I. INTRODUCTION

The National Board of Veterinary Medical Examiners (NBVME) produces and provides the North American Veterinary Licensing Examination (NA VLE®) 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 2012-2013 NAVLE.

II. TEST DEVELOPMENT

The NBVME identified 22 veterinarians to write new items for the NAVLE; 19 veterinarians submitted items. An item-writing workshop was conducted at the NBME office in Philadelphia on February 23, 2011 to train 10 new NAVLE item writers and five new NBVME Qualifying Examination writers, as well as participants from the American College of Veterinary Emergency and Critical Care. 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. At the workshop, each new committee member received an item-writing guide and was asked to write items to be reviewed during the workshop. Following the workshop, NBME staff prepared item-writing assignments based on species, organ system, and
activity code. An item-writing assignment and a list of guidelines for completing assignments were mailed to each item writer two weeks after the meeting.

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 9-10, 2011.

At the item-review meeting, the committee was divided into four groups: two small animal groups, an equine/professional behavior, communication, and practice management group, and a bovine/porcine/ovine/caprine/public health group. A total of 734 new items were reviewed. Of the new items, 689 items were approved for addition to the pool. A total of 32 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 test construction engine using content and statistical constraints. Each form also contained 60 pretest items selected randomly by species and organ system. Twenty-four participants, including NBVME members, members of the NBVME’s Examination Development Advisory Board (EDAB), representatives of the American Association of Veterinary State Boards, and recent veterinary graduates, met on April 18-19, 2012 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 meeting 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 2012-13 NAVLE Bulletin of Information for Candidates was produced by the NBVME and a copy was sent to each NAVLE candidate who applied to take the examination through a US licensing board.

Seat Blocking: The NBME worked with the NBVME and Prometric to reserve seats at selected testing centers domestically and internationally. Data from the previous year were used to forecast the number of seats required for candidates during the 2012-13 testing cycle. Seats were reserved at centers in Australia, Brazil, Canada, Chile, China, the Dominican Republic, Egypt, Germany, Great Britain, Guam, Hawaii, Hong Kong, India, Ireland, Italy, the Netherlands, New Zealand, Philippines, Puerto Rico, Singapore, South Korea, Taiwan, Turkey, and United Arab Emirates for the November-December testing window. Seats were reserved at centers in Australia, Brazil, Canada, Chile, China, the Dominican Republic, Egypt, France, Germany, Great
Eligibility Processing: The NBVME and the NBME processed eligibilities for 4,003 candidates for the November-December 2012 testing window and for 1,005 candidates for the April 2013 testing window. Special accommodations according to the ADA requirements were approved by the licensing boards and the NBVME and were processed by NBME for 71 individuals for November-December and 32 individuals for April.

B. Examination Summary

November-December 2012: Of the 4,003 eligible candidates, 3,978 examinees tested at Prometric test centers during the November-December 2012 NAVLE administration. The majority of these examinees (3,972) took the examination during the scheduled testing window of November 12 - December 8, 2012. Six examinees who were granted eligibility extensions by the NBVME tested after the official testing window closed. All examinees completed testing by January 15, 2013.

April 2013: Of the 1,005 eligible candidates, 978 examinees tested at Prometric test centers during the April 2013 NAVLE administration. The majority of these examinees (974) took the examination during the published testing window of April 8 - 20, 2013. Four examinees who were granted eligibility extensions by the NBVME tested after the published testing window closed. All examinees completed testing by May 3, 2013.

A summary of Prometric test center information from the November-December 2012 and April 2013 NAVLE testing windows is presented in Appendix C.

C. Test Administration Issues

Proctor Reports: Prometric test center staff filed 521 Center Problem Reports (CPRs) for the November-December testing window and 171 CPRs for the April 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.

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 was provided to the NBVME in August 2013.

IV. SCORING AND ANALYSIS

A. Key Validation

Prior to the administration, four items were deleted from scoring on the French forms. Based on the results of an item analysis, live and pretest items were selected to be reviewed at the annual
January pool review. After reviewing the items and their associated statistics, content experts had the opportunity to either re-key items before scoring or delete the items from scoring.

