Technical Report
NBVME Qualifying Examination
September 2008, January 2009, and May 2009 Test Administrations

National Board of Veterinary Medical Examiners
P.O. Box 1356
Bismarck, ND 58502
701-224-0332
www.nbvme.org

I. Introduction

The primary objective of the NBVME's Qualifying Examination (QE) is to provide a comprehensive objective examination in basic veterinary medical sciences for use by the Program for the Assessment of Veterinary Education Equivalence (PAVE) of the American Association of Veterinary State Boards in evaluating the education equivalence of veterinarians who are graduates of veterinary schools not accredited by the Council on Education of the American Veterinary Medical Association. In addressing this objective, the QE also protects the public by ensuring that veterinarians demonstrate a specified level of knowledge and skills before entering veterinary practice, and provides a common standard in the evaluation of candidates that will be comparable from jurisdiction to jurisdiction.

II. Test Development

Qualifying Examination test development is done by the NBVME in cooperation with the National Board of Medical Examiners (NBME). The NBVME identified 14 content experts to write items for examinations to be administered on September 11, 2008, January 15, 2009, and May 14, 2009 (the 2008-2009 test cycle) (Appendix A). An item-writing workshop was conducted at the NBME offices in Philadelphia on March 1, 2007. The purpose of the workshop was to provide the new item writers with guidelines for writing well-structured items and to hold a practice item-writing and review session.

After the workshop, NBME staff prepared item-writing assignments based on each item writer's specialty and the content categories. These assignments as well as an item-writing guide and instructions for submitting items were sent to each item writer.

Thirteen of the 14 item writers submitted items. All new items received from the item writers were edited and reviewed for technical item flaws by NBME staff. The edited and annotated items were returned to the item writers for initial revision and approval. All of the newly written items and associated pictorials were reviewed by the item writers at a meeting at the NBME offices on September 25-26, 2007. At that meeting, 534 new items, 100 revised pool items, and 149 new pictorials were reviewed. A total of 506 new items, 73 revised pool items, and 142 new pictorials were approved for use.

After the meeting, the newly-approved items were reviewed again by NBME staff and added to the item pool for the QE. Three new 300-item examination forms were generated using content and statistical constraints. Eleven participants, including seven item writers for the 2008-2009 cycle and four new writers for the 2009-2010 cycle, met on February 27, 2008 to review the forms (Appendix B). Small groups of writers reviewed items within their areas of expertise, evaluating the quality of the items, identifying content overlap between items, and assessing the content equivalence of the three forms. NBME staff incorporated the committee suggestions and prepared updated forms. Final versions of the examination forms were reviewed, revised as

necessary, and approved by the Executive Director of the NBVME in April 2008.

After the forms were finalized, items were prepared for web-based presentation, and files containing item text, pictorials, and associated information were created for delivery by Internet Testing Systems, LLC. Quality control procedures were implemented at each stage of the test development process to ensure that standards were being met.

III. Test Administration

A. Examination Summary

September 11, 2008: The QE was administered on September 11, 2008 to 130 of the 133 eligible PAVE candidates at 19 test sites, including: Alabama, California, Georgia, Florida, Illinois, Indiana, Louisiana, Kansas, Minnesota, North Carolina, Oregon, Tennessee, Texas, Washington, Grand Cayman, Grenada, Guam, St. Kitts, and Sydney (Australia).

January 15, 2009: The QE was administered on January 15, 2009 to 156 of the 159 eligible PAVE candidates at 23 test sites, including: Alabama, Arizona, California, Florida, Georgia, Indiana, Illinois, Iowa, Kansas, Louisiana, Massachusetts, Minnesota, New Jersey, New York, North Carolina, Oregon, Pennsylvania, Texas, Wisconsin, Grand Cayman, Grenada, Guam, London (UK), and St. Kitts.

May 14, 2009: The QE was administered on May 14, 2009 to 148 of the 158 eligible PAVE candidates at 15 test sites including: Alabama, Georgia, Illinois, Indiana, Louisiana, Massachusetts, Michigan, New Jersey, Oklahoma, Oregon, Texas, Washington, Grand Cayman, Grenada, and St. Kitts.

B. Test Administration Incidents

Calls for Test Day Support: NBME staff members received seven calls from proctors during the September administration, five calls during the January administration, and three calls during the May administration; each reporting problems experienced at the test center during the administration. Most of the calls were regarding the Secure Browser, launching examinations for examinees with incorrect biographic data, or issuing examination restarts to examinees experiencing technical issues.

