U.S. Department
of Transportation

**Federal Railroad
Administration**

SELECTION OF LOCOMOTIVE ENGINEER TRAINEES

Office of Research and
Development
Washington, D.C. 20590

FINAL REPORT

Merri-Ann Cooper
Charles Uhlman
B. Lynn Frost

University Research Corporation
7200 Wisconsin Avenue
Bethesda, MD 20814

October 1993

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle
Selection of Locomotive Engineer Trainees

5. Report Date
October, 1993

6. Performing Organization Code

7. Author(s)
M. Cooper, C. Uhlman, & B.L. Frost

8. Performing Organization Report No.

9. Performing Organization Name and Address
University Research Corporation
7200 Wisconsin Avenue
Bethesda, MD 20814

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
DTFR 53-86-C-00032

12. Sponsoring Agency Name and Address
U.S. Department of Transportation
Federal Railroad Administration
Office of Research and Development
Washington, DC 20590

13. Type of Report and Period Covered
FINAL REPORT
10/87 - 9/93

14. Sponsoring Agency Code
RRD-11

15. Supplementary Notes

16. Abstract

The objective of this study was the development of a battery of tests for predicting an applicants' potential success for the job of locomotive engineer. In preparation for this battery, a job analysis was conducted.

The job analysis for the locomotive engineer's job involved seven steps: 1) review of existing research on locomotive engineer job requirements; 2) a site visit to one of the participating railroads, Union Pacific; 3) review of the task list by subject matter experts; 4) ratings of the tasks by Union Pacific engineers; 5) identification of the required knowledges, skills, and abilities (KSAs) by Union Pacific engineers; 6) review of the task list by additional participating railroads, (Amtrak, Burlington Northern, Conrail, and Santa Fe) including a site visit to Amtrak; and) identification of required KSAs by the other participating railroads. This thorough and comprehensive job analysis procedure resulted in a list of critical KSAs that were considered appropriate for subsequent test development. These requirements included reading, memorizing, understanding oral instructions, decision making, attention, and conscientiousness.

Test development proceeded directly from the job analysis results and was designed to be as practical as possible, to facilitate their use by the railroads, and to look the following six cognitive ability tests were developed to measure the important KSAs for the selection and promotion of locomotive engineers: Memory, Reading Comprehension, Perception, Listening, Logical Reasoning, and Dichotic Listening.

17. Key Words
Selection, Tests, Performance Measures, Validation

18. Distribution Statement

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

16. **Abstract (continued)**

Internal consistency analyses on the tests indicated that they had high reliability; coefficient alpha estimates ranged from .77 to .98. Furthermore, the tests correlated with one another in a meaningful, interpretable pattern. These analyses further substantiated the technical adequacy of the predictor battery.

The next step in the project involved determining if the tests predict engineers' job performance. This step is called validation. Validation is demonstrated by a statistical relationship between tests scores and ratings of job performance. In other words, a test is valid to the extent that tests scores predict job performance ratings.

The method used for validation in this study involved having a sample of engineers from 11 participating companies take the tests and be rated on their job performance, using a specially developed rating form. The relationship between the scores on the tests and the job performance ratings were statistically examined. No significant relationship was found between them which would be necessary for validating the tests. We hypothesized that the use of subjective ratings across varying conditions caused these results. A second validation study was conducted using a more objective measure of job performance, performance on a simulator.

Burlington Northern offered to have several classes of engineer trainees take the predictor tests. The trainees were also evaluated on three separate simulator runs, and on two end of training multiple-choice written knowledge tests involving general operating rules and air brakes.

Statistical analyses were performed on data from 97 engineer trainees to assess the relationship between the six cognitive ability tests, the two multiple-choice training tests, and scores on the three simulator runs. The major results are summarized as follows: 1) Scores on the cognitive ability tests were not significantly related to simulator performance; 2) Scores on three of the cognitive ability tests - Reading, Logical Reasoning, and Dichotic Listening - were significantly related to performance on the two training exams; 3) Performance on the training tests were significantly corrected with simulator performance.

EXECUTIVE SUMMARY

The first step in the study was a systematic analysis of the engineer's job. Project staff accomplished the following: observed engineers operate trains; interviewed engineers, road foremen, and trainers; and collected ratings of tasks and abilities. The result was the identification of the task and ability requirements of the engineer job.

The second step in the study was the development of a battery of tests to assess the abilities which had been rated as required for performance of the important engineer tasks. The ability tests which we developed measured: reading; logical reasoning; attention to detail; listening; memory; and dichotic listening (a measure of the ability to focus attention). Some of the tests (reading, logical reasoning, and memory) measure cognitive or reasoning abilities. The dichotic listening test and the attention to detail test both measure attention. The listening test measures both cognitive ability and perceptual ability (i.e., the ability to listen).

The tests were initially pretested with a sample of engineers and road foremen to identify confusing instructions and test items. A second pretest was undertaken with a sample of college students to determine appropriate time limits and to evaluate the technical quality of the tests. Statistical analyses of the pretest results and comparison of these tests with other selection tests indicated that the tests developed for the project were comparable to other tests which have proven effective for selecting applicants for a wide variety of jobs.

The third step in the project involved determining if the tests predict engineers' job performance. This step is called validation. Validation is demonstrated by s statistical relationship between tests scores and ratings of job performance. In other words, a test is valid to the extent that tests scores predict job performance ratings.

The method used for validation in this study involved having a sample of engineers, from the participating railroads, take the tests and be rated on their job performance. Project staff administered the ability tests at each of the companies to small groups of engineers. Later the engineers were rated on their job performance using a specially developed rating form. The engineers were observed during a normal run and then rated by their road foremen on their performance during that run. The relationship between the scores on these ability tests and the job performance ratings were then examined. There was no significant relationship between the tests and the ratings which would be necessary for validating the tests.

We concluded that there are two reasons for the disappointing findings. First, is the variability in railroads, road foremen who made the ratings, territory, equipment, train consist, and work conditions. This variability likely introduced use of different rating standards and expectations for the engineers. With such diversity, it is difficult to obtain significant relationships between test scores and performance ratings. The second factor concerned the degree to which the road foremen do not have an opportunity to consistently observe the performance of the engineers they supervise. In conclusion, we think that the most important cause of these findings was the use of a measure of job performance which was not given under consistent conditions and the results of which could be affected by the engineer's attempts to be careful when being observed.

We decided to conduct a second study using a more objective measure of job performance, performance on a locomotive simulator. Burlington Northern uses a combination of classroom instruction, on-the-job experience, and performance on an IITRI-built simulator to train their locomotive engineers. The trainees are then evaluated using two multiple choice written tests, one on operating rules and the second on air brakes, and performance on three-run simulator exercise. Burlington Northern offered to have their engineers take the tests during classroom training (administered by project staff) and send URC their scores on the written training tests and scores on the simulator exercises (both given at the conclusion of classroom training and on-the-job experience).

We statistically analyzed the relationship between the scores on the ability tests (those developed for selection), the written training tests (one on air brakes and one on operating rules), and the mean score on the simulator exercises. The results were:

- Three of the ability tests (reading, logical reasoning, and dichotic listening) were significantly related to performance on the two written training tests.

- Scores on the written training tests were significantly related to performance on the simulator.

- Scores on the ability tests were not significantly related to performance on the simulator.

Basic abilities (as measured by the tests we developed) are necessary to learn the job of locomotive engineer. People who have these abilities can successfully complete training and will be able to perform the engineer job under normal circumstances. Based on the results of this study, we can conclude that the tests we developed on this project are appropriately used to select applicants for engineer training. Engineer training is both necessary to learn the job and required according to FRA regulations. The use of the ability tests for selection into training can reduce training dropouts or failures. We, therefore, recommend that these tests be used to select candidates for engineer training.

TABLE OF CONTENTS

LIST OF EXHIBITS

LIST OF EXHIBITS - (CONTINUED)

LIST OF APPENDICES

INTRODUCTION

The locomotive engineer's job is a highly demanding one. Not only must engineers have the skill and feel necessary for the smooth and safe handling of the train, but they must also display this skill under varying track and road conditions. Efficient management of train operations, including safe handling of the train and keeping on schedule, is the responsibility of engineers.

The consequences of error in performance of the engineer's job, in terms of loss of life and destruction of property and equipment, are extraordinary. Improper or unsafe train handling can lead to severe consequences such as delays, equipment and cargo damage, derailments, collisions, and casualties. Therefore, railroads need to ensure that the individuals selected for (or promoted into) the engineer position have the skills and aptitudes necessary to learn how to perform engineer duties in a safe, efficient, and reliable manner.

Selection is the most effective and legally defensible way to identify competent engineer candidates. Selection procedures can have several benefits. First, such procedures can increase the productivity of workers since the capabilities of the workers adequately meet the demands of the job. These productivity gains can be translated into dollar savings for the organization. Hunter (1979) estimated that if the Philadelphia Police Department were to drop its use of a cognitive ability test to select entry-level police officers, the cost to the city would be 170 million dollars over a ten year period. Arnold, Rauschenberger, Soubel, and Guion (1982) concluded that the dollar savings due to increased productivity resulting from use of a selection test was $5,000 per steel worker during their first year on the job, or savings of $9.1 million for the company. The use of well-designed selection tests also reduces threats of legal action and provides organizations with the means to deal with charges of discrimination as they replace subjective estimates of an applicant's employability with more objective evaluations.

The Federal Railroad Administration (FRA) has recognized that, like other industries, the railroad industry faces legal, economic, and public pressures to improve the selection of its operating personnel, especially engineers. The FRA undertook this study to develop and validate a battery of selection tests that can be used for the identification of internal or external candidates for the engineer job. The purpose of the project is to provide a battery of tests to the railroads which they can use in their selection of engineer candidates.

JOB ANALYSIS

The first step of the project was a job analysis. A job analysis defines and documents the behaviors that are performed in a job. It is the most effective and appropriate procedure for identifying job requirements and for establishing job-related standards of effective performance. Therefore, job analysis is the appropriate basis on which to develop and validate selection procedures. In addition, the 1978 *Uniform Guidelines on Employee Selection Procedures* (Section 15 B [3]) requires job analysis except under very restricted circumstances that are not applicable to this study.

The job analysis conducted in this project provides an empirical link between the job requirements and the selection procedures. It involved seven steps: (1) Review of previous research; (2) A site visit to Union Pacific; (3) Subject matter expert review of the tasks list; (4) Ratings of tasks by Union Pacific engineers; (5) Identification of required knowledge, skills, and abilities (KSAs) by Union Pacific engineers; (6) Review of the task list by additional railroads (those railroads which became involved in the study after the beginning of the project) and site visits to Amtrak; and (7) Identification of KSAs by the additional railroads.

Review of Previous Research

Initially, the project staff reviewed earlier studies related to locomotive engineer job performance. Several previous studies identified tasks performed by locomotive engineers and determined the requirements for effective engineer performance. The most relevant study is the *Railroad Industry Job Analysis of Locomotive Engineer* (Railroad Personnel Association 1981) prepared by C.H. Lawshe. This job analysis resulted in lists of job activities (or tasks) rated for importance, critical job requirements, and work conditions. Hale and Jacobs (1975) reported the results of a study which determined the knowledge, skill, and training requirements for safe performance of engineer duties. McDonnell Douglas (1972) prepared detailed descriptions of the tasks performed by locomotive engineers during over-the-road freight operations. Project staff also reviewed a study URC conducted concerning the task and job requirements of commuter rail engineers (Myers, Harding, Hunter, and Fleishman, 1985).

A preliminary list of locomotive engineer tasks and work conditions was prepared using the task information available in these reports. The initial task list contained 31 tasks.

Site Visit

The project staff visited Union Pacific's Training Division in Salt Lake City in November, 1986. During this visit, the project staff discussed the initial task list with Union Pacific (UP) staff who conduct engineer training. The trainers eliminated, revised, and clarified some tasks, and added others to the provisional task list. In addition, the project staff, with a Union Pacific engineer trainer, rode in the engine cab on an over-the-road trip to observe engineer performance, equipment, and work conditions. Based on the meeting and observations, the preliminary task list was revised.

Review of Task List

In December, 1986, the revised list of job tasks was reviewed by seven locomotive engineers attending UP's Advanced Engineers Seminar. Generally, the task statements were well received although some clarifications and changes were suggested.

The revised task list was then reviewed by nine UP general road foreman. Further revisions and additions were made as a result of this step. The rating instructions and cover letter were also reviewed and revised by the road foremen and by UP management. The final task list contains 39 tasks and is presented in Exhibit 1.

Task Ratings

Job tasks frequently differ in their relative importance. The next step was undertaken to identify the tasks which are considered critical for the job. These tasks are the focus of the study since it is our concern that engineers have the capabilities to learn to perform these critical tasks.

Six UP road foremen were each sent five task rating forms to be filled out by locomotive engineers they supervise. The engineers were asked to rate each task using two ratings scales, Time Spent and Task Importance. The scales are presented in Exhibits 2 and 3.

Sample

Twenty-six engineers of the thirty to whom the forms were distributed (86.7%) returned the task ratings forms. The mean age of the engineers in the sample was 42.4 years. The engineers had an average of 13.35 years of education, 15.15 years experience as an engineer, and 19.54 years of railroad experience. The sample included 23 males and 3 females; 23 Whites, 2 Hispanics, and 1 African American engineer. Thirteen engineers were assigned to through-freight service, two to local service, two to yard service, and nine to other types of service. The locomotive engineer sample included 4 engineers from The Southern region, 15 from the Central region, and 7 from the Western region. Exhibit 4 includes more detailed descriptive statistics on the sample.

Data Analysis

The goal of the data analysis of the task rating was to identify a set of critical tasks that would be the foundation for identifying ability and skill requirements. As a first step in the data analysis, descriptive statistics were run on the two task ratings: Importance and Time Spent.

Exhibit 1

Task List
Locomotive Engineer

1. Obtain information required for trip including: train orders; special notices; general orders; work orders; special orders; load consist information; check train register.

2. Transmit information and/or instructions (in person or by electronic equipment) to other train crew members, dispatchers, mechanical force personnel, and other railroad employees.

3. Conduct job briefing by talking with crew about what needs to be done and how crew will operate to accomplish the job.

4. Inspect locomotive before run to verify quantity of fuel, sand, water, flagging equipment, and other supplies, as required by federal and company rules.

5. Sign daily inspection report, if no mechanical force personnel are available.

6. Start engines by operating switches, valves, and circuit breakers in proper sequence.

7. Perform initial terminal and other air brake tests as required by federal power brake law and company rules.

8. Receive and understand hand and radio signals.

9. Receive proceed signal from appropriate person (e.g., yard master, conductor), operate controls such as throttle and air brakes to move train.

10. Read and comply with train orders, signals, and railroad rules and regulations while operating locomotive.

11. Use knowledge of territory and train makeup to plan in advance how to synchronize throttle and brakes in order to operate train safely and efficiently.

12. Call out (wayside) train signals as they come up and receive verification from other crew members in cab.

13. Check accuracy of speed indicator by using watch to measure time between mileposts.

14. Observe track and surrounding area to detect obstructions and to anticipate operating problems.

Exhibit 1 (continued)

Task List

Locomotive Engineer

15. Identify malfunctions and reset protective devices.
16. Inspect locomotive and train during run to detect damage or defective equipment.
17. Prepare 2A Engine Work Report.
18. Notify proper authorities and, if necessary, prepare reports to explain accidents, unscheduled stops or delays, and advise designated personnel as specified by federal or company rules.
19. Operate locomotive between various shop locations, service tracks, and switching areas.
20. Operate locomotive in yard to switch cars between tracks.
21. Pilot or supervise operation of trains where engineer is unfamiliar with territory.
22. Start train from stretched or bunched condition and on varying grades.
23. Stop train in stretched or bunched condition and on varying grades.
24. Control speed and slack of train by use of throttle, dynamic braking, and/or air brakes.
25. Change operating ends of locomotive consist.
26. Set out or pick up units on line including connecting hoses or change hose mu cables.
27. Respond to unintentional application of automatic brakes.
28. Control throttle so as to avoid unnecessary stress on the engine, generator, traction motor and draw bars.
29. Operate helper locomotive under direction and in coordination with unit lead engineer.
30. Direct operation of helper locomotive by giving instructions to engineer.
31. Control operation of remote controlled engines.

Exhibit 1 (continued)

Task List
Locomotive Engineer

32. Modify train handling techniques in response to operating problems, malfunctions and changing conditions.

33. Observe condition of passing train and report results.

34. Operate pace setter system.

35. Sound whistle and ring bell when approaching crossing and during impaired visibility conditions.

36. Operate pulse equipment for cabooseless operations.

Exhibit 2

Time Scale

(1) How much time do you spend performing this task?

5 ____	I generally spend a great deal of time performing this task.
4 ____	
3 ____	I generally spend a moderate amount of time performing this task.
2 ____	
1 ____	I generally spend a small or no amount of time performing this task.

Exhibit 3

Importance Scale

(2) How important is the successful completion of this task for the smooth safe, and timely operation of the train?

5 ____	of critical importance to the smooth, safe and timely operation of the train
4 ____	
3 ____	of moderate importance to the smooth, safe and timely operation of the train.
2 ____	
1 ____	of little importance to the smooth, safe and timely operation of the train.

Exhibit 4

Description of the Sample of UP Engineers Rating the Tasks

Age (years)	N	Percentage
31-35	7	26.9
36-40	8	30.8
41-45	2	7.7
46-50	3	11.5
51-55	3	11.5
above 55	3	11.5
Education		
10-12	13	50.0
13-14	5	19.2
15-16	8	30.8
Experience as an Engineer (years)		
10-below	8	30.0
11-15	10	38.5
16-20	3	11.5
21-25	2	11.5
above 25	3	7.7
Railroad Experience (years)		
10-below	1	3.8
11-15	13	50.0
16-20	5	19.2
21-25	1	3.8
above 25	6	23.0

The inter-rater reliability of the ratings were then analyzed. Inter-rater reliability concerns the extent to which the raters agree in their ratings. For these ratings, inter-rater reliability provides an estimate of the degree to which the engineer participants agree in the relative importance and time requirements of the tasks. The inter-rater reliability of each rating scale was estimated using intraclass correlation coefficients. For each scale, an analysis of variance was performed and an intraclass coefficient calculated (Shrout & Fleiss, 1979, p. 423). This coefficient assumes random effects from raters and for tasks. Further, effects due to between rater differences and rater-by-task interaction residual are considered error variance, while between effects due to task differences are assumed to estimate true variance.

