

2009-2010 TECHNICAL REPORT

North American Veterinary Licensing Examination

National Board of Veterinary Medical Examiners

P.O. Box 1356
Bismarck, ND 58502
701-224-0332
www.nbvme.org

I. INTRODUCTION

The National Board of Veterinary Medical Examiners (NBVME) produces and provides the North American Veterinary Licensing Examination (NAVLE®) to assist the State Boards of Veterinary Medicine, the Canadian National Examining Board, and the individual Provincial Boards of Veterinary Medicine in assessing the practice competence of candidates for licensure as veterinarians. The primary objectives of the NAVLE are as follows:

- To provide a comprehensive objective examination to state or provincial boards charged with the licensing of veterinarians;
- To protect the public by ensuring that veterinarians demonstrate a specified level of knowledge and skills before entering veterinary practice;
- To assess the professional competency of veterinarians in terms of their qualifications to enter practice;
- To provide a common standard in the evaluation of candidates that will be comparable from jurisdiction to jurisdiction;
- To contribute to the veterinary profession through the development of improved definitions of the relationship between knowledge and professional practice; and
- To facilitate interstate/interprovincial licensing reciprocity for practicing veterinarians.

The NAVLE is administered as a computer-based examination. Beginning in November 2000, it replaced the National Board Examination (NBE: Part A) and the Clinical Competency Test (CCT: Part B) as the uniform licensing examination for veterinary medicine in North America.

The NAVLE is owned and operated by the NBVME. The NBVME has contracted with the National Board of Medical Examiners (NBME®) for testing services. This report documents the development, administration, and psychometric analysis of the 2009-2010 NAVLE.

II. TEST DEVELOPMENT

The NBVME identified 22 veterinarians to write new items and revise pool items previously identified for revision for the NAVLE; all 22 veterinarians submitted items. An item-writing workshop was conducted at the NBME office in Philadelphia on February 26, 2008 to train six new NAVLE item writers as well as the new writers for the NBVME's Qualifying Examination. The purpose of the workshop was to provide new item writers with guidelines for writing well-structured items and to hold a mock item-review meeting to demonstrate how to review items effectively. Prior to the workshop, each new committee member received an item-writing guide and was asked to write items to be reviewed at the workshop. After the workshop, NBME staff prepared item-writing assignments based on species, organ system, and veterinary practice role.

An item-writing assignment and a list of guidelines for completing assignments were sent to each item writer.

All items received from the committee members were edited and reviewed for technical flaws by NBME staff. Each author received his/her edited items for review and approval prior to inclusion in the meeting draft. A draft of edited materials was sent to the committee members for use at the item review meeting, which was held at the NBME offices on November 11-12, 2008.

At the item-review meeting, the committee was divided into five groups: two small animal groups, an equine/public health group, a bovine/ovine/caprine/porcine group, and a pet bird/other small animal/professional behavior, communication, and practice management group. A total of 911 new items were reviewed, in addition to 246 pool items. Of the new items, 836 items were approved for addition to the pool, 166 of the pool items reviewed at the meeting were kept in the pool, and the remainder were designated for deletion or future revision. A total of 96 new pictorials associated with the approved items were also approved for use. The list of item writers and reviewers is shown in Appendix A.

After the meeting, new items were updated by NBME staff and entered in the test item library. Multiple 300-item examination forms were generated by automated test assembly (ATA) using content and statistical constraints. Each form also contained 60 pretest items selected randomly by species and organ system. Twenty-eight participants, including NBVME members, representatives of the American Association of Veterinary State Boards, members of the NBVME's Examination Development Advisory Board (EDAB), and recent veterinary graduates, met on April 6-7, 2009 to review the forms. Small groups of the committee reviewed complete forms for quality and content overlap and to ensure content equivalence. Approximately 10% of the items in each form were replaced to accomplish this goal. Following the meeting, NBME staff replaced the items and created updated forms. The list of participants for the form-review meetings is shown in Appendix B.

After the forms were finalized, live and pretest items sufficient to prepare French forms of the NAVLE were proofread and sent to the Canadian National Examining Board for translation. The final step in the test development process was the creation of resource files for both the English and French versions of the examination for delivery by Prometric. Quality control procedures were implemented at each stage of the test development process to ensure that standards were being met.

III. TEST ADMINISTRATION

A. Pre Administration

Bulletin of Information: The 2009-10 NAVLE Bulletin of Information for Candidates was produced and shipped to licensing boards in the United States and the NBVME during May of 2009.

