VALIDATION REPORT FOR

Written Examination for and Simulator Evaluation of Puget Sound Pilots

SUBMITTED TO

State of Washington
Board of Pilotage Commissioners
2901 Third Avenue, Suite 500
Seattle, Washington 98121

FINAL REPORT

comira

PREPARED BY

COMIRA
Psychometric Services Division
110 Blue Ravine Road, Suite 160
Folsom, California 95630
916.358.8525

July 20, 2009

110 BLUE RAVINE ROAD, SUITE 160
FOLSOM, CALIFORNIA 95630

T 916.358.8525
F 916 358.8528

www.comiratesting.com

July 20, 2009

Captain Harry Dudley
Board of Pilotage Commissioners
2901 Third Avenue, Suite 500
Seattle, Washington, 98121

Dear Captain Dudley:

I have attached the report describing the process for developing and administering the written examination and simulator evaluation. The report describes several procedures: how the practice analysis was conducted, how specifications were developed to support the content of the written examination and simulator evaluation, and how they were developed and administered.

In this process, several focus groups of licensed pilots served as subject matter experts to develop the practice analysis survey, develop multiple-choice questions, develop content for the written examination and simulator evaluation, and, establish passing scores for the written examination and simulator evaluation.

I am available for questions regarding the report and our overall description of the process. I can be reached at 916.358.8525 or at nhertz@comiratesting.com.

Sincerely,

Norman R. Hertz

Norman R. Hertz, Ph. D.
Director, Psychometric Services Division

ACKNOWLEDGEMENTS

We wish to acknowledge all of the Puget Sound Pilots who participated as subject matter experts and evaluators in the examination process. Working pilots were involved in every part of the process to ensure that the written examination and simulator evaluation were valid. A number of pilots were asked to identify the critical tasks performed by pilots and the knowledge required in order to perform the tasks safely and effectively. The information obtained from these interviews was used to develop the practice analysis survey, the results of which would serve as the foundation of the written examination and simulator evaluation. Virtually all of the pilots completed the survey from which the results were used to develop examination specifications for the written examination and simulation evaluation. Not only were the pilots involved in writing, reviewing, and vetting the written examination questions, they were involved in developing and administering the simulation examination.

The participation of Puget Sound Pilots contributed significantly to the development of a fair, job-related selection process to ensure that the pilot trainees will be selected according to their qualifications. The Puget Sound Pilots should be commended for their efforts in developing the written examination and simulator evaluation.

EXECUTIVE SUMMARY

This report describes the procedures used to develop a reliable and valid examination program for Puget Sound Pilots. First, a practice (occupational) analysis of current job competencies was conducted. Several focus groups of licensed pilots were convened to develop the written examination and simulator evaluation. Several procedures were used to construct the content specifications, develop multiple-choice questions, and develop scenarios and rating scales for the simulator evaluation. The scope of work included:

- Conducting a practice analysis of current pilotage practice
- Developing detailed content specifications for the written examination
- Developing detailed content specifications for the simulator evaluation
- Developing multiple-choice items (questions) based on the content specifications and job relatedness for new trainees
- Providing critical review of questions by evaluating accuracy of the content and making editorial revisions
- Selecting questions for the published examination based on content specifications
- Developing scenarios and rating scales for the simulator evaluation
- Developing standardized protocols to administer the simulator evaluation
- Establishing a criterion-referenced passing score (Angoff) for the written examination
- Establishing a competency score for the written examination
- Establishing a criterion-referenced passing score (Angoff) for the simulator evaluation

The entire validation process incorporates the Standards for Educational and Psychological Testing (1999). Each aspect of the examination program is linked to the content specifications. The specifications establish the content-related validity of the examination program by identifying the important subject matter areas involved in practice. The content specifications are the foundation for all aspects of the examination development process including item writing, item review, and test publication. All documentation necessary to verify that the validation process has been implemented in accordance to professional standards is included in the report.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

APPENDICES

SECTION 1: OVERVIEW

This validation report is structured into sections that address background of the examination program and the processes involved in content specifications, examination development, passing score for the written examination, competency score for the written examination and passing score for the simulator evaluation. Whenever possible, the appropriate standards are cited from the Standards for Educational and Psychological Testing (1999).

SECTION 2: INTRODUCTION

BACKGROUND

The Board of Pilotage Commissioners was created in 1935 by the Washington State Legislature to prevent the loss of human life, prevent loss of or damage to property and vessels, and protect the marine environment in the state's inland waters through sound application of compulsory pilotage provisions in certain state waters. These waters include waters in Puget Sound which are located inside the international boundary line between Washington and the province of British Columbia; and, all inland waters, channels, waterways, and navigable tributaries within Grays Harbor and Willapa Harbor.

To accomplish this end the board examines proficiency of potential pilots, licenses pilots, regulates pilots, enforces the use of pilots, sets pilotage rates, receives and investigates reports of accidents involving pilots, keeps records of various matters affecting pilotage and fulfills other responsibilities enumerated in the Revised Code of Washington (RCW).

The legislative policy and intent in establishing the Board was to maintain an efficient and competent pilotage service on the inland waters to ensure the protection of shipping, safety of human life and property, and protection of the marine environment.

Thus, the Board's duties can be summarized as follows:

- To adopt rules and regulations necessary for the enforcement and administration of marine pilots;
- To issue training licenses and pilot licenses to pilot candidates meeting the qualifications and such additional qualifications as may be determined by the Board;
- To establish a comprehensive training program to assist in the training and evaluation of pilot candidates before final licensing;
- To establish additional training requirements as required to maintain a competent pilotage service;
- To maintain a register of pilots, records of pilot accidents and other history pertinent to pilotage, along with a roster of vessels, agents, owners, operators, and masters necessary for the maintenance of a roster of persons to ensure pilotage and maritime safety;

- To determine the number of pilots necessary to be licensed in each district of the state to optimize operation of a safe, fully regulated, efficient, and competent pilotage service in each district;

- To establish the pilotage tariffs for pilotage services performed aboard vessels;

- To file a report that includes background information about current pilots and their assignments in each pilotage district; and,

- To provide all relevant statutes which affect pilotage without cost to pilots, agents, owners, operators, and masters.

AUTHORITY

The Revised Code of Washington Chapter 88 RCW § 16 specifically governs marine pilotage. Washington Administrative Code Chapter 363 WAC § 116-185 and § 116-300 are specific to fees established for all pilotage services rendered in the Grays Harbor and Puget Sound pilotage districts, respectively.

APPLICANT QUALIFICATIONS

To become a Puget Sound Pilot, an applicant must have at least one to two years of experience as the captain of an ocean-going vessel, tugboat, passenger ferry, or a military vessel.

Qualified applicants (candidates) sit for an extensive written examination followed by evaluation of the candidate's shiphandling skills in a bridge simulator. A variety of real world complications and hazards are presented, and the candidates are evaluated on how well they handle unforeseen events.

TRAINING LICENSE

Candidates are selected based on their combined scores on the written examination and the simulator evaluation. Successful candidates may be issued a Training License by the state and authorized to begin hands-on training specifically tailored for the pilot trainee's background, skills, and needs. The initial pilot training program takes between eight months and three years. During this time, under the careful supervision of veteran pilots, trainees perform pilotage on board working vessels in the various ports of Puget Sound. Their performance on each trip is evaluated in writing as to the candidate's basic shiphandling skills, knowledge of local waters, and ability to handle the requirements of the job. These evaluations are collected and reviewed by the Washington State Board of Pilotage Commissioners before the initial license is issued.

If approved, candidates receive their initial, restricted pilot's license which limits the size and type of vessels they can handle. They must successfully demonstrate competence handling smaller vessels before moving on to larger vessels. For example, during the first year, the new pilot is not allowed to work on loaded tankers

or on passenger vessels. During this first year, the pilots are also required to attend a marine simulator for more instruction. The second year, the new pilot can work on small loaded tankers, etc., gradually increasing the size and type of vessels as the years progress.

UNRESTRICTED LICENSE

The restricted period lasts five years, after which time the trainee pilot receives an unrestricted Washington State marine pilot's license. Only after the end of this five-year period is a pilot authorized to handle the largest tankers, container, cargo, and cruise ships.

UTILIZATION OF SUBJECT MATTER EXPERTS

The Board identified licensed, working pilots to provide subject matter expertise in interviews and focus groups. The subject matter experts (SMEs) who participated had varying degrees of experience ranging from newly licensed to very experienced.

CONTENT VALIDATION STRATEGY

To ensure that the examination reflected the actual tasks performed by Puget Sound Pilots, a content validation strategy was employed to establish the link between the job tasks and the content of the examination. Therefore, persons who were licensed pilots were consulted to identify the major subject matter areas and to develop the examinations.

PSYCHOMETRIC STANDARDS

The Standards for Educational and Psychological Testing (1999) set forth by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (hereafter called the "Standards"), serve as the standards for development of all aspects of a test, including test development, passing score, administration of tests, and reporting of results. The Standards are used by the measurement profession as the psychometric standards for validating all examinations, including licensing and certification examinations.

SECTION 2: GENERAL OVERVIEW

GENERALIZED METHODOLOGY

The general methodology for developing the written examination and simulator evaluation relied heavily on the expertise of SMEs who participated in the focus groups. The SMEs worked with members of the Pilot Examination Committee, staff at the Pacific Maritime Institute, and Drs. Norman Hertz and Roberta Chinn of Comira.

STANDARDIZED PROCEDURES AND PROTOCOLS

Standardized procedures and protocols to administer both the written examination and simulator evaluation were developed to ensure that every candidate had the same examination experience.

