of a particular minority group to conduct such a study.

- The employer should consider test or selection alternatives which will achieve its business purposes with lesser adverse impact.

The concept of validation as used in personnel psychology involves the establishment of the relationship between a test instrument or other selection procedure and performance on the job. Federal equal employment opportunity law has added a requirement to the process of validation. In conducting a validation study, the employer should consider available alternatives which will achieve its legitimate business purpose with lesser adverse impact (Albermarle Paper Co. v. Moody, 1975). The employer cannot concentrate solely on establishing the validity of the instrument or procedure which it has been using in the past.

This same principle of using the alternative with lesser adverse impact is applicable to the manner in which an employer uses a valid selection procedure. The guidelines assume that there are at least three ways in which an employer can use scores on a selection procedure:

- To screen out of consideration those who are not likely to be able to perform the job successfully,
- To group applicants in accordance with the likelihood of their successful performance on the job,
- To rank applicants, selecting those with the highest scores for employment (Robinson v. Lorillard Corp., 1971).

- The employer who uses cut-off scores or rank ordering must justify use of those methods.

The setting of a "cutoff score" to determine who will be screened out may have an adverse impact. If so, an employer is required to justify the initial cutoff score in terms of its relationship to the selection of a trustworthy and efficient work force. Similarly, use of results for grouping or for rank ordering is likely to have a greater adverse effect than use of scores solely to screen out unqualified candidates. If the employer chooses to use a rank order method, the evidence of validity must be sufficient to justify that method of use.

#### How is the Validation Conducted?

Validation has become highly technical and complex, and yet is constantly changing as a set of concepts in industrial psychology. What follows here is a simple introduction to a highly complex field. There are three concepts which can be used to validate a selection procedure. These concepts reflect different approaches to investigating the job relatedness of selection procedures and may be interrelated in practice. They are:

- (1) Criterion-related validity
- (2) Content validity
- (3) Construct validity

#### Criterion-Related Validity

Evidence of the validity of a test or other selection procedure by a criterion-related validity study should consist of empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance. (Miner & Miner, 1979)

#### Content Validity

Evidence of the validity of a test or other selection procedure by a content validity study should consist of data showing that the content of the selection procedure is representative of important aspects of performance on the job for which the candidates are to be evaluated.... (Miner & Miner, 1979)

#### Construct Validity

Evidence of the validity of a test or other selection procedure through a construct validity study should consist of data showing that the procedure measures the degree to which candidates have identifiable characteristics which have been determined to be important in successful performance in the job for which the candidates are to be evaluated. (Miner & Miner, 1979)

## References

Equal Employment Opportunity Commission, Uniform guidelines on employee selection procedure. Federal Register, Washington D.C. August 26, 1978.

Fuchs, E.F. Fairness evaluation. Final Report on Purchase Order DAHC-19-77-M-D0ID, January, 1978.

Miner, M.G., & Miner, J.B., Uniform guidelines on employee selection procedures. Bureau of National Affairs, Inc., Washington D.C., 1979.

U.S. Supreme Court: Griggs v. Duke Power Co., March 8, 1971.

U.S. Supreme Court: Washington v. Davis, June 7, 1976.

U.S. Supreme Court: McDonnell Douglas v. Green, 1973.

U.S. Supreme Court: Albermarle Paper Co. v. Moody, 1975.

U.S. Supreme Court: Furnco v. Waters, 1978.

## Recommended Reading List

Uniform Guidelines on Employee Selection Procedures.

## Includes:

An analysis by Mary G. Miner and John B. Miner.

Questions and answers on the Guidelines prepared by a joint committee.

Bibliography on Testing.

Miner and Miner, Employee Selection Within the Law. BNA Books, 1978, Chapter 3.

These pamphlets can be obtained from:

BNA Books  
1231 25th Street N.W.  
Washington D.C. 20037



Appendix  
Sample Job Performance Measures

The following four sample Job Performance Measure (JPM) test items are intended to serve as examples of how JPM items can be constructed for different kinds of law enforcement tasks. They are not intended to be all inclusive or a completely accurate representation of performance measurement of a task in any particular Florida law enforcement job. You should use them only as guidelines for test item construction. The JPMs are presented on a Job Performance Specification Sheet derived from Task Summary Sheets.

Equipment Related JPM - JPM I

The first example is an item constructed for the task "Apply hand irons," which involves the use of a common piece of law enforcement equipment. As you can see from the sample, test items for equipment-related tasks should include use of the actual equipment in the test for high predictive validity and physical fidelity. In job performance testing situations, you would most likely standardize the conditions for performance of the task. That is, an officer on the job would have to be able to apply hand irons in a variety of situations with a subject who does not cooperate and attempts to get away.