B. Scoring Procedures

The automated test assembly 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 2012 and the April 2013 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 in accredited veterinary schools who have previously taken the NAVLE or took the examination with special accommodations, and graduates of accredited schools; and
3. Non-Accredited Group: senior students and graduates of 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 0.73. This means that, on average, items were answered correctly by 73% of criterion-group examinees. The average item difficulties for criterion candidates ranged from 0.67 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 0.92).
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. The same standard was applied to scoring the 2009-2010 and 2010-2011 examinations. Application of this standard resulted in a failure rate of 5.3% for the criterion group and 14.7% for the total group of candidates who took the 2010-2011 NAVLE.

The most recent standard setting study was undertaken on December 12, 2011. Fifteen judges participated in the study. Based on the results of the study and other information, the NBVME set a new passing standard at .83 logits that failed 7% of the November-December 2011 criterion group. Failure rates resulting from applying the standard to the 2012-2013 NAVLE administrations 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 2012-2013 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 2012 NAVLE by Examinee Group

<table>
<thead>
<tr>
<th>Examinee Group</th>
<th>Mean Scale Score</th>
<th>SD Scale Score</th>
<th>Number of Examinees Failing</th>
<th>Percent of Examinees Failing</th>
<th>Total Examinees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion Group(^1)</td>
<td>517.6</td>
<td>66.6</td>
<td>272</td>
<td>8.0</td>
<td>3410</td>
</tr>
<tr>
<td>Non-Criterion Group(^2)</td>
<td>444.9</td>
<td>69.1</td>
<td>109</td>
<td>42.7</td>
<td>255</td>
</tr>
<tr>
<td>Non-Accredited Group(^3)</td>
<td>412.8</td>
<td>67.7</td>
<td>189</td>
<td>60.4</td>
<td>313</td>
</tr>
<tr>
<td>Total Group</td>
<td>504.7</td>
<td>74.2</td>
<td>570</td>
<td>14.3</td>
<td>3978</td>
</tr>
</tbody>
</table>

### Table 1.2
Performance on April 2013 NAVLE by Examinee Group

<table>
<thead>
<tr>
<th>Examinee Group</th>
<th>Mean Scale Score</th>
<th>SD Scale Score</th>
<th>Number of Examinees Failing</th>
<th>Percent of Examinees Failing</th>
<th>Total Examinees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion Group</td>
<td>492.6</td>
<td>62.3</td>
<td>27</td>
<td>10.6</td>
<td>255</td>
</tr>
<tr>
<td>Non-Criterion Group</td>
<td>440.7</td>
<td>57.9</td>
<td>150</td>
<td>37.9</td>
<td>396</td>
</tr>
<tr>
<td>Non-Accredited Group</td>
<td>402.7</td>
<td>68.3</td>
<td>211</td>
<td>64.5</td>
<td>327</td>
</tr>
<tr>
<td>Total Group</td>
<td>441.5</td>
<td>71.5</td>
<td>388</td>
<td>39.7</td>
<td>978</td>
</tr>
</tbody>
</table>

### Table 1.3
Performance on Both Administrations by Examinee Group

<table>
<thead>
<tr>
<th>Examinee Group</th>
<th>Mean Scale Score</th>
<th>SD Scale Score</th>
<th>Number of Examinees Failing</th>
<th>Percent of Examinees Failing</th>
<th>Total Examinees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion Group</td>
<td>515.8</td>
<td>66.6</td>
<td>299</td>
<td>8.2</td>
<td>3665</td>
</tr>
<tr>
<td>Non-Criterion Group</td>
<td>442.4</td>
<td>62.5</td>
<td>259</td>
<td>39.8</td>
<td>651</td>
</tr>
<tr>
<td>Non-Accredited Group</td>
<td>407.6</td>
<td>68.1</td>
<td>400</td>
<td>62.5</td>
<td>640</td>
</tr>
<tr>
<td>Total Group</td>
<td>492.2</td>
<td>77.8</td>
<td>958</td>
<td>19.3</td>
<td>4956</td>
</tr>
</tbody>
</table>

\(^1\) 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 had previously taken the NAVLE or who took the NAVLE with test accommodations, and graduate veterinarians from accredited veterinary schools.