Seat Blocking: The NBME worked with Prometric to reserve seats at selected testing centers in Australia, Canada, Great Britain, Guam, Ireland, the Netherlands, New Zealand, Puerto Rico and the US Virgin Islands for both testing windows. Data from the previous year were used to forecast the number of seats required for candidates during the 2009-10 testing cycle.

Eligibility Processing: The NBVME and the NBME processed eligibilities for 3,680 candidates for the Fall 2009 testing window and 785 candidates for the Spring 2010 testing window. Special accommodations according to the ADA requirements were approved by the licensing boards and the NBVME and were processed by NBME for 42 individuals for the Fall and 23 individuals for the Spring.

B. Examination Summary

Fall 2009: Of the 3,680 eligible candidates, 3,632 examinees tested at Prometric test centers in Australia, Canada, Great Britain, Ireland, New Zealand, the Netherlands, and the United States during the Fall 2009 NAVLE administration. The majority of these examinees (3,614) took the examination during the scheduled testing window of November 16 - December 12, 2009. Eighteen examinees who were granted eligibility extensions by the NBVME tested after the official testing window closed. All examinees completed testing by December 29, 2009.

Spring 2010: Of the 785 eligible candidates, 749 examinees tested at Prometric test centers in Australia, Canada, Great Britain, Ireland, New Zealand, the Netherlands, and the United States during the Spring 2010 NAVLE administration. The majority of these examinees (744) took the examination during the published testing window of April 12 through 24, 2010. Five examinees who were granted eligibility extensions by the NBVME tested after the published testing window closed. All examinees completed testing by May 7, 2010.

Appendix C lists the number of candidates who tested at each test center and a list of the dates that candidates took the NAVLE.

C. Test Administration Issues

Proctor Reports: Prometric test center staff filed 649 Center Problem Reports (CPRs) to the NBME for the fall testing window and 149 CPRs for the spring testing window, each reporting problems experienced by examinees on test day. These reports were forwarded to NBME for review and follow-up before forwarding to the NBVME. A spreadsheet summarizing the incidents reported during the administration was forwarded to the NBVME after each administration.

Remote Procedure Call (RPC) Error: The RPC error was first encountered when NBME examinations began to be delivered using a new test driver in August 2008, and the NBME has been working with Prometric to resolve the problem since it was first discovered. Until November 2009, the error rate was approximately 1% of all NBME exam days. Beginning in November, and accelerating in December, there was a spike in the number of occurrences that triggered the error message received by NAVLE candidates. When this error occurs, the outcome file cannot be sent, and the appointment cannot be marked as complete. An additional side effect is that the examinee's end-of-session notice can't be printed at the test center. Outcome files for candidates who encountered the RPC error were sent to the NBME via a manual process. Anything that prevents an examination from launching, continuing, or ending correctly is classified as a "Fatal Error." During the fall 2009 testing window, 142 NAVLE candidates saw the "Fatal Error" message at the end of their examination. The nature of the message caused concern for many candidates that resulted in numerous calls to NBVME and request for confirmation that the candidate's records were complete. Due to the concerns over the language, the message was changed after the fall administration to read, "Attention: Please contact the proctor to shut down your machine. Your end of session notice may not print. Your exam data have been saved and will be submitted."

Workstation Issues: There were issues with two workstations at the Montreal Prometric testing center during the November-December administration that resulted in delays launching the examination for five candidates. The workstations have since been replaced with new workstations that meet the minimum hardware requirements for administering the NAVLE.

Software Load Issue: Three candidates scheduled to test at the Melbourne Prometric center on November 24 were unable to test due to a technical problem involved with loading their examination. The affected candidates were rescheduled.

D. Exit Survey Information

Examinees were asked to complete an optional post-test survey after completing the last block of examination items. Test administration statistics and selected survey results are shown in Appendix E. The full report of survey results is provided to the NBVME after each testing window.

IV. SCORING AND ANALYSIS

A. Key Validation

Prior to the administration, two items were deleted from scoring on the French forms. Key validation was carried out in January 2010, providing a final check that test items are keyed correctly and free from defects. Based on the results of an item analysis, items were selected to be reviewed by EDAB members with expertise in the appropriate content areas. After reviewing the items and their associated statistics, the experts had the opportunity to either re-key items before scoring or delete the items from scoring. An additional five items were deleted from scoring as a result of the key validation.

B. Scoring Procedures

The ATA procedures used to generate the test forms ensured that forms were as similar as possible in difficulty. However, because forms had different items, their difficulties varied to some small degree. Therefore, it was necessary to estimate examinee proficiency on a common scale using equating methods. This included a calibration of responses given by senior students of accredited schools taking the examination for the first time, in English, under standard testing conditions.