The intraclass correlation coefficient for the Time Spent rating is .93 for the mean across 22 raters (for whom complete data were available). The coefficient for the Importance ratings was .87 for the mean across 22 raters. These results indicate substantial agreement among raters in their assessment of the tasks. After determining the reliability of the ratings, we studied the ratings on each of the two scales.

The grand mean of the Time Spent rating was 2.9 on a five-point scale, with the task rating means ranging form 1.2 through 4.4 (see Exhibit 5). These mean task ratings cover the gamut of rating levels. In contrast, the mean task importance ratings are generally higher, with a grand mean of 4.2 and task means ranging from 2.6 to 5.0 (Exhibit 5). Few tasks were rated at or below moderate importance (three on the five-point scale).

In order to get an overall criticality rating for the tasks, the two ratings were combined using the following formula:

$$\text{Task Critically} = \frac{\text{Mean Time Spent}}{\text{Standard Deviation of the Mean Time Spent Ratings Across Tasks}} + \frac{\text{Mean Importance}}{\text{Standard Deviation of the Mean Importance Ratings Across Tasks}}$$

The mean task criticality values are presented in Exhibit 5. These indices provide the best estimates of task criticality.

Generally, it is acceptable to eliminate tasks from further study when they are rated as unimportant and infrequently performed. In this study, because of the generally high importance ratings, no tasks could be eliminated based on such a rule. Therefore, we concluded that all the tasks were critical.

Identification of Job Requirements

The next step in the job analysis was the determination of job requirements or characteristics which engineers must have in order to perform the job effectively. These job requirements cover the knowledges, skills, abilities, and other characteristics required for successful performance of job tasks.

Exhibit 5

UP Task Ratings

Task	N	Time Spent Mean	Time Spent S.D.	Importance Mean	Importance S.D.	Criticality Mean	Criticality S.D.
1	26	3.58	1.39	4.85	0.61	8.42	1.70
2	26	2.69	1.23	3.96	1.15	6.65	1.90
3	26	2.77	0.82	4.19	0.80	6.96	1.22
4	26	2.50	0.95	3.85	0.93	6.35	1.67
5	26	2.08	1.06	3.31	1.16	5.39	1.96
6	26	2.00	1.27	3.96	1.00	5.96	1.82
7	26	3.15	1.35	4.89	0.20	8.04	1.13
8	26	3.96	1.11	4.96	0.20	8.92	1.13
9	26	3.92	1.32	4.65	0.63	8.58	1.63
10	26	4.39	0.80	4.92	0.27	9.31	0.88
11	26	4.35	1.02	4.96	0.20	9.31	1.16
12	26	3.27	1.22	4.23	1.11	7.50	1.79
13	26	1.50	0.76	2.58	1.30	4.08	1.67
14	26	2.85	1.35	4.19	1.02	7.04	1.95
15	26	3.04	1.46	4.31	1.05	7.35	2.23
16	26	3.77	1.11	4.58	0.64	8.35	1.57
17	26	2.62	1.27	3.85	0.88	6.46	1.79
18	26	3.31	1.29	4.62	0.57	7.92	1.57
19	26	2.08	1.02	3.39	1.36	5.46	2.02
20	26	2.12	1.03	4.27	0.87	6.39	1.53
21	26	2.42	1.27	3.31	1.32	5.73	2.18
22	26	2.73	1.61	3.77	1.42	6.50	2.67
23	26	1.31	0.74	4.04	1.31	5.35	1.60
24	26	3.50	1.21	4.46	0.76	7.96	1.78
25	26	3.77	1.03	4.46	0.76	8.23	1.66
26	26	4.42	0.95	4.77	0.51	9.19	1.33
27	26	2.08	0.85	4.00	1.17	6.08	1.70
28	26	2.27	1.04	4.00	1.20	6.27	1.85
29	26	2.27	1.28	4.73	0.53	7.00	1.50
30	26	4.08	1.09	4.50	0.71	8.58	1.53
31	25	1.48	0.92	4.00	1.35	5.48	1.92
32	26	1.58	1.10	3.89	1.42	5.46	2.06
33	23	1.22	0.74	3.48	1.70	4.70	1.99
34	26	3.50	1.21	4.54	0.76	8.04	1.76
35	26	3.46	1.10	4.62	0.70	8.08	1.52
36	23	1.30	0.82	2.91	1.47	4.22	1.83
37	26	4.31	1.05	4.92	0.27	9.23	1.11
38	26	3.23	1.51	4.27	1.12	7.50	2.05
39	26	2.89	1.34	4.81	0.63	7.69	1.59

The procedure used to identify job requirements is based on the research of the project consultant, C.H. Lawshe. Dr. Lawshe developed an inventory of job requirements which covers: education proficiency requirements (e.g., understanding writtenmaterial, performing calculations); other proficiency requirements (e.g., understanding oral communication; understanding graphic information); decision making and information processing requirements (e.g., making choices); and physical/sensory requirements (e.g., using hands in work activity).

The job requirements studied by Lawshe were compared to the job requirements that previous studies identified as needed by a locomotive engineer (Hale & Jacobs, 1975; Myers, Harding, Hunter, & Fleishman, 1985; Railroad Personnel Association, 1981) and to more general analyses of abilities and skills (Ekstrom, French, & Harman, 1976; Fleishman & Quaintance, 1984; McCormick, Jeanneret, & Mecham, 1972). The goal of the review was to prepare a list of job requirements that included all job requirements that could be assessed in a written selection test. The result, presented in Exhibit 6, is a list of 51 skills and abilities categorized into 11 sets of job requirements. The complete list of job requirements is in Appendix A.

Procedure to Identify Job Requirements

A group of 14 engineers attending Union Pacific's Advanced Engineering Training Course participated in this study. They were given the list of 39 locomotive engineer tasks that were prepared in the job analysis phase of the study and a rating form which listed the 51 job requirements.

The rating form presented each of the 11 job requirement areas on a separate page with the specific job requirements from each area listed below. The raters were asked to review each of the job requirements and determine if it is/is not essential in order to safely and competently perform each job task. If the requirement was considered essential, a check was written, and if it is not, a zero was written.

All 14 of the engineer participants were White males. Their mean age was 39.5 years. The engineers had an average of 13.1 years of education, 15.8 years of experience working in railroads, and 12.4 years of experience as an engineer. Seven engineers were assigned to through-freight service, three to local service, two to yard service, one to extra boards and one had another assignment. Six of the engineers worked in the Southern region, two in the Central region, and five in the Western region of the Union Pacific System.

Reliability

The inter-rater reliability in the ratings made by these engineers was assessed using intraclass correlation coefficients (Shrout & Fleiss, 1979, p. 423). The average agreement, across the raters, as to whether a job requirement was or was not necessary to perform a task was estimated at .79, which is adequate for our purposes. The number of raters in the analysis was 10, because these were the number of raters who made all of the ratings.

Exhibit 6

Job Requirement Categories

Understanding printed/written material

Performing calculations

Understanding oral communications

Making oneself understood in writing

Understanding graphic information

Exercising mechanical insight

Making estimates

Making choices and/or solving problems

Making visual or auditory discrimination

Making gross body movements

In addition, the inter-rater reliability for each job requirement separately was estimated using the same analysis. The results are presented in Exhibit 7. Some of the reliabilities are quite low, although most indicate acceptable levels of agreement in job requirement judgments.

Identification of Necessary Job Requirements

The job requirements data were analyzed in two ways. The first method has been used by Lawshe (1987, personal communication). According to Lawshe, a job requirement is necessary for performing a task if significantly more than half of the raters indicate that it is required to perform the task. Using this approach, we analyzed the ratings data to identify whether each job requirement was necessary to perform each task. Exhibit 8 presents, for each job requirement, the number of tasks for which the job requirement was identified as necessary.

Using this analysis, the following job requirements were rated as essential to perform the greatest number of tasks: reading, memorizing, understanding oral communication, making oneself understood orally, maintaining attention, and using ones hands and arms for reaching.

The ratings data were then analyzed with a consideration given to the relative criticality of the task. Using this analysis, both the number of tasks and the criticality of the tasks linked to the job requirement are used in the analysis. In this analysis, the mean percentage of raters who identified a job requirement as necessary for the tasks was determined. This percentage was weighted in terms of task criticality (normalized so that the sum = 1), for each of the tasks linked to the job requirement, so that the more critical tasks had a greater impact on the percentages. The percentages are presented in Exhibit 9.

These latter percentages were then rank ordered (see Exhibit 9). It is reasonable to assume that the top ranked job requirements are the more important requirements for the study. We decided to consider the 25 highest ranked (from a total of 51) job requirements. The weighted mean percentage of the 25 highest job requirements is 26.8%. This percentage is slightly greater than expected if half the raters judged the job requirement as needed for one-half of the tasks. Using this rule, the following job requirements are needed to be an engineer: memorizing, adding/subtracting, understanding oral communication, making oneself understood orally, making choices/solving problems, discriminating visual detail, recognizing colors, maintaining attention and reaching.

We also reviewed the findings of three studies which analyzed the requirements of the locomotive engineer job. Hale and Jacobs (1975) determined, through expert judgment, the job requirements needed for safe operations of a train. The Railroad Personnel Association (1981) identified the job requirements of engineers, as rated by engineers and supervisors. Myers, Harding, Hunter, and Fleishman (1985) had a group of engineers and psychologists select the more important cognitive, perceptual, psychomotor, and physical ability locomotive engineer job requirements.

Exhibit 7

Inter-rater Reliability of the UP Job Requirement Rating

General Requirements	Job Requirements	Intra Class Correlation
Understanding printed/ Written Material	1	.62
	2	.65
	3	.68
	4	.53
	5	.72
	6	.32
Performing calculations	7	.65
	8	.61
	9	.53
	10	.57
	11	.40
Understanding oral communication	12	.83
	13	.70
	14	.24
	15	.51
Making oneself understood orally	16	.80
	17	.66
	18	.35
	19	.60
Making oneself understood in writing	20	.80
	21	.73
	22	.90
	23	.90
	24	.24
Understanding graphic information	25	.26
	26	.48
Exercising mechanical insight	27	.25
	28	.23
	29	.33
	30	.14

Note: The numbers of these job requirements refer to the list of job requirements in Appendix A

Exhibit 7 (continued)

Inter-rater Reliability of the UP Job Requirement Rating

General Requirement	Specific Ability/Skill	Intra Class Correlation
Making estimates	31	.41
	32	.48
	33	.52
Making choices/solving problems	34	.09
	35	.23
	36	.19
	37	.33
Making visual/auditory discrimination	38	.52
	39	.49
	40	.53
	41	.59
	42	.18
	43	.12
	44	.44
	45	.15
	46	.45
Using hands/fingers	47	.84
	48	.47
Gross body movements	49	.82
Climbing balancing	50	.82
	51	.80

Exhibit 8

The Number of Task Needing Each Job Requirement

Number of Task Requiring General Requirement	Job Requirements	Job Requirements
Understanding printed/ Written Material	1	9
	2	6
	3	5
	4	4
	5	1
	6	9
Performing calculations	7	1
	8	0
	9	0
	10	0
	11	0
Understanding oral communication	12	15
	13	12
	14	0
	15	1
Making oneself understood orally	16	11
	17	6
	18	0
	19	0
Making oneself understood in writing	20	0
	21	0
	22	2
	23	2
	24	0
Understanding graphic information	25	0
	26	0
Exercising mechanical insight	27	0
	28	0
	29	0
	30	0

Note: The numbers of these job requirements refer to the list of job requirements in Appendix A

Exhibit 8 (continued)

The Number of Task Needing Each Job Requirement

Number of Task Requiring General Requirement	Job Requirements	Job Requirements
Making estimates	31	1
	32	0
	33	0
Making choices/solving problems	34	0
	35	0
	36	0
	37	0
Making visual/auditory discrimination	38	0
	39	0
	40	1
	41	2
	42	0
	43	0
	44	0
	45	0
	46	5
Using hands/fingers	47	15
	48	0
Gross body movements	49	1
Climbing balancing	50	1
	51	0

Exhibit 9

The Weighted Mean Percentage of UP Raters Identifying An Ability As Required

General Requirement	Job Requirements Specific Ability/Skill	Mean Percentage	Ranking
Understanding printed/ Written Material	1	53.8	5
	2	42.9	12
	3	34.1	20
	4	40.7	12
	5	24.6	26
	6	65.4	1
Performing calculations	7	32.9	21
	8	7.2	47
	9	5.1	50
	10	5.8	49
	11	8.3	43
Understanding oral communication	12	55.5	4
	13	63.6	3
	14	38.8	15
	15	26.8	25
Making oneself understood orally	16	52.0	7
	17	51.1	8
	18	21.3	30
	19	10.5	40
Making oneself understood in writing	20	7.7	46
	21	7.8	45
	22	13.7	35
	23	12.9	37
	24	2.9	51
Understanding graphic information	25	6.3	48
	26	12.3	38
Exercising mechanical insight	27	20.2	31
	28	21.5	29
	29	8.0	44
	30	11.8	39

Exhibit 9 (continued)

The Weighted Mean Percentage of UP Raters Identifying An Ability As Required

General Requirement	Job Requirements Specific Ability/Skill	Mean Percentage	Ranking
Making estimates	31	14.6	34
	32	18.2	32
	33	24.5	27
Making choices/solving problems	34	43.6	10
	35	44.7	9
	36	37.8	16
	37	29.8	23
Making visual/auditory discrimination	38	36.6	17
	39	23.1	28
	40	39.6	14
	41	34.8	18
	42	29.0	24
	43	31.1	22
	44	34.2	19
	45	17.6	33
	46	52.9	6
Using hands/fingers	47	63.9	2
	48	42.9	11
Gross body movements	49	13.2	36
Climbing balancing	50	9.8	42
	51	9.9	41

We combined the results from the three studies and the two analyses of the data from the present study. These results are presented in Exhibit 10. Included in the table are the job requirements identified in the three previous studies, the ranking of the job requirements based on the weighted percentages, and the number of raters who identified each job requirement as essential.

Using all these sources of information, we conclude that the following are the most important job requirements for the locomotive engineer: reading, memorizing, making choices, understanding oral communications, making oneself understood orally, recognizing colors, judging distance, maintaining attention, and using one's hands and arms for reaching. The least important job requirements are: performing calculations, understanding graphic information, exercising mechanical insight, making gross body movements, and climbing/balancing.

Exhibit 10

Combined Analyses of Locomotive Engineer Job Requirements

General Requirement	Specific Job Requirement	Ranking	No. of Tasks Needed for Job Requirements	Railroad Personnel Assoc.	Hale & Jacobs	Myers etc,
Understanding printed/written material				X		X
	1	5	9			
	2	12	6			
	3	20	5			
	13	4				
	5	26	1			
	6	1	9	X	X	X
Performing calculations						X
	7	21	1			
	8	47				
	9	50				
	10	49				
	11	43				
Understanding oral communication				X		X
	12	4	15			
	13	3	12			
	14	15				
	15	25	1			
Making oneself understood orally					X	X
	16	7	11			
	17	8	6			
	18	30				
	19	40				

Exhibit 10 (continued)

Combined Analyses of Locomotive Engineer Job Requirements

General Requirement	Specific Job Requirement	Ranking	No. of Tasks Needed for Job Requirements	Railroad Personnel Assoc.	Hale & Jacobs	Myers etc.
Making oneself understood in writing						X
	20	46				
	21	45				
	22	35	2			
	23	37	2			
	24	51				
Understanding graphic information						
	25	48				
	26	38				
Exercising mechanical insight						X
	27	31				
	28	29				
	29	44				
	30	39				
Making estimates						
	31	34	1			
	32	32				
	33	27		X		
Making choices/solving problems				X		X
	34	10				
	35	9		X		
	36	16				
	37	23				

Exhibit 10 (continued)

Combined Analyses of Locomotive Engineer

Job Requirements

General Requirement	Specific Job Requirement	Ranking	No. of Tasks Needed for Job Requirements	Railroad Personnel Assoc.	Hale & Jacobs	Myers etc.
Making visual/auditory discrimination				X		X
	38	17				
	39	28				
	40	14	1	X		
	41	18	2			X
	42	24				
	43	22				
	44	19				
	45	33				
	46	6	5			X
Using hands/fingers						X
	47	2	15			
	48	11				
Gross body movements						
	49	36	1			
Gross body movement						
Climbing/balancing						
	50	42	1			
	51	41				

INCLUSION OF NEW RAILROADS

The President of the Railroad Personnel Association (RPA), and the Assistant Vice President of the American Association of Railroads, were contacted concerning increasing the number of railroads participating in the study. They arranged a meeting representatives of Norfolk Southern (representing RPA), Burlington Northern and Santa Fe to discuss the project. In addition to these companies, Amtrak and Conrail began discussions about the project. At the conclusion of the negotiations, Amtrak, Burlington Northern, Conrail, and Santa Fe decided to join the study. Because of the differences between freight and passenger trains, project staff made two trips on Amtrak trains to observe engineer performance.

Because the job analysis was undertaken with only one railroad, we decided to perform an abbreviated job analysis with the new railroads. During December, 1987 and the first several months of 1988, the railroads were sent task review forms.

