

# Judging Competency

A study of in-training evaluation of veterinary students

A thesis presented in partial fulfilment of the requirements  
for the degree of

Doctor of Education

at Massey University, Manawatū, New Zealand

Elizabeth J Norman

2016

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

# Abstract

---

In-training evaluations are a common but highly criticised method of assessing the competency of veterinary students completing training. They involve assessment of on-going performance in the workplace, performed by the supervisor. They are highly feasible and one of the few ways that a student's performance in an authentic context can be evaluated. Psychometric research has suggested, however, that in-training evaluations are unreliable, do not discriminate aspects of performance, and do not predict performance on other assessments, casting doubt on the credibility of scores. Research on rater judgement processes suggests, in contrast, that multiple aspects are discriminated and that accounting for context and inferred reasons for behaviour contributes to rater variability. Very little research has considered in-training evaluation in a veterinary context.

In a mixed method study this research investigated how well the in-training evaluation used during clinical placements in one veterinary school captured the aspects of student performance it was designed to capture. It explored the supervisor's view of student performance, and how that related to the dimensions being assessed in in-training evaluation, and to the constructs of competency articulated in frameworks. Complementary research strands involved analysis of semi-structured interviews with supervisors, common factor analysis of in-training evaluation scores, ordinal logistic regression relating factors to overall judgement, and thematic comparisons of findings with competency frameworks.

Together, the nature of what supervisors considered, the dimensional structure of scores, and the relationship of dimensions with the overall judgement suggested that the in-training evaluation is both holistic and discriminating, and that important aspects of performance are student engagement and trustworthiness. The aspects captured by the evaluation aligned well with the design of the instrument, and generally well with the veterinary competency frameworks. However, some areas were highlighted where concepts of veterinary competency and the competencies required in different subdisciplines need further consideration by the profession. The findings give insights into the process of judgement of competency by veterinary supervisors that will inform further research. They support some aspects of a validity argument in relation to scoring processes, and inform the design of evaluation instruments by underscoring the construct-relevance of interrelated dimensions.



# Acknowledgements

---

Firstly, I would like to thank my supervisors, Dr Peter Rawlins and Dr Linda Leach. You have been a tremendous support and source of encouragement from the beginning. Thank you for being so available, for listening, and for your wise and gentle guidance. I have always enjoyed discussing things with you, and hope that may continue.

I would also like to thank those who have encouraged me on this research journey in one way or another. In particular, Professor Mark Brown, who first suggested I should do an EdD. You were so right! Thank you for your mentorship and support. Thanks also to Professor Tim Parkinson, Dr Marg Gilling, and Dr Dianne Gardner who encouraged me to research in education, and Dr Jenny Poskitt, who believed in me enough to let me move from science to the EdD programme.

A big note of thanks goes to all the interview participants who enthusiastically gave up their time and shared their thoughts with me. Without you, the research could not have happened, and I am tremendously grateful.

Thanks also to all those who helped me with the research in various ways: Andrew Rowatt who anonymised data for me so carefully; Simon Verschaffelt who set up computers for me; Sue Leathwick and Georgie Cowley, who gathered up data for me and answered my hundreds of questions; Professor Cord Heuer who helped me with SAS code while I was still getting my head around GLIMMIX. Thanks also to all those workmates who also helped with this by taking on work themselves or waiting patiently for things to be done. You all helped me have time to do this research.

Thank you also to Massey University for the financial support that accompanied the Vice Chancellor's Award for Sustained Commitment to Teaching Excellence that I received in 2012. This has covered all the costs of the research and enabled me share it at conferences.

To my walking buddies, breakfast buddies, and other friends: Naomi, Wendi, Jenny, Eloise, Kirsty, Mandy, Bridget, Johanna, Rose, Elizabeth, and Susan. You have all, at various times, listened with interest while I told you about my research. It has been so helpful to have people to tell, and I thank you for your friendship and support.

To my husband Richard, thank you for your unwavering love and support, and for listening and discussing things with me. To my children, Catie and Rob, thank you for your patience and understanding with a mother welded to a computer, and for still managing to growing up into such fine people. And especial thanks to Catie for your help with grammar when I was stuck.

Finally, I would like to dedicate this thesis to my late mothers, Marianne and Verity; I know you would both be proud.