These proficiency estimates were then translated to the reported scale scores that ranged from 200 to 800. The proficiency estimate corresponding to a minimum passing score had a scale score value of 425. These three-digit scores were also translated into two-digit, locally derived scores that ranged from 0 to 99. The three-digit passing score (425) was set equal to 70 on one locally derived scale and 75 on the other.

C. Summary Statistics

Performance of Examinee Groups: Summary statistics describing the performance of candidates on the November-December 2009 and the April 2010 NAVLE administrations are shown in Tables 1.1, 1.2, and 1.3. These tables present the mean three-digit scaled scores and the standard deviations for three primary groups:

(1) Criterion Group: senior students in accredited veterinary schools who took the NAVLE for the first time under standard testing conditions;

(2) Non-Criterion Group: senior students of accredited veterinary schools who have previously taken the NAVLE or took the examination with special accommodations, and graduates of accredited schools; and

(3) Foreign Group: senior students and graduates of foreign veterinary schools that are not accredited by the American Veterinary Medical Association's Council on Education.

Test characteristics: Mean percent correct scores reflect the average percentage of items answered correctly by the candidates. They may also be interpreted as average difficulties of the examination. They are influenced by the inherent difficulty of the items as well as the proficiency of the examinees. The average difficulty (p-value) of all scored items in the test was 0.75. This means that, on average, items were answered correctly by 75% of criterion-group examinees. The average item difficulties for criterion candidates ranged from 0.72 to 0.79 across content areas.

The reliability coefficient (KR_{20}) is a measure of internal consistency that provides an estimate of the accuracy or stability of scores. Scores of an examination are reliable to the extent that administration of a different, random sample of items of the same size (number of items) and from the same content area would result in no significant change in a candidate's rank order in the group. The reliability coefficient depends, among other things, on number of the items and the homogeneity of the group. The mean reliability for the total test was 0.90 (range 0.89 to 0.91).

D. Passing Standard

Two standard-setting exercises were held in Philadelphia in January 2001. Nineteen panelists used a content representative subset of items from one of the NAVLE forms administered in November and December 2000, in a modified Angoff procedure. The results of the standard-setting exercises were presented to the NBVME on January 20, 2001. The NBVME decided on a passing standard of 425 on the reported three-digit score scale (.50 logits), and continued to apply this standard through the 2003-2004 NAVLE cycle.

On December 14, 2004, another standard setting exercise was held in Philadelphia. Twenty-two panelists reviewed one form of the 2003-2004 examination in a modified Angoff procedure. The results of the standard setting procedure were presented to the NBVME Executive Committee during a conference call in January 2005. The NBVME decided to apply a standard of 425 on the reported three-digit score scale (.64 logits) for the 2004-2005 cycle. For the 2005-2006 cycle, the standard was set at .67 logits and the scores were rescaled so they would be equal to 425. For the 2006-2007 cycle, the standard was set at .72 logits and the scores rescaled so they would be equal to 425. For the 2007-2008 cycle, the standard was set at .72 logits and the scores rescaled so that they would be equal to 425.

On July 9, 2008, another standard setting exercise was held in Philadelphia. Fourteen panelists reviewed one form of the 2007-08 examination in a modified Angoff procedure. The results of

the standard setting procedure were presented to the NBVME at its January 2009 meeting in Alexandria, Virginia. The NBVME decided to apply a standard of 425 on the reported three-digit score scale (.72 logits) for the 2008-09 cycle. Applying this standard resulted in a failure rate of 4.1% for the criterion group and 14.6% for the total group of candidates who took the 2009-10 NAVLE. Failure rates resulting from applying this standard are indicated in Tables 1.1, 1.2, and 1.3.

V. SCORE REPORTING

Performance reports for candidates were generated by NBME and distributed to licensing boards. Reports were sent to agencies in all 50 states, the U.S. Virgin Islands, Puerto Rico, and Canada. Appendix D provides samples of the following 2009-2010 NAVLE reports:

- NAVLE Candidate Performance Report: 3-Digit Score
- NAVLE Candidate Performance Report: 70 Score
- NAVLE Candidate Performance Report: 75 Score
- NAVLE Diagnostic Performance Report
- Canadian Diagnostic Performance Report

Table 1.1
Performance on November/December 2009 NAVLE by Examinee Group

	Mean Scale Score	SD Scale Score	Number of Examinees Failing	Percent of Examinees Failing	Total Examinees
Criterion Group ¹	537.2	67.6	116	3.9	2939
Non-Criterion Group ²	455.5	77.1	70	38.0	184
Foreign Group ³	453.5	73.4	169	33.2	509
Total Group	521.3	76.3	355	9.8	3632