EXAMINATION SECURITY

All persons involved in the written examination or simulator evaluation signed a nondisclosure agreement which specified that they would keep all examination materials secure, discuss the examination only during department-sponsored meetings, and avoid involvement in any examination-oriented review program for prospective pilot candidates. This included Comira staff, staff at the Pacific Maritime Institute, and pilots that served as subject matter experts to develop the examinations.

MAJOR ACTIVITIES

Several focus groups accomplished the following tasks:

- Conducting a practice analysis of current pilotage practice
- Developing detailed content specifications for the written examination
- Developing detailed content specifications for the simulator evaluation
- Developing multiple-choice items (questions) based on the content specifications and job relatedness for new trainees
- Providing critical review of questions by evaluating accuracy of the content and making editorial revisions
- Selecting items for the published examination based on content specifications
- Developing scenarios and rating scales for the simulator evaluation

- Developing standardized protocols to administer the simulator evaluation

- Establishing criterion-referenced passing scores (Angoff) for the written examination and simulator evaluation

SECTION 3: OVERVIEW OF SELECTION PROCESS

RATIONALE

The selection process involved multiple evaluations of candidate qualifications. A candidate must meet the minimum requirements to sit for the written examination. If candidates pass the written examination based on minimum competency criteria and, in addition, achieve a sufficiently high competency score, they are allowed to sit for the simulator evaluation.

The written examination and simulator evaluation were equally weighted (150 points each). The rationale for equal weighting was that each examination assessed different types of knowledge and skills and contributed equally to trainee success.

MAJOR COMPONENTS

The major components of the selection process were:

MINIMUM REQUIREMENTS	• Minimum license as a master of steam or motor vessels of not more than 1,600 GRT or 3,000 GT (ITC) • Minimum amount of service on one of the following vessels as a master: a) One year on a cargo or tank of 5,000 GRT or 10,000 GT (ITC) in ocean or near coastal waters, b) Two years on a cargo or tank of 700 GRT or 1,400 GT (ITC) in ocean waters or near coastal waters, c) Two years on a cargo or tank of 1,600 GRT or 3,000 GT (ITC) in inland waters, d) Two years on a passenger or ferry vessel of 1,600 GRT or 3,000 GT (ITC) in ocean waters, near coastal or inland waters, or, e) Two years on a towing vessel of 150 GRT or 300 GT (ITC) in ocean waters, near coastal or inland waters. • Either of the following may be used as a substitute for minimum service: a) Three years of service as an active member of an organized professional pilot association or as a government employed pilot , or, b) Two years of service as a commanding officer or master of U.S. flag government vessels of not less than 3000 displacement tons.

WRITTEN EXAMINATION	• The multiple-choice examination consists of 150 items covering eight subject matter areas: Prevoyage planning (8%), Docking and undocking (20%), Use of anchors (9%), Waterway transit (7%), Shiphandling (24%), Navigation (12%), Pilot-master exchange (11%), and Safety (9%) • There is a criterion-referenced passing score based on minimum qualified criteria • In addition to the passing score, a competency score was established as a cutoff to identify candidates who have achieved a greater degree of knowledge. The written examination functions as a selection examination in which it is acceptable to select from the higher scoring candidates. • Candidates who passed the written examination and achieved the selected competency score were eligible to take the simulator evaluation
SIMULATOR EVALUATION	• The simulator evaluation assessed skills required for ship navigation in piloting waters that cannot be derived from experience or fully assessed in a written examination • Included are common navigational situations: fixed hazards, other vessels in meeting or crossing situations, situations with reduced visibility, and situations requiring emergency response • There are thirty (30) measurement opportunities covering seven subject matter areas: a) Fundamental piloting and ship handling b) Assimilates and prioritizes data to safely maneuver ship c) Responds appropriately in routine situations d) Responds appropriately in non-routine or emergency situations e) Communicates well and projects proper bridge presence f) Bridge resource management g) Rules of the road • There is a criterion-referenced passing score based on minimum qualified criteria

SECTION 4: PRACTICE ANALYSIS

PURPOSE

The purpose of the practice analysis was to define current practice in terms of the tasks that would be performed and the knowledge base that would be used by trainees and, to provide the foundation of the specifications for the written examination and simulator evaluation.

APPLICABLE STANDARDS

The most relevant Standards that apply to practice analysis for credentialing examinations are:

Standard 14.8 *"Evidence of validity based on test content requires a thorough and explicit definition of the content domain of interest." (p 160)*

Standard 14.10 *"When evidence of validity based on test content is presented, the rationale for defining and describing a specific job content domain in a particular way (e.g., in terms of tasks to be performed or knowledge, skills, abilities and other personal characteristics) should be stated clearly." (p 160)*

Standard 14.14 *"The content domain to be covered by a credentialing test should be defined clearly and justified in terms of the importance of content for credential-worthy performance in an occupation or profession. A rationale should be provided to support a claim that the knowledge or skills being assessed are required for credential-worthy performance in an occupation and are consistent with the purpose for which the licensing or certification program was instituted." (p 161)*

METHODOLOGY

General approach

The methodology employed by Comira is analogous to the process involved in a review of the literature for a research study in which references are searched for until no new references can be identified. Similarly, Comira continued data gathering in background research and conducting interviews for the practice analysis until no new tasks or knowledge were identified.

Comira conducted the practice analysis with the assumption that all tasks and knowledge in pilotage practice would be identified. The level of specificity of the tasks and knowledge was controlled by creating tasks and knowledge with a conceptual "size" that required about the same amount of work. For example, each task is a unit of work rather than individual steps in a procedure. The task must be a "stand alone" unit of work that does not depend on qualifiers, e.g., properly, correctly, or, upon the context of other tasks to be meaningful.

Subject matter experts, in consultation with Comira's psychometric staff, derived the underlying knowledge by considering what organized bodies of knowledge are required to perform one or more tasks. For example, the task "Plan transit by reviewing current, tide, and weather" requires a knowledge base such as "Knowledge of tide and current tables." Other bodies of knowledge could include rules, procedures, systems, methods, or requirements.

Rating scales

Rating scales are important to the outcome of the practice analysis so that the resultant data reflects the importance of tasks and knowledge along a continuum. The behavioral descriptors on the anchors of the rating scales are constructed to ensure that there are equal intervals between descriptors, (e. g., very easy, easier than most, average difficulty, more difficult, most difficult).

Equal intervals are essential to the application of mathematical operations such that there is a meaningful numeric "weight" assigned to that task. This is because the task ratings determine the weight of each subject matter area of the examination. For example, the mean importance of the ratings was calculated for the tasks. The average of the weights for the tasks in a subject matter area can be used to determine the percentage of items to be developed for a subject matter area.

One scale was used in the survey (see Appendix A). Practitioners were asked to consider the relative importance of tasks. A non-response option ("Not Relevant") was provided in the rating scale so that respondents could indicate that the task or knowledge did not apply to their job.

Background research

An important step in conducting a practice analysis is to gain a conceptual understanding of the profession to be evaluated. For this practice analysis, Comira conducted a thorough review of available documentation and relevant resource textbooks and manuals. By reviewing these materials, Comira became familiar with terms and concepts involved in pilotage practice, e.g., type of line placement and their function, maneuvering characteristics, International Maritime Organization regulations, etc.

PROCEDURES

Interviews with practitioners

Several subject matter experts were interviewed in person at the Board's office in Seattle. During the interviews, the experts were asked to identify major subject matter areas of their practice and the job tasks performed in each subject matter area. They were also asked to identify specific knowledge and abilities necessary to perform each job task safely and competently.

Transcription of tasks and knowledge

The information gathered from the interviews was transcribed and combined with information from background materials to develop a preliminary list of job tasks and knowledge statements with a consistent format and language.

Then, the preliminary list was reviewed by a focus group of subject matter experts. The task and knowledge statements were refined until each task and knowledge was the same conceptual "size", structured in a consistent format and phrased in consistent technically and conceptually accurate terminology. Every task was associated with at least one knowledge and every knowledge was associated with at least one task. The intent of the focus group review was to provide the depth and breadth in the tasks and knowledge such that they were sufficient to develop items for an examination.

The task statements were used as the basis of the survey questionnaire; the knowledge statements were linked to the task statements when the content specifications were developed.

Survey questionnaire

The survey questionnaire consisted of two parts. The first part asked practitioners to provide demographic information about themselves and their practice. The second part asked practitioners to rate the relative importance of 149 tasks.

Distribution and data collection

The survey questionnaire was available to all Puget Sound Pilots via an Internet link listed on the Comira website. The online version allowed respondents to log on multiple times until the questionnaire was completed. The respondent ratings were automatically saved in a database each time the respondent logged off the system. When respondents were satisfied with their ratings, they selected "submit" and completed the questionnaire. Once the questionnaire was submitted to the server, no further ratings could be made. Practitioners could log on and obtain a randomly generated password in order to complete the survey.

RESULTS

Response rate

There were a total of 56 (56/58 = 96%) Puget Sound Pilots who responded to the survey.

Reliability of ratings

All ratings from the questionnaire were evaluated with a standard index of reliability called coefficient alpha (α). Coefficient alpha is an estimate of internal consistency reliability of the respondents' ratings of job task and knowledge/abilities in the questionnaires. Coefficient alpha for task ratings were highly significant ($\alpha < .01$; tasks: $r = .97$ to .96).