# Contents

---

<b>Chapter 1: Introduction</b> .....	<b>1</b>
Research questions .....	4
Structure of the thesis .....	4
Terminology .....	5
In-training evaluations .....	6
Student, supervisor, and rater .....	6
Competency .....	6
Discipline-specific and non-discipline-specific aspects of competency .....	7
Workplace-based assessment .....	8
Summary .....	8
<b>Chapter 2: Literature review</b> .....	<b>11</b>
Veterinary competency.....	11
The importance of veterinary competency .....	11
Defining competency .....	13
Evaluating competency in the context of the workplace .....	19
In-training evaluation .....	20
Other workplace-based assessment instruments .....	22
Rating scales for workplace-based assessment instruments .....	23
Summary of the background to veterinary competency.....	26
Criticisms of the in-training evaluation.....	26
The reliability of in-training evaluation scores .....	27
Leniency-stringency .....	29
Halo error and dimensionality of in-training evaluations.....	31
How in-training evaluation scores relate to other measures of performance.....	36
Summary of problems with the in-training evaluation .....	37
Elucidating the process of evaluation .....	38
Expectations of student performance .....	40
Forming a picture of student performance .....	42
Comparing expectations and pictures to make an evaluation .....	45
Effect of cognitive load on making an evaluation .....	48

Translating a holistic narrative evaluation to a numerical score .....	51
Summary of the process of evaluation .....	55
The in-training evaluation in veterinary medicine .....	56
Evidence of leniency.....	57
Internal structure .....	57
Relationship with other assessments.....	58
Relationship of in-training evaluation scores with conceptions of and approaches to practice.....	59
Student and supervisor perspectives on the in-training evaluation .....	60
Comparison with self-assessment.....	61
Summary of research on the veterinary in-training evaluation.....	61
Conclusion .....	62
<b>Chapter 3: Research design .....</b>	<b>63</b>
Research perspective .....	63
The mixed method research design .....	66
Validity of the research .....	69
Ethical considerations.....	71
Setting for the research.....	72
Summary.....	74
<b>Chapter 4: What supervisors value.....</b>	<b>75</b>
Research method.....	75
Interviews .....	75
Participants.....	77
Thematic analysis procedures.....	78
Findings.....	79
Themes described by supervisors .....	79
Frequency of student strengths and weaknesses.....	82
Use of themes by supervisors .....	83
Engagement .....	83
Trustworthiness .....	89
Discipline-specific knowledge and skills.....	91
Relating to others.....	97
Personal functioning .....	100



---

Caring for animals.....	101
Other aspects.....	102
Interrelatedness of themes .....	107
Differences between what was important in excellent and weak students .....	108
Supervisor use of observation, explanation, and inferences .....	109
Strengths and limitations of Phase 1 .....	110
Summary .....	112
<b>Chapter 5: Dimensionality of in-training evaluation .....</b>	<b>117</b>
Research method .....	117
Nature of the in-training evaluation instrument.....	117
Sampling and data preparation .....	119
Missingness analysis and management.....	120
Justification for common factor analysis over other techniques .....	121
Factor analysis method.....	122
Higher-order factor analysis method .....	124
Method for construction of three-dimensional graphs.....	125
Findings .....	125
Sample .....	125
Factor analysis .....	126
Suitability of the matrix for factoring .....	126
Extraction and rotation of factors .....	127
Factor solution.....	130
Higher-order factor analysis .....	131
Three-dimensional correlation matrices .....	132
Interpretation of the number of factors.....	137
Strengths and limitations of Phase 2 .....	140
Strengths and limitations of the factor analysis methods.....	140
Comparison of results when assumptions were violated .....	141
Other limitations.....	144
Sampling .....	144
Method effects .....	145
Summary .....	146

<b>Chapter 6: Relationship of factors and overall grade .....</b>	<b>147</b>
Research method.....	147
Sample.....	147
Deriving factor scores.....	147
Missingness analysis and management .....	148
Ordinal logistic regression method .....	148
Findings.....	150
Distribution of overall grade .....	150
Results of model selection procedures.....	151
Regression findings .....	151
Illustrating the findings .....	155
Strengths and limitations of Phase 3.....	162
Summary.....	164
<b>Chapter 7: Alignment of practice with intentions .....</b>	<b>165</b>
Research method.....	165
Findings.....	167
Themes in the mark scheme .....	167
Overall.....	167
By dimension (factor).....	168
Themes in the two competency frameworks: the BVSc learning outcomes and the VCNZ standards.....	172
Strengths and limitations of Phase 4.....	173
Interpreting the findings.....	175
Summary.....	177
<b>Chapter 8: General discussion .....</b>	<b>179</b>
Supervisor judgement as holistic and discriminating.....	179
The importance of engagement and trustworthiness .....	184
Alignment of what supervisors value with the assessment criteria and competency frameworks .....	189
General alignment.....	189
Balance of nontechnical and discipline-specific aspects.....	190
Alignment of standards .....	192
Summary .....	193