Test Center Incident Reports: Each proctor is asked to complete an incident report at the conclusion of the administration to document issues, if any, encountered by examinees at the testing center. Incident reports were forwarded to NBVME for review shortly after each examination administration.

C. Post Test Survey

Examinees were asked to complete an optional post-test survey after completing the examination. Results of the survey for each administration were provided to the NBVME.

IV. Scoring and Analysis

A. Summary Statistics

Summary statistics for all forms of the QE administered to date are provided in Table 1.

Statistics for the 2002 and 2003 administrations were based on the total group; statistics for subsequent administrations are based on the reference group. The reference group is defined as candidates taking the examination for the first time under standard conditions.

The mean P-value is an indication of the difficulty of the test, and represents the proportion of candidates who correctly answered the average item. The standard deviation represents the variability of item difficulties around the mean. P-values are influenced both by the inherent difficulty of the items and by the ability of the candidates. Because changes in mean P-value from one year to the next could reflect item difficulty, candidate ability, or both, comparisons across years have limited value and should be made with caution.

The mean discrimination index of an item is the point-biserial correlation coefficient ($r_{p\text{-}bis}$) between the item score and the total test score. It is used to indicate how well an item separates high scoring from low scoring candidates. The standard deviation of $r_{p\text{-}bis}$ represents the variation in item discriminations around the mean value.

The reliability coefficient (KR_{20}) is a measure of internal consistency that provides an estimate of the accuracy or stability of scores. An examination is reliable to the extent that administration of a different, random sample of items of the same size and from the same content area would result in little or no change in a candidate's rank order in the group. Reliability is affected by the homogeneity of the items and candidates, as well as by the length of the examination. In general, long examinations of items with similar content administered to a diverse group of candidates yield high reliabilities. Possible values of the coefficient range from 0 to 1.

Key validation takes place after the examination is administered and before scores are derived. Items that are flagged by the computer as potentially flawed or mis-keyed are reviewed by content experts, and such items are re-keyed or deleted from the scoring key, as appropriate.

B. Examinee Performance

Starting with the September 2008 administration, the QE scores were placed on a fixed reference scale. This scale was based on the performance of a Base Reference Group. This group comprised all candidates who took the QE for the first time under standard conditions beginning with the September 2005 administration through the May 2008 administration. Scores of administrations after September 2008 were equated and placed on the reference scale.

A content-based standard setting study was conducted at the NBME on July 8, 2008. After considering results of the study and other information and considerations, the NBVME set a minimum passing score (MPS) on the new equated scale of .07 logits. This MPS was translated into a reported score of 203.

Table 2 provides the history of failure rates on all forms of the QE administered to date.

C. Score Reporting

A sample score report and a sample candidate diagnostic report are included in Appendix C.

Table 1
Summary Statistics

Administration	N	Number of Items Scored (Deleted)	Mean P-Value (Standard Deviation)	Mean Discrimination Index: $r_{p\text{-}bis}$ (Standard Deviation)	KR_{20} Reliability Coefficient
August 2002[1]	33	290 (10)	.60 (.26)	.13 (.19)	.84
January 2003[1]	36	287 (13)	.55 (.25)	.10 (.18)	.81
August 2003[1] Form 1	11	292 (8)	.59 (.24)	.14 (.32)	.87
August 2003[1] Form 2	7	297 (3)	.59 (.26)	.17 (.36)	.91
January 2004	29	297 (3)	.59 (.23)	.22 (.21)	.93
August 2004	116	286 (14)	.61 (.21)	.21 (.13)	.92
January 2005	49	282 (18)	.64 (.20)	.19 (.16)	.90
May 2005	49	277 (23)	.62 (.22)	.19 (.17)	.92
September 2005	125	272 (28)	.60 (.21)	.17 (.14)	.90
January 2006	65	279 (21)	.60 (.21)	.17 (.14)	.90
May 2006	75	283 (17)	.60 (.22)	.19 (.15)	.92
September 2006	77	278 (22)	.56 (.21)	.17 (.15)	.90
January 2007	56	277 (23)	.60 (.21)	.17 (.15)	.90
May 2007	87	276 (24)	.60 (.22)	.18 (.13)	.91
September 2007	105	288 (12)	.58 (.18)	.20 (.13)	.93
January 2008	114	285 (15)	.58 (.19)	.21 (.14)	.93
May 2008	84	284 (16)	.60 (.22)	.15 (.12)	.88
September 2008	87	290 (10)	.59 (.19)	.22 (.13)	.94
January 2009	119	294 (6)	.61 (.18)	.20 (.12)	.93
May 2009	109	288 (12)	.59 (.20)	.20 (.14)	.93

[1] Summary statistics are based on the total group of candidates. All others are based on the reference group (candidates taking the examination for the first time).