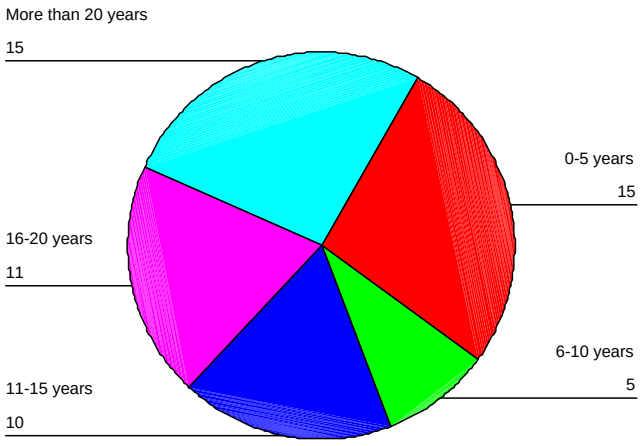
Respondent demographics

a) Have you been involved in the navigation of a vessel while holding or working under the authority of a pilot license?

All 56 respondents answered "Yes" to this question.

b) How long have you been a licensed pilot in Puget Sound?

Figure 1 – Years licensed in Puget Sound



c) Did you pilot prior to being licensed as a Puget Sound Pilot?

Table 1 – Worked prior to being licensed as PSP

Yes	23
No	33

d) If you responded "Yes" in the previous item, under what authority did you perform piloting duties?

Table 2 – Authority

State license	4
Federal license	15
Not applicable	37

e) What is your maritime background?

Respondents were asked to mark all categories that applied to their job functions.

Table 3 – Maritime background

Tugboats (inland, towing vessel and vessels towed > 1600 GT)	20
Tugboats (ship assist)	18
Tugboats (offshore)	19
Freighters	27
Tankers	15
Passenger ships	7
Other	26

"Other" included:

- Charter yachts
- Coastal freighters under 500 tons
- Coastwise
- Container
- DP vessels
- Factory trawlers
- Ferries (8)
- Ferries and high speed cats
- Fishing industry (4)
- LNG carriers (2)
- Marine surveyor (2)
- Navy vessels (2)
- Operations manager (2)
- Owner's rep on salvage jobs
- Pilot in Alaska – 16 years

- Seismic research vessels
- Survey vessel
- Tugs
- US Coast Guard

SECTION 5: CONTENT SPECIFICATIONS

APPLICABLE STANDARDS

The most relevant Standards that apply to specifications for credentialing examinations are:

Standard 3.3 *"The test specifications should be documented, along with their rationale and the process by which they were developed. The test specifications should define the content of the test, the proposed number of items, the item formats, the desired psychometric properties of the items and the item and section arrangement." (p 43)*

Standard 3.5 *"When appropriate, relevant experts external to the testing program should review the test specifications. The purpose of the review, the process by which the review is conducted, and the results of the review should be documented." (p 43-44)*

Standard 14.4 *"The content domain to be covered …should be defined clearly and justified in terms of the importance of the content for credential-worthy performance in an occupation or profession. A rationale should be provided to support a claim that the knowledge or skills being assessed are required for credential-worthy performance in an occupation and are consistent with the purpose for which the licensing or certification program was instituted." (p. 161)*

PROCESS

Licensed Puget Sound Pilots served as SMEs in a focus group that was convened to evaluate the major subject matter areas that should be included in the content specifications. The group considered relevant WAC and RCW statutes in developing the content specifications.

Only those tasks that were deemed testable by the SMEs were included in the (see Appendix B) content specifications for the written examination and simulator evaluation.

SUBJECT MATTER AREAS IN WRITTEN EXAMINATION

The focus group evaluated the importance of major subject matter areas of practice, their relative importance to practice, and validated the weights (percentages) to each area based on the task ratings in the practice analysis. For example, if the weight is 20%, there would be 30 questions on a 150-item examination. The content specifications for the written examination covering eight subject matter areas are listed below.

I. Prevoyage planning (8%)
II. Docking and undocking (20%)
III. Use of anchors (9%)
IV. Waterway transit (7%)
V. Shiphandling (24%)
VI. Navigation (12%)
VII. Pilot-master exchange (11%)
VIII. Safety (9%)

SUBJECT MATTER AREAS IN SIMULATOR EVALUATION

The focus group also determined which tasks were to be measured in the simulator evaluation. They evaluated the relative importance of these tasks to practice and assigned them to each subject matter area. The subject matter areas of the simulator evaluation are listed below.

Subject matter area	Number of measurement opportunities
I. Fundamental piloting and ship handling	8
II. Assimilates and prioritizes data to safely maneuver ship	11
III. Responds appropriately in routine situations	3
IV. Responds appropriately in non-routine or emergency situations	2
V. Communicates well and projects proper bridge presence	1
VI. Bridge resource management	3
VII. Rules of the road	2
Total	30

SPECIFICATIONS FOR WRITTEN EXAMINATION

The content specifications for the written examination are presented on pages 18-25. The specifications are organized in terms of subject matter area, tasks that a new trainee must be able to perform, and the knowledge base that they must possess to perform the tasks.

SPECIFICATIONS FOR SIMULATOR EVALUATION

The content specifications for the simulator evaluation are presented on pages 26-28. The specifications are organized in terms of categories of piloting duties to be performed.

CONTENT SPECIFICATIONS FOR WRITTEN EXAMINATION

I. Prevoyage planning (8%)	
Tasks	Knowledge
T2. Plan transit by reviewing charts, current, tide and weather. T3. Plan transit in consideration of berth side to. T4. Evaluate dispatched tugs by considering type, capabilities, limitations and bollard pull for assigned job. T7. Calculate underkeel clearance based upon depth of waterway and predicted height of tide. T8. Calculate air draft clearance based upon depth of waterway and predicted height of tide.	K1. Knowledge of information contained on the dispatch sheet. K2. Knowledge of tide and current tables. K3. Knowledge of effect of environmental conditions (e.g., height of tide, direction and speed of current, wind, sea state). K4. Knowledge of different configurations of ships. K5. Knowledge of propulsion/steerage maneuvering systems. K6. Knowledge of type and bollard pull/horsepower of available tugs. K9. Knowledge of Local Notice to Mariners. K10. Knowledge of vessels at anchor and their locations. K16. Knowledge of procedures for calculating air draft and overhead clearance. K17. Knowledge of procedures for calculating underkeel clearance. K79. Knowledge of Chart One. K97.Knowledge of actions to take when encountering restricted visibility.

II. Docking and undocking (20%)	
Tasks	Knowledge
T9. Plan placement of tugs by considering tug capabilities, weather, current, vessel characteristics, berth and maneuverability required to dock/undock ship. T10. Confer with tugs regarding tug position and lines, and docking/undocking plans prior to reaching/departing berth. T11. Establish sequence and placement of lines in making fast or letting go. T12. Determine safe speed by considering proximity to other vessels, underkeel clearance, weather conditions and capabilities of tugs in making approach to berth or leaving berth. T13. Check in/out with Vessel Traffic Service (VTS) prior to departing dock or when docking is complete. T14. Evaluate effect of underkeel clearance on handling characteristics of ship. T15. Make adjustments for cushion and bank effect when docking/undocking ship. T16. Anticipate ship's response to effect of rudder and rotation of propeller prior to ordering an astern bell. T18. Confer with master to obtain information regarding ship's handling characteristics and readiness for docking/undocking. T19. Assess effect of underkeel clearance and berth configuration on ship's movement when planning docking/undocking. T20. Determine number, type of tugs and bollard pull while considering effect of current, wind, and availability of bow thrusters.	K18. Knowledge of limitations of tugs assigned to ship movement. K19. Knowledge of vessel characteristics, e.g., propulsion, thrusters, type of rudder, engine speeds, bollard safe working load (SWL), propeller type. K21. Knowledge of docking/undocking maneuvers. K22. Knowledge of types of line placement and their function. K23. Knowledge of stopping ability relative to type of propulsion. K24. Knowledge of maneuvering characteristics, e.g., steam, diesel propulsion. K25. Knowledge of effects of speed in confined and shallow waters. K26. Knowledge of inertia and its effects. K27. Knowledge of effects of hydrodynamics on docking/undocking. K28. Knowledge of effects of propeller forces when going ahead or astern. K29. Knowledge of information contained on the pilot card. K30. Knowledge of effects of sail area on maneuverability of ship. K31. Knowledge of amount of bollard pull/horsepower to overcome effects of wind loading. K32. Knowledge of sequence of taking lines in, e.g., proximity to thruster and propeller. K33. Knowledge of differences between horsepower and bollard pull as applied to tugs.

III. Use of anchors (9%)	
Tasks	Knowledge
T22. Assess anchorage for potential hazards. T23. Verify ship's position by using ranges and bearings. T25. Determine amount of anchor chain needed after considering wind/current conditions, depth of water, quality of holding ground and maneuvering room. T26. Confer with master regarding procedures for dredging the anchor. T29. Account for actual position of the anchor when calculating swinging radius.	K34. Knowledge of natural hazards of anchorages, e.g., depths, obstructions. K35. Knowledge of navigational practices for determining location of anchorage area. K36. Knowledge of chart information relative to anchoring. K37. Knowledge of Coast Guard anchorage regulations. K38. Knowledge of factors to consider in determining scope of chain when anchoring. K39. Knowledge of watchkeeping standards for vessels at anchor. K40. Knowledge of anchor dredging techniques. K41. Knowledge of ship's anchoring equipment. K42. Knowledge of deep water anchor procedures.

IV. Waterway transit (7%)	
Tasks	Knowledge
T30. Determine water depths, applied tides and underkeel clearance for transiting route. T32. Compensate for effects of currents and weather on maneuverability of ship. T33. Consider air draft interactions, moored vessels, bank effect and effective use of tugs during transit.	K43. Knowledge of effects of predicted depth vs. observed depth of water (barometric pressure). K44. Knowledge of vessel's draft and squat in relation to speed and depth of water. K45. Knowledge of effects of wind and/or visibility on ability to maneuver in waterway. K46. Knowledge of visual cues to assess wind and current. K48. Knowledge of safe speed to proceed through waterway. K49. Knowledge of waterway dimensions and width of horizontal clearance when calculating available room to navigate vessel. K98. Knowledge of procedures to mitigate bank effect with tugs and ship's power. K99. Knowledge of procedures to mitigate effects of wind with tugs and ship's power. K100. Knowledge of methods to manage rate of turn and pivot point in waterways. K101. Knowledge of hydrodynamics involved in transiting narrow channels. K102. Knowledge of methods to use tugs to back vessel in a waterway.

