

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.

REFERENCE REPORTS:

**A META-ANALYTIC REVIEW OF PREDICTIVE
VALIDITY AND AN EXPERIMENTAL STUDY OF
RATING ACCURACY**

A thesis presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
in Psychology at Massey University

KARL BRANDON PAJO

1996

**Massey University Library
Thesis Copyright Form**

Title of Thesis:

(1) (a) I give permission for my thesis to be made available to readers in Massey University Library under conditions determined by the Librarian.

~~(b) I do not wish my thesis to be made available to readers without my written consent for ... months.~~

(2) (a) I agree that my thesis, or a copy, may be sent to another institution under conditions determined by the Librarian.

~~(b) I do not wish my thesis, or a copy, to be sent to another institution without my written consent for ... months.~~

(3) (a) I agree that my thesis may be copied for Library use.

~~(b) I do not wish my thesis to be copied for Library use for ... months.~~

Signed



Date

31/5/96

The copyright of this thesis belongs to the author. Readers must sign their name in the space below to show that they recognise this. They are asked to add their permanent address.

NAME AND ADDRESS

DATE

Abstract

Reference reports are a commonly used selection method in New Zealand and overseas. Although popular with practitioners, they have attracted little attention from researchers. To ascertain the predictive validity of reference reports a meta-analytic review was conducted. Results of the preliminary analysis provided a bare-bones estimated mean validity of 0.15, and a fully corrected estimated mean validity of 0.32. Substantial variance remained unaccounted for following corrections for sampling error. Subsequent moderated meta-analyses, based on degree of structure and psychometric soundness of the reference reports, was found to account for much of the variation in observed validity coefficients. Highly structured reports were found to be consistently superior to unstructured reports. Improvements in the validity of highly structured reports can be attributed to the control of leniency in ratings. However, no studies to date have evaluated the accuracy of referees' ratings. Drawing on the performance rating literature, an experimental study examining the influence of scale format, ratee characteristics, rating purpose, and rater affect on the accuracy of ratings was implemented. Asymmetrical, positively toned scales were found to reduce leniency in ratings compared to unstructured and Likert-type rating forms. Raters who expressed liking for the ratee were more lenient in their ratings compared to raters who expressed neutral or antagonistic feelings toward the ratee. No significant effects for rating purpose and ratee characteristics were apparent. Overall, the investigation points to deficiencies in the way reference reports are presently employed, and highlights the need for a more rigorous approach in their development and application.

Acknowledgments

I am indebted to my supervisors, John Podd and Mike Smith, for the support and assistance they offered during the research and writing of this dissertation. The common sense and sound advice they have provided has been much appreciated. I would like to express my special thanks to John, who willingly took on the major supervisory duties when Mike left Massey University to take up another position.

I am very grateful for the encouragement and aid forthcoming from my colleagues in the Departments of Psychology, and Human Resource Management. Not only are they a great bunch of people to work with, but they have proved to be good friends as well. My thanks also go to John Spicer who helped clarify some of the statistical issues that arose throughout the course of the research.

I would also like to express my gratitude to my friends and my family who have been a supportive and understanding crew, especially on those occasions when I couldn't "come out to play."

Finally, it is with deep appreciation that I wish to acknowledge the contribution of my partner, Jacqui. I am very grateful for her constant encouragement, patience, and support during the long nights and weekends when she would have much preferred to see me at home. Thank you.

TABLE OF CONTENTS

	Page
Abstract	ii
Acknowledgements	iii
Chapter 1	
Overview	1
Chapter 2	
Literature Review: Reference Reports	5
Surveys of Use	6
Content of Reference Reports	11
Psychometric Issues	15
<i>Reliability</i>	15
<i>Validity</i>	20
<i>Source of the Recommendation</i>	20
<i>Target Population</i>	22
<i>Format and Content</i>	23
<i>Miscellaneous Threats to Validity</i>	27
<i>Meta-Analytic Studies</i>	31
The Meta-Analysis	32
Chapter 3	
Meta-Analysis	34
Method	34
Literature Search	34
Decision Rules for Coding Studies into the Meta-Analysis	35