Process JPM - JPM II

The second example is a JPM constructed for the task "Advise Persons of rights," which is a process-related task because how the task is performed is the most critical output of the task performance. For this kind of task the breakdown of performance into specific steps is essential so that an individual may be evaluated on performance of each step. Very often, as is true in this task, the sequence of performance is also important. From this JPM you could easily develop a checklist built on each

step with check-off points for each standard of performance listed for the step. In a job performance testing situation, however, you would again want to build in realistic conditions for performance of the task such as having to advise a subject who is drunk or doesn't understand English.

#### Forms related JPM - JPM III

The third example is a JPM constructed for the task "Prepare a Rights Warning Waiver Form," which is intended to be an example of a forms-related test item; this particular task may be performed with a wide variety of kinds of forms within different Florida law enforcement agencies. Such a forms related JPM is the basis of a paper-and-pencil test of the task performance. That is, you would test the individual on how well he or she actually completes the form in a test or job situation. Again, you may use a checklist as an evaluation tool as well as comparison of the individual's form to a correctly completed form.

#### Product JPM - JPM IV

The last example is a JPM constructed for the task "Fingerprint persons," for which the most important final outcome is a product, ie., a set of clear, useable prints. Process (how the prints are taken) may be important if you are testing in a training situation. Therefore, all the steps in fingerprinting are broken down in the example. For job performance testing, however, you would most likely base successful performance on the final product only. It's doubtful that an officer on the job would be negatively evaluated for rolling fingers right-to-left instead of left-to-right if the quality of the final prints were acceptable.

JPM I

Equipment Related JPM

JPM I task sheet here

JPM II

Process Related JPM

## JPM I - Equipment Related JPM

## JPM Specification Sheet

NUMBER: Task 098  
Equipment piece 217

TASK: Apprehend Suspects

ELEMENT OR  
SKILL: Apply Hand Irons

SIMULATION  
REQUIREMENTS: The "suspect" will be role played by an available person.  
The "suspect" should not put up resistance.

QUESTION  
CHARACTERISTICS: The person being evaluated will be told that following:  
The suspect has been searched and his weapon has been  
confiscated. The suspect is under investigation and is  
to be transported to the station for further questioning.  
He will make the trip in hand irons.

RESPONSE  
CHARACTERISTICS: The person will follow the steps listed in the TSS in  
that order.

JPM POOL  
CHARACTERISTICS: All of the JPMs will require the person to apply hand  
irons. In one the "suspect" will not resist; in one he  
will resist; in a third the "suspect" will wear a cast on  
one arm.

## JPM I - Equipment Related JPM

## Apply Hand Irons

## Instructions to JPM administrator for JPM 098

1. Have a person role play the suspect. He is not to resist.
2. The hand irons should be in their case and given to the person being tested to put on his equipment belt.
3. Have an "Apply Hand Irons" scoring sheet on which to record the results.
4. Read the following instructions to the person.  
"This suspect has been searched and his weapon has been confiscated. The suspect is under investigation and is to be transported to the station for further questioning. He is to make the trip in hand irons. Put them on him."
5. Record the person's actions on the scoring sheet.

## JPM I - Equipment Related JPM

## Apply Hand Irons

## Scoring Sheet for JPM 098

Step:	Standard:	Pass	Fail
1. Hand irons are positioned on the equipment belt.	According to agency regulations.	<input type="checkbox"/>	<input type="checkbox"/>
2. Assume correct position for applying hand irons.	The officer must center him or herself behind the subject and remove one pair of hand irons from their case.  The officer must not turn away from the subject or lose visual contact with the subject and must be able to remove the hand irons from the case using his or her non-firing hand	<input type="checkbox"/>	<input type="checkbox"/>
3. Instruct the subject to place one hand behind his back.	The officer must give the instruction in a loud and clear voice.  The officer must specify which hand the subject is to move first.  The officer must repeat the instruction if the subject does not comply.	<input type="checkbox"/>	<input type="checkbox"/>
4. Grasp the subject's hand and apply the hand iron to that hand.	The officer must holster the revolver.  The officer must have a controlling grasp of the subject's hand such that if the subject makes a sudden movement, the officer can forcibly control him through leverage of the subject's arm.  The subject's hand must be grasped by the officer in a manner such that the subject's palm faces outward.  The hand iron must be applied in a manner that does not injure the subject; ie., it cannot be "snapped" on the subject's wrist.  The fit of the hand iron must not be so tight as to restrict circulation in the subject's hand or prevent any movement nor must it be so loose as to allow too much movement of the subject's hand or the possibility of escape from the hand iron.	<input type="checkbox"/>	<input type="checkbox"/>