V. Shiphandling (24%)	
Tasks	Knowledge
T39. Evaluate ship's responsiveness when executing maneuvers, e.g., turns, slowing or speeding up. T42. Monitor ship's performance by using visual and other natural cues. T43. Adjust amount of rudder applied as pivot point and speed changes. T44. Monitor advance and transfer. T46. Consider effects of sail area on track line of ship. T47. Determine effectiveness of bow thruster in relation to wind or speed of ship. T48. Evaluate change in pivot point location resulting from different external forces. T49. Evaluate effect of engine type and propulsion system on maneuverability. T50. Evaluate effect of trim on steering. T51. Evaluate effect of propeller transverse thrust going astern. T52. Evaluate effects of different types and locations of rudders on handling of ship.	K58. Knowledge of vessel maneuvering characteristics as posted. K59. Knowledge of slowdown time and distance necessary for reducing speed. K60. Knowledge of emergency procedures for slowing ship. K61. Knowledge of limitations and discrepancies of instrumentation calculations. K64. Knowledge of location of pivot point of a vessel underway. K65. Knowledge of advance and transfer and the elements and forces that affect it. K66. Knowledge of effects of speed, trim and draft on bow thruster. K67. Knowledge of effects of water speed and engine speed on rudder control. K68. Knowledge of effect of engine types, propellers and rudders on maneuverability of ship. K69. Knowledge of trim and its effect on pivot point. K70. Knowledge of propeller thrust while maneuvering astern. K71. Knowledge of specialized rudders and their effect. K103. Knowledge of handling characteristics with sternway.

VI. Navigation (12%)	
Tasks	Knowledge
T37. Follow international rules of the road (COLREGS). T38. Monitor Vessel Traffic Service (VTS) and bridge-to-bridge frequencies. T53. Confirm position from navigational instruments (e.g., gyro, fathometer, radar, GPS). T54. Obtain information regarding names of ships and other vessels by using Automatic Identification Systems and Vessel Traffic Service. T55. Obtain information regarding collision avoidance by using all available means. T56. Monitor targets to determine crossing or close quarters situations. T57. Identify vessels by their light characteristics at night or day shapes by day. T58. Conduct vessels in accordance with international rules of the road. T59. Use all available means to keep vessel's position on desired track line.	K50. Knowledge of factors that affect ETA calculations. K51. Knowledge of the difference between speed over ground and speed through the water. K52. Knowledge of calculations for set and drift. K53. Knowledge of engine slow down procedures. K54. Knowledge of capabilities of navigational equipment. K55. Knowledge of International Rules of the Road (COLREGS). K56. Knowledge of VHF radio protocols. K72. Knowledge of basic navigation equipment. K73. Knowledge of errors that can occur with navigational equipment. K74. Knowledge of how to recognize and compensate for instrument errors. K75. Knowledge of information displayed by AIS and its limitations. K76. Knowledge of collision avoidance procedures. K78. Knowledge of different methods to assess risk of collision, e.g., constant bearing, radar and visual bearings. K81. Knowledge of aids to navigation.

VII. Pilot-master exchange (11%)	
Tasks	Knowledge
T61. Confer with master regarding bridge team expectations. T62. Confer with master regarding condition of propulsion systems, navigation systems and anchors. T63. Confer with master regarding voyage plan, including route and ETA. T64. Determine ship's maneuvering characteristics. T65. Review ship's particulars (pilot card). T66. Confer with master regarding number and types of tugs needed. T67. Establish contingency plan when escort tugs are utilized for controlling ship's movement in event of engine/rudder failure.	K82. Knowledge of crew duties and bridge procedures when assuming the conn. K83. Knowledge of rate of turn vs. rudder commands. K84. Knowledge of responsibilities to fix position. K85. Knowledge of routing, current, weather and expected traffic. K86. Knowledge of deficiencies that would not allow ship to depart berth or enter port. K87. Knowledge of information master passes to pilot prior to pilot taking conn.

VIII. Safety (9%)	
Tasks	Knowledge
T68. Evaluate placement and condition of pilot ladder to assess safety for embarking and disembarking. T71. Consider effects of environmental and traffic conditions when establishing safe speed. T72. Anticipate conditions that could result in an error chain. T74. Determine if environmental conditions are safe for movement of the ship.	K88. Knowledge of proper way to rig a pilot ladder. K89. Knowledge of International Maritime Organization regulations. K90. Knowledge of escort tug's capabilities. K91. Knowledge of information provided by electronic equipment and VTS. K92. Knowledge of wind and visibility when establishing safe speed. K94. Knowledge of procedures for identifying potential error chain. K95. Knowledge of multiple distractions that may cause lack of situational awareness. K96. Knowledge of environmental conditions that affect safe movement of ship.

CONTENT SPECIFICATIONS FOR SIMULATOR EVALUATION

	Task to be assessed	Subject Matter Areas						
		FUNDAMENTAL PILOTING AND SHIP HANDLING ABILITY	ASSIMILATES AND PRIORITIZE DATA TO SAFELY MANEUVER SHIP	RESPONDS APPROPRIATELY IN ROUTINE SITUATIONS	RESPONDS APPROPRIATELY IN NON-ROUTINE OR EMERGENCY SITUATIONS	COMMUNICATES WELL AND PROJECTS PROPER BRIDGE PRESENCE	BRIDGE RESOURCE MANAGMENT	RULES OF THE ROAD
T56	Monitor targets to determine crossing or close quarters situations.	X	X		X		X	X
T55	Obtain information regarding collision avoidance by using all available means.	X	X		X		X	X
T58	Conduct vessels in accordance with international rules of the road.			X	X			X
T37	Follow international rules of the road (COLREGS).			X	X			X
T71	Consider effects of environmental and traffic conditions when establishing safe speed.		X					X
T57	Identify vessels by their light characteristics at night or day shapes by day.		X					X
T41	Verify commands have been carried out by monitoring ship's equipment.	X				X	X	
T54	Obtain information regarding names of ships and other vessels by using Automatic Identification Systems and Vessel Traffic Service.		X			X	X	
T72	Anticipate conditions that could result in an error chain.		X		X		X	
T60	Modify voyage plan based on traffic/berth conflicts or environmental conditions.		X		X		X	
T17	Advise master to have anchors at the ready.				X		X	

CONTENT SPECIFICATIONS FOR SIMULATOR EVALUATION

	Task to be assessed	Subject Matter Areas						
		FUNDAMENTAL PILOTING AND SHIP HANDLING ABILITY	ASSIMILATES AND PRIORITIZE DATA TO SAFELY MANEUVER SHIP	RESPONDS APPROPRIATELY IN ROUTINE SITUATIONS	RESPONDS APPROPRIATELY IN NON-ROUTINE OR EMERGENCY SITUATIONS	COMMUNICATES WELL AND PROJECTS PROPER BRIDGE PRESENCE	BRIDGE RESOURCE MANAGMENT	RULES OF THE ROAD
T12	Determine safe speed by considering proximity to other vessels, underkeel clearance, weather conditions and capabilities of tugs in making approach to berth or leaving berth.	X	X					
T32	Compensate for effects of currents and weather on maneuverability of ship.	X		X	X			
T45	Consider effects of current on track line of ship.	X	X	X				
T42	Monitor ship's performance by using visual and other natural cues.	X		X				
T59	Use all available means to keep vessel's position on desired track line.	X	X					
T53	Confirm position from navigational instruments (e.g., gyro, fathometer, radar, GPS).	X	X					
T44	Monitor advance and transfer.	X	X					
T40	Assess rate-of-turn from visual or electronic means.	X	X					
T39	Evaluate ship's responsiveness when executing maneuvers, e.g., turns, slowing or speeding up.	X	X					
T36	Monitor closest point of approach with other traffic during voyage.	X	X					
T3	Plan transit in consideration of berth side to.	X	X					

CONTENT SPECIFICATIONS FOR SIMULATOR EVALUATION

	Task to be assessed	Subject Matter Areas						
		FUNDAMENTAL PILOTING AND SHIP HANDLING ABILITY	ASSIMILATES AND PRIORITIZE DATA TO SAFELY MANEUVER SHIP	RESPONDS APPROPRIATELY IN ROUTINE SITUATIONS	RESPONDS APPROPRIATELY IN NON-ROUTINE OR EMERGENCY SITUATIONS	COMMUNICATES WELL AND PROJECTS PROPER BRIDGE PRESENCE	BRIDGE RESOURCE MANAGMENT	RULES OF THE ROAD
T25	Determine amount of anchor chain needed after considering wind/current conditions, depth of water, quality of holding ground and maneuvering room.	X	X					
T24	Establish target location in the anchorage.	X	X					
T23	Verify ship's position by using ranges and bearings.	X						
T2	Plan transit by reviewing current, tide and weather.	X						
T1	Plan transit by reviewing dispatch sheet of ship's particulars.	X						
T7	Calculate underkeel clearance based upon depth of waterway and predicted height of tide.		X					
T5	Plan route based on potential congestion, other traffic and tides/currents.		X					
T22	Assess anchorage for potential hazards.		X					

SECTION 6: DEVELOPMENT OF WRITTEN EXAMINATION

APPLICABLE STANDARDS

Several of the Standards apply to the validation of examinations:

Standard 3.1	*"Tests and testing programs should be developed on a sound scientific basis. Test developers and publishers should compile and document adequate evidence bearing on test development." (p. 43)*
Standard 3.7	*"The procedures used to develop, review, and try out items, and to select items from the item pool should be documented." (p. 44)*
Standard 3.11	*"Test developers should document the extent to which the content domain of a test represents the defined domain and test specifications." (p. 45)*
Standard 14.14	*"The content domain to be covered…should be defined clearly and justified in terms of the importance of the content for credential-worthy performance in an occupation or profession. A rationale should be provided to support a claim that the knowledge or skills being assessed are required for credential-worthy performance in an occupation and are consistent with the purpose for which the licensing or certification program was instituted…. In tests used for licensure, skills that may be important to success but are not directly related to the purpose of licensure, e.g., protecting the public, should not be included." (p. 161)*

OVERVIEW

The written examination was developed according to the Standards to ensure that the examination met all psychometric standards. The written examination consisted of 150, four-option multiple-choice items (questions). The examination was constructed according to content specifications developed from the 2008 practice analysis of Puget Sound Pilots. A strong methodological approach was used to develop the questions in the examination. All questions were developed and reviewed by multiple focus groups of practicing pilots. Great care was taken by the focus groups to ensure that the content of the questions was technically accurate and was based on a published reference source.