Meta-Analysis Method Used	39
Results	43
Post hoc Analyses	45
Discussion	51

Chapter 4

Literature Review: Performance Rating	60
Rater Training	60
Cognitive Processes in Performance Rating	64
Rating Scale Format.....	65
<i>Summary</i>	69
Rater Acquaintance and Affect	71
<i>Summary</i>	79
Ratee Characteristics.....	79
<i>Race</i>	80
<i>Sex</i>	82
<i>Summary</i>	84
Purpose of Rating	86
<i>Rater Motivation</i>	92
<i>Summary</i>	95
Rater Accuracy	96

Chapter 5

Experimental Study	99
---------------------------------	----

Chapter 6

Method	103
Procedural Overview	103
Participants.....	103
Procedure.....	104

Experimental Methods	105
<i>Vignette</i>	105
<i>Likability Ratings</i>	108
<i>Demographic Questionnaire</i>	108
<i>Rating Forms</i>	109
<i>Leniency Scale</i>	110
True Scores	112
Dependent Variables.....	114
<i>Accuracy Measures</i>	114
<i>Error Measures</i>	116
Data Analysis	117
<i>Rescaling of Ratings and Measures</i>	117
<i>Analyses</i>	119

Chapter 7

Results	121
Demographic Questionnaire	121
Tests of Hypotheses	127
Accuracy Measures.....	131
Error Measures.....	134
Comparisons Between Measures	136
Regression Analysis of Error Scores.....	138
Prediction of Leniency	140

Chapter 8

Discussion	147
The Rating Instrument	148
Rater Affect.....	151
Rating Purpose	157
Ratee Characteristics	161

	Page
Criterion Measures	163
Survey of Reference Reports	165
Limitations of the Current Study	166
Future Research	170
Summary and Conclusions	175
 References	 179
 APPENDIX 1	
Letter Requesting Participation in the Study.....	205
 APPENDIX 2	
Teaching Vignette and Likability Scale	208
 APPENDIX 3	
Demographic Questionnaire	217
 APPENDIX 4	
Unstructured Rating Forms Developed for the Referee's Report and for the Performance Appraisal.....	219
 APPENDIX 5	
Likert-Type Rating Forms Developed for the Referee's Report and for the Performance Appraisal.....	222
 APPENDIX 6	
Asymmetrical Rating Forms Developed for the Referee's Report and for the Performance Appraisal.....	227

APPENDIX 7

Gender-Typed versions of the Leniency Scale 230

APPENDIX 8

Comparison of Rating Form Accuracy Using the Original
Rating Metric 233

LIST OF TABLES

	Page
TABLE 1	
<i>Studies, sample type, total number of participants and validity coefficients contributed by each study to the meta-analysis</i>	35
TABLE 2	
<i>Meta-analysis of the predictive validity of reference reports</i>	44
TABLE 3	
<i>Meta-analysis of the predictive validity of reference reports moderated by structure</i>	48
TABLE 4	
<i>Meta-analysis of the predictive validity of reference reports moderated by structure and criterion type.....</i>	49
TABLE 5	
<i>Mean reliability (average correlations) of judges' ratings for responses on the unstructured rating form.....</i>	117
TABLE 6	
<i>Summary of responses to the yes/no items in the demographic questionnaire.....</i>	122
TABLE 7	
<i>Descriptive statistics for questions from the demographic questionnaire.....</i>	122
TABLE 8	
<i>Significant main effects for type of form from the overall ANOVAs calculated for each teaching dimension.....</i>	128

TABLE 9	
<i>Means and standard deviations of ratings of teaching dimensions from three different rating forms.....</i>	129
TABLE 10	
<i>Results from t-test comparing mean ratings from three different forms.....</i>	130
TABLE 11	
<i>Mean accuracy values as a function of purpose and rating form</i>	131
TABLE 12	
<i>Results of t-tests comparing mean accuracy values for ratings from three different forms.....</i>	132
TABLE 13	
<i>Mean error values for ratings from three different forms used for performance appraisal and reference report purposes.....</i>	135
TABLE 14	
<i>Correlations among rater error measures</i>	136
TABLE 15	
<i>Correlations among rater accuracy measures</i>	137
TABLE 16	
<i>Correlations between accuracy and error measures.....</i>	138

TABLE 17	
<i>Results from a standard multiple regression analysis using four error measures to predict each of four accuracy scores</i>	139