Issues for the meaning of veterinary competency and its assessment .....	193
Personal functioning .....	194
Impact on the supervisor .....	194
Prospects of the student .....	195
Caring for animals .....	196
Summary .....	197
The influence of veterinary subdiscipline on in-training evaluation .....	197
Insights into the process of judgement .....	200
Determining the factor structure .....	203
Summary .....	206
<b>Chapter 9: Conclusion .....</b>	<b>207</b>
Summary of the research findings .....	207
Research contribution .....	209
Limitations of the research .....	211
Future research directions .....	212
Implications and recommendations .....	213
Implications for validity of interpretations of in-training evaluation scores .....	214
Recommendations for future development and use of in-training evaluation .....	215
Final thoughts .....	218
<b>References .....</b>	<b>219</b>

## Appendices

<b>Appendix A: Interview protocol .....</b>	<b>251</b>
<b>Appendix B: In-training evaluation instrument .....</b>	<b>253</b>
<b>Appendix C: Missingness analysis .....</b>	<b>257</b>
The problem of missing data .....	257
The problem of unbalanced data .....	261
Conclusion .....	262

Missingness analysis method .....	262
Missingness analysis results .....	263
Evaluation-level missingness.....	264
All-item missingness.....	264
Some-item missingness.....	265
Relationship of item value with other variables .....	266
Missingness arising from the different numbers of evaluations performed on each student.....	267
Missingness in factor scores.....	268
Implications of missingness for analysis.....	269
Factor analysis.....	269
Evaluation-level and all-item missingness .....	269
Some-item missingness.....	270
Unbalanced data .....	270
Regression analysis.....	270
Evaluation-level and all-item missingness .....	271
Some-item missingness.....	271
Unbalanced data .....	272
Conclusion regarding the effect of missingness on this research .....	273
<b>Appendix D: Details of factor analysis method .....</b>	<b>275</b>
Suitability of the data for common factor analysis .....	275
Preparation of the adjusted correlation matrix for factor analysis .....	277
Suitability of sample size for factoring .....	278
Suitability of the matrix for factoring.....	279
Extraction and rotation of factors.....	279
Determining the number of factors to retain .....	280
Eigenvalues greater than one rule .....	282
Scree test.....	283
Parallel analysis .....	283
Velicer’s minimum average partial (MAP) .....	285
Measures of model fit.....	285
Variance explained.....	288
Interpretation of the importance of the factors.....	288
<b>Appendix E: Additional results for factor analysis.....</b>	<b>291</b>
Additional factor matrices.....	292

<b>Appendix F: Additional details for regression analysis</b> .....	<b>295</b>
Additional details for regression method .....	295
Additional results from the regression analysis .....	296
<b>Appendix G: Ethics Committee letter of approval</b> .....	<b>299</b>
<b>Appendix H: Request for access – Director, Student Management, Massey University</b> .....	<b>300</b>
<b>Appendix I: Request to conduct research – Pro-Vice Chancellor, College of Sciences</b> .....	<b>302</b>
<b>Appendix J: Request to conduct research – Head of Institute, IVABS</b> .....	<b>304</b>
<b>Appendix K: Information sheet for potential participants</b> .....	<b>306</b>
<b>Appendix L: Consent form for interview participants</b> .....	<b>308</b>
<b>Appendix M: Transcriber confidentiality agreement</b> .....	<b>309</b>
<b>Appendix N: Authority for release of transcripts</b> .....	<b>310</b>

## List of tables

Table 2.1: Taxonomy of veterinary competencies.....	15
Table 2.2: Details of studies investigating the internal structure of in-training evaluations. ....	33
Table 4.1: Themes arising from supervisors’ descriptions of excellent, weak, and marginal veterinary students, and their definitions.....	80
Table 4.2: Number of supervisors using each theme in their initial spontaneous descriptions or within all (both spontaneous and elaborated) descriptions of any student.....	81
Table 4.3: Presence of positive, negative, and mixed themes in descriptions of excellent, weak, and marginal students: spontaneous descriptions. ....	83
Table 4.4: Presence of positive, negative, and mixed themes in descriptions of excellent, weak, and marginal students: spontaneous and elaborated descriptions. ....	83
Table 5.1: Domains and items on the in-training evaluation.....	118
Table 5.2: Distribution of scores awarded for each item in 3215 evaluations in which at least some items were scored. ....	126
Table 5.3: Number of factors suggested by each method used to determine the number of factors to retain.....	127