Step:	Standard:	Pass	Fail
4. Continued	The hand iron must be applied so that the double lock of the iron is up and the keyhole faces away or outward from the subject.	<input type="checkbox"/>	<input type="checkbox"/>
5. Instruct the subject to place his head on the support surface.	The officer must give the instruction in a loud and clear voice.  The officer must repeat the instruction or explain it if the subject does not comply.  The officer must allow the subject to move forward slightly if the subject is unable to place his head on the support surface.	<input type="checkbox"/>	<input type="checkbox"/>
6. Instruct the subject to place his remaining free hand behind his back.	The officer must give the instruction in a loud and clear voice.  The officer must repeat the instruction if the subject does not comply.	<input type="checkbox"/>	<input type="checkbox"/>
7. Grasp the subject's remaining free hand and apply the remaining hand iron.	The same standards of performance described for step #3 apply to this step.	<input type="checkbox"/>	<input type="checkbox"/>
8. Lock the hand irons.	The officer must retrieve the key from its position in his or her uniform with one hand while retaining a controlling hold on the subject with the other hand.  The officer must double lock the hand irons and return the key to its position..	* <input type="checkbox"/> *	* <input type="checkbox"/> *
9. Assist the subject in standing up.	The officer must not yank the subject up by pulling on the hand irons chain.  The officer must retain control by holding the arm of the subject and then pulling the subject away from the wall into a standing position.	<input type="checkbox"/>	<input type="checkbox"/>
TOTAL PASS PASSING SCORE IS 10 OR MORE BUT 3 * BOXES MUST BE RATED PASS		<input type="checkbox"/>	<input type="checkbox"/>
Final Score _____ Evaluators Signature _____ Date _____		Pass <input type="checkbox"/>	Fail <input type="checkbox"/>

## TASK SUMMARY SHEET

## JPM II - Process-Related JPM

## JPM Specification Sheet

NUMBER: Task 099

TASK: Advise Persons of Rights (per Miranda)

ELEMENT  
OR SKILL:

## SIMULATION

REQUIREMENTS: The "suspect" will be role played by an available person. He will agree to sign the waiver. The person being evaluated will be given a brief oral description of the offense for which the suspect has been apprehended.

## QUESTION

CHARACTERISTICS: The person being evaluated will be told the following: "The suspect has been detained by the department store personnel for shop lifting. The store personnel have accused him of taking two watches and three diamond rings. You are the first law enforcement officer to talk to the suspect. Before you question him, advise him of his rights."

## RESPONSE

CHARACTERISTICS: The person will follow the steps listed in the TSS. The officer should not ask any questions not listed on the TSS and should not answer questions about other matters from the suspect.

## JPM POOL

CHARACTERISTICS:

1. All the JPMs will require the person to advise persons of rights.
2. The setting of the crime or reason for apprehension can vary.
3. The "suspect" may try to get the officer off the subject.

JPM II Process Related JPM  
Advise Persons of Rights

Instructions to JPM Administrator for JPM 099

1. Have a person role play the subject.  
He is to be cooperative.
2. Have an "Advise Persons of Rights" scoring sheet on which to record the results.
3. Read the following instructions to the person.  
"The suspect has been detained by department store personnel for shoplifting. The store personnel have accused him of taking two watches and three diamond rings. (These were items of over \$2,000 value.) You are the first law enforcement officer to talk to the suspect. Before you question him, advise him of his rights."

JPM II - Process Related JPM  
Advise Persons of Rights

Scoring Sheet for JPM 099

Step:	Standard:	Pass	Fail
1. Inform the subject of the official position.	The officer should identify him or herself in a loud and clear voice giving name, rank, and law enforcement agency.	<input type="checkbox"/>	<input type="checkbox"/>
2. Inform the subject of the nature of the offense(s) under investigation and the fact that he is a suspect or accused of that offense and under apprehension.	The officer should speak in a loud and clear voice and concisely inform the subject of the suspected offense and the fact of apprehension.  The officer should not answer any questions of the subject at this point or discuss any other matters.	<input type="checkbox"/>	<input type="checkbox"/>
3. Inform the subject of his rights.	The officer must read the rights warning card verbatim without skipping any parts or changing the order. The officer must use a loud and clear voice and must go slowly enough for comprehension.	<input type="checkbox"/>	<input type="checkbox"/>
4. Determine whether or not the subject understands his rights.	The officer should ask the subject if he understands his rights. The officer must use a loud and clear voice, and must repeat the question if the subject does not respond. The officer must ask the subject to state out loud that he understands his rights if the subject only nods in answer to the question.	<input type="checkbox"/>	<input type="checkbox"/>
5. Repeat the rights warning and determine whether or not the subject understands his rights.	The officer should again read the warning card word-for-word, slowly enough so that the subject can comprehend the meaning. If the subject again indicates a lack of understanding, the officer may not further question the subject, and the JPM is completed.	<input type="checkbox"/>	<input type="checkbox"/>