Task Review

The purpose of the task review was to determine the comprehensiveness and adequacy of the Union Pacific-derived task list for the new railroads and to develop company-specific task lists. Each company was sent a set of task review forms and asked to have them filled out by a sample of personnel knowledgeable about the engineer job. The sample is presented in Exhibit 11. The Union Pacific task list was used for the task review and the raters were asked if each task did/did not describe an important part of the engineer's job at their company. They were also asked how well the task list covered the important parts of the engineer's tasks. Finally, they were given the opportunity to revise and to add tasks to the list. The results for each added railroad were analyzed to determine if the railroad engineer jobs were behaviorally similar enough to be considered the same as the engineer jobs at other railroads. For a task to be considered important to the engineer's job, 80% of the raters must have given that task a "yes" rating. Similarly, 80% of the original task list had to have been considered an important part of the engineer's job for that railroad to be included in the study. Summary of the responses follows in Exhibit 12. Of the four railroads that were evaluated, all had an 81% overlap or greater with the UP task list. This indicates that the engineer jobs at the four additional railroads are all similar to the UP engineer jobs.

Summary of Data from the New Railroads

Despite variations in railroad rules and conditions, the original task list accurately describes the engineer's job in the four additional railroads. Most tasks included in the original Union Pacific task list were rated as an important part of the engineer's job in each of the new railroads.

Exhibit 11

Task Review

Description of Sample of Raters from Added Railroads

	Amtrak	Burlington Northern	Conrail	Santa Fe
Demographics				
Number of Raters	5	13	9	26
Mean Age	54.6	43.3	46.8	43.9
Sex Mix	5 males	13 males	9 males	26 males
Ethnic Mix	4 White 1 African American	12 White 1 American Indian	9 White	25 White 1 American Indian
Mean Years Experience as Engineer	18.2	9.1	16.7	12.9

Exhibit 12

Percent of Raters that Considered the Task to be an Important Part of Their Job

Tasks	RAILROADS			
	Amtrak n=5	Burlington N=13	Conrail n=9	Santa Fe n=26
1.	100	100	100	100
2.	80	73	89	100
3.	80	45	78	88
4.	100	92	100	100
5.	100	100	100	92
6.	80	83	89	85
7.	100	100	100	100
8.	100	100	100	100
9.	100	100	100	100
10.	100	100	100	100
11.	100	100	100	100
12.	100	100	100	100
13.	80	25	22	69
14.	100	92	100	92
15.	100	100	100	100
16.	100	100	100	100
17.	80	92	100	100
18.	100	92	89	100
19.	80	72	78	88
20.	100	83	89	88
21.	80	60	89	88
22.	80	100	89	100
23.	100	92	89	92
24.	100	100	100	100
25.	100	100	100	100
26.	80	100	100	100
27.	100	100	100	100
28.	100	100	100	100
29.	100	83	100	92
30.	80	72	89	92
31.	20	35	67	100
32.	100	100	100	100
33.	100	92	100	96
34.	0	92	78	96
35.	100	100	100	100
36.	0	92	100	100
37.	100	100	100	100

When given the opportunity to modify the task list, most of the changes concerned terminology. The task, "Prepare 2A Engine Work Report," was modified in terms of the report number or title used in each railroad. Modifications were also made to the task "Relay wayside or cab signals to the dispatcher using the radio." This task was also the one least likely to be rated as an important part of the engineer's job. Each company did, additionally, make some modifications to the task list, including the elimination and revision of tasks.

Under the section of the questionnaire regarding task additions most of the recommendations concerned job requirements or abilities including attention to changing conditions. Several respondents also mentioned training new engineers as a requirement of the engineer's job.

It is not surprising that the Amtrak respondents made the greatest number of changes to the task list. What is surprising, given the differences is their assignments (briefer trips, shorter trains, faster trips, operations of passenger rather than freighter trains), almost all the tasks performed by freight engineers at Union Pacific were performed by Amtrak engineers. Three tasks were not performed: "Operate pace setter system", "Operate telemetry equipment for cabooseless operations," and "Control operation of remote controlled engines."

As a result of this review, company-specific task lists were prepared for each of the four railroads. These task lists were used for the determination of the job requirements in the new railroads.

Job Requirements Ratings

The purpose of obtaining additional job requirements ratings was to determine if the job requirements needed at the new railroads were different from those needed at Union Pacific. In addition, Conrail suggested including additional job requirements on the rating form since their validation studies indicated that measures of certain personality traits predicted job performance.

The company coordinators at each of the new railroads--Amtrak, Burlington Northern, Conrail, and Santa Fe--were each sent job requirement rating forms. They were asked to select individuals who were knowledgeable about the locomotive engineer job respond to the forms. The raters were asked to indicate if each job requirement was/was not essential to perform the tasks on the list. The sample of raters for the new companies is described in Exhibit 13.

Exhibit 13

Description of the Sample of Raters from the New Railroads Rating Job Requirements

	Amtrak	Burlington Northern	Conrail	Santa Fe
Background				
Number of Raters	13	9	16	21
Mean Age	38.6 yrs.	41.0 yrs.	39.8 yrs.	45.9 yrs.
Sex Mix	13 males	9 males	16 males	19 males, 1 female 1 no I.D.
Ethnic Mix	12 Whites, 1 African American	9 Whites	16 Whites	20 Whites
Mean Years Experience as Engineer	12.08 yrs.	10.22 yrs.	12.69 yrs.	17.1 yrs.

Because the job requirements ratings were to be conducted using a mail-out procedure, we concluded that fewer job requirements could be included in the rating form that had been used previously at Union Pacific. We selected job requirements from those rated by Union Pacific engineers, using the following criteria:

1. Ranked in the top 10 in terms of the number of tasks needed for the job requirement, or
2. Required for several (three or more) tasks by significantly more than half of the raters, or
3. Identified as required in two or more previous studies of engineers, or
4. Required by at least one task and identified as required in one previous study, or
5. Identified by Conrail as a predictor of engineer job performance in their validation study.

The final list of 20 job requirements is presented in Exhibit 14.

In analyzing the data from the four new companies and reanalyzing the data from Union Pacific, we first determined the number of tasks identified as needing a job requirement by significantly more than half of the sample (see Exhibit 15). Since the tasks listed and number of tasks differed among the companies, we could not compare the companies in terms of the number of tasks linked to each job requirement. Instead, the job requirements were ranked in terms of number of tasks rated as needing the job requirements.

Two analyses were carried out to compare the rankings across the companies. In one analysis, the 14 job requirements common to the ratings in all five railroads were compared. Kendall's coefficient of concordance (Walker & Lev, 1953, p.284) was used to compare the ratings of the job requirements. The coefficient of concordance is a measure of the agreement in the rankings of job requirements across raters (in this case, the railroads). The coefficient W is .53 ($df = 13$, $p < .001$). The coefficient, W, was also calculated for the 20 job requirements common to the four new companies. In this analysis, W equals .76 ($df = 19$, $p < .001$). Both analyses indicate significant agreement in the ranking of job requirements. We can conclude that there is a consistency across the railroads in the job requirements identified as necessary for performing important engineer tasks.

Exhibit 14

Job Requirement for Review by New Railroads

Understanding Printed/Written Material

1. Reading simple words, such as position signs on machine equipment (e.g., "On/Off", or "Start/Stop").
2. Reading simple sentences, such as posted signs or directions (e.g., "Keep boxes out of aisles").
3. Reading complex sentences, such as written material on work tickets or printed material on containers (e.g., "This material may explode if it gets wet").
4. Reading paragraphs which describe a thing or event or present multiple instruction in sequence, such as instructions in operating.
5. Memorizing and recalling specific information learned from printed materials.

Making Oneself Understood Orally

6. Understanding oral instructions or work procedures information provided by supervisors or others.
7. Coordinating work with co-workers through conversation/discussion where effectiveness depends upon being understood.
8. Providing routine oral status or progress reports to supervisor or others, in person, by phone, or by radio.

Making Choices

9. Making choices/decisions in which the risks or consequence are slight, such as: sorting materials or parts.
10. Making choices/decisions affecting the security or well-being of others and/or which involve serious risk or consequences.

Exhibit 14 (continued)

Job Requirement for Review by New Railroads

Perceptual Abilities

11. Recognizing colors, such as: light signals, containers, or electrical parts.
12. Maintaining attention to a task over long periods of time.
13. Judging distance from observer to objects and/or between objects.

Using Hands in Work Activity

14. Reaching-extending hand(s) and arm(s) in any direction.
15. Exercising hand-eye coordination.

Personality Traits

16. Agreeable - good natured and cooperative.
17. Conscientious - responsible, careful and dependable.
18. Calm - composed under stress.
19. General Activity busy, active in projects.
20. Outgoing - likes to be with people.

Exhibit 15

Number of Tasks Rated as Requiring a Job Requirement by Significantly More than Half of the Raters in each Company

	AMTRAK	BURLINGTON NORTHERN	CONRAIL	SANTA FE	UNION PACIFIC
Reading Words	10	8	15	12	9
Reading Sentences	8	8	7	7	6
Reading Complex	2	2	5	4	5
Reading Para	4	4	4	3	4
Memorizing	30	23	38	32	9
Understanding Oral	12	16	16	18	12
Coordinating	12	13	21	19	11
Providing Rpts.	11	2	19	11	6
Slight Choices	4	20	27	31	0
Serious Choices	2	8	8	21	0
Colors	9	5	7	11	1
Attention	17	11	23	20	5
Distance	17	11	23	19	2
Reaching	15	0	18	21	15
Hand-eye	18	1	20	23	--
Agreeable	1	0	13	4	--
Conscientious	34	9	38	37	--
Calm	11	7	33	20	--
Activity	0	0	0	0	--
Outgoing	0	0	0	0	--

Amtrak rated **34** tasks
Burlington Northern rated **37** tasks
Conrail rated **39** tasks
Santa Fe rated **37** tasks
Union Pacific rated **39** tasks

Given this consistency, the most highly ranked job requirements were identified as appropriate for test development. These are:

- Reading
- Memorizing
- Understanding Oral Instructions
- Speaking
- Decision Making
- Recognizing Colors
- Attention
- Judging Distance
- Reaching with Hands/Arms
- Being Conscientious
- Being Calm

TEST DEVELOPMENT

Test Development Goals

In preparing the test plan, we considered the following objectives:

1. Practical - To the extent feasible, the tests should be written and chosen so they can be easily used by the railroads. Considerations in making this decision include:

    - ease of administration;
    - objectivity and ease in scoring;
    - amenability to group or individual administration; and
    - use of simple or minimal equipment.

2. Face valid - To the extent possible, we should include selection procedures that look relevant to the job.

3. Appropriate for both entry selection and for promotion.

4. Relevant to the job requirements of the job.

Preliminary Test Plan

Using the test development goals for guidance, we prepared a plan concerning the measure of important job requirements identified in the job analysis. The preliminary test plan is presented in Exhibit 16.

We decided to develop several tests that would include engineer-relevant material. We thought that the reading, memory, logic and attention tests could be feasibly developed using engineer-relevant content. Published tests measuring other job requirements were already available. These tests included measures of understanding oral instructions, conscientiousness and calmness.

When reviewing the literature on measures of attention, we found a testing procedure that predicts pilot training success (Gopher, 1982; Gopher & Kahneman, 1971) and reduced accident rates in bus drivers (Kahneman, Ben-Ishai, & Lotan, 1973). The test measures selective attention and involves dichotic listening or listening to a different message in each ear. The test requires participants to maintain attention to vocal information presented to a designated ear and to ignore information simultaneously presented to the other ear. Dr. Glen R. Griffin, of the Naval Aerospace Medical Research Laboratory provided URC with a tape recording of the dichotic listening test he was using in research on Navy pilots.

Exhibit 16

Preliminary Test Plan

Job Requirement	Recommended Test
Reading	URC Reading Test
Memorizing	URC Memory Test
Understanding Oral Instructions	PSI - Basic Skills Test Following Oral Directions
Speaking	Interview
Making Choices and Decisions	URC Logic Test
Recognizing Colors	Recommendations for Physician's Examination
Attention	URC Perception Test Dichotic Listening
Judging Distance	Cannot be assessed feasibly
Reaching with Hands/Arms	Recommendation for Physician's Exam
Hand-Eye Coordination	URC Computerized Test of Coordination
Conscientious	Hogan Reliability Scale
Calm	Personality Test

The literature on judging distance was extensive but we were unable to find any standard measures of distance perception that would be useful for selection testing. One obvious measure would be the distance perceived over a number of yards. This measure of distance would require a field or space not generally available during test administration. The use of pictures and film of objects at some distance were problematic methods for measuring perceived distance. Perceptual psychologists, with whom we spoke, concluded that simulated distance measures do not correlate with the perception of real distance. Since the assessment of actual distance is problematic in most test situations, we decided to refrain from attempting to measure this ability in the test battery.

Test Review

In November, 1988, representatives from the participating railroads (Amtrak, Burlington Northern, Conrail, Santa Fe, and Union Pacific) and Garold Thomas, the FRA project monitor, met with URC project staff to review the draft test plan and examples of test items.

The company representatives generally liked the test battery. They made specific suggestions regarding the content, format and instructions of specific tests. The group suggested that the format of the Logical Reasoning and Perception tests should be modified. The group did not like the Understanding Oral Directions Test. They thought it was too clerical in nature and suggested that it be revised with engineer content. The representative from Burlington Northern volunteered to assist us in the development of a new test on listening skills. The group also decided not to include any measures of personality, including a test measuring antisocial behavior, because the tests were considered intrusive and excessively personal. The meeting participants also decided to refrain from using a computerized test of hand-eye coordination because of limited availability of personal computers in the railroads. The group also recommended assessing the reading level of the reading test to ensure it was consistent with the reading level required for the job. Several suggestions were also made for improving the test instructions.

Test Completion

The tests were revised and completed consistent with the recommendations given by the railroad representatives. The revised test plan is presented in Exhibit 17.

We decided to evaluate the reading level of the reading test, both to meet the request of the railroad representatives and to ensure that the verbal complexity of this test was no more difficult than the materials that must be read and understood on the job. In order to assess the relative difficulty of the reading test passages and locomotive engineer materials, we requested that the railroads provide us with documents that represented what the engineers needed to read and understand. The companies sent rules and regulations.

Exhibit 17

Revised Test Plan

Job Requirement	Recommended Testing Procedure
Reading	URC Reading Test
Memorizing	URC Memory Test
Understanding Oral Instructions	URC Listening Test
Speaking	Interview
Making Choices and Decisions	URC Logic Test
Recognizing Colors	Physical Examination
Attention	Dichotic Listening URC Perception Test
Reaching with Hands/Arms	Physical Exam

We evaluated the complexity of the test passages and company materials using the Flesch Reading Ease Index, which is calculated using Grammatik II, a computer program which analyzes writing. The Flesch formula is based on sentence length and the number of syllables per hundred words in samples from prose passages (Flesch, 1948). The formula was used to calculate the reading level of the reading passages and selected passages (of approximately two pages in length) from materials sent by each railroad.

The reading level of the test passages was lower than that of the reading material used on the job. The reading test complexity level was seventh grade level. The materials from Amtrak and Santa Fe required an 11th grade reading level; those from Conrail required a 12th grade reading level; and the passages from Burlington Northern and Union Pacific required a college sophomore reading level.

Pretesting the Battery

During February, 1989, we conducted two pretests of the test battery with railroad engineers and road foremen. At Conrail, two road foremen, two engineers, and two engineer instructors participated in the pretest. All were male and had a mean of 8.5 years of experience as engineers. At Amtrak, six road foremen and one transportation manager participated in the pretest. Again, all were white males and had an average of 15.4 years of experience as engineers.

The pretests were conducted to identify confusing test items and instructions as well as to estimate the time required for each test in the battery. The procedure for the pretests involved explaining the project to the participants, administering a test, discussing the test, and going on to the next test. As a result of the pretest: one reading paragraph was eliminated and replaced, several logic questions were eliminated, the content of the listening test was simplified, some questions on the memory test were changed, and the instructions for all of the tests were made clearer and simpler.

During March and April, 1989, a second set of pretests were undertaken. The purpose of these pretests were to evaluate the individual items on each test in the test battery. Because a larger sample is required for item analysis, we decided to use a college sample where obtaining a large sample was more feasible than at a railroad or a technical school.

Participants were volunteers from three colleges: the University of Virginia, the University of Maryland, and Piedmont Virginia Community College. The number of participants who took each test at each school is in Exhibit 18.

The pretest was conducted, in part, to determine if the proposed time limits were reasonable. It is our goal to have only the perception test as a speeded test and, therefore, we want the respondents to finish all of the tests. Most of the proposed time limits were appropriate, with the exception of the time limits for the reading and logic tests for which all participants finished with 10 minutes to spare.

Exhibit 18

Pretest Participants

TEST	UNIVERSITY OF VIRGINIA	UNIVERSITY OF MARYLAND	PIEDMONT VALLEY COMMUNITY COLLEGE
Reading	76		30
Memory	76		30
Listening	76		30
Logic	56		30
Dichotic Listening	20	11	--
Perception	76		30

The participants also identified unclear items. They found the listening test contained too much information and made suggestions to improve the clarity of the perception test.

Item analyses were also conducted. The criteria we used to evaluate the items were: item-test correlations; whether the correct answer was the most frequently chosen; whether one incorrect answer was frequently chosen; and item difficulty. Items were considered for elimination or revision if: the item was not correlated with total test score; if one wrong answer was frequently selected; and, if the item was exceedingly hard or easy.

The memory test and the logic test had adequate internal consistency, reliability (.85 and .82 respectively), and only a limited number of items that needed to be revised. The reading test required a great deal of revision. In consultation with Dr. Lawshe, we decided to lengthen this test and to replace the items that did not correlate with the total reading test score. We also decided to replace the perception test items that did not correlate with the total test score. Some changes were also made in the listening test questions.

Because of the extensive revisions in the reading test, we again evaluated its reading level using the Flesch Test. This time the reading level was the 11th grade, which is at or below the reading levels of the engineer documents.

DEVELOPMENT OF CRITERION MEASURE

The goal of this study was to develop and validate a battery of tests for selecting candidates most likely to be able to learn to perform the locomotive engineer job. In order to empirically validate this test battery, it was necessary to develop a measure of job performance or criterion measure. The criterion measure would be the standard against which the test battery would be validated.

The major factor complicating the development of the criterion was the lack of close and consistent supervision of engineers. Road foremen ride with engineers at regular intervals, but these intervals can be every six months or longer. The other personnel (switchman, brakemen) who consistently ride with engineers are not necessarily trained or experienced to evaluate engineer job performance.