Table 5.4: Factor pattern coefficients produced by extraction of 1-, 2-, 3-, and 4-factor solutions on the adjusted Pearson correlation matrix, using maximum likelihood estimation and direct oblimin rotation.....	130
Table 5.5: Interfactor correlations produced by extraction of 2-, 3-, and 4-factor solutions on the adjusted Pearson correlation matrix, using maximum likelihood estimation and direct oblimin rotation.....	131
Table 5.6: Factor pattern coefficients produced by extraction of one higher-order factor on the intercorrelation matrix from the 4-factor solution, using maximum likelihood estimation. ....	131
Table 6.1: Distribution of overall grades across all evaluations. ....	151
Table 6.2: Fixed effects of the independent variables on overall grade. ....	153
Table 6.3: Effect of individual factor scores to raise overall grade for four representative placements – changes in probabilities ( $p$ ).....	156
Table 6.4: Effect of individual factor scores to lower overall grade for four representative placements – changes in probabilities ( $p$ ).....	157
Table 6.5: Effect of individual factor scores to raise overall grade for four representative placements – changes in relative risk (RR).....	160
Table 6.6: Effect of individual factor scores to lower overall grade for four representative placements – changes in relative risk (RR).....	161
Table 7.1: Side by side comparison of the item descriptors for the in-training evaluation and the themes from the interviews represented in each dimension.....	170

## List of tables in appendices

Table C.1: Definitions of types of missingness.....	259
Table C.2: Proportion of missingness in evaluations, overall grades, and item scores. ....	264
Table C.3: Item-level missingness in 3215 evaluations that had some items scored.....	265
Table C.4: Percentage of evaluations with missingness in factor scores.....	268
Table C.5: Complete cases remaining when evaluations with any missing factor scores were deleted.....	269
Table D.1: Expected performance of various procedures used to determine the number of factors to retain .....	281
Table E.1: Adjusted correlation matrix used for the main factor analysis .....	291

Table E.2: Initial and final communality estimates for each item in the in-training evaluation.....	291
Table E.3: Factor structure coefficients produced by extraction of 1-, 2-, 3-, and 4-factor solutions on the adjusted Pearson correlation matrix, using maximum likelihood estimation and direct oblimin rotation. ....	292
Table E.4: Reference structure correlations produced by extraction of 1-, 2-, 3-, and 4-factor solutions on the adjusted Pearson correlation matrix, using maximum likelihood estimation and direct oblimin rotation. ....	293
Table F.1: Distribution of overall grade according to covariate grouping .....	296
Table F.2: Parameter estimates and their standard errors and 95% confidence intervals for the main model.....	297
Table F.3: Parameter estimates and their standard errors and 95% confidence intervals for the regression model .....	297

## List of figures

Figure 2.1: Concept map for the process of judgement in in-training evaluation .....	39
Figure 3.1: Sequence and relationship of qualitative and quantitative research phases designed to investigate how well the scores on in-training evaluation capture the aspects of student performance the instrument is intended to assess.....	67
Figure 4.1: Number of supervisors using each theme in their initial spontaneous descriptions of excellent, marginal, and weak veterinary students .....	81
Figure 4.2: Frequency of positive, negative, and mixed depictions of themes occurring in all descriptions (both spontaneous and elaborated) of excellent, marginal, and weak students.....	93
Figure 5.1: Graphs depicting the results of procedures to determine the number of factors to retain.....	128
Figure 5.2: Three-dimensional graph of the adjusted correlation matrix.....	133
Figure 5.3: Comparison of the adjusted correlation matrix displayed in three dimensions from above and in a table form .....	134
Figure 5.4: Three-dimensional graphs of partial correlations remaining after extraction of factors .....	135
Figure 5.5: Comparison of the factor structure with the adjusted correlation matrix displayed in three dimensions from above .....	136
Figure 6.1: Distribution of standardised factor scores for each overall grade.....	153

Figure 6.2: Distribution of the overall grades within each placement showing the variation between placements in the proportion of each level of overall grade awarded ..... 154

Figure 6.3: Distribution of GPA for each overall grade ..... 154

Figure 6.4: Distribution of the overall grades within placements based at the University (academic) and placements based in external veterinary practices (non-academic)..... 155

Figure 6.5: The effect of each factor score on raising (A) or lowering (B) overall grade for four representative placements..... 158

Figure 7.1: Proportion of words devoted to each theme in the in-training evaluation (ITE) mark scheme, the spontaneous descriptions by supervisors in the interviews and all (both spontaneous and clarified and prompted) descriptions in the interviews ..... 167

Figure 7.2: Proportion of words in the mark scheme represented by each theme, for each item in the in-training evaluation ..... 169

Figure 7.3: Proportion of words in the mark scheme represented by each theme, for each factor in the in-training evaluation ..... 169

Figure 7.4: Proportion of words devoted to each theme in the VCNZ standards, BVSc learning outcomes, spontaneous descriptions by supervisors in the interviews and all (both spontaneous and clarified and prompted) descriptions in the interviews..... 173

Figure C.1: Frequency distribution of the number of evaluations for each student ..... 267

## List of boxes

Box 1: Three vignettes of students who were primarily interested in large (farm) animal work, during their time on a small animal (dog and cat) placement ..... 88

Box D.1: SAS code for preparation of the adjusted correlation matrix ..... 278

Box F.1: SAS code for generalised linear mixed modelling of overall grade with various independent variables..... 295