STRATEGY

Licensed pilots served as SMEs in multiple focus groups whose task was to create questions (items) for the examination and select items for the examination. By doing so, the content of the items was created and reviewed in a controlled setting, where SMEs could exchange feedback from other SMEs regarding the technical content of items.

PROCEDURES

Item development

SMEs were provided a formal orientation in the principles of good item construction, opportunities to familiarize themselves with the content specifications, and opportunities to work with fellow SMEs to create the items. For each item, considerable emphasis was placed on specifying the linkage of item content to the content specifications and providing a citation from an authoritative reference source. Therefore, each item was linked to a specific topic in the specifications and to a page or section of an authoritative reference source. There were numerous opportunities for individual assistance with item development as well as opportunities for review by other SMEs.

Item review

Formal orientation was provided to benefit SMEs who were previously involved in item development and SMEs who were new to the process. SMEs worked individually or in pairs in a face-to-face focus group to provide initial review of the items and then reviewed the items as a group. Items were evaluated for clarity, technical accuracy, readability, and applicability to actual job situations.

Test publication

SMEs were given a formal orientation to the publication process. In the orientation, they were given an overview of the process. Then, they were asked to consider several factors in selecting items for the written examination, the percentage of items designated for each area in the content specifications, depth and breadth of content coverage in each area, and similarity/dissimilarity of item content.

The process for selecting items was as follows. The SMEs considered the content of items, subject matter area by subject matter area and then made their selections, e.g., all items for shiphandling, then, items for master-pilot relationship, items for pilot safety, and so on. After each set of items was selected, SMEs mutually agreed upon the selection of the items that met the percentages of items prescribed by the content specifications.

As a final step, a draft form of the written examination was critically reviewed by SMEs to ensure that the examination was without error in content, wording, and was consistent with the content specifications for the written examination.

RESULTS OF STATISTICAL ANALYSIS

Statistical analysis, e.g., item analysis, of the examination indicates that the questions performed well and that the examination is a very reliable and robust measure of candidates' knowledge. Thus, candidates who achieve a high score on the written examination were determined to have a greater amount of knowledge than those who achieved low scores.

SECTION 7: PASSING SCORE FOR THE WRITTEN EXAMINATION

MEANING OF A PASSING SCORE

A criterion-referenced passing score, or cut score, is the score that reflects minimum standards of competent practice for pilot trainees. The passing score is based upon the difficulty of the items in an examination, not the scores of candidates who sat for an examination.

UNDERLYING PREMISE

The underlying premise of passing score ratings is minimum qualifications criteria required for safe, competent practice. The criteria defines minimum qualifications in terms of critical, job-related work behaviors and take into account the training and experience that candidates would bring to the examination.

APPLICABLE STANDARDS

The most relevant Standards applicable to passing scores are:

Standard 3.4	*"The procedures used to interpret test score, and, when appropriate the normative or standardization samples or the criterion used should be documented." (p. 43)*
Standard 14.17	*"The level of performance required for passing a credentialing test should depend on the knowledge and skills necessary for acceptable performance in the occupation or profession and should not be adjusted to regulate the number or proportion of persons passing the test." (p. 162)*

EXAMINATION SECURITY

The passing score was established at the same time the examination was administered to ensure that there was no possibility that prior knowledge of the questions would be available to the candidates.

PROCEDURES

A criterion-referenced passing scoring methodology (modified Angoff) was implemented to establish the lowest score that would qualify candidates to be minimally qualified. The ratings are based upon the question, "What percentage of minimally competent candidates would answer this item correctly?" There is emphasis on "would" versus "should" because minimum qualifications standards are based on what would actually happen rather than what a minimally competent candidate should be doing. The ratings for each item ranged from 25% (guessing) to 95% (very easy).

The major steps were:

a) Reviewing the purpose of the examination so that the SMEs in the focus group understood that the examination was designed to identify candidates who possessed the minimum qualifications to practice without harming the public health, safety, or welfare.

b) Reviewing the content specifications for the written examination so that SMEs could review the critical tasks and knowledge required of pilot trainees. The content specifications were discussed until the SMEs had a clear understanding of the purpose of the specifications and the practice analysis in the overall validation process.

c) Reviewing minimum qualifications for taking the examination to gain an understanding of what training and experience new trainees bring to the examination.

d) Reviewing the content specifications for the examination to understand the breadth and depth of the content covered in the items.

e) Identifying minimum qualifications behaviors that represented highly effective and ineffective performance so that SMEs had a common understanding of the behaviors that could be exhibited by minimally competent candidates. The SMEs were instructed to consider these behaviors during the rating process of the passing score workshop.

f) Taking the examination to assist SMEs in gaining an understanding of the difficulty of the items and the structure of the examination from the perspective of the candidates.

g) Providing an orientation to the rating process so that SMEs were basing their ratings on the minimum qualifications criteria. First, SMEs were asked to make independent ratings (Round 1) for a few items and provide their rationales for their ratings to the group. Then, SMEs were asked to consider the rationales of others in the group and make final ratings (Round 2). SMEs were not required to reveal their final ratings to the group.

h) Proceeding with the rating process for blocks of items until ratings were obtained for all items in the examination.

CALCULATION OF PASSING SCORE

The results from the passing score focus group established 105 as the lowest passing score. The passing score was calculated by summing all of the Round 2 ratings in the examination and dividing it by the number of SMEs. This method involves only one mathematical calculation and avoids rounding errors.

Candidates who achieved scores at or above 105 were identified as those who demonstrated sufficient knowledge to meet minimum qualifications criteria. Candidates who scored below 105 were identified as those who lacked sufficient knowledge for further consideration according to the minimum qualifications criteria established by the SMEs.

.

SECTION 8: COMPETENCY SCORE

PHILOSOPHY

While the passing score for the written examination thoroughly established the minimum level of knowledge required of pilot trainees, the competency score was used to identify those candidates with a greater understanding of the knowledge required for piloting vessels. The written examination was designed to array candidates along a continuum. Therefore, candidates who achieved higher scores on the written examination demonstrated greater knowledge.

The overall philosophy is to select the most competent candidates who have the potential to be employed. A list was created based on the sum of the scores from the written examination and simulator evaluation. Therefore, higher scores, represent a higher degree of knowledge. It should be noted that the use of these examinations is for selection rather than licensure purposes. In a licensing examination, every candidate who achieves a higher score than the minimum passing score is deemed to be equally qualified.

In this situation, the examinations are used as a selection procedure for pilot trainees who must be selected based upon the scores achieved in both examinations. A low score on the written examination implies that the candidate has less knowledge than a candidate with a high score. Therefore, the higher scorers on the written examination should be selected to participate in the simulator evaluation. While the sum of the written and simulator scores may be compensatory, the process of combining the scores should be considered as a conjunctive methodology implying that the high scorers have a greater degree of knowledge than low scorers.

SELECTION OF COMPETENCY SCORE

A minimum score of 119 was selected based on the probability that the candidate would be a viable trainee if selected according to rank and provisions of WAC 363-116-077 that specify that candidates who pass a written examination on or after July 1, 2008, and whose scores are among the top twenty (or such other number as may be set by the board) of those taking the written examination (plus any pilot applicants who tie a qualifying score) shall be eligible to take the simulator evaluation set forth in this section.

SECTION 9: DEVELOPMENT OF SIMULATOR EVALUATION

RATIONALE

The simulator evaluation was designed to be a fair and objective measure of candidate skills required for ship navigation in piloting waters. This includes the candidate's ability to pilot and handle ships, assimilate and prioritize data, respond appropriately, communicate well, project proper bridge presence, implement bridge resource management, and apply rules of the road. Local knowledge was not required or tested.

APPLICABLE STANDARDS

A number of the Standards apply to the development of performance examinations. The simulator evaluation developed for the pilot trainee program was standardized in terms of format and content, so that every candidate would have the same examination experience. Standardization is especially critical in performance examinations because there is a greater risk for human error in all aspects of administration and scoring.

Standard 3.15 *"When using a standardized testing format to collect structured behavior samples, the domain, test design, test specifications and materials should be documented as for any other test. Such documentation should include a clear definition of the behavior expected of the test takers, the nature of expected responses, and any materials or directions that are necessary to carry out the testing."*
(p. 46).