Since traditional supervisory or peer ratings were not appropriate as a criterion measure, we considered the use of company accident and incident records. Company managers were reticent to allow us access to personal records of this sort. In addition, these records would not differentiate adequate from good performances and, at best, would identify specific problematic incidents.

The third option considered was the use of the records made by event recorders mounted on locomotives. Although the records from these recorders can be used to evaluate speed, braking, and stalling, the track and route must be considered in evaluating each pulse tape. Since there were no consistent rules for evaluating these tapes and linkage of the tape with a specific engineer might be problematic, we choose not to use these as a criterion measure.

Instead, we decided to develop an observational rating form that would be used by road foremen. Since the road foreman would rate the engineers after observing their performance, the ratings would likely reflect engineer performance.

Development of the Rating Form

As a first step in preparing the observational rating form, we reviewed the reports which described specific engineer behaviors. Among the documents reviewed were the report of the L & N/BLE demonstration projection for locomotive engineer training (Department of Transportation, 1982), *Train Track Dynamics* (2nd edition, undated), *Railroad Engineman Task and Skill Study* (McDonnell Douglas Corp. 1972) and *Railroad Engineer Task and Skill Analysis* (Louisville & Nashville Railroad, 1981). In addition, the engineer rating forms for Norfolk Southern, Conrail, and Union Pacific were reviewed. We also spoke to Union Pacific road foreman concerning the key components of engineer performance, the behaviors they are able to observe and the common errors that engineers make.

Based on the literature review and interviews, we compiled a list of pre-start and enroute tasks performed by engineers. Special emphasis was given to train handling tasks as it was felt that differences between engineers would most likely occur in this aspect of the job.

Based on this information, a road foreman observation guide was prepared. The purpose of the guide was to provide examples of good and poor train handling in over-the-road situations such as: starting train, accelerating; maintaining speed; slowing; stopping; and switching. We attempted to include examples from the important, and commonly occurring situations, that only an engineer would likely meet.

The observation rating form included ratings of specific tasks. Each task was rated whether or not the engineer used acceptable procedures. Sets of tasks, for one type of train operations, e.g., prestart, were also evaluated on a five-point rating scale.

The draft observation guide was initially reviewed by several Union Pacific road foreman and trainers. Representatives from each of the other participating railroad also reviewed and made recommendations for revisions to draft versions of the rating form. Based on all of their comments, revisions and additions were made to the rating form. The revised rating form is in Appendix B.

VALIDATION OF THE TEST BATTERY

The objective of the validation was to secure evidence that the test battery is useful for the selection of locomotive engineers. The evidence which we think is most appropriate to use in support of this test battery is the results of a criterion-related validation study. Evidence for criterion-related validity typically consists of a demonstration of a useful relationship between the scores on the test battery and ratings on a measure of job performance, or a criterion. The design chosen for obtaining criterion-related evidence is a concurrent one. In a concurrent study, one obtains test and criterion data simultaneously, with a sample of job incumbents.

Selection of the Sample of Engineers

In designing the validation, consideration was given to selecting the validation sample, collecting test data, and obtaining the criterion data. The plan for the validation sample was based on technical and practical considerations. One goal was to select a sample which was both large and diverse enough so that there was a high probability of detecting the true validity of the test battery. A study undertaken with a sample which is small or homogeneous is not likely to produce positive results, regardless of the intrinsic quality of the test.

A second consideration in designing the sample was the cost for the railroads and the contractor. Because of engineer contracts and schedules, it was costly to have engineers participate in the validation. The engineers often had to be paid an entire day's wages for participation in a three hour data collection. The cost of participation limited the number of engineers that any railroad could provide for the study. The railroads also determined that only engineers who volunteered should participate. Because of the costs of administration, URC decided that test battery data could only be collected at a limited number of locations.

These considerations led to a plan for sample selection in which railroads were asked to identify the number of engineers they could provide in a limited number of testing locations. They were also asked to select locations with different kinds of routes or other characteristics which might affect engineer performance. After selecting feasible locations, the railroads asked for volunteers, at the locations, who would be willing to participate in the validation. Although URC requested that the railroads try to include a representative sample of engineers, in terms of race, sex, and competence, the participation of only volunteers did not allow us to ensure a diverse sample of engineers.

Collection of Test Data

URC prepared a plan so that the test data could be collected in a standardized way. In order to increase uniformity, URC project staff undertook all test administration and processing. Instructions for each test administrator were prepared so that instructions, timing, and explanations were consistent across the sessions.

Company coordinators were asked to identify specific times and locations for test administration. They were also asked to inform the engineers about the study. The participants, however, did not receive any instructions or information regarding the specific tests prior to the testing session.

At the onset of each testing session, the project staff member who served as test administrator provided a brief overview about the project, summarizing the earlier phases of the study, and describing the purpose of the test validation. The participants were assured that the test and job performance data would remain confidential, and be used only for the purposes of the validation study.

The participants then filled out background information forms. The tests were administered in the following order: memory, reading, perception, listening, logic, and dichotic listening. The test administrator read the instructions for each test and the participants were given only those test materials. All of the tests were timed. However, except for the perception and memory tests, all the participants were given all the time they needed to complete each test. Sometimes this resulted in several additional minutes for the reading comprehension test. The test administrator was asked to record any instances when additional time was required.

Collection of Criterion Data

The railroad representative who arranged for the validation was given a set of engineer observation forms, after the testing, and asked to distribute them to the road foremen who supervised each of the engineer participants. The observation form contained instructions for its completion. The railroad representatives were asked to contact URC project staff if they, or the road foremen, had any questions regarding the observation form or the validation study.

The railroad representatives were asked to have the road foremen use the observation form during a normal ride with the engineers. The instructions on the observation form request that the road foreman read the form before the trip and fill out the form immediately upon completion of the ride. The forms were to be returned to the railroad representative who was asked to send the entire set to URC.

Addition of New Railroads

The five participating railroads were having difficulty meeting the sample size requirements of the validation study, due to the costs of providing engineers for testing and the difficulties in coordinating the collection of the criterion data. The data collection was taking place more slowly and with fewer engineers than had been anticipated. The delay in the receipt of the observation rating forms was particularly problematic.

We decided that the inclusion of additional railroads in the validation study would reduce the burden on the companies currently participating in the study and increase the size of the sample of engineers. We contacted the Railroad Personnel Association to help us identify additional railroads willing to participate in the study. The President of the Railroad Personnel Association, described the project during the June 1989 meeting of the group. As a result, four additional railroads, Canadian Pacific, Chicago Northwestern, CSX, and Norfolk Southern, agreed to participate in the validation phase of the study.

These railroads were particularly suited to participate in this phase of the study because they were all Class 1 railroads and did not require any additional task analysis to determine their similarity to the UP engineer's job.

Later in the project, a group of smaller railroads: Bessemer and Lake Erie; Elgin, Joliet, and Eastern; Union; and Duluth, Missabe, and Iron Range; became interested in the project and volunteered to participate. Because these were not Class 1 railroads, we were concerned that the job of locomotive engineer might be different from the engineer's job in the other railroads. In order to determine if the job was similar, the railroads were asked to rate the task list originally prepared for Union Pacific. The railroads were sent the same form that was used for the inclusion of the railroads during the job analysis phase of the project.

Exhibit 19 displays the background information for the sample of new railroads. The number of raters was five for all four of the new companies. The mean age of engineers ranged from 41.2 years to 51.6 years. The sample was completely composed of white men. The mean years of experience was as little as 7.6 years and as much as 19.8 years.

The new railroads were evaluated using the same approach as was used for the railroads that entered the study in the job analysis phase. For the tasks to be considered important to the engineer's job, four of the five raters (80%) must have given that task a "yes" rating. Similarly, 80% of the original task list had to have been considered an important part of the engineer's job for that railroad to be included in the study. Of the four railroads that provided ratings, all four met the 80% requirement. Bessemer and Lake Erie, Union, and Elgin, Joliet and Eastern, railroads all had an 81% overlap of important tasks with the UP task list. The engineers from Duluth, Missabe, and Iron Range rated 86% of the UP tasks as being important to their job.

The raters had an opportunity to revise the task list and add tasks that were missing. Engineers from the Elgin, Joliet and Eastern Railroad changed two tasks. Task 10 was revised to "Read and comply with track warrants and bulletins, signals, and railroad rules and regulations, while operating locomotive." and Task 13 was revised to "Relay wayside of cab signals to crew members using the radio." Working conditions were mentioned on the rating forms by engineers from Bessemer and Lake Erie. These included long hours and extreme weather conditions. Engineers from Duluth, Missabe, and Iron Range added tasks: such as "Instruct the crew on safety and customer accommodations;" "Handle tonnage trains on steep grades;" "Handle trains and instruct crew members in the operation and procedure for a straight air retainer system;" and, "Handle tonnage trains in sub-zero weather with composition brake shoes." Finally, engineers from Elgin, Joliet, and Eastern added "Control the speed of a train with the dynamic brake," and "Prevent overcharging of the brake equipment."

As a result of the inclusion of the eight new railroads, we were able to complete the validation study. The railroads participating in the validation data collection, the dates of testing, and the number of engineers involves are listed in Exhibit 20.

Exhibit 19

Task Review
Description of the Sample of Raters from Railroads

	Bessemer Lake Erie	Duluth Missabe Iron Range	Elgin, Joliet & Eastern	Union
Bacground				
Number of Raters	5	5	5	5
Mean Age	47.6 yrs.	45.4yrs.	41.2 yrs.	51.6 yrs.
Sex Mix	5 men	5 men	5 men	5 men
Ethnic Mix	5 Whites	5 Whites	5 Whites	5 Whites
Mean Years Experience as Engineer	13.8 yrs.	7.6 yrs.	12.2 yrs.	19.8 yrs.

Exhibit 20

FRA Locomotive Engineer Validation Study

Railroad	Testing Site	Date	Engineers Tested
Amtrak	Washington, DC	June 27, 1989	7
	New York, NY	June 28, 1989	6
	Chicago, IL	March 30, 1990	3
Burlington Northern	Overland Park, KS	August 2, 1989	10
Canadian Pacific	Agincourt, ON	September 6-7, 1989	13
	Toronto, ON	March 26, 1990	14
Chicago Northwestern	Naperville, IL	July 25, 1989	8
Conrail	Greentree, PA	May 22, 1989	6
	Indianapolis, IN	May 26, 1989	3
	Harrisburg, PA	June 2, 1989	4
	Dearborn, MI	June 6, 1989	3
	Philadelphia, PA	June 8, 1989	6
	Selkirk, PA	June 13, 1989	6
CSX	Evansville, IN	August 15-16, 1989	17
Duluth, Missabe & Iron Range	Duluth, MN	July 16-17, 1990	10
Elgin, Joliet & Eastern	Joliet, IL	July 18-19, 1990	12
Norfolk Southern	Atlanta, GA	August 2, 1989	10
	Chattanooga, TN	April 2-3, 1990	16
Union	Pittsburgh, PA	July 9-10, 1990	10
Union Pacific	Salt Lake City, UT	March 23, 1990	16

Data Analysis

Description of the Sample

One hundred and eighty engineers from eleven different railroads participated in the study. The largest number of engineers from any single railroad was twenty-seven, and the smallest number was five. Complete data sets, including both test and criterion data, were available for 143 engineers. The latter sample was the sample used for the validation.

The sample consisted of 141 males, 1 female, and 1 engineer who did not indicate gender. One hundred and twenty-six white engineers, thirteen African American, and three Hispanic engineers participated in the study; one engineer did not indicate ethnic group. The engineers ranged from 23 to 66 years old, with a mean age of 40.5 years. Experience as an engineer ranged from less than 1 year to 39 years, with an average tenure of 12.0 years.

Analysis of the Predictor Tests: Internal Consistency

Initially, the items in each of the tests were correlated with the total score on the test to identify items which were not internally consistent. Two items on the reading comprehension test (items 25 and 36) and one item on the logical reasoning test (item 18) were negatively correlated with the total test score, and therefore were eliminated. After these items were eliminated, the tests were rescored and internal consistency analyses were performed.

The internal consistency reliability of the tests was assessed using coefficient alpha. Coefficient alpha is a numerical index of the extent to which the items on a test measure a single trait. The internal consistency results presented in Exhibit 21 indicate that the five tests show adequate levels of reliability. It should be noted that the reliability of the perception test could not be estimated using coefficient alpha, because the perception test is a speeded test. A speeded test is one in which no examinee has time to attempt all of the items. Internal consistency indices of reliability are spuriously high for speeded tests (Anastasi, 1988).

Analysis of the Predictor Tests: Inter-test Correlations

The correlations of the six tests in the predictor battery are displayed in Exhibit 22. Not surprisingly, all of the cognitive ability tests (memory, reading comprehension, listening, and logical reasoning) were significantly correlated. The perception and dichotic listening tests were less highly correlated with this set of tests and were not correlated with each other.

In order to refine our understanding of the relationships among the tests, a nonorthogonal principal components analysis with varimax rotation was used. (Appendix C presents a brief explanation of principal components analysis.) This procedure is intended to identify the underlying factors that account for the correlations among the tests. The results indicate two factors: (1) a cognitive ability factor, underlying performance on the memory, reading, listening, and logic tests; and (2) a perception factor. The dichotic listening test loaded on the cognitive ability factor, but to a lesser extent than did the memory, reading, listening, and logic tests. The memory test loaded on both factors. This rather unexpected finding may be explained by considering the content of the memory test. The test requires the applicant to remember codes (which are figures) and the operating principle associated with the code. Exhibit 23 presents the factor loadings. Together, the two factors account for 62.2 percent of the variance.

Analysis of the Criterion Measure

The criterion measure we used was a supervisory rating of train handling ability. The engineer was evaluated on nine dimensions of job performance: prestart, rules compliance, operation of equipment, starting, acceleration, controlling speed, negotiating a cresting grade, stopping, and switching. For each dimension, there were specific behaviors which were rated according to whether or not the engineer followed acceptable procedures. These ratings were called "specific behavior" ratings.

In addition, for each dimension, there was a five point rating scale for overall performance, ranging from unsatisfactory to outstanding. A sixth rating option was available if it was not possible to observe behaviors relevant to the dimension. These ratings were labeled "dimension ratings." In addition to the evaluations of specific behaviors and to dimension ratings on the dimensions, the supervisor was asked to rate the difficulty of the trip, considering factors such as the territory, time of day, and weather. This rating was made on a five point scale from easy to difficult.

For each engineer, two sets of ratings scores were used to evaluate performance. These scores are presented in Exhibit 24. First, were the scores of the specific behaviors. For each of the nine dimensions, these ratings were summed, with a "yes' rating (indicating adequate performance) scored +1, a "no" rating (indicating unacceptable performance) scored -1, and a rating of "not observed" or "not applicable" scored 0. The ratings for all of the behaviors within each dimension were then summed. In order to interpret these scores, one needs to consider the number of behavior ratings for each dimension. The number of behaviors for each dimension are: Prestart (6); Rules Compliance (15); Operation of Equipment (10); Starting the Train (10); Accelerating (4); Controlling Speed (17); Negotiating a Cresting Grade (6); Stopping a Train (12); and for Switching (6).

These summed ratings do not provide information about the relative use of the three rating options. In order to better understand the ratings, we analyzed the frequency with which the" yes," "no," and "not observed/not applicable" response options were observed. The frequency distributions are presented in Exhibit 25.

Of the 85 specific behavior ratings, the mean number of "yes" ratings was 49.9, the mean number of "no" ratings was 1.4, and the mean number of "not observed/not applicable" ratings was 33.7. It appears that on the specific behaviors, engineers were rated as either performing acceptably, or a rating of

Exhibit 21

Test Means Internal Consistency Estimates

Test	Test Mean	S.D.
Dichotic Listening	168.88	21.74
Logical Reasoning	19.51	6.21
Memory	21.66	7.19
Listening	27.19	5.08
Reading Comprehension	31.21	5.38
Perception	13.92	3.66

Exhibit 22

Inter-test Correlations

	Memory	Reading	Perception	Listening	Logical Reasoning	Dichotic Listening
Memory	1.00					
Reading	.44	1.00				
Perception	.23	.05	1.00			
Listening	.31	.50	-.01	1.00		
Logical Reasoning	.49	.70	.14	.57	1.00	
Dichotic Reasoning	.20	.11	.11	.09	.18	1.00

Exhibit 23

Principal Components Analysis on the Predictor Tests

	Rotated Factor Matrix	
	Factor 1	Factor 2
Memory	.54	.60
Reading	.88	.02
Perception	-.02	.94
Listening	.72	.06
Logical Reasoning	.86	.21
Dichotic Listening	.33	.06

Exhibit 24

Descriptive Statistics on the Criterion Measure

		Specific Behaviors		Rating	
	N	Mean	SD	Mean	SD
Prestart	137	5.31	1.08	3.15	0.44
Rules Compliance	138	9.68	1.92	3.18	0.44
Operation of Equipment	132	2.48	2.15	3.15	0.43
Starting	141	5.73	2.15	3.15	0.43
Acceleration	139	2.71	1.19	3.15	0.52
Controlling Speed	139	10.43	3.83	3.18	0.51
Negotiating Crest Grade	123	4.13	2.02	3.23	0.49
Stopping	131	8.46	3.37	3.18	0.53
Switching	99	3.84	2.79	3.28	0.47
Sum of All Dimensions		49.97	11.54	25.60	2.76

Notes:

N = number of subjects

Specific behaviors refer to the specific behaviors which are rated on the Engineers Rating Form in Appendix B.

The overall rating is the rating for the entire dimension or category of behaviors, e.g., Prestart.

Exhibit 25

Percentage of Engineers Receiving Each Rating

Dimension	Performed Acceptably	Performed Unacceptably	Not Performed	Missing
Prestart	89.16	1.75	8.68	0.40
Rules Compliance	68.33	1.76	28.28	1.63
Operation of Equipment	34.68	0.30	64.72	0.30
Starting the Train	59.11	1.76	38.99	0.18
Acceleration	70.63	2.18	26.60	0.60
Controlling Speed	65.22	2.82	31.85	0.11
Negotiating Crest Grade	70.73	0.98	28.28	0.00
Stopping Train	72.23	1.90	24.42	1.45
Switching	58.57	0.40	38.63	2.40

not observed/not applicable was given. Very few responses of inadequate performance were given. However, we should not conclude from these data that the raters evaluated the engineers as performing uniformly very well, since the mean ratings for each dimension were in the satisfactory range rather than the superior or outstanding range.