Table 1.2
Performance on April 2010 NAVLE by Examinee Group

	Mean Scale Score	SD Scale Score	Number of Examinees Failing	Percent of Examinees Failing	Total Examinees
Criterion Group	512.1	65.7	9	7.4	122
Non-Criterion Group	443.3	63.0	78	38.4	203
Foreign Group	425.7	72.9	196	46.2	424
Total Group	444.5	75.6	283	37.8	749

Table 1.3
Performance on Both Administrations by Examinee Group

	Mean Scale Score	SD Scale Score	Number of Examinees Failing	Percent of Examinees Failing	Total Examinees
Criterion Group	536.2	67.7	125	4.1	3061
Non-Criterion Group	449.1	70.2	148	38.2	387
Foreign Group	440.9	74.4	365	39.1	933
Total Group	508.2	81.5	638	14.6	4381

¹ The criterion group consists of senior students in accredited veterinary schools who took the NAVLE for the first time under standard testing conditions.

² The non-criterion group consists of senior students in accredited veterinary schools who have previously taken the NAVLE, and candidates testing with special accommodations.

³ The foreign group consists of senior students and graduates of foreign veterinary schools that are not accredited by the American Veterinary Medical Association's Council on Education

APPENDIX A

North American Veterinary Licensing Examination Item Writers/Reviewers for the 2009-2010 Examination Cycle

Dr. Alan Corber (Canine)
Dr. Paul Coe (Bovine/Beef)
Dr. Karen Cornell (Professional Behavior, Communications, and Practice Management)
Dr. Cheryl Dhein (Canine and Feline)
Dr. Alison Eddy (Equine)
Dr. Maria Fahie (Canine)
Dr. Ned Gentz (Other Small Animal)
Dr. Nadine Hackman (Professional Behavior, Communications, and Practice Management)
Dr. Johanna Heseltine (Canine)
Dr. Amanda House (Equine)
Dr. Veronika Kiklevich (Feline)
Dr. Adam Langer (Public Health)
Dr. Isabelle Langlois (Pet Bird)
Dr. Susan Little (Feline)
Dr. Jeanne Lofstedt (Bovine/Dairy)
Dr. Fernando Marques (Equine)
Dr. Dianne Mawby (Canine and Feline)
Dr. Morgan Morrow (Swine)
Dr. Fred Scott (Feline)
Dr. Saralyn Smith-Carr (Canine)
Dr. Jeff Tyler (Bovine/Beef)
Dr. Frank Welcome (Bovine/Dairy)

APPENDIX B

**North American Veterinary Licensing Examination
April 6-7, 2009 Form Review Meeting Participants**

Dr. Jonathan Betts (AAVSB)
Dr. Linda Blythe (NBVME)
Dr. Benjamin Darien (EDAB)
Dr. Jeff Dill (Recent Graduate)
Dr. Dennis Feinberg (NBVME)
Dr. Benjamin Franklin (NBVME)
Dr. Gary Gackstetter (NBVME)
Dr. Meg Glattly (NBVME)
Dr. Jay Hedrick (NBVME)
Dr. Meghan Herron (Recent Graduate)
Dr. Francis Kallfelz (NBVME)
Dr. Norman LaFaunce (EDAB)
Dr. Sylvie Latour (NBVME)
Dr. John Lawrence (AAVSB)
Dr. Karen Lehe (EDAB)
Dr. Susan Little (EDAB)
Dr. Lila Miller (NBVME)
Dr. Patricia Provost (EDAB)
Dr. Curt Rehling (Recent Graduate)
Dr. Darcy Shaw (EDAB)
Dr. Brea Smith (Recent Graduate)
Dr. Mike Thomas (NBVME)
Dr. Rick Tubbs (NBVME)
Dr. Helen Tuzio (EDAB)
Dr. Jeff Tyler (EDAB)

APPENDIX E

Test Administration Statistics	Fall 2009	Spring 2010
Eligible examinees (permits sent)	3,680	785
Examinees tested	3,632	749
Examinees with test accommodations	42	23
Number in Criterion Group	2,939	122
Percent in Criterion Group	81%	16%
Proctor-reported issues	649	149
Examinee Responses to Post-Test Survey	Fall 2009	Spring 2010
Received first choice of test site	89%	84%
Received first choice of test date	73%	59%
Traveled more than 50 miles to test site	23%	19%
Spent one or more nights away to take examination	23%	18%
Good or excellent registration services at test site	86%	83%
No computer difficulties were encountered	76%	69%
Proctor needed during sessions	4%	5%
Would recommend test site	87%	74%
No distractions in testing area, site, or outside site	46%	49%
Not enough time for test	18%	28%
Found test “very difficult”	45%	36%