<u>Step:</u>	<u>Standard:</u>	Pass	Fail
6. Ask the subject whether or not he wishes to answer any questions at this time.	The officer should ask the subject in a loud and clear voice.	<input type="checkbox"/>	<input type="checkbox"/>
7. Ask the subject to sign a waiver.	If the subject says he is willing to answer questions, the officer should ask in a loud and clear voice without any implied threat or coercion.  If the subject is not willing to answer questions, this step is omitted, and the officer may not further question the subject.	<input type="checkbox"/>	<input type="checkbox"/>

ALL 7 BOXES MUST BE "PASS." THE CARD MUST BE READ VERBATIM.

Final Score \_\_\_\_\_  
 Evaluator's Signature \_\_\_\_\_  
 Date \_\_\_\_\_

Pass ☐ Fail ☐

JPM III

Forms Related JPM



Task summary sheet here

FHP Interview Report goes here

## JPM III - Forms Related JPM

## JPM Specification Sheet

NUMBER: Task 099  
Form 2

JPM 099A

TASK: Advise Person of Rights (per Miranda)

ELEMENT  
OR SKILL: Prepare a Rights Warning Waiver Form

SIMULATION  
REQUIREMENTS: The "suspect" will be role played by an available person.  
The "suspect" will indicate that he is willing to sign the waiver.

QUESTION  
CHARACTERISTICS: The person being evaluated will be told the following:  
1. The subject has been apprehended for an offense and advised of his constitutional rights. He now indicates he understands his rights and is willing to answer questions.

RESPONSE  
CHARACTERISTICS: The person will follow the steps listed in the TSS. The order of steps 1 and 2 are important. For items 3 through 6 the order is not important but must follow step 2 and precede step 7. Steps 7, 8, and 9 should be performed in order.

JPM POOL  
CHARACTERISTICS: 1. All of the JPMs will require the person to fill in the form.  
2. Witnesses can be another option.  
3. Other agencies can adapt the JPM to their own waiver form.

## JPM III - Forms Related JPM

## Prepare a Rights Warning Waiver Form

## Instructions to JPM administrator for JPM 099A

1. Have a person role play the subject. He is to be cooperative.
2. The FHP Interview Report (FHP-9) and a pen or pencil should be readily available.
3. Have a "Prepare a Rights Warning Waiver" scoring sheet on which to record the results.
4. Read the following instructions to the person:  

"The subject has been apprehended for an offense. He's been advised of his constitutional rights. He has said he understands his rights and is willing to answer questions.  
 Prepare the rights warning waiver form."
5. Give the person the waiver form and a pen or pencil.

Step:	Standard:	Pass	Fail
9. Note any witnesses to the subject's signing of the waiver.	The officer should legibly print the full name and address of any witnesses.	<input type="checkbox"/>	<input type="checkbox"/>

Process Score	Pass	Fail
Are the scored steps performed correctly and in the right order?	<input type="checkbox"/>	<input type="checkbox"/>

B. MUST BE A PASS

Product Score	Pass	Fail
1. Is the printing legible?	<input type="checkbox"/>	<input type="checkbox"/>
2. Are both dates correct?	<input type="checkbox"/>	<input type="checkbox"/>
3. Is the hour correct using the 24 hour clock?	<input type="checkbox"/>	<input type="checkbox"/>
4. Is the offense type, date, and location correct?	<input type="checkbox"/>	<input type="checkbox"/>
5. Is the subject's name correct?	<input type="checkbox"/>	<input type="checkbox"/>
6. Is all the information placed in the correct boxes?	<input type="checkbox"/>	<input type="checkbox"/>

C. ALL MUST BE PASSES

Final Score	Pass	Fail
Evaluator's Signature _____ Date _____	<input type="checkbox"/>	<input type="checkbox"/>