As can be seen in Exhibit 25, a number of engineers were not rated on the switching dimension. We decided to drop this dimension--both the specific behaviors and the dimension rating--from further analysis.

One consideration in determining a way to score the criterion was the extent to which we could combine the specific behavior ratings and the dimension ratings, within each dimension. We correlated the sum of the behaviors ratings, for each dimension, with the dimension ratings. The correlations are presented in Exhibit 26. These correlations are modest to moderate and do not suggest that the specific behavior ratings and dimension rating should be combined.

A second consideration was whether we should combine each of the ratings across dimensions. Internal consistency analyses of the dimension ratings and the specific behavior ratings were undertaken. Coefficient alpha was used as the indicator of internal consistency. The coefficient alpha for the behavior ratings was .89 and the coefficient alpha for the dimension was .71. Based on the extent of these ratings, we decided to use the sum of the overall performance ratings and the sum of the behavior ratings as the criteria for this study.

A small study was undertaken to evaluate the reliability of the criterion ratings. A test-retest strategy was used to evaluate reliability. This approach was used since the railroads indicated that it was neither feasible nor appropriate to have two supervisors rate an engineer's performance.

Twenty-five engineers at Union Pacific Railroad were evaluated twice using the criterion measure. The intervals between the data collection ranged from one week to over one month. The test-retest reliabilities of the sum of the performance ratings and the sums of the behaviors within a dimension are presented in Exhibit 27. The reliabilities are modest for the specific behavior ratings and moderate for the dimension ratings.

Analysis of the Validation Results

The first analysis of the validation data was the examination of the correlations between the tests in the predictor battery and criterion data (the sum of the dimension ratings and the sum of the specific behavior ratings). Exhibit 28 displays these correlations. None of the correlations are statistically significant. These correlations indicate that the tests in the selection battery are not effective predictors of the performance ratings. Subsequent multiple regression analyses support this conclusion. (Brief explanations of multiple regression and statistical significance are presented in Appendix C.) The multiple correlation between the six tests and the sum of the specific behavior ratings was .22 ($df=6, 115$; $F=.98$, $p=.44$). The multip8le correlation of the six tests and the sum of dimension ratings was also not significant (multiple $R=.21$, $df=6, 102$, $F=.75$, $p=.62$).

Exhibit 26

Correlations by Dimension of the Sum of Specific Behaviors with Overall Ratings

	Correlation	N
Prestart	.21*	137
Rules Compliance	.28**	138
Operation of Equipment	.14	132
Starting	.39**	141
Acceleration	.39**	139
Controlling Speed	.28**	139
Negotiating Cresting Grade	.13	123
Stopping	-.06	131
Switching	.14	99

Correlation Between the Sum of Specific Behaviors and the Sum of Overall Ratings (Excluding Switching)

r = .26**

Notes:

* $p < .05$
** $p < .01$
N = number of subject

Exhibit 27

Test-Retest Reliabilities for Criterion Scales
Sum of Ratings of Behaviors

	Test-Retest Reliability	N
Prestart	-.15	25
Rules Compliance	-.01	18
Operation of Equipment	.33	25
Starting	.13	24
Acceleration	.50*	25
Controlling Speed	.80*	25
Negotiating Cresting Grade	.57	25
Stopping	.54*	24
Sum Across Dimensions	.42	17

Ratings

	Test-Retest Reliability	N
Prestart	.47	24
Rules Compliance	.64*	24
Operation of Equipment	.50*	24
Starting	.46	24
Acceleration	.58*	25
Controlling Speed	.45	25
Negotiating Cresting Grade	.64	23
Stopping	.02	25
Sum of Dimension ratings	.62	25

Notes:

* $p < .05$
N = number of subject

Exhibit 28

Correlations Between Predictor Tests and Criterion Ratings

	Sum of Specific Behaviors	Sum of Overall Ratings
Memory	-.04	-.10
Reading	-.04	.11
Perception	-.03	-.04
Listening	.00	.11
Logical Reasoning	-.16	.05
Dichotic Listening	-.02	-.02

In an attempt to shed insight on why the tests were not predicting the criterion scores, we examined the test and criterion intercorrelations, undertaken separately for each of the participating railroads. Exhibits 29 and 30 display these correlations for the individual railroads. The correlations vary substantially in size across railroads, suggesting that aggregating the data across railroads might be attenuating the correlations.

Potential explanations of these low correlations could focus on the tests, the criterion scores, or both. Examination of the predictor tests reveals no substantial evidence that the tests failed to capture the cognitive characteristics they purport to measure. As depicted in earlier sections of this report, the tests displayed adequate internal consistencies and reasonable means and standard deviations. Both the pattern of test intercorrelations and the factor analysis results are interpretable. Furthermore, much previous research attests to the efficacy of cognitive ability tests for predicting performance across a wide variety of jobs and organizations (Hunter & Hunter, 1984). Hence, it seems reasonable that we could expect these tests to predict job performance for locomotive engineers.

Consideration of the criterion scores suggests several potential explanations for the low correlations. First, the moderate test-retest reliabilities suggest the presence of measurement error in the ratings. Variance due to differences among raters, railroads, and the nature of the run (e.g., terrain, time of day, difficulty, duration of trip, weather) introduces error into the criterion scores. Furthermore, each of the criterion scores demonstrated very few examples of poor performance.

For the specific behaviors, there were very few instances of "no" ratings, indicative of unacceptable performance. For the dimension ratings, raters displayed a strong tendency to label the engineers performance as satisfactory, rather than extremely good or extremely poor. The large number of moderate ratings could serve to lower the correlations between the predictor tests and criterion scores.

Exhibit 29

Correlations Between Predictor Tests and Sum of Behavior Ratings By Railroad

Railroad	N	Memory	Reading	Perception	Listening	Logical Reasoning	Dichotic Listening
1	10	.60	.62	-.08	.29	.71	-.22
2	8	-.45	-.16	-.32	.28	-.40	.50
3	13	-.03	-.20	.16	-.37	.52	.16
4	22	-.39	.12	-.24	.25	-.05	-.03
5	1						
6	9	.02	-.73	-.28	-.65	-.56	-.29
7	9	.32	-.05	.07	-.09	.04	.45
8	3	-.80	.87	.95	.98	-.91	-.99
9	10	.35	-.61	.44	-.18	-.21	-.05
10	7	-.39	-.21	-.04	.39	.10	-.01
11	2						

Notes:

N = number of subject

It was not possible to analyze the data from the railroads with only one and two subjects.

Exhibit 30

Correlations Between Predictor Tests and Sum of Overall Ratings by Railroad

Railroad	N	Memory	Reading	Perception	Listening	Logical Reasoning	Dichotic Listening
1	10	-.26	.23	.02	-.34	.30	.12
2	8	-.70	-.47	-.03	.02	-.44	.35
3	13	-.14	.25	-.26	-.21	-.10	.27
4	22	-.37	-.19	-.10	.23	-.25	-.34
5	1						
6	9	.48	-.01	.13	.05	-.13	.12
7	9	.52	.25	.04	.75	.69	.59
8	3	-.56	.65	1.00	.87	-.99	-.89
9	10	.09	-.23	-.17	-.37	-.40	-.40
10	7	-.48	-.55	.29	-.12	-.04	.31
11	2						

Notes:

N = number of subject

It was not possible to analyze the data from the railroads with only one and two subjects.

SECOND VALIDATION STUDY

The project staff thought that one major factor in the poor validation results was the criterion measure. We proposed that a second study using a different criterion measure would be a better means to determine the validity of the test battery.

The alternative criterion had to be more standardized and likely to produce more variance in scores. We concluded that use of a locomotive simulator built by Illinois Institute of Technology looked most promising. Burlington Northern Railroad uses this simulator for students attending first-time engineer training. Burlington Northern offered to assist us in undertaking a validation study using these students.

URC project staff administered the battery of tests to engineer trainees at Burlington Northern (BN) during their classroom training. Their training starts with 30 days of on-the-job training in which the trainee rides with an engineer trainer, observes the trainer operate the train, and receives instruction in train handling procedures, safety and mechanics, and sometimes, operates the train. The second phase of training is three weeks of classroom instruction. During this time, the trainees are given the test battery, using the same procedures as used for the other validation study.

At the end of the three weeks of classroom instruction, trainees are assigned blocks of time on the simulator, including an evaluation run. Then the trainees return to their home territory and spend approximately 10 weeks of time operating trains under the supervision of a qualified engineer.

Twenty-two weeks after the beginning of the training program, the trainees take the final written examinations and the simulator evaluations. These written examinations and the simulator evaluation served as the criterion measures for the study.

Criterion Procedures and Scoring

The criterion measure scores were provided by Burlington Northern Railroad Technical Training Center. The criterion measure included a combination of the final simulator exam and two final knowledge exams. Burlington Northern provided URC with these scores for the trainees who took the battery of selection tests.

Final Examinations

In the last two weeks of their engineer apprenticeship, the trainees take the final written examinations which evaluate the knowledge acquired through training. The examinations consist of two tests, each with 300 multiple choice questions. The exams are administered on two consecutive days. The first day the trainees take the General Code of Operating Rules test and on the second day they take the Air Brake and Train Handling Rules test. Burlington Northern provided URC with the total percentage correct on each of the tests, and then the total percentage correct on a composite of the two tests.

Simulator Examinations

The simulator examinations include three separate runs, each administered on one of three consecutive days. The simulator examinations are designed to measure the trainees' skill in train handling. The trainees are provided with a written profile for each of the runs in advance of the examination. The profiles include information such as the objective of the run, speed restrictions, route profile and characteristics and train configuration. To perform well on the simulator examination, the trainees must observe all speed limits, comply with signal indications and instruction, keep in-train forces within acceptable limits, blow whistle, make gradual throttle changes, etc.

The first run (Orin) contributes 40% toward the total simulation evaluation score. The purpose of this run is to test responses to Centralized Traffic Control (CTC) block signal indications while handling a 110-car loaded coal train. The trainees are randomly assigned to one of three variations of this run which takes one hour and 15 minutes.

The second run (Billings) contributes 40% toward the total simulator examination score. The purpose of this run is to test response to CTC block signal indications and track flagging situations while operating a 58-car intermodel train in 60 MPH territory. The trainees are randomly assigned to one of three variations of this run which takes one hour and 15 minutes.

The third run (Generic) contributes 20% toward the total simulator examination score. The purpose of this run is to evaluate the application of specific train handling rules and methods while operating a mixed freight in planned slowdown, stop, and acceleration situations. All trainees complete the same run which is scheduled for 60 minutes.

All three simulation runs are scored electronically based upon starting, stopping, speed control, timely whistle blows, etc. Each run has between 8 and 15 items that contribute to the total score for that run. The items are weighted in relative proportion to the importance of that item to the run as a whole and percentage of points earned is multiplied by the respective weight. The three weighted percentages are summed to create the composite score for the three runs. Simulator segments were prioritized by experienced Burlington Northern engineers to establish weights for the simulation.

Description of the Sample

A total of 141 engineers participated in the second validation study. The participants' ages ranged from 21 to 64, with an average age of 36.1 years. The sample included 9 females and 132 males. One hundred and seventeen participants were white, eleven were African American, eight were Hispanic, four Asian, and one person failed to identify his/her ethnic group. Years of experience ranged from less than 1 year to 42 years, with a mean of 11.2 years.

Description of the Test and Criterion Data

Exhibit 31 shows the mean test scores for the sample.

Of the 141 participants in Study 2, only 123 completed the two paper and pencil training tests. Scores on these tests were analyzed in terms of percentages of items correct. Scores on the General Rules test ranged from 77% to 99%, with a mean of 93.9% and a standard deviation of 4.25. Scores on the Airbrake test ranged from 63% to 99%, with a mean of 93.8% and standard deviation of 5.29. For each test, a grade of 90% or higher was the cutoff for passing the exam. Of the 123 individuals who completed these tests, 10 did not pass the General Rules Test and 8 did not pass the Air Brake test.

Simulator scores were available for 111 participants. The procedure for calculating simulator scores was slightly more complex. Two of the simulator runs--Orin and Billings--had three versions (A, B, and C) while the other one (Generic) had one version.

The first issue facing us was whether or not we could treat the three versions of the Orin run and Billings run as functionally equivalent. If so, we could collapse across the A, B, and C varieties of these runs. An analysis of variance on the total percentage score means of Orin A, B, and C revealed no statistically significant differences ($F=.59$, $df=2, 108$; $p=.55$). Similarly, an analysis of variance on the Billings run means also revealed no differences ($F=.80$, $df=2,108$, $p=.45$). (Appendix C contains a brief explanation of analysis of variance). Therefore, in all subsequent analyses, we treat the A, B, and C versions of Orin and Billings runs as equivalent.

Total percentage scores on the Orin simulation run ranged from 39 to 94, with a mean of 77.9 and a standard deviation of 11.23. The scores on the Billings run ranged from 48 to 90, with a mean of 76.9, and a smaller standard deviation of 7.63. The scores on the Generic run ranged from 21 to 84, with a mean of 64.2, and a standard deviation of 10.9.

Both the paper and pencil tests and the simulator runs exhibited reasonable descriptive statistics and represent fairly broad differences in performance from inadequate to excellent. Furthermore, the scores on the two criterion measures were significantly correlated as shown in Exhibit 32, indicating that the written training test were significantly correlated with performance on the simulators. (A brief description of the correlation coefficient is presented in Appendix C.)

Validation Results for Study 2

Exhibit 32 shows the correlations between each of the predictor tests and the various criterion measures for Study 2. None of the six tests is correlated highly with the simulator runs. Reading comprehension, logical reasoning, and dichotic listening, however, were significantly correlated with the two paper and pencil instruments administered at the end of training. In turn, the paper and pencil training tests were correlated with performance on the simulator. Based on these findings, we conclude that the selection tests predict performance in training (as indicated by their correlation with the Air Brake and General Rules tests) and that

Exhibit 31

Predictor Tests Means (Study 2)

Test	Test Mean	S.D.
Memory	22.12	7.11
Reading Comprehension	28.66	5.55
Perception	13.56	3.68
Listening	26.89	4.94
Logical Reasoning	17.96	6.12
Dichotic Listening	165.25	30.49

Exhibit 32

Correlations Among the Study 2 Criterion Scores

	Simulator Run 1	Simulator Run 2	Simulator Run 3	General Rules	Airbrake
Simulator Run 1	1.00				
Simulator Run 2	.27*	1.00			
Simulator Run 3	.36**	.12	1.00		
General Rules	.03	.41**	.27*	1.00	
Airbrake	.20	.37**	.34**	.68**	1.00

* $p < .01$

** $p < .001$

performance in training predicts performance on the simulator. Although the selection tests do not directly predict job performance, they predict training which is critical for performance as an engineer. In other words, the tests predict training performance and training performance predicts simulator performance. There is, however, no direct relationship between test scores and performance on the simulator.

The findings of Study 2 are consistent with the results of prior research indicating that basic cognitive abilities are critical for successful training. Hunter (1983) sheds insight on the links between ability tests, training measures, and job performance. Hunter's meta-analysis of 14 validation studies suggests that cognitive abilities help determine the extent to which an individual masters the knowledge and skills required for job performance. Thus, job knowledge mediates the relationship between cognitive abilities and job performance. On relatively routine jobs with fairly constant job requirements, such as the locomotive engineer, ability tests scores may not be directly related to job performance, but rather indirectly related via job knowledge.

RECOMMENDATIONS

What implications do these results have for the usefulness of the ability tests? First, three of the tests - reading, logical reasoning, and dichotic listening - predict scores on written training exams. This finding is consistent with extensive previous research demonstrating that cognitive abilities are important predictors of training success. Because training of locomotive engineers is necessary for safe and efficient train handling, is legally mandated, is time consuming and expensive, it would benefit the railroads to predetermine who is most likely to pass training. The three cognitive ability tests do indeed identify applicants who are most likely to successfully complete training. We, therefore, recommend using these three tests to select applicants for engineer training.

PLANS FOR ADDITIONAL VALIDATION WORK

The validation of the cognitive ability tests developed in this project for the locomotive engineering job revealed that these tests are positively correlated with training performance and that training performance is positively correlated with performance on a simulator. This finding was based on a predictive validation in one railroad, Burlington Northern. Burlington Northern engineers took the tests while they were in initial classroom training, and at the end of their training, were evaluated on two training knowledge tests and a three simulator runs.

Although this study is important in supporting the validity of the test battery, it was undertaken in only one railroad. The design of this study limits the generalizability of its findings. Although we think that the results support the use of the tests for prediction purposes throughout the industry, we recommend that other railroads undertake validation studies.

We recommend that either individual railroads undertake a validation study or that a group of railroads work together in a consortium to undertake a validation study. We also recommend that the railroads undertake a predictive validation study, since we think this study is most appropriate considering the results of this project.

Predictive Validation Design

A predictive validation design uses applicants who are given the tests prior to being hired and then are followed through training and on the job to determine their job performance over a specified period of time. The predictive validation design proposed for these tests is fairly straightforward. Implementation of the design requires the following five steps:

Step 1. Give job applicants the three tests.

Step 2. Hire the required number of applicants as locomotive engineer trainees based on the recommended cut-off score.

Step 3. Obtain training tests scores for the trainees. If simulator performance data are available, include such data.

Step 4. Obtain performance data for the trainees six months after they are on the job.

Step 5. Analyze the data.

Although each railroad will establish its own procedure for determining which applicants will be given these tests, test scores for all applicants who take the battery should be kept on file. In addition to test scores, some basic demographic data for each applicant should be kept. This should include name, age, ethnic background, sex, and prior company experience. This information will provide the data for determining restriction in range of test scores as well as for EEOC-type analyses.