\(^3\) The non-accredited group consists of senior students and graduates of 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 2012-2013 Examination Cycle

Dr. Rodney Auffet (Bovine/Beef)
Dr. Dale Boyle (Public Health)
Dr. Ruthann Chun (Professional Behavior, Communications, and Practice Management)
Dr. Alan Corber (Canine)
Dr. Lais Costa (Equine)
Dr. Marion Desmarchelier (Pet Bird)
Dr. Alex Gallagher (Canine)
Dr. Tom Graham (Bovine/Dairy)
Dr. Locke Karriker (Swine)
Dr. Lauren Kleine (Equine)
Dr. Joe Klopfenstein (Bovine/Dairy)
Dr. Fernando Marqués (Equine)
Dr. Sandra Mitchell (Other Small Animal)
Dr. B. Anthony Nanton (Feline)
Dr. Elizabeth Snead (Canine)
Dr. Vicki Thayer (Feline)
Dr. David Van Metre (Small Ruminant)
Dr. Joie Watson (Professional Behavior, Communications, and Practice Management)
Dr. Drew Weigner (Feline)
APPENDIX B

North American Veterinary Licensing Examination
April 18-19, 2012 Form Review Meeting Participants

Dr. Linda Blythe (NBVME)
Dr. Dale Boyle (EDAB)
Dr. Jason Coggeshall (Recent Graduate)
Dr. Benjamin Darien (EDAB)
Dr. Dennis Feinberg (NBVME)
Ms. Joyceanne Fick (NBVME)
Dr. Julie Fixman (EDAB)
Dr. Benjamin Franklin (NBVME)
Dr. Gary Gackstetter (NBVME)
Dr. Meg Glattly (NBVME)
Dr. Jay Hedrick (NBVME)
Dr. Joanna Hughes (Recent Graduate)
Dr. Tom Kendall (AAVSB)
Dr. Norman LaFaunce (EDAB)
Dr. Karen Lehe (EDAB)
Dr. Susan Little (EDAB)
Dr. Beckey Malphus (AAVSB)
Dr. Lila Miller (NBVME)
Dr. Joan Norton (Recent Graduate)
Dr. Patricia Provost (EDAB)
Dr. Darcy Shaw (EDAB)
Dr. Charly Stansbery (Recent Graduate)
Dr. Rick Tubbs (NBVME)
Dr. Helen Tuzio (EDAB)
# Test Administration Statistics

<table>
<thead>
<tr>
<th></th>
<th>Nov-Dec 2012</th>
<th>April 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligible examinees (permits sent)</td>
<td>4,003</td>
<td>1,005</td>
</tr>
<tr>
<td>Examinees tested</td>
<td>3,978</td>
<td>978</td>
</tr>
<tr>
<td>Examinees tested outside US and Canada</td>
<td>209</td>
<td>221</td>
</tr>
<tr>
<td>Examinees with test accommodations</td>
<td>71</td>
<td>32</td>
</tr>
<tr>
<td>Number in Criterion Group</td>
<td>3,410</td>
<td>255</td>
</tr>
<tr>
<td>Percent in Criterion Group</td>
<td>86%</td>
<td>26%</td>
</tr>
<tr>
<td>Proctor-reported issues</td>
<td>521</td>
<td>171</td>
</tr>
</tbody>
</table>

# Examinee Responses to Post-Test Survey

<table>
<thead>
<tr>
<th>Response</th>
<th>Nov-Dec 2012</th>
<th>April 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Received first choice of test site</td>
<td>89%</td>
<td>82%</td>
</tr>
<tr>
<td>Received first choice of test date</td>
<td>75%</td>
<td>63%</td>
</tr>
<tr>
<td>Traveled more than 50 miles to test site</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td>Spent one or more nights away to take examination</td>
<td>20%</td>
<td>17%</td>
</tr>
<tr>
<td>Good or excellent registration services at test site</td>
<td>83%</td>
<td>84%</td>
</tr>
<tr>
<td>No computer difficulties were encountered</td>
<td>74%</td>
<td>72%</td>
</tr>
<tr>
<td>Would recommend test site</td>
<td>85%</td>
<td>76%</td>
</tr>
<tr>
<td>No distractions in testing area, site, or outside site</td>
<td>47%</td>
<td>53%</td>
</tr>
<tr>
<td>Not enough time for test</td>
<td>19%</td>
<td>26%</td>
</tr>
<tr>
<td>Found test “very difficult”</td>
<td>49%</td>
<td>45%</td>
</tr>
</tbody>
</table>