## JPM III - Forms Related JPM

## Prepare a Rights Warning Waiver Form

## Scoring Sheet for JPM 099A

Step:	Standard:	Pass	Fail
1. Explain the purpose of the form to the subject.	The officer should tell the subject that the form provides written confirmation of his (the subject's) willingness to waive his rights and answer questions without a lawyer being present.	<input type="checkbox"/>	<input type="checkbox"/>
2. Ask the subject if he will sign the waiver form.	The officer should ask the subject in a non-threatening manner.	<input type="checkbox"/>	<input type="checkbox"/>
3. Complete the date, time, and location spaces.	The officer should legibly print the date (month, day, year, ie., 6/11/80) the time (using the 24 hour time clock, ie., 2317 hours), and the location of the interview.	<input type="checkbox"/>	<input type="checkbox"/>
4. Complete the personal identification space in the form.	The officer should legibly print his or her name and rank in the appropriate spaces.	<input type="checkbox"/>	<input type="checkbox"/>
5. Specify the exact offense for which the subject will be questioned.	The officer should legibly print the type of offense, date of occurrence, and location if appropriate.	<input type="checkbox"/>	<input type="checkbox"/>
6. Fill in the name of the subject.	The officer should legibly print the subject's name (first, middle, and last names) in the appropriate space.	<input type="checkbox"/>	<input type="checkbox"/>
7. Ask the subject to read the completed portion of the form and sign where indicated.	The officer should allow the subject time to read the completed portion of the form and then indicate where the subject is to sign.	<input type="checkbox"/>	<input type="checkbox"/>
8. Note the date and time the subject signed.	The officer should legibly print the date and time in the same format as used in the upper portion of the form.	<input type="checkbox"/>	<input type="checkbox"/>

**CONTINUED**

**1 OF 2**

JPM IV

Product Related JPM

TASK SUMMARY SHEET GOES HERE

## JPM Specification Sheet

NUMBER: Task 109

TASK: Fingerprint Persons

ELEMENT  
OR SKILL:SIMULATION  
REQUIREMENTS: The "suspect" will be role played by an available person. The "suspect" will be cooperative.QUESTION  
CHARACTERISTICS: The person being evaluated will be told to fingerprint the suspect. All the materials will be available.RESPONSE  
CHARACTERISTICS: The person will follow the steps listed in the TSS but will not be evaluated on the process. For evaluation the card will be compared with a model card on the listed criteria.JPM POOL  
CHARACTERISTICS: All of the JPMs will require the person being evaluated to fingerprint a person. The alternatives could be an uncooperative person, a person with a crippled hand, or a person with missing fingers.  
The JPM could be increased in length by adding the fingerprinting of both hands, (Four fingers together, then the thumb).

## Instructions to JPM Administrator for JPM 109

1. Have a person role play the subject. He is to be cooperative.
2. Have an available place to work as much like the agency's fingerprinting facilities as possible.
3. Have two fingerprint cards, fingerprint card holder, printer's ink, ink roller, and inking surface ready and easily accessible to the person being evaluated.
4. Allow the person two attempts at fingerprinting. Allow him to choose the better card for you to judge.
5. Compare the card with the model card, check off the print attributes as listed on the JPM scoring sheet.

## JPM IV - Product related JPM

## Fingerprint Persons

## Scoring Sheet for JPM 109

<u>Step:</u>	<u>Standard:</u>	Pass	Fail
1. Are all fingers and thumb in the proper place on card?	All must be in their proper block. All blocks must be filled.	<input type="checkbox"/>	<input type="checkbox"/>

2. Are all the information blocks properly filled in?		<input type="checkbox"/>	<input type="checkbox"/>
---	--	--------------------------	--------------------------

3. Compare the card to FBI standards. Rate each print.

Check if a "pass"

Complete      Clear

Right Hand

Thumb	<input type="checkbox"/>	<input type="checkbox"/>
Index finger	<input type="checkbox"/>	<input type="checkbox"/>
Middle finger	<input type="checkbox"/>	<input type="checkbox"/>
Ring finger	<input type="checkbox"/>	<input type="checkbox"/>
Little finger	<input type="checkbox"/>	<input type="checkbox"/>

Left Hand

Thumb	<input type="checkbox"/>	<input type="checkbox"/>
Index finger	<input type="checkbox"/>	<input type="checkbox"/>
Middle finger	<input type="checkbox"/>	<input type="checkbox"/>
Ring finger	<input type="checkbox"/>	<input type="checkbox"/>
Little finger	<input type="checkbox"/>	<input type="checkbox"/>

ALL BLOCKS MUST BE  
CHECKED FOR A  
"PASS"

**END**