Applicants who meet the minimum cut-off score for the test battery will be selected as trainees, when there is a railroad need for such hires. All individuals hired using these tests are automatically used for the validation study. Training performance data on the trainees must be collected. The railroad can use its own training tests. We also recommend that the railroads use a standard written engineer training test developed for this project by Ralph and Lyn Haber, which is presented in Appendix D. The test is a 95 item multiple choice test pertaining to train handling procedures of over-the-road through freight trains. The test was developed in consultation with engineers and has been reviewed by both railroads and the FRA. If the railroads use a simulator to evaluate training performance, we strongly recommend that all the trainees take one or two standard runs and that scores on these runs be retained.

After completing training, the new engineers should be followed for six months and evaluated on their job performance. We recommend the use of a standard measure of job performance, like the one used in this study (see Appendix B). When data for 100 engineers have been collected, the data can be analyzed. Since few railroads will hire that many engineers in a reasonable period of time, we recommend that several railroads work together in a consortium study. By pooling the data across several companies, such a validation can be conducted and data analyzed much earlier than if a company had to wait until enough individuals were hired to result in reliable analyses. We should note that if a consortium study is planned, it is important to collect training and job performance data using the same forms in all the participating companies.

Data Collection

The predictive validation strategy described above requires that four types of data be collected. These are demographic data, test score data, training performance data, and job performance data. Basic demographic or background data should be collected from each applicant who takes the tests. These data should be kept on file regardless of whether the applicant is hired.

In addition to the basic demographic data, test scores on all three of the tests should be recorded for each applicant who takes the test. Again, these scores should be retained whether or not the applicant is hired. It is important to note that all data collected in a predictive validation study are extremely sensitive and should be kept strictly confidential by those conducting the validation. The applicant's test scores must not be shared with the applicant or the applicant's trainer or supervisor. Unless these data are kept confidential, the results of the validation may be biased. For example, if a trainer or supervisor finds out that an applicant achieved a very high score, then this knowledge could affect treatment of this person and ultimately may affect the person's training and job performance.

Careful attention should also be given to the standardized collection of criterion measures. For each person hired into an engineer trainee position, the following data should be collected:

1. Training performance data. Test scores used during and at the completion of training should be retained. We recommend use of the test prepared for this project and available in Appendix D.

2. Simulator performance data (if available)

3. Supervisor performance ratings. The person selected to evaluate the new engineer should be in a position to have observed the engineers performance, preferably several times. This rater should be instructed not to discuss the rating with the engineer or anyone else in the railroad. The rater should be assured that the rating will not influence the engineer's job and will not be placed in the engineer's personnel file. The ratings are for the exclusive use of validating the tests. If the railroad does not use a standardized rating form, the supervisor rating form prepared for this project can be used.

Data Analysis

Several sets of analyses should be conducted as part of this validation study:

1. Correlations among the three selection tests.

2. Correlations between scores on the selection tests and scores on the training tests.

3. Correlations between scores on the selection tests and scores on the supervisory ratings.

4. Correlations between scores on the training tests and scores on the supervisory ratings.

5. (If feasible) Correlations between scores on the selection tests and scores on the simulator runs. Correlation between scores on the simulator runs and scores on the supervisory ratings.

REFERENCES

Anastasi, A. (1988). *Psychological testing*. (6th ed.) New York: McMillan.

Arnold, J.D., Raushenberger, J.H., Soubel, W.G., & Guion, R. M. (1982). The validation and utility of a strength test for selecting steelworkers. *Journal of Applied Psychology, 67,* 588-604.

Ekstrom, R.B., French, J.W., & Harmon, H.H. (1976). *Manual for kit of factor - referenced cognitive tests.* Princeton, NJ: Educational Testing Service.

Fleishman, E.A., & Quaintance, M.K. (1984). *Taxonomies of human performance: the description of human tasks.* New York, NY: Academic Press.

Flesh, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32,* 221-233.

Gopher, D. (1982). A selective attention test as a predictor of success in flight training. *Human Factors, 24,* 173-183.

Gopher, D., & Kahneman, D. (1971). Individual differences in attention and the prediction of flight criteria. *Perceptual and Motor Skills, 33,* 1335-1342.

Hale, A., & Jacobs, H.H. (1975). *Proposed qualification requirements for selected railroad jobs.* Darien, CT: Dunlap and Associates.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictions of job performance. *Psychological Bulletin, 96,* 72-98.

Kahneman, D., Ben-Ishai, R., & Lotan, M. (1973). Relation of a test of attention to road accidents. *Journal of Applied Psychology, 58,* 113-115.

Louisville and Nashville Railroad (1981). *Railroad engineer task and skill analysis.* Louisville, KY: Louisville and Nashville Railroad.

McCormick, E. J., Jeanneret, P.R., & Mecham, R.C. (1972). A study of job characteristics and job dimensions as learned on the position analysis questionnaire. *Journal of Applied Psychology, 56,* 347-367.

McDonnell Douglas (1972). *Railroad engineman task and skill study.* Saint Charles, MO: McDonnell Douglas.

Myers, D.C., Harding, F.D., Hunter, R.R., & Fleishman, E.A. (1985). *Development of selection requirement and assessment procedures for the position of locomotive engineer.* Bethesda, MD: Advanced Research Resources Organization.

Railroad Personnel Association (1981). *The railroad industry job analysis project: a final report.* Washington, DC : Association of American Railroads.

Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Review, 86*, 420-428.

Appendix A

Job Requirements List

Understanding Printed/Written Material

1. Reading simple words, such as position signs on machine equipment (e.g., "On/Off", or "Start/Stop").

2. Reading simple sentences, such as posted signs or directions (e.g., "Keep boxes out of aisles").

3. Reading complex sentences, such as written material on work tickets or printed material on containers (e.g., "This material may explode if it gets wet").

4. Reading paragraphs which describe a thing or event or present multiple instruction in sequence, such as instructions in operating.

5. Reading computer print-outs, computer screens, or other material that is primarily numerical in nature.

6. Memorizing and recalling specific information learned from printed materials.

Performing Calculations

7. Adding and/or subtracting whole numbers.

8. Multiplying and/or dividing whole numbers.

9. Adding, subtracting, multiplying, and/or dividing fractions.

10. Adding, subtracting, multiplying, and/or dividing involving decimals and/or percentages.

11. Using a simple formula to solve for an unknown.

Understanding Oral Communications

12. Coordinating work with co-workers through conversation/discussion where effectiveness depends on understanding others.

13. Understanding oral instructions or work procedures information provided by supervisors or others.

14. Receiving on-the-job training provided by supervisor or others.

15. Participating in group meetings or training sessions where effectiveness depends on understanding others.

Making Oneself Understood Orally

16. Coordinating work with co-workers through conversation/discussion where effectiveness depends upon being understood.

17. Providing routine oral status or progress reports to supervisor or others, in person, by phone, or by radio.

18. "Breaking in" a new employee or otherwise instructing others.

19. Making informal reports to small groups.

Making Oneself Understood in Writing

20. Entering simple information or data on forms such as: recording temperature, pressure, thickness, quality, or number or errors.

21. Copying information (from print-outs, reports, etc.) into hand-written form.

22. Preparing simple records, such as: work reports, logs, or information for next shift.

23. Preparing written reports, such as: equipment malfunctions, performance results, or accident data.

24. Entering information by computer terminal, typewriter, or other keyboard or data entry device.

Understanding Graphic Information

25. Reading simple blueprints, sketches, diagrams, or shop drawings.

26. Reading numerical information in graphic form.

Exercising Mechanical Insight

27. Understanding mechanical relationships in practical situations, such as: understanding leverage, pulleys, or the direction gear arrangements turn.

28. Understanding the relationship of physical objects to one another in order to visualize a number of such objects acting together.

29. Visualizing objects in three dimensions.

30. Making visual comparisons between objects or pictures or diagrams.

Making Estimates

31. Estimating weight or objects.

32. Estimating size of large objects or areas relative to other objects or areas, such as: when parking a car or moving a crate between machines.

33. Estimating the speed or distance of moving objects or parts.

Making Choices and/or Solving Problems

34. Making choices/decisions in which the risks or consequence are slight, such as: sorting materials or parts.

35. Making choices/decisions affecting the security or well-being of others and/or which involve serious risk or consequences.

36. Solving problems involving limited options by applying common sense understandings, such as: selecting the correct tool for a job.

37. Solving problems involving a few relatively concrete options or variables by applying principles or methodologies, such as: troubleshooting malfunctions or breakdowns in familiar equipment.

Making Visual or Auditory Discriminations

38. Discriminating visual detail at distances within arm's reach.

39. Discriminating fine visual detail at distances within arm's reach.

40. Recognizing colors, such as: light signals, containers, or electrical parts.

41. Judging distance from observer to objects and/or between objects.

42. Recognizing changes in sounds.

43. Recognizing audible signals, such as: bells, whistles, or sirens.

44. Recognizing objects or signals under conditions of limited visibility, such as: seeing signals in fog or recognizing a sound in the presence of other noises.

45. Memorizing and recalling visual information such as maps or scenes.

46. Maintaining attention to a task over long periods of time.

Using Hands in Work Activity

47. Reaching - extending hand(s) and arm(s) in any direction.

48. Seizing, holding, grasping, turning, or otherwise working with hand(s) when fingers are not involved.

Making Gross Body Requirements

49. Stooping - bending body downward and forward by bending spine at waist.

Climbing or Balancing

50. Ascending and/or descending ladders, stairs, scaffolding, or poles, using feet and arms or hands.

51. Maintaining body equilibrium to prevent falling when walking, standing, crouching, or running on irregular, slipper, or erratically moving surface.

Appendix B

Engineer Rating Form

TRAIN HANDLING AND OPERATING PROCEDURES OBSERVATION FORM

Instructions

Use this form to record how the engineer you are riding with performed during the trip. The form is built around several situations in which an engineer could be expected to demonstrate train handling proficiency, for example, starting or stopping the train. For each of the situations that the engineer encountered during the trip, record your observation of the engineer's performance by placing check marks in front of all statements that reflect what the engineer did or what happened to the train as a result of the engineer's actions.

After making these ratings, make an evaluation of the engineer's performance in that situation. Use the scale at the bottom of each page to indicate your appraisal of the performance that you observed during the trip.

Please review this form before the trip in order to familiarize yourself with the ratings you will be making. Taking notes during the trip may be helpful. Fill out the rating form immediately after the trip to ensure the ratings are accurate.

Engineer: ______________________________

Rater: ______________________________

Date of Trip: ______________________________

Duration of Observation: ______________________________

Train Symbol: ______________________________

Train Consist (No. of locomotives,
loads and empties): ______________________________

Special Train Makeup (if appropriate): ______________________________

Territory: ______________________________

Rate the diffuculty of the trip (considering the train consist, territory, weather, time of the day, and other factors.)

__________ __________ __________ __________ __________

easy average difficult

Prestart

		YES	NO	Not observed or Not applicable
1.	Reports to work on time and is fit for service.	___	___	___
2.	Wears proper clothing and footwear.	___	___	___
3.	Has and reviews required documents (time table, train bulletin, track warrants, general orders, etc.)	___	___	___
4.	Performs locomotive inspections.	___	___	___
5.	Checks for safety equipment.	___	___	___
6.	Performs radio check.	___	___	___

Did the engineer follow acceptable procedures? (YES / NO / Not observed or Not applicable)

Based on observations during this trip, my evaluation of the engineer's performance for Prestart is checked below:

___	___	___	___	___	___
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Rules Compliance

Did the engineer follow acceptable procedures?

		YES	NO	Not observed or Not applicable
1.	Uses bell, whistle, headlight.	___	___	___
2.	Verifies speedometer accuracy.	___	___	___
3.	Acknowledges and complies with fixed signals.	___	___	___
4.	Speed compliance.	___	___	___
5.	Reports signal malfunctions to dispatcher.	___	___	___
6.	Reports defective track conditions.	___	___	___
7.	Inspects train while moving.	___	___	___
8.	Observes passing trains and reports defects.	___	___	___
9.	Fills out necessary paperwork.	___	___	___
10.	Communicates instructions and information with others.	___	___	___
11.	Uses proper radio procedure.	___	___	___
12.	Performs air brake tests, as required.	___	___	___
13.	Protects the train, as required (e.g., if train derails, ensures adjacent track is protected.)	___	___	___
14.	Observes other crew members for proper job performance.	___	___	___
15.	Other rules, as applicable (specify).	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Rules Compliance is checked below:

___	___	___	___	___	___
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Operation of Equipment

		Did the engineer follow acceptable procedures? YES	NO	Not observed or Not applicable
1.	When alarm sounds, takes action to identify problem.	___	___	___
2.	Isolates a malfunctioning unit, as required.	___	___	___
3.	Resets devices, as required.	___	___	___
4.	Restarts the engine, as required.	___	___	___
5.	Notifies proper authorities of equipment malfunctions.	___	___	___
6.	Writes proper work reports.	___	___	___
7.	Properly conditions unit for lead or trail including connecting/disconnecting hoses and cables, and positions other controls when picking up or setting out units.	___	___	___
8.	Monitors gauges and devices (including end of train device) and takes corrective action when required.	___	___	___
9.	Drains unit during cold weather operations.	___	___	___
10.	Recognizes if locomotive is not performing properly.	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Operation of Equipment is checked below:

___	___	___	___	___	___
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Starting the Train

	YES	NO	Not observed or Not applicable
Did the engineer follow acceptable procedures?			
<u>All Conditions</u>			
1. Properly charges brake system before moving.	___	___	___
2. Allows enough time for brakes to fully release before starting the train.	___	___	___
3. Uses independent brake valve to control locomotive acceleration, when required .	___	___	___
4. Reacts to wheel slippage.	___	___	___
5. Applies lead truck sanding on lead locomotive if required.	___	___	___
6. Advances throttle one notch at a time while observing load current meter in such a manner to provide smooth operation under proper slack control.	___	___	___
<u>Ascending Grade</u>			
7. Advances throttle before releasing automatic and independent brakes.	___	___	___
8. On heavier grade, sets automatic brake and backs locomotive into train to bunch it. Release brakes and starts forward movement so rear cars do not roll back.	___	___	___
<u>Descending Grade</u>			
9. After releasing automatic brake, gradually releases independent brake until entire train is moving.	___	___	___
10. With dynamic brake applied, gradually releases independent brake to keep slack bunched until dynamic brake becomes effective.	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Starting the Train is checked below:

___	___	___	___	___	___
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Accelerating

Did the engineer follow acceptable procedures?

		YES	NO	Not observed or Not applicable
1.	Advances throttle one notch at a time as load current meter stops increasing or begins dropping.	___	___	___
2.	Reduces throttle at indication of wheel slippage.	___	___	___
3.	Gradually reduces dynamic brake while observing load meter.	___	___	___
4.	Does not accelerate until brake is released throughout the train.	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Accelerating is checked below:

______	______	______	______	______	______
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Controlling Speed

		Did the engineer follow acceptable procedures?		
		YES	NO	Not observed or Not applicable
All Conditions				
1.	Varies throttle movements slowly, only one notch at a time.	___	___	___
2.	Adjusts dynamic brakes slowly, while observing load meter.	___	___	___
3.	Keeps speed and train forces fairly constant with dynamic brake adjustments.	___	___	___
4.	Uses automatic brake to supplement dynamic brake, when required.	___	___	___
5.	Maintains speeds required in curves, turnouts and restricted zones, without building up excessive train forces.	___	___	___
6.	Uses retainers as needed.	___	___	___
7.	Where possible, adjusts throttle for maintaining speed.	___	___	___
8.	Uses dynamic brake rather than automatic brake, where practical.	___	___	___
9.	Uses cycle braking to control speed.	___	___	___
10.	Makes initial automatic brake service reduction plus additional reductions, as required (including blended braking).	___	___	___
Slack Bunched				
11.	Uses maximum permissible dynamic brakes before applying automatic brake (or blended brakes).	___	___	___
12.	Releases automatic brake in order to maintain desired speed.	___	___	___

	Did the engineer follow acceptable procedures?		
	YES	NO	Not observed or Not applicable
13. Keeps train bunched using dynamic or independent brake.	___	___	___
<u>Slack Stretched</u>			
14. Gradually reduces throttle as automatic brake becomes effective, while observing load meter.	___	___	___
15. Uses throttle to keep train stretched.	___	___	___
16. Releases automatic brake before reaching desired speed.	___	___	___
17. Does not advance throttle until brakes are released.	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Maintaining Speed is checked below:

___	___	___	___	___	___
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Negotiating Cresting Grade

		Did the engineer follow acceptable procedures? YES	NO	Not observed or Not applicable
1.	Has train moving at an appropriate speed at crest.	___	___	___
2.	Reduces throttle as locomotive crests grade in order to maintain speed and reduce coupler forces in train at crest.	___	___	___
3.	Controls speed of train by further reductions of throttle.	___	___	___
4.	Uses dynamic brake to control speed of train in descent. Transition to dynamic brake is made smoothly.	___	___	___
5.	Use automatic brake to supplement dynamic brake.	___	___	___
6.	Makes initial automatic brake service reduction plus additional reductions as required .	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Negotiating a Cresting Grade is checked below:

___	___	___	___	___	___
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Stopping Train

	Did the engineer follow acceptable procedures? YES	NO	Not observed or Not applicable
Slack Stretched			
1. Maintains sufficient throttle to keep train stretched.	___	___	___
2. Makes initial brake service reduction plus additional reductions as required .	___	___	___
3. Reduces throttle one notch at a time, while observing load meter.	___	___	___
4. Train decelerates smoothly and efficiently to stop at desired location.	___	___	___
5. Applies independent brakes when stopped.	___	___	___
6. Places throttle in idle at appropriate time.	___	___	___
7. Makes final automatic brake application and keeps brake applied while stopped.	___	___	___
Slack Bunched			
8. Uses dynamic brake as primary braking source.	___	___	___
9. When automatic brakes are needed to supplement dynamic brake, makes initial brake reduction plus additional reductions as required.	___	___	___
10. Train decelerates smoothly and efficiently to stop at desired location.	___	___	___
11. Makes final automatic brake application.	___	___	___
12. Applies independent brake as dynamic brake fades.	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Stopping Train is checked below:

___	___	___	___	___	___
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Switching

	YES	NO	Not observed or Not applicable
	Did the engineer follow acceptable procedures?		
1. Starts releasing independent brake and applying light power until slack is adjusted and all cars in block are moving.	___	___	___
2. Speed is appropriate for switching conditions.	___	___	___
3. Bunches and stretches slack smoothly.	___	___	___
4. Makes moves at consistent speed so crew can anticipate stopping and starting distances.	___	___	___
5. Care is used in applying the independent brake to avoid wheel slide.	___	___	___
6. Observes and responds to hand or radio signal.	___	___	___

Based on observations during this trip, my evaluation of the engineer's performance for Switching is checked below:

______	______	______	______	______	______
Not observed or not applicable	Unsatisfactory: Does not perform this part of the job at a competent level	Marginal: Performs this part of the job at a minimally competent level	Satisfactory: Competently performs this part of the job	Superior: Performs this part of the job at a more than competent level	Outstanding: Performance exceeds requirements for this part of the job

Other Comments

Appendix C

Explanation of Statistical Terms

Pearson Product - Moment Correlation Coefficient is a statistical technique to specify the degree of association between two variables. The correlation coefficient ranges from -1.00 through +1.00. The size and sign of the correlation indicates the relationship between the variables. The closer the relationship between the variables, the greater than size of the correlation coefficient (the closer in approximates +1.00 or -1.00). The sign (+ or -) of the correlation coefficient indicates whether the variables are positively related (a positive correlation) or negatively related (a negative correlation). If two variables are positively correlated, then higher scores on one variable are associated with higher scores on the second variable. If two variables are negatively correlated, then higher scores on one variable are associated with lower scores on the second variable.