Standard 3.20 *"The instructions presented to test takers should contain sufficient detail so that test takers can respond to a task in the manner that the test developer intended. When appropriate, sample material, practice or sample questions…should be provided to test takers prior to the administration of the test or included in the testing material as part of the standard administration instructions." (p. 47)*

Standard 3.22 *"Procedures for scoring and, if relevant, scoring criteria should be presented by the test developer in sufficient detail and clarity to maximize the accuracy of scoring. Instructions for using rating scales or for deriving scores obtained by coding, scaling, or classifying constructed responses should be clear." (p. 47)*

Standard 3.23 *"The process for selecting, training, and qualifying scorers should be documented by the test developer. The training materials, such as the scoring rubrics and examples of test takers' responses that illustrate the levels on the score scale, and the procedures for training scorers should result in a degree of agreement among scorers that allows for the scores to be interpreted as originally intended by the test developer. Scorer reliability and potential drift over time in the raters' scoring standards should be evaluated and reported by the person(s) responsible for conducting the training session." (p. 47-48)*

Standard 5.9 *"When test scoring involves human judgment, scoring rubrics should specify criteria for scoring. Adherence to established scoring criteria should be monitored and checked regularly. (p. 65)*

The critical incident technique (Flanagan, 1954) was also used as a standard to develop the content of the scenarios and the rating scales. Specifically, the critical incident technique involved identifying behaviors that contributed to the success or failure of pilot trainees in the context of a bridge simulator environment. The criticality of the behavior could be defined in terms of what conditions led up to the situation, what observable behaviors would be effective or ineffective, or what outcomes were the results of the observable behaviors.

MEASUREMENT OPPORTUNITIES

There were a total of 30 measurement opportunities in the simulator evaluation organized according to the abilities assessed. As stated earlier, tasks selected for the simulator evaluation were derived from the tasks in the practice analysis, thus, establishing validity from the results of the practice analysis (see Appendix B). As stated earlier, the measurement opportunities were distributed across seven subject matter areas listed below.

Subject matter area	Number of measurement opportunities
I. Fundamental piloting and ship handling	8
II. Assimilates and prioritizes data to safely maneuver ship	11
III. Responds appropriately in routine situations	3
IV. Responds appropriately in non-routine or emergency situations	2
V. Communicates well and projects proper bridge presence	1
VI. Bridge resource management	3
VII. Rules of the road	2
Total	30

SCENARIOS

Licensed pilots served as SMEs. They worked with members of the Pilot Examination Committee, psychometric staff at Comira, and personnel at the Pacific Maritime Institute to develop scenarios that would be similar to situations encountered by pilot trainees. Several scenarios were developed and pretested on various members of the Pilot Examination Committee prior to pretesting the scenarios on licensed pilots who were naïve to the development process.

RATING SCALES

SMEs worked with members of the Pilot Examination Committee, psychometric staff of Comira, and personnel at the Pacific Maritime Institute to develop behaviorally anchored rating scales to evaluate candidate performance in the simulator.

There were four possible levels of candidate performance. The point values of the ratings were as follows:

- Highly Effective = + 2 points
- Effective = +1 points
- Not Effective = 0 (zero) points
- Unsafe/Unsatisfactory = (- 2 points)

Using these point values, the highest possible score on the simulator evaluation was 60 points (2 points times 30 measurement opportunities).

PRETESTING OF SCENARIOS

SMEs worked with members of the Pilot Examination Committee and psychometric staff at Comira to pretest the simulator evaluation multiple times prior to completing the development of the examination.

STANDARDIZED ADMINISTRATIVE PROTOCOLS/SCRIPTS

The following aspects of examination administration were carefully scripted to ensure that every candidate had the same examination experience. All protocols and scripts were read verbatim to the candidates.

Aspect	Purpose
General orientation, demonstration, and study materials prior to examination	• Provides overview of the ship, its course, and instruments on the bridge, e.g., rudder angle indicator, rate of turn indicator, engine order telegraph indicator, rpm gauge, as well as study materials about the ship's course, handling characteristics and other vessel particulars
Radar briefing script	• Provides demonstration of scales, electronic bearing locator, range and bearing to target on ship's radar equipment
Instructions to candidates	• Provides information regarding candidate's role as captain, available bridge team members, overall grading criteria, seven subject matter areas to be assessed in the examination
Preparation room script	• Explains purpose of the period for candidate review of a written bridge simulator scenario, relevant charts, vessel particulars and handling characteristics
Review of instruments and charts (script)	• Provides another overview of instruments on the bridge, points out specific landmarks on the chart
Watch change (script)	• Explains existing environmental conditions, vessel's course and speed, radar contacts, and other vessels that may be encountered along route
Script for computer operators in simulator	• Provides standardized language for all radio communications internal and external to ship
Mate's announcements	• Provides standardized language for mate to announce distances and vessel speeds

CANDIDATE MATERIALS FOR SIMULATOR

Prior to the orientation and demonstration, all candidates were provided copies of the chart for the route in the simulator, instructions to candidates for simulator evaluation, handling characteristics of the vessel, and ship's particulars. The candidates were instructed to study them prior to reporting for the simulator evaluation.

CANDIDATE ORIENTATION AND DEMONSTRATION OF SIMULATOR

Members of the Pilot Examination Committee, psychometric staff of Comira, and personnel at the Pacific Maritime Institute developed a standardized candidate orientation and a demonstration of the simulator that included an explanation of the instrumentation and radar monitors *during the week prior to the examination*.

The demonstration was designed to give candidates an opportunity to experience the ship traversing the actual route but without adverse environmental conditions (e.g., wind, current reduced visibility) and without other vessel traffic.

During the orientation, all candidates received full-sized charts for the route in the simulator, instructions to candidates for simulator evaluation, handling characteristics of the vessel, and ship's particulars.

ADMINISTRATIVE PROCEDURES AND PROTOCOLS

Standardized administrative procedures and protocols were developed by the Pilot Examination Committee, psychometric staff at Comira, and personnel at the Pacific Maritime Institute to administer the simulator evaluation.

The procedures and protocols specified that examination would begin with an overview of the instruments and equipment on the bridge, and review of the chart for the route in the bridge simulator. There was a review of instruments and equipment designed to refresh candidates' familiarity with the instruments and equipment on the bridge, in addition to the overview already presented during the orientation and demonstration.

Once the overview was complete, candidates would be shown the landmarks on a full-sized chart and read the watch change by one of the personnel at the Pacific Maritime Institute. Candidates were provided a written copy of the watch change that they could follow and make notes as the script was read to them.

For security purposes, the watch change document and any notes were collected from candidates at the conclusion of each examination.

TRAINING OF EVALUATORS

The Pilot Examination Committee worked with psychometric staff at Comira, and personnel at the Pacific Maritime Institute to design a training session for the licensed pilots who served on an evaluator panel. The evaluator panel for each candidate included a member of the Pilot Examination Committee, licensed pilots from the Puget Sound pilotage district, and a retired Puget Sound Pilot who was on staff at the Pacific Maritime Institute.

Evaluators were asked to review the seven subject matter areas, e.g. Rules of the Road, as well as all procedures and scripts. They were also provided several opportunities to experience the actual route of the bridge simulator with and without environmental conditions and to use the rating scales.

For training purposes, an evaluator would take the role of a candidate traversing the actual route in the bridge simulator with environmental conditions. The other

evaluators scored the evaluator-candidate independently and discussed their ratings for calibration purposes.

MONITORING OF EVALUATORS

The Pilot Examination Committee worked with psychometric staff at Comira, and personnel at the Pacific Maritime Institute to develop a protocol for monitoring the performance of evaluators during the examination. According to the protocol, each evaluator of a three-person team would review any irregularities in their ratings after each candidate to ensure that they were using the same criteria.

SECTION 10: PASSING SCORE FOR SIMULATOR EVALUATION

RATIONALE

As discussed for the written examination, the intent of a criterion-referenced methodology is to establish a passing score, or cut score, that differentiates qualified from unqualified pilot trainees. The concept underlying criterion-referenced methodologies, such as the Angoff, is that there is some level of performance, a minimum level of competence that must be achieved to pass the simulator evaluation.

There is a decision rule with a criterion that separates trainees who can demonstrate essential competencies involved in ship navigation from those who cannot. Therefore, the passing score represents the lowest possible score that candidates can achieve and still demonstrate the essential competencies that trainees must perform in the pilot training program.

APPLICABLE STANDARDS

The most applicable of the Standards is Standard 4.21 to establishing a passing score with performance examinations. It states that expert judges (here, evaluators) can establish a defensible passing score if they have a sound basis for making the judgments requested.

> "When cut scores defining pass-fail or proficiency categories are based on direct judgments about the adequacy of item or test performances or performance levels, the judgmental process should be designed so that judges can bring their knowledge and experience to bear in a reasonable way." (p. 60)

The comment following Standard 4.21 states:

> "…The procedures used to elicit judgments about test performances or performance levels (e.g., the level that would characterize a borderline examinee) should result in reasonable defensible standards that accurately reflect the judges' values and intentions. Special care must be taken to assure that judges have a sound basis for making the judgments requested. Thorough familiarity with descriptions of different proficiency categories, practice in judging task difficulty with feedback on accuracy…and other forms of information may be beneficial in helping judges to reach sound and principled decisions." (p. 60)

RATING SCALES

As mentioned earlier, the rating scales were behaviorally anchored for each of the 30 measurement opportunities and ranged from “Highly Effective” to “Unsafe/Unsatisfactory.” The “Effective” rating in each measurement opportunity was designed to reflect minimum competence. The highest score that could be achieved was 60 points if a candidate received a “Highly Effective” rating in each measurement opportunity.

Rating	Points
Highly Effective	+ 2
Effective	+ 1
Not Effective	0
Unsafe/Unsatisfactory	- 2

PROCEDURES

In this process, the evaluators worked independently to ensure that their ratings reflected their beliefs and was not influenced by the other raters. The evaluators were presented with a listing of the measurement opportunities within each of the seven subject matter areas. They were asked to identify the minimum number of “Effective” ratings that a candidate should receive within each subject matter area. The number of “Effective” ratings were summed across the seven subject matter areas and combined across evaluators in order to establish the passing score.

