

2008-2009 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 2008-2009 NAVLE.

II. TEST DEVELOPMENT

The NBVME identified 27 veterinarians to write items for the NAVLE; 24 submitted items. An item-writing workshop was conducted at the NBME office in Philadelphia on March 1, 2007 to train 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 13-14, 2007.

At the item review meeting, the committee was divided into six groups: two small animal groups, an equine group, a bovine/ovine/caprine group, a porcine/poultry/public health group, and a small animal/pet bird/practice management group. A total of 908 new items and 273 revised pool items were reviewed in addition to 307 pool items. Of the new and revised items, 843 new items and 223 revised pool items were approved for addition to the pool, 103 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 114 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-three 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 9-10, 2008 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 2008-09 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: NBME worked with Prometric to reserve seats at selected testing centers in Australia, Canada, Great Britain, New Zealand, and Puerto Rico for both testing windows. Data from the previous year were used to forecast the number of seats required for candidates during the 2008-09 testing cycle.

Eligibility Processing: The NBVME and the NBME processed eligibilities for 3,478 candidates for the Fall 2008 testing window and 799 candidates for the Spring 2009 testing window. Special accommodations according to the ADA requirements were approved by the licensing boards and the NBVME and were processed by NBME for 47 individuals for the Fall and 11 individuals for the Spring.

B. Examination Summary

Fall 2008: Of the 3,478 eligible candidates, 3,447 examinees tested at Prometric test centers in Australia, Canada, Great Britain, New Zealand, The Netherlands, and the United States during the Fall 2008 NAVLE administration. The majority of these examinees (3,439) took the examination during the scheduled testing window of November 17 - December 13, 2008. Eight examinees who were granted eligibility extensions by the NBVME tested after the official testing window closed. All examinees completed testing by December 30, 2008.

Spring 2009: Of the 799 eligible candidates, 764 examinees tested at Prometric test centers in Australia, Canada, Great Britain, New Zealand, The Netherlands, and the United States during the Spring 2009 NAVLE administration. The majority of these examinees (761) took the examination during the published testing window of April 13 through 25, 2009. Three examinees who were granted eligibility extensions by the NBVME tested after the published testing window closed. All examinees completed testing by May 8, 2009.

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 287 Center Problem Reports (CPRs) to the NBME for the fall testing window and 119 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.

Unseen Items: One candidate for the November-December NAVLE encountered technical issues that prevented her from seeing all of the items on the examination. NBME staff provided the NBVME with score analyses for this candidate based on the items seen. A decision was made to score the examinee based on the items seen and the examinee was offered a free retest in April if the examination was not passed.

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 test window.

IV. SCORING AND ANALYSIS

A. Key Validation

Key validation was carried out in January 2009, 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 who are experts 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. No items were deleted from scoring.

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

1. Performance of Examinee Groups

Summary statistics describing the performance of candidates on the November-December 2008 and the April 2009 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.

2. 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 0.74. This means that, on average, items were answered correctly by 74% of criterion-group examinees. The average item difficulties for criterion candidates ranged from 0.70 to 0.78 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 0.87 to 0.93).

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 6.9% for the criterion group and 15.8% for the total group of candidates who took the 2008-09 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 2008-2009 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 2008 NAVLE by Examinee Group

	Mean Scale Score	SD Scale Score	Number of Examinees Failing	Percent of Examinees Failing	Total Examinees
Criterion Group ¹	522.0	67.4	190	6.6	2880
Non-Criterion Group ²	447.4	70.4	66	39.1	169
Foreign Group ³	454.3	76.2	127	31.9	398
Total Group	510.5	73.3	383	11.1	3447

Table 1.2
Performance on April 2009 NAVLE by Examinee Group

	Mean Scale Score	SD Scale Score	Number of Examinees Failing	Percent of Examinees Failing	Total Examinees
Criterion Group	492.6	61.8	18	14.0	129
Non-Criterion Group	438.4	58.3	115	41.7	276
Foreign Group	435.2	76.4	149	41.5	359
Total Group	446.0	71.0	282	36.9	764

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	520.7	67.4	208	6.9	3009
Non-Criterion Group	441.8	63.2	181	40.7	445
Foreign Group	445.2	76.8	276	36.5	757
Total Group	498.8	77.0	665	15.8	4211

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 2008-2009 Examination Cycle**

Dr. Kevin Anderson (Bovine)
Dr. John August (Feline)
Dr. R. B. Baker (Swine)
Dr. Paul Coe (Bovine)
Dr. Cheryl Dhein (Canine and Feline)
Dr. Alison Eddy (Equine)
Dr. Maria Fahie (Canine)
Dr. Robert Froehlich (Professional Behavior, Communications, and Practice Management)
Dr. Ned Gentz (Other Small Animal)
Dr. Eric Gingerich (Poultry)
Dr. Johanna Heseltine (Canine)
Dr. Veronika Kiklevich (Canine and Feline)
Dr. Adam Langer (Public Health)
Dr. Isabelle Langlois (Pet Bird)
Dr. Karen Lehe (Swine)
Dr. Susan Little (Feline)
Dr. Jeanne Lofstedt (Equine)
Dr. Dianne Mawby (Canine and Feline)
Dr. Tatjana Mirkovic (Canine)
Dr. Daryl Nydam (Bovine)
Dr. Joel Schrader (Equine)
Dr. Saralyn Smith-Carr (Canine)
Dr. Jeff Tyler (Bovine)
Dr. David Van Metre (Bovine)

APPENDIX B

**North American Veterinary Licensing Examination
April 9-10, 2008 Form Review Meeting Participants**

Dr. Linda Blythe (NBVME)
Dr. Kerry Collins (Recent Graduate)
Dr. Steve Colquhoun (AAVSB)
Dr. James Dalley (NBVME)
Dr. Benjamin Darien (EDAB)
Dr. Emilio DeBess (AAVSB)
Dr. Dennis Feinberg (NBVME)
Dr. Julie Fixman (EDAB)
Dr. Meg Glattly (NBVME)
Dr. Malia Ireland (Recent Graduate)
Dr. Francis Kallfelz (NBVME)
Dr. Sylvie Latour (NBVME)
Dr. Kimberly MacDonald (Recent Graduate)
Dr. Annette McCoy (Recent Graduate)
Dr. Patricia Provost (EDAB)
Dr. David Reeves (EDAB)
Dr. Darcy Shaw (EDAB)
Dr. Rick Tubbs (NBVME)
Dr. Helen Tuzio (EDAB)
Dr. Jeff Tyler (EDAB)
Dr. Frank Welcome (EDAB)
Dr. Jennifer White (Recent Graduate)

APPENDIX E

Test Administration Statistics	Fall 2008	Spring 2009
Eligible examinees (permits sent)	3,478	799
Examinees tested	3,447	764
Examinees with test accommodations	47	11
Number in Criterion Group	2880	129
Percent in Criterion Group	84%	17%
Proctor-reported issues	287	119
Examinee Responses to Survey	Fall 2008	Spring 2009
Received first choice of test site	93%	81%
Received first choice of test site and date	78%	62%
Traveled more than 50 miles to test site	21%	27%
Good or excellent registration services at test site	94%	90%
One or more computer difficulties were encountered	18%	17%
More than 50% of images were poor quality	14%	11%
Proctor needed during sessions	8%	17%
Would recommend test site	86%	81%
Distractions in testing area, site, or outside site	36%	38%
Not enough time for test	15%	26%
Test more difficult than expected	40%	44%