Factor Analysis - is a statistical technique used to determine the structure of a set of variables. In using factor analysis, a researcher is concerned with determining whether there are underlying factors which account for the relationship among the variables. The result of factor analysis is a smaller set of variables, or factors. The correlation between the original variables and the factors is the loading. The pattern of loadways can be used to explain the nature of the factor. For example, if only reading tests have high loadings on a factor, then we could interpret this factor in terms of reading ability.

Principal components analysis is similar to factor analysis and that it is a technique to reduce the number of variables and make the relationship simpler.

Statistical Significance - research is almost always undertaken with a sample, rather than with a population of individuals. Having undertaken the research and analyzed the data, the researcher is left to ask how likely is it that these research results were a result of chance factors. Statistical significance considers the likelihood that the data from research could have happened by chance. In order to reject the explanation of chance, or what is called the null hypothesis, the data from the study is compared to the distribution of result that would arise by chance alone if the research results are very unlikely to have occurred by chance (1 out of 20 times is the usually used) then the researcher concludes that the results are not a chance phenomenon. The term "$p < .05$" is used to represent the finding that these data could have occurred with a probability of less than .05 or 1/20. This term indicates that the results are statistically significant.

Analysis of Variance - is a statistical procedure to determine the statistical significance of the differences between mean scores in three or more groups. For example, this procedure is used to determine if the mean math achievement scores differs in three classes.

Appendix D

Test of Train Handling Procedures

A Test of Train Handling Procedures for Freight Locomotive Engineers
Test Prepared and Copyrighted by Lyn R. Haber and Ralph Norman Haber
Human Factors Consultants, Highland Park, Illinois 60035
August 10, 1991

Instructions to the Testee

The following are 100 multiple choice questions pertaining to train handling procedures of over-the-road through freight trains. Each question has four alternative answers. Select the best alternative and mark it on the attached answer form. The questions are written so that one of the alternatives is the best answer for the given train, territory, and track configuration specified. A few of the questions may have to be interpreted in terms of the operating stategies used by the particular railroad.

Train Configuration: Unless otherwise stated, all the questions in the test refer to the handling of a freight train of about 100 mixed freight cars, a total weight about 6,000 tons, and a matched locomotive consist of 4 locomotives generating 2,500 hp each. Only a few of the cars are empty, and these are distributed evenly throughout the train. None of the cars are over-long or over-high. The automatic brakes are pressure-maintaining. The locomotives were made in 1976 or more recently, and are equipped with dynamic brakes. The automatic brakes use 26L equipment. There is no brake flow indicator, and no rear end telemetry device.

Territory: Nearly all questions will pertain to a particular kind of territory, which is specified in the question itself. If no territory is mentioned, then the question pertains to all territories.

The territories specified in the questions include five uniformly changing grades, and four non-uniform grades.

The uniform grades are:

1. Level terrain: no changes in grade greater than 1/4% (1/4 foot per 100 feet of track, or 12 feet per mile). Normally no in-train forces are generated by the terrain when it is level.

2. Light ascending: the grade increases at a rate between 1/4% to 1% (12 feet to 50 feet per mile)

3. Heavy ascending: the grade increases at a rate greater than 1% (50 feet per mile).

4. Light descending: the grade decreases at a rate between 1/4% and 1%

5. Heavy descending: the grade decreases at a rate greater than 1%.

The non-uniform grades include crests, hogbacks, sags, and undulating terrain. All are sufficiently steep to affect train handling procedures as the train negotiates the territory.

6. Crest: the grade has a long single rise followed by a long single descent, both at a rate exceeding the surrounding terrain by at least 1/4%. The distance from the beginning of the rise to the end of the descent is at least double the train length.

7. Hogback (also called a hump or knoll): the grade has a single short rise followed by a single descent. The prevailing terrain before and after the hogback may be level, generally ascending, or generally descending.

8. Sag (also called dip): the grade has a single descent followed by level, generally ascending, or generally descending terrain.

9. Undulating: a grade changing so often that an average train has some cars on three or more alternating ascending and descending grades.

Curve Configurations: Unless specifically mentioned, all track is assumed to be straight, without any curves greater than a half degree.

1. Dynamic braking
   a. produces a braking effort similar to an independent brake, especially at low speeds
   b. is generally less effective than the independent brake beause it is operative only on powered wheels
   c. is more efficient than the independent brake because it does not depend on air pressure for its operation
   d. produces a retarding force proportional to the load shown on the ammeter

2. Running releases made at low speed are dangerous because
   a. some brakes may stick
   b. the next brake pipe reduction may not apply properly
   c. the train may make an unintended stop
   d. excessive draft forces are likely due to insufficient time for rearend release

3. On level terrain with straight track, when crossing an intersecting track,
   a. the speed of the consist should be reduced to 45 mph or track speed (whichever is less) until the consist has crossed
   b. a high throttle should be momentarily reduced until the locomotive consist has crossed
   c. the speed of the entire train should be reduced to 45 mph or track speed (whichever is less) until the entire train has crossed
   d. no adjustment is needed unless required by special instructions

4. You are starting a train on level terrain. What do you do first?
   a. advance the throttle to run 1
   b. release all brakes
   c. release only the automatic brakes
   d. release only the independent brakes

5. A split reduction
   a. occurs when a high initial brake reduction is followed by a much smaller reduction
   b. increases the time it takes to stop the train
   c. reduces excessive slack action
   d. occurs whenever any combination of two or more of the braking systems are engaged at the same time.

6. Accelerating on level can produce undesirably high in-train forces if
   a. the train stays below the minimum continuous speed too long
   b. the throttle is advanced from notch to notch before the ammeter stablizes
   c. the short-time rating is exceeded for at least five seconds
   d. all of the above

7. Normally, slowdowns or stops should be completed with no more than a 15 psi total brake pipe reduction

a. to avoid split reductions

b. to avoid excessive run-in

c. because a full service reduction leaves no reserve braking power (except emergency brake)

d. to avoid sticking brakes after a subsequent release

8. When starting a train on undulating terrain with the headend on a descending grade, the engineer should follow the same procedures as when starting on

a. a heavy descending grade

b. a sag with the headend on the descending portion

c. a crest with the headend over the crest

d. none of the above: starting on undulating terrain is unique

9. Stretch braking

a. is only a good practice on level terrain

b. is used in slowing or maintaining speed to control slack in the train

c. requires that throttle reduction begins before brake application

d. can only be done when dynamic braking is available

10. The proper procedure to start on level terrain is

a. release independent brakes, advance throttle to run 1, and then release automatic brakes, and continue until rearend moves

b. release all brakes, advance throttle to run 1 to get locomotive consist moving, remain in run 1 (or 2 if necessary) until rearend moves

c. release all brakes, advance throttle several notches, but keep ammeter below 500 amps, until rearend moves

d. advance throttle gradually, but ammeter should not exceed 800 amps until speed reaches 5 mph

11. To slow in undulating territory,

a. the dynamic brakes are preferred to the automatic brakes

b. running releases of the automatic brakes, even at high speed, should be done with extra care, if at all

c. if automatic brakes are used, they may be supplemented with the independent brakes

d. if automatic brakes are used, the initial reduction should be at least 10 psi

12. You are starting a slack-stretched train on a light ascending grade. What do you do first?

a. advance the throttle slowly to run 3 or 4

b. release the automatic brake

c. advance the throttle to run 2

d. release the independent brake

13. You are negotiating level terrain at track speed with your throttle in notch 5. You know there is a slow speed zone ahead. You decide to slow your train keeping slack stretched. What should you do first?

a. Gradually reduce the throttle to idle
b. Gradually reduce the throttle to 1
c. Make a minimum brake pipe reduction
d. Make at least a 10 psi brake pipe reduction

14. To start a backup movement on level terrain:

a. allow sufficient time for the automatic brakes to fully release, then apply throttle
b. release the automatic brakes and immediately apply the throttle
c. release the independent brake, leave the automatic brakes applied, and throttle up until the locomotive consist moves at 1 mph
d. release the automatic brakes fully, throttle up gradually using the independent brakes to limit the locomotive consist speed to 1 mph

15. To control speed on a curved descending grade, where the curvature exceeds 2^{o}, the engineer should

a. depend more on the automatic brakes than on the dynamic brakes
b. depend more on the dynamic brakes than on the automatic brakes.
c. depend on the same combination of braking that he would use on straight descending track
d. avoid the use of the independent brake, if at all possible

16. When accelerating after slowing through a sag

a. a light application of the automatic brakes until the locomotive consist starts uphill will help control slack
b. a steady advance of the throttle throughout the sag is proper as long as the ammeter stays below 800 amps
c. it may be necessary to reduce the throttle after the rearend of the train passes the sag to permit slack to adjust
d. it is best to hold the throttle constant until the rearend of the train clears the sag, then advance the throttle

17. You wish to slow while approaching a hogback. Your throttle is in notch 3 and you do not have dynamic brakes. What should you do first?

a. Immediately reduce the throttle one notch
b. Immediately make a minimum brake pipe reduction
c. Reduce the throttle one notch when the headend reaches the summit of the hogback
d. Make a minimum brake pipe reduction when the headend reaches the summit of the hogback

18. You are negotiating a sag travelling below track speed. What is the best train handling method to prevent run-in?

a. Travel through, making no changes in throttle or brakes.
b. Manipulate the throttle
c. Manipulate the throttle and automatic brakes
d. Manipulate the automatic brakes

19. The dynamic and independent brakes should not be used at the same time except when

a. changing from dynamic to air braking during stopping
b. starting a train on a crest
c. sand is not available or is impractical
d. all of the above

20. To stop on a heavy descending grade,

a. dynamic brakes should be used if available
b. automatic brakes should be used, even if dynamic brakes are available
c. the engineer is free to choose whatever brakes he prefers, depending on train makeup and terrain
d. the independent brake can be used to supplement dynamic braking until the dynamic brakes begin to lose their effectiveness

21. With doubleheading, extreme care in both road and helper consists must be used in the manipulation of the throttle to avoid

a. exceeding safe coupler limits and high L/V ratios
b. excessively high lateral forces due to coupler or car-body angularity
c. excessively high buff in-train forces, especially on curves, crossovers, and turnouts
d. all of the above

22. You wish to slow on an upcoming light descending grade. You wish to bunch your train. You do not have dynamic brakes. What do you do first?

a. Make a minimum brake pipe reduction
b. Make a brake pipe reduction of at least 10 psi
c. Gradually reduce the throttle
d. Make a minimum reduction and apply the independent brake

23. You are about to negotiate a hogback traveling at track speed. What is the best train handling procedure to minimize slack action and in-train forces?

a. Travel through, making no changes in throttle or brakes
b. Manipulate the throttle setting
c. Manipulate the throttle and automatic brakes on descent
d. Manipulate the automatic brakes

24. You are starting a stretched train on a light descending grade. What is the correct starting procedure?

a. release all brakes, allow the automatic brakes to release fully, and then advance the throttle to run 1

b. release the automatic brake fully. Then advance the throttle to run 1. Use the independent brake to control the speed of the locomotive consist.

c. release the independent brake, place throttle in run 1, and release the automatic brake

d. release the automatic brake fully. Then release the independent brake. Advance the throttle only after the entire train is moving

25. You are gradually approaching track speed over light descending terrain. What is the single best procedure to maintain correct speed?

a. Use the dynamic brake to control train speed

b. Apply the locomotive brakes

c. Make a minimum automatic brake reduction

d. Reduce the throttle setting one notch at a time

26. To start a fully stretched train stopped on a crest, where the middle of the train is on the summit, the first step is to

a. run out all slack

b. advance the throttle to 1 before releasing the independent brake

c. release the independent brake before releasing the automatic brakes

d. release the automatic brakes fully before releasing the independent brakes and adding power

27. To stop on a heavy descending grade, when the train has an uneven distribution of heavy and empty cars

a. if the heavy cars are at the headend, the independent brakes should be applied before the final stop

b. if the heavy cars are at the headend, keep some power applied until the final stop

c. if the heavy cars are at the tailend, the independent brakes should be applied before the final stop

d. use the same procedure as with a balanced weight train, but allow more distance to slow and stop

28. With a mid-train helper, extreme care in both road and helper consists must be used in the manipulation of the throttle to avoid

a. exceeding safe coupler limits and high L/V ratios

b. excessively high lateral forces due to coupler or car-body angularity

c. excessively high buff forces, especially on curves, crossovers, and turnouts

d. high headend tractive forces

29. You want to slow your train on a heavy descending grade, and you are certain that dynamic braking alone will not be sufficient. What should you do to slow your train?

a. Start with at least a 10 psi brake pipe reduction and supplement with the dynamic brakes

b. Start with a minimum brake pipe reduction, supplement with the dynamic brakes, and add further automatic brake reductions as needed.

c. Start with a minimum brake pipe reduction and allow the locomotive brakes to set as well, adding further automatic reductions as needed.

d. Supplement the dynamic brakes with a single service reduction

30. You are about to negotiate a crest. Which of the following is the best train handling procedure?

a. Manipulate the throttle and automatic brakes

b. Manipulate the dynamic and automatic brakes

c. Manipulate the automatic and independent brakes

d. Manipulate the throttle

31. You are starting a slack-bunched train on a light ascending grade. What is the correct starting procedure?

a. release the automatic brakes, immediately advance the throttle one notch at a time until the locomotive consist moves, then release the independent brakes

b. release the automatic brakes, release the independent brakes, then advance the throttle one notch at a time until the cars begin moving one at a time while brakes are releasing

c. release the automatic brakes, wait until fully released. Then add throttle sufficient to hold train and release the independent brake

d. release the independent brake and add sufficient throttle to hold the locomotive consist. Then release the automatic brake and gradually add further power to move one car at a time

32. You are approaching track speed on a heavy descending grade. Your train does not have dynamic brakes. Your throttle is in notch 3. How should you avoid overspeed?

a. Make a minimum brake pipe reduction. Immediately begin to throttle down one notch at a time

b. Make a minimum brake pipe reduction. When effective throughout the train, begin to throttle down one notch at a time

c. Make a minimum brake pipe reduction. When effective throughout the train, begin to throttle down and also apply the independent brake

d. Reduce the throttle one notch at a time. Then make a minimum brake pipe reduction

33. When planning to stop on level terrain

a. it is nearly always desirable to stop with slack stretched

b. the major consideration of how the slack is treated is the way in which the train is to be subsequently started

c. whether slack is to be bunched or stretched in stopping depends primarily on the distribution of heavy and light cars

d. the engineer can freely choose whatever method of stopping he wants, as long as he does it properly

34. When stopping while the train is crossing a sag or dip

a. it is usually preferred to keep slack bunched regardless of the overall average grade

b. it is usually important to keep slack stretched, regardless of the overall average grade

c. if the overall average grade is ascending, throttle reduction alone can often be used to stop the train

d. if the overall average grade is descending, slack should be kept stretched when using dynamic brakes

35. When braking against power, so as to keep slack stretched

a. do not let the independent brake apply

b. the independent brake should be applied along with the automatic brakes

c. the dynamic brake, when available, should be used along with the automatic brakes

d. the dynamic brakes, when available, are preferable to use rather than the independent brake

36. If power is applied too soon after making a running release, the train may separate. This is more likely if

a. the reduction being released was light (7 psi or less) and the train is stretched

b. the reduction being released was 10 psi or greater and the train is stretched

c. the reduction being released was 10 psi or greater and the train is long

d. the reduction being released was light, and the train is short

37. Stopping on an undulating terrain is best done with

a. automatic brake and throttle manipulation, even if the dynamic brakes are available

b. the dynamic brake, plus the independent brake as the train slows

c. throttle manipulation only, unless the train entered the undulating territory at high speed, and the stop was unexpected

d. the dynamic brakes, plus the automatic brakes as the train slows

38. When stopping on a light ascending grade,

a. if the stop can be made entirely with throttle reduction, this is usually preferable to the use of automatic or dynamic brakes

b. whatever method is used, the independent brake should be applied about 50 feet before the final desired stopping point

c. whatever method is used, the throttle should be at idle at least 50 feet before the desired stopping point

d. whatever method is used, avoid sand if possible

39. You are traveling just below track speed, with the headend just on the summit of the crest and the rearend still on the uphill side. In order to negotiate the crest properly, what should you do?

a. Make a minimum brake pipe reduction

b. Apply the dynamic brake

c. Gradually reduce the throttle

d. Do nothing until the locomotive consist is well over the crest

40. You wish to slow on a light ascending grade. What is the best method?

a. make a total brake pipe reduction of at least 10 psi

b. Make a minimum service reduction and apply the independent brake

c. Keep the same throttle setting and the let the terrain slow the train

d. Gradually reduce the throttle

41. You are starting a slack-bunched train on a light descending grade. What is the correct starting procedure?

a. release the independent brake, then advance the throttle to run 1 and release the automatic brakes

b. release the independent and the automatic brakes, then advance the throttle to run 1

c. release the automatic brakes, then release the independent brakes gradually, using the independent brakes to control frontend speed until the entire train is moving

d. release the independent brake gradually, release the automatic brakes and allow the entire train to stretch as the automatic brakes release

42. When accelerating after starting a train in a sag, once the entire train is moving

a. and the locomotive consist is still heading downhill, the independent brake can be used to control in-train forces

b. and the locomotive consist is heading uphill, hold the throttle in the lowest notch capable of maintaining movement until the tailend of the train clears the sag

c. regardless of where the locomotive consist is located on the sag, the throttle has to be advanced more slowly then when accelerating on level, so that the slack can adjust through the sag

d. regardless of where the locomotive consist is located, this task is similar to accelerating on level

43. When negotiating a heavy descending grade with dynamic brakes, the best train handling procedure is to

a. use the dynamic brakes alone, or supplement them with the independent brakes, if necessary

b. start with the automatic brakes and supplement them with the dynamic brakes, if necessary

c. set retainers before entering the grade

d. use the dynamics brakes alone, or supplement them with a minimum automatic brake reduction, if necessary

44. You want to start your train on a hogback. The locomotive consist and the first third of the train are on the descending part of the hogback. The first step is

a. advance the throttle to 1 until the entire train is in motion and then advance the throttle notch by notch

b. release the automatic brakes to let the slack run out, and then add throttle slowly

c. gradually release the independent brake and add throttle to keep the locomotive speed slow until the entire train is moving

d. any of the above, depending on the conditions.