Table 2
History of Failure Rates

	Total Group		Reference Group	
Administration	**N**	**Failure Rate**	**N**	**Failure Rate**
August 2002	5/33	15.2%	5/33	15.2%
January 2003	11/36	30.6%	9/31	29.0%
August 2003 Form 1	1/11	9.1%	1/11	9.1%
August 2003 Form 2	1/7	14.3%	1/7	14.3%
January 2004	9/30	27.6%	8/29	27.6%
August 2004	29/123	23.6%	26/116	22.4%
January 2005	18/75	24.0%	5/49	10.2%
May 2005	9/57	15.8%	7/49	14.3%
September 2005	29/135	21.5%	26/125	20.8%
January 2006	21/85	24.7%	13/65	20.0%
May 2006	19/79	24.1%	15/75	20.0%
September 2006	25/90	27.8%	16/77	20.8%
January 2007	19/65	29.2%	13/56	23.2%
May 2007	38/100	38.0%	27/87	31.0%
September 2007	49/129	38.0%	35/105	33.3%
January 2008	52/148	35.1%	37/114	32.5%
May 2008	45/117	38.5%	19/84	22.6%
September 2008	41/124	33.1%	25/87	28.7%
January 2009	57/146	39.0%	36/119	30.3%
May 2009	43/154	27.9%	23/109	21.1%

Appendix A
2007 Qualifying Examination Item Writers

Dr. Thomas Caceci, Histology
Virginia-Maryland Regional College of Veterinary Medicine, Blacksburg, VA

Dr. Paul Davenport, Physiology
University of Florida College of Veterinary Medicine, Gainesville, FL

Dr. Anton Hoffman, Anatomy
Texas A&M University College of Veterinary Medicine, College Station, TX

Dr. Gayle Johnson, Pathology
University of Missouri College of Veterinary Medicine, Columbia, MO

Dr. Ron Johnson, Pharmacology
Ontario Veterinary College, University of Guelph, Guelph, ON

Dr. Lynne Kushner, Pharmacology
University of Pennsylvania School of Veterinary Medicine, Philadelphia, PA

Dr. Susan Little, Parasitology
Oklahoma State University College of Veterinary Medicine, Stillwater, OK

Dr. Abdelfattah Nour, Physiology
Purdue University School of Veterinary Medicine, West Lafayette, IN

Dr. Jon Patterson, Pathology
Michigan State University College of Veterinary Medicine, East Lansing, MI

Dr. Tom Phillips, Virology
Western University College of Veterinary Medicine, Pomona, CA

Dr. Marc Ratzlaff, Anatomy
Washington State University College of Veterinary Medicine, Pullman, WA

Dr. Beth Spangler, Clinical Pathology
Auburn University College of Veterinary Medicine, Auburn, AL

Dr. Jean Whichard, Bacteriology
U.S. Centers for Disease Control and Prevention, Atlanta, GA

Appendix B
2008 Qualifying Examination Form Reviewers

Dr. Kevin Anderson, Anatomy
University of Florida College of Veterinary Medicine, Gainesville, FL

Dr. Robert (Pete) Bill, Pharmacology
Purdue University School of Veterinary Medicine, West Lafayette, IN

Dr. Paul Davenport, Physiology
University of Florida College of Veterinary Medicine, Gainesville, FL

Dr. Ron Green, Radiology
Animal Radiology Clinic, Dallas,TX

Dr. Mary Hondalus, Bacteriology
University of Georgia College of Veterinary Medicine, Athens, GA

Dr. Ron Johnson, Physiology
Ontario Veterinary College, University of Guelph, Guelph, ON

Dr. Sanjay Kapil, Virology and Immunology
Oklahoma State University College of Veterinary Medicine, Stillwater, OK

Dr. Susan Little, Parasitology
Oklahoma State University College of Veterinary Medicine, Stillwater, OK

Dr. Abdelfattah Nour, Physiology
Purdue University School of Veterinary Medicine, West Lafayette, IN

Dr. Jon Patterson, Pathology
Michigan State University College of Veterinary Medicine, East Lansing, MI

Dr. Beth Spangler, Clinical Pathology
Auburn University College of Veterinary Medicine, Auburn, AL