The evaluators were trained in the modified Angoff procedure to determine the passing score. In this procedure, the evaluators participated in a discussion of minimum qualifications to gain an understanding of the difference between minimum qualifications skills and higher level skills. They determined the number of measurement opportunities within each subject matter area that were necessary in order to be classified as minimally competent. Throughout the procedure, the evaluators used the concept of minimum qualifications in making their determinations.

It should be noted that the evaluators who established the passing score for the simulator evaluation were the same persons who served as evaluators for the simulator evaluation. They were able to observe the consequences of candidates’ actions to determine whether a particular behavior (measurement opportunity) was necessary for minimally competent performance.

PASSING SCORE

The passing score was the average of the number of measurement opportunities considered necessary for minimum competence, rounded down to eliminate

fractions of points. Here, the collective judgment of the SMEs determined the passing score to be 20 points.

DESCRIPTIVE STATISTICS

The highest score was 35 and the lowest score was 2 points. The average score for all candidates was 16 points, and the standard deviation was 8.48 points.

SECTION 11: BACKGROUND REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). Standards for Educational and Psychological Testing. Washington, DC: Author.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 508-600). Washington, DC: American Council of Education.

Chinn, R. N. (2006). Considerations in setting cut scores. Lexington, Kentucky: Council on Licensure, Enforcement, and Regulation, Resource Brief.

Chinn, R. N., & Hertz, N. R. (2006, April) Practice analysis for licensing and certification programs. Presented as a pre-conference training session, National Council on Measurement in Education, San Francisco, California, CA.

Chinn, R. N., & Hertz, N. R. (2002). Alternative approaches to standard setting for licensing and certification examinations. Applied Measurement in Education, 15, 1-14.

Flanagan, J. C. (1954). The critical incident technique. Psychological Bulletin, 51, 327-358.

Hertz, N. R., & Chinn, R. N. (2005). Practical guidelines for constructing examinations. Folsom, CA: HZ Assessments.

Hertz, N. R., & Chinn, R. N. (2002, April). The role of deliberation style in standard setting for licensing and certification examinations. Paper presented at the National Council on Measurement in Education, New Orleans, LA.

Hurtz, G., & Hertz, N. R. (1999). How many raters should be used for establishing cutoff scores with the Angoff method? A generalizability study. Educational and Psychological Measurement, 59(6), 885-897.

SECTION 12: TECHNICAL REFERENCES

33 CFR Part 110 - Anchorage Regulations (General - 110.1 only)

33 CFR Part 117 Subpart A - Drawbridge Operation Regulations (General)

33 CFR Part 157 - Tanker Regulations [Mostly this is hull and system requirements, but there is a section at 157.415 that talks about Bridge Resource Management and a reference to the navigation watch working with pilots.

33 CFR Part 160 - Ports and Waterways Safety

33 CFR Part 161 - Vessel Traffic Systems (Subparts A (except 161.12(c)) and B)

33 CFR Part 164 - Navigation Safety Regulations

33 CFR Part 165 - Regulated Navigation Areas and Limited Access Areas (Subparts A, B, C, D, E and G)

33 CFR Part 168 - Escorts Requirements for Certain Tankers [all pertinent parts except areas of enforcement]

33 CFR Part 26 - Bridge-to-Bridge Radiotelephone Regulations

33 CFR Part 334 - Danger Zones and Restricted Access Areas (334.1 to 334.6)

Adams, M. R. (2006). Shipboard Bridge Resource Management. Eastport, ME: Nor'easter Press.

Adams, Michael R. (2006). Shipboard Bridge Resource Management. East Port, Maine. Nor'Easter Press.

Chart No. 1, (2000). Nautical Chart Symbols, Abbreviations and Terms (10TH Ed). Annapolis, MD: Lighthouse Press, a division of Pro Star Publications, Inc.

Crawford, W. P. (1992). Mariner's Weather. New York, NY: W. W Norton & Co.

Hensen, H. (2003). Tug Use in Port, A Practical Guide (2nd Ed). London: Nautical Institute.

Hooyer, H. H. (1983). Behavior and Handling of Ships, Centreville, MD: Cornell Maritime Press, Inc.

Hooyer, H. H. Behavior and Handling of Ships. (1983). Centreville, MD: Cornell Maritime Press.

International Convention of the Safety of Life at Sea (SOLAS) Consolidated Edition. (2004). London: International Maritime Organization.

Light List (2007). Volume VI: Pacific Coast and Pacific Islands. Annapolis, MD: Lighthouse Press, a Division of ProStar Publications Inc.

MacElrevey, D. H. & MacElrevey, D. E. Shiphandling for the Mariner. Centreville, MD: Cornell Maritime Press, Fourth Edition, 2004. ISBN 0-87033-558-8

MacElrevey, D. H., & MacElrevey, D. E. (2004). Shiphandling for the Mariner (4th ed.). Centreville, MD: Cornell Maritime Press, Inc.

Pacific Coast Current Tables of North America and Asia. (2008). Annapolis, MD: Lighthouse Press, division of Pro Star Publications, Inc.

Publication 1310. (2001). Radar Navigation Manual and Maneuvering Board Manual (7th Ed). Bethesda, MD.: National Imagery and Mapping Agency.

Reid, G. H. (1986). Shiphandling with Tugs, Centreville, MD: Cornell Maritime Press, Inc.

Reid, G. H. (2004). Primer of Towing (3rd Ed). Centreville, MD: Cornell Maritime Press, Inc.

Revised Code of Washington, Chapter 88.16 (2008). Pilotage Act.

Rowe, R. W. (2004). The Shiphandler's Guide (2nd Ed). London, England: The Nautical Institute.

Tide Tables: West Coasts of North and South America (2008). Annapolis, MD: Lighthouse Press. A division of Pro Star Publications, Inc.

United States Coast Guard Navigation Rules: International-Inland, with all revisions and corrections through Notice to Mariners No. 22/07. Annapolis, MD: Lighthouse Press, a Division of ProStar Publications Inc.

Washington Administrative Code, Chapter 363-116 (2008).

APPENDIX A: RATING SCALES

Instructions: Prior to providing ratings for the tasks, read the description of the rating scales carefully. Your understanding of the rating scales and how each rating description applies to a task is essential to the accuracy of your ratings. You should base your ratings on your experience as a licensed pilot in Puget Sound.

CRITICALITY

How critical is effective performance of this task relative to overall job performance?

1 - Not critical

2 - Moderately critical

3 - Very critical

DIFFICULTY

How difficult is it to acquire proficiency in this task?

1 - Very easy

2 - Easier than most

3 - Average difficulty

4 - More difficult

5 - Most difficult

APPENDIX B: TASK RATINGS

Item	Task	Criticality	Difficulty	Overall Importance Σ (Crit X Diff)	Written	Simulator
	I. Prevoyage Planning					
T1	Plan transit by reviewing dispatch sheet of ship's particulars.	2.18	1.61	3.66	NT[1]	X
T2	Plan transit by reviewing current, tide and weather.	2.57	1.91	5.07	X	X
T3	Plan transit in consideration of berth side to.	2.41	1.91	4.77	X	X
T4	Evaluate dispatched tugs by considering type, capabilities, limitations and bollard pull for assigned job.	2.57	2.68	7.02	X	NT
T5	Plan route based on potential congestion, other traffic and tides/currents.	2.39	2.61	6.39	NT	X
T6	Identify ships in pilotage waters under operational control by pilots.	1.71	1.64	3.02	NT	NT
T7	Calculate underkeel clearance based upon depth of waterway and predicted height of tide.	2.89	2.66	7.70	X	X
T8	Calculate air draft clearance based upon depth of waterway and predicted height of tide.	2.84	2.64	7.61	X	NT
	II. Docking and undocking				W	S
T9	Plan placement of tugs by considering tug capabilities, weather, current, vessel characteristics, berth and maneuverability required to dock/undock ship.	2.79	3.20	9.05	X	NT
T10	Confer with tugs regarding tug position and lines, and docking/undocking plans prior to reaching/departing berth.	2.52	2.43	6.43	X	NT
T11	Establish sequence and placement of lines in making fast or letting go.	2.21	2.07	4.88	X	NT
T12	Determine safe speed by considering proximity to other vessels, underkeel clearance, weather conditions and capabilities of tugs in making approach to berth or leaving berth.	2.84	3.21	9.29	X	X
T13	Check in/out with Vessel Traffic Service (VTS) prior to departing dock or when docking is complete.	2.05	1.34	2.89	X	NT
T14	Evaluate effect of underkeel clearance on handling characteristics of ship.	2.80	3.27	9.34	X	NT
T15	Make adjustments for cushion and bank effect when docking/undocking ship.	2.68	3.38	9.21	X	NT

[1] NT means not testable. X means that the task was included as part of the content specifications.

Item	Task	Criticality	Difficulty	Overall Importance Σ (Crit X Diff)	Written	Simulator
T16	Anticipate ship's response to effect of rudder and rotation of propeller prior to ordering an astern bell.	2.70	3.05	8.32	X	X
T17	Advise master to have anchors at the ready.	2.36	1.52	3.61	NT	X
T18	Confer with master to obtain information regarding ship's handling characteristics and readiness for docking/undocking.	2.64	2.02	5.39	X	NT
T19	Assess effect of underkeel clearance and berth configuration on ship's movement when planning docking/undocking.	2.70	2.86	7.86	X	X
T20	Determine number, type of tugs and bollard pull while considering effect of current, wind, and availability of bow thrusters.	2.86	3.30	9.57	X	NT
T21	Verify when all lines are clear or made fast.	2.64	1.61	4.32	NT	NT
	III. Use of anchors				W	S
T22	Assess anchorage for potential hazards.	2.55	2.32	6.13	X	X
T23	Verify ship's position by using ranges and bearings.	2.66	2.73	7.39	X	X
T24	Establish target location in the anchorage.	2.71	2.71	7.43	NT	X
T25	Determine amount of anchor chain needed after considering wind/current conditions, depth of water, quality of holding ground and maneuvering room.	2.79	2.64	7.41	X	X
T26	Confer with master regarding procedures for dredging the anchor.	2.54	2.98	7.61	X	NT
T27	Confer with master regarding procedures for walking out or letting go with the brake.	2.77	2.68	7.48	NT	NT
T28	Confer with master regarding condition of brake and windlass.	2.66	2.55	6.75	NT	NT
T29	Account for actual position of the anchor when calculating swinging radius.	2.68	2.88	7.84	X	NT
	IV. Waterway transit				W	S
T30	Determine water depths, applied tides and underkeel clearance for transiting route.	2.77	2.59	7.25	X	NT
T31	Evaluate congestion in waterway when planning transit and establishing ship speed.	2.66	2.59	7.02	NT	NT

Item	Task	Criticality	Difficulty	Overall Importance Σ (Crit X Diff)	Written	Simulator
T32	Compensate for effects of currents and weather on maneuverability of ship.	2.89	3.63	10.64	X	X
T33	Consider air draft interactions, moored vessels, bank effect and effective use of tugs during transit.	2.84	3.39	9.84	X	NT
	V. Main ship channels				W	S
T34	Monitor ship's position in traffic lanes by using ranges and bearings from fixed objects.	2.41	2.38	5.91	NT	NT
T35	Monitor speed to prevent wake damage.	2.70	2.55	6.98	NT	NT
T36	Monitor closest point of approach with other traffic during voyage.	2.80	2.59	7.36	NT	X
T37	Follow international rules of the road (COLREGS).	2.89	2.41	6.98	X	X
T38	Monitor Vessel Traffic Service (VTS) and bridge-to-bridge frequencies.	2.66	1.98	5.34	X	NT
	VI. Shiphandling				W	S
T39	Evaluate ship's responsiveness when executing maneuvers, e.g., turns, slowing or speeding up.	2.82	3.23	9.20	X	X
T40	Assess rate-of-turn from visual or electronic means.	2.71	2.88	8.00	NT	X
T41	Verify commands have been carried out by monitoring ship's equipment.	2.95	2.32	6.84	NT	X
T42	Monitor ship's performance by using visual and other natural cues.	2.82	3.14	8.91	X	X
T43	Adjust amount of rudder applied as pivot point and speed changes.	2.66	3.05	8.30	X	NT
T44	Monitor advance and transfer.	2.75	3.36	9.36	X	X
T45	Consider effects of current on track line of ship.	2.77	3.09	8.73	NT	X
T46	Consider effects of sail area on track line of ship.	2.77	3.45	9.77	X	NT
T47	Determine effectiveness of bow thruster in relation to wind or speed of ship.	2.82	3.27	9.29	X	NT
T48	Evaluate change in pivot point location resulting from different external forces.	2.70	3.55	9.70	X	NT
T49	Evaluate effect of engine type and propulsion system on maneuverability.	2.70	3.05	8.36	X	NT

Item	Task	Criticality	Difficulty	Overall Importance Σ (Crit X Diff)	Written	Simulator
T50	Evaluate effect of trim on steering.	2.55	3.12	8.12	X	NT
T51	Evaluate effect of propeller transverse thrust going astern.	2.62	2.88	7.59	X	NT
T52	Evaluate effects of different types and locations of rudders on handling of ship.	2.59	3.05	7.98	X	NT
	VII. Navigation				W	S
T53	Confirm position from navigational instruments (e.g., gyro, fathometer, radar, GPS).	2.75	2.45	6.75	NT	X
T54	Obtain information regarding names of ships and other vessels by using Automatic Identification Systems and Vessel Traffic Service.	2.18	2.00	4.59	X	X
T55	Obtain information regarding collision avoidance by using all available means.	2.91	2.87	8.41	X	X
T56	Monitor targets to determine crossing or close quarters situations.	2.96	2.89	8.61	X	X
T57	Identify vessels by their light characteristics at night or day shapes by day.	2.50	2.36	6.04	X	X
T58	Conduct vessels in accordance with international rules of the road.	2.89	2.59	7.34	X	X
T59	Use all available means to keep vessel's position on desired track line.	2.52	2.55	6.66	NT	X
T60	Modify voyage plan based on traffic/berth conflicts or environmental conditions.	2.70	2.77	7.54	NT	X

Item	Task	Criticality	Difficulty	Overall Importance Σ (Crit X Diff)	Written	Simulator
	VIII. Pilot-Master Exchange				W	S
T61	Confer with master regarding bridge team expectations.	2.45	2.23	5.63	X	NT
T62	Confer with master regarding condition of propulsion systems, navigation systems and anchors.	2.80	2.21	6.27	X	NT
T63	Confer with master regarding voyage plan, including route and ETA.	2.45	2.02	5.07	X	NT
T64	Determine ship's maneuvering characteristics.	2.71	2.45	6.68	X	NT
T65	Review ship's particulars (pilot card).	2.64	1.87	5.04	X	NT
T66	Confer with master regarding number and types of tugs needed.	2.52	2.23	5.93	X	NT
T67	Establish contingency plan when escort tugs are utilized for controlling ship's movement in event of engine/rudder failure.	2.93	3.16	9.29	X	NT
	IX. Safety				W	S
T68	Evaluate placement and condition of pilot ladder to assess safety for embarking and disembarking.	2.68	2.55	6.93	X	NT
T69	Evaluate condition of ship to identify potential safety hazards and seaworthiness.	2.34	2.95	7.23	NT	NT
T70	Evaluate traffic conditions in boarding area prior to embarking/disembarking.	2.80	2.80	8.02	NT	NT
T71	Consider effects of environmental and traffic conditions when establishing safe speed.	2.79	2.91	8.20	X	X
T72	Anticipate conditions that could result in an error chain.	2.84	3.82	10.96	X	X
T73	Comply with required regulations regarding personal safety gear.	2.32	1.82	4.39	NT	NT
T74	Determine if environmental conditions are safe for movement of the ship.	2.91	3.59	10.52	X	NT

APPENDIX C: CONSULTANT BACKGROUND

NORMAN R. HERTZ, PH.D.
DIRECTOR OF PSYCHOMETRIC SERVICES

Dr. Hertz is a licensed psychologist with more than 25 years of experience in the measurement field. He received his Bachelor of Arts degree from Baylor University in psychology, and his Master of Science degree in psychology and his Ph.D. in industrial-organizational psychology from the University of Memphis.

He was the managing partner of HZ Assessments, a private psychometric consulting firm that he co-founded after his retirement from the California Department of Consumer Affairs in 2001. He has provided psychometric expertise to national and international organizations and has developed licensing and certification examinations for several western states including California, Washington, Oregon, and Arizona. He has extensive experience in private industry and government settings and has conducted validation studies, developed licensing and certification examinations, and established cut scores for more than 30 professions, ranging from the construction trades to medical specialties. He specializes in conducting psychometric audits of examination programs.

Prior to HZ Assessments and Comira, Dr. Hertz was the Chief of the Office of Examination Resources at the California Department of Consumer Affairs for 15 years. During his tenure at Consumer Affairs, he handled the most sensitive aspects of examination programs for more than 30 boards including expert witness testimony for legislative committees.

He has chaired and presented at the annual meetings of the Council on Licensure, Enforcement and Regulation and the National Council on Measurement in Education and has also co-authored several technical papers and journal articles. He is a member of the American Psychological Association, the Society for Industrial Organizational Psychology, the American Educational Research Association, the National Council on Measurement in Education, and the Council on Licensure, Enforcement and Regulation.

ROBERTA N. CHINN, PH.D
SENIOR PSYCHOMETRIC SPECIALIST

Dr. Roberta Chinn has more than 18 years of experience in the measurement field. She received her Bachelor of Science degree from the University of California at Davis in psychology, her Master of Arts degree from the University of the Pacific in experimental psychology, and her Ph.D. in experimental and cognitive psychology from Louisiana State University.

She was a general partner in HZ Assessments, a private psychometric consulting firm that she co-founded in 2001. Prior to HZ Assessments and Comira, Dr. Chinn was a senior psychometric consultant at the Office of Examination Resources at the California Department of Consumer Affairs for over 11 years. During her tenure at Consumer Affairs, she handled sensitive aspects of examination programs for more than 30 boards and was

instrumental in the development of standardized practical examinations, applied law and ethics examinations, and standardized oral examinations.

She has developed licensing and certification examinations for several western states including California, Washington, Oregon, and Arizona. She has extensive experience in government settings and has conducted validation studies, developed licensing and certification examinations, and/or established cut scores for numerous professions, including appraisers, dentists, dental auxiliaries, dietitians, engineers, engineering geologists, environmental site assessors, fiduciaries, hydrogeologists, pest control personnel, psychologists, ship pilots, and veterinarians. She specializes in the development of multiple-choice, performance and oral examinations and has developed innovative methods to streamline procedures for job analyses and examination development.

She has chaired and presented at the annual meetings of the Council on Licensure, Enforcement and Regulation and the National Council on Measurement in Education and has also co-authored several technical papers and journal articles. She is a member of the American Psychological Association, the American Educational Research Association, the National Council on Measurement in Education, and the Council on Licensure, Enforcement and Regulation.