45. You are negotiating level terrain at track speed. Ahead is a sag. What should you do?

a. Make no change in the throttle or brake

b. Reduce speed

c. Keep the train speed constant

d. Add throttle gradually

46. When stopping on a heavy ascending grade, the train should be

a. stretched

b. bunched

c. either stretched or bunched, depending on train makeup and territory

d. stretched only when dynamic braking is available

47. When negotiating a heavy descending grade without dynamic brakes, the best train handling procedure to control speed is to

a. hold throttle as constant as possible and make several automatic brake applications

b. use primarily throttle manipulation and minimum automatic brake application

c. supplement throttle manipulation with independent brake applications

d. set retainers before entering the grade

48. Remote locomotive units are particularly useful in cold temperatures because they reduce
- a. brake pipe gradient
- b. brake pipe charging time
- c. brake release time
- d. all of the above

49. To slow a train on a heavy descending grade, cycle braking of the automatic brakes is acceptable
- a. if dynamic brakes are not available
- b. if speed is reduced sufficiently before the brakes are released to provide adequate time to recharge before the next cycle
- c. if when speed drops below 10 mph the engineer is prepared to stop the train rather than release the brakes
- d. only if all of the above conditions are met

50. The time required for the automatic brakes to achieve release on the last car depends on
- a. train length
- b. brake pipe leakage
- c. the amount of reduction
- d. all of the above

51. If one or more locomotives in the consist starts to lurch or slip during a start, what is the safest procedure?
- a. add another notch of power
- b. add sand
- c. close throttle, come to stop, and determine the cause
- d. close throttle, take slack, and start over again

52. To stop on a heavy descending grade with dynamic brakes available
- a. the dynamic brakes should be fully applied by the time the first application of the automatic brakes becomes effective
- b. automatic brake applications should precede the dynamic brakes
- c. the independent brake should be used to supplement the automatic brake if needed
- d. avoid using sand if possible

53. You are negotiating undulating territory in which the prevailing grade is ascending. Your speed is 10 mph below track speed. What is the best train handling method to maintain speed?
- a. Travel through, make no changes in the throttle or brakes
- b. Manipulate the throttle
- c. Manipulate the throttle and automatic brakes
- d. Manipulate the automatic brakes

54. You wish to slow on a heavy ascending grade. What is the best method?
a. Make a minimum brake pipe reduction
b. Keep a high throttle setting and allow the terrain to slow the train
c. Make a minimum brake pipe reduction and apply the independent brake
d. Gradually reduce the throttle

55. You are beginning to accelerate on light descending terrain with a fully bunched train without dynamic brakes. When the entire train is moving you should
a. fully release the independent brake before adding any throttle until the train is fully stretched
b. advance the throttle at least to run 1 before fully releasing the independent brake
c. gradually release the independent brake while slowly adding throttle to stretch the train
d. continue to work the independent brake to keep the train bunched as long as possible as it gains speed

56. Your train is accelerating rapidly and approaching track speed over level terrain. What should you do to achieve correct speed?
a. Maintain your present throttle setting until you reach track speed, and then notch back quickly to a setting that should provide a balance speed
b. Start notching back now gradually
c. Maintain your present throttle setting until your speed is slightly in excess of track speed, and then notch back quickly to a setting that should provide a balance speed
d. Maintain your present throttle setting until you reach track speed, and then notch back gradually to a setting that maintains balance speed

57. You are starting a slack-stretched train on a heavy ascending grade. What is the correct procedure?
a. release the automatic and independent brakes, then advance the throttle slowly to a position sufficient to hold the train
b. sand, then release the automatic brakes and advance the throttle slowly to a position to hold train
c. advance the throttle to notch 1, then release the automatic and independent brakes
d. release the automatic brakes, advance the throttle slowly, sufficiently to hold the train, then release the independent brakes

58. When beginning to accelerate a fully bunched train on a light descending grade, and dynamic brakes are being used
a. the independent brake is used to keep slack bunched until desired speed on descent is achieved, or bottom is reached
b. The independent brake is gradually released as the dynamic brakes become effective
c. Both dynamic and independent brakes must be released by at least 10 mph so the train can be stretched
d. Throttle is added before dynamic brakes become effective to reduce in-train forces

59. You are negotiating undulating terrain. The best train handling goal is to

a. keep at track speed
b. vary speed with the ups and downs of the terrain
c. keep the locomotive consist speed constant
d. avoid heavy braking

60. To plan a start on a hogback, the engineer should consider

a. where the train is located in relation to the summit of the hogback
b. the slack condition of the train when it stopped
c. the severity of the curvature of the track over the hogback
d. all of the above

61. When accelerating on level terrain, pausing between each throttle advance insures that

a. in-train forces do not become excessive
b. wheel slips do not occur
c. excessive amperage is avoided
d. all of the above

62. Retainers are used on heavy descending grades when

a. dynamic brakes alone would not be sufficient to control train speed
b. automatic and dynamic brakes together would not be sufficient to control train speed
c. the grade exceeds 3%
d. the total train weight exceeds 10,000 tons

63. Which of the following describe the impact of curves on train handling?

a. Curves increase rolling resistence and therefore produce faster slowing and greater buff forces than do straight track
b. Curves increase lateral forces and therefore produce a greater chance of rail turnover and wheel climb
c. Curves increase draft forces when starting and make string-lining more likely
d. All of the above

64. With large locomotive consists, there is the potential for too much dynamic braking capacity. Dynamic braking capacity must be limited when

a. the independent brakes are also used
b. crossing turnouts and in sharp curves
c. there are high, long, or unloaded cars near the frontend of the train
d. all of the above

65. On a heavy descending grade, using primarily the dynamic brake to retard the train, if the dynamic brakes suddenly become ineffective, the engineer should

a. stop the train quickly with an emergency application
b. immediately add independent brake to replace the dynamic
c. reduce throttle immediately to idle if not already there
d. make a full service reduction, allowing the independent brake to set

66. Once the entire train is moving, accelerating on curved heavy ascending terrain

a. can be done at the same rate of notching up the throttle as on straight track, as long as the curve is 2^0 or less.
b. requires a slower rate of notching up to prevent string lining of heavy cars
c. requires a slower rate of notching up to prevent separation
d. is safer if the train has a higher hp/ton ratio

67. When starting a train in a sag when the locomotive consist is in the ascending portion, you should

a. release the automatic brakes and the independent brake to bunch the train before applying power
b. release the automatic brakes and apply minimum power as the independent brake is released to prevent the headend from rolling back
c. advance the throttle gradually several notches to stretch the train before releasing the automatic brakes
d. keep the automatic brakes applied until after the independent brakes are released to prevent high in-train forces

68. Train separation is likely when starting on a crest

a. at the headend on the descending side of the crest, just behind the locomotive consist, where drawbar forces are maximum
b. at the headend, especially when the crest is also on a curve
c. at the middle of the train at the summit, where drawbar forces are maximum
d. none of these: train separation is not a serious problem on crests

69. You are traveling at track speed. You are about to enter undulating terrain. What should you do?

a. It depends on the prevailing grade of the undulations
b. Make no change in throttle or brakes
c. Keep the train speed constant
d. Reduce speed

70. You have been traveling over level terrain. Ahead is a cresting grade, and you plan to slow while negotiating through the cresting territory. What is the best method?

a. Gradually reduce the throttle while the locomotive consist is on ascending portion of the grade up the crest

b. Gradually reduce the throttle as soon as the locomotive consist comes over the crest

c. Make a minimum reduction as soon as the locomotive consist comes over the crest

d. Make no change, allowing the ascending portion of the crest to slow the train

71. Which single factor has the biggest influence on stopping distance for a freight train?

a. Speed
b. Train weight
c. Train length
d. Weather conditions

72. To stop a train on a cresting grade

a. begin throttle reduction while the locomotive consist is cresting the grade

b. do not use automatic brakes when dynamic brakes are available, to prevent high draft forces at the rear of the train

c. the highest draft forces will occur just behind the locomotive consist after the consist clears the crest

d. the independent brakes should be set about 50 feet before reaching the final desired stopping point

73. To stop on a light descending grade with slack bunched, the first step is to

a. make a throttle reduction
b. make a light application of the automatic brakes
c. make a heavy application of the automatic brakes
d. the engineer can freely choose to start with throttle manipulation or with a brake application

74. The time needed for the automatic brakes to begin to apply after an application is initiated by the engineer is

a. the same for a 6 psi as for a 12 psi reduction
b. greater for a 6 psi than for a 12 psi reduction
c. less for a 6 psi than for a 12 psi reduction
d. whether it is the same, greater or less depends on the brake pipe pressure before the reduction

75. When stopping on a hogback

a. a slack-stretched method of stopping should be used to avoid large buff in-train forces

b. if the locomotive consist is on the ascending portion of the hogback, increase the throttle while the automatic brakes apply

c. if the locomotive consist has cleared the summit of the hogback, advance the throttle sufficiently to maintain speed as the brakes apply.

d. all of the above

76. Sanding on curves

a. is less effective in increasing adhesion than on straight track

b. should be minimized because it increases lateral forces

c. is needed more than on straight track due to the centrifugal forces generated by the train

d. has the same impact on train handling as on straight track

77. To stop a train on level terrain and keep the slack stretched, you should

a. apply the automatic brakes before you make any throttle reduction

b. reduce the throttle to idle before you begin any brake application

c. use the independent brake to control in-train forces

d. allow the independent brakes to set when applying the automatic brakes

78. Accelerating through undulating territory should be done

a. with power added every time the locomotive consist starts down to maintain a stretched condition

b. more quickly than over uniform terrain to keep slack stretched at all times.

c. not at all: speed should be held as constant as possible until the undulations end

d. at a slow rate

79. You are starting a slack-stretched heavy train with 3 hp/ton on a 2% ascending grade. If the locomotive consist does not begin to move forward as the independent brakes are released and power is added, you should:

a. keep adding more power, at least until the ammeter goes above 1200

b. add sand

c. consider doubling the hill

d. allow the locomotive consist to roll back a few feet to pick up some slack and then try again

80. You are starting a slack-bunched train on a heavy descending grade. The initial control of the speed of the locomotive consist is made by:

a. throttle manipulation

b. releasing the automatic brakes

c. gradual release of the independent brakes

d. gravity

81. To accelerate after starting on a hogback

a. the independent brake should be used to control speed to prevent excess stretching

b. high throttle position should be used as soon as the locomotive consist crosses the hogback to prevent run-in

c. the throttle should be advanced carefully, using the ammeter to indicate when severe in-train forces are likely on the hogback

d. all of the above

82. If you have to initiate a slowdown just as the headend reaches the bottom of a sag, your first action should be to make:

a. a minimum reduction of the automatic brakes

b. a more rapid than usual reduction in throttle

c. a minimum reduction of the automatic brakes, allowing the independent brakes to set as well

d. a normal reduction of the throttle before setting any brakes

83. You wish to slow on a light descending grade, keeping your train stretched. Your throttle is in notch 3. You do not have dynamic brakes. What do you do first?

a. Reduce the throttle gradually

b. Make a minimum brake pipe reduction

c. Make a brake pipe reduction of at least 10 psi

d. Make a minimum brake pipe reduction and apply the independent brake

84. When stopping on light ascending grades,

a. it is advantageous to have slack stretched for the next start

b. it is advantageous to have slack bunched because there is less run-in forces during the stop

c. it is advantageous to have slack bunched to reduce rollout during starting

d. Either slack stretched or bunched is acceptable, depending on train makeup

85. When stopping a train on a 2^{o} curve, and dynamic braking has been used to slow the train

a. an automatic brake application should be made to supplement the dynamic brake

b. an automatic brake application should be made and a corresponding reduction in dynamic braking be made

c. the dynamic brakes should be completely released and replaced by the independent brakes

d. the dynamic brakes should be completely released and replaced by the automatic brakes

86. When it is desirable to take slack in a train stopped on an ascending grade, in order to start the train (and no communication with the rear of train is available), the engineer should

a. make a brake pipe reduction sufficient to just hold the train, reverse the train with low throttle until all the slack is bunched, release the brakes and advance the throttle only enough to move the train forward

b. release all brakes and allow the train to roll back until the locomotive consist has bunched into the cars, then advance the throttle to run 1 to move forward

c. release all brakes, reverse the locomotive and advance throttle to run 1 until the train is bunched. Then set all brakes, reverse again, advance the throttle and release the brakes.

d. all of these are proper

87. In planning to slow on a light ascending grade, when past experience suggests that throttle reduction alone will not be sufficient, you should

a. make a minimum brake pipe reduction and then begin to throttle back

b. reduce the throttle to idle before beginning a minimum brake pipe reduction

c. reduce the throttle gradually and supplement with the independent brake

d. begin to throttle down and supplement with a minimum brake pipe reduction

88. You are negotiating level terrain at track speed and decide to slow your train allowing slack to bunch. What should you do first?

a. Make a minimum brake pipe reduction

b. Apply the independent brake

c. Gradually reduce the throttle

d. Make at least a 10 psi brake pipe reduction

89. When accelerating a fully stretched train on heavy ascending terrain

a. to avoid the minimum continuous speed limit, the engineer should get to a high throttle position as soon as feasible

b. sanding should be used whenever wheel slip is anticipated

c. high in-train forces occur from too high a throttle position when still at low speed

d. all of the above

90. In hogback territory, when accelerating after a slow down

a. buff forces among the cars on the hogback is the greatest concern

b. control of slack action within the train at the point where it crosses the summit of the hogback is the greatest concern

c. control of slack must be handled by manipulation of the throttle along with the independent brake

d. the rate of acceleration can be increased once the locomotive consist begins descent off the hogback

91. You are starting a slack-bunched train on a heavy descending grade. The independent brake had been fully applied to hold the train. What should you do first?

a. apply the automatic brake
b. release the independent brake
c. advance the throttle to run 1
d. ascertain that the automatic brake system is fully recharged

92. You have started your train and stretched it. In order to accelerate properly, you should:

a. advance the throttle one notch at a time, pausing a minimum of 20 seconds before making the next advance
b. advance the throttle one notch at a time, pausing until the speedometer registers at least a 2 mph gain
c. advance the throttle as many notches as you can until the ammeter exceeds 800 amps on any advance.
d. advance the throttle one notch at a time, watching for the ammeter to stabilize or drop before the next advance

93. You wish to slow while crossing over a cresting grade. What do you do first?

a. Gradually reduce the throttle
b. Make a minimum brake pipe reduction
c. Make a total brake pipe reduction of at least 10 psi
d. No change--allow the ascending grade to slow the train

94. To stop a train on level terrain with slack bunched, what is the first step?

a. Reduce the throttle gradually to idle to gather the slack
b. Apply light independent braking to gather the slack
c. Make a light automatic brake application while beginning to reduce the throttle
d. It depends upon the makeup of the train

95. In general, when stopping on a light descending grade, the engineer

a. should usually stop with slack bunched, as it makes it easier to start again
b. should usually stop with slack stretched, as this reduces the in-train forces during the stop
c. should usually stop with slack stretched if dynamic brakes are available, otherwise bunched.
d. is free to bunch or stretch the slack as dictated by train makeup and territory

ANSWER KEY

1. d
2. d
3. b
4. b
5. c
6. b
7. c
8. b
9. b
10. b
11. b
12. b
13. c
14. a
15. a
16. c
17. d
18. b
19. a
20. b
21. a
22. c
23. b
24. c
25. d
26. d
27. a
28. c
29. b
30. d
31. b
32. b
33. c
34. c
35. a
36. a
37. a
38. a
39. c
40. d
41. c
42. c
43. d
44. c
45. b
46. a
47. b
48. d
49. d
50. d
51. c
52. a
53. b
54. d
55. c
56. b
57. d
58. b
59. c
60. d
61. d
62. b
63. d
64. d
65. b
66. c
67. b
68. c
69. d
70. a
71. a
72. a
73. a
74. a
75. d
76. b
77. a
78. d
79. c
80. c
81. c
82. a
83. b
84. a
85. b
86. a
87. d
88. c
89. c
90. b
91. d
92. d
93. a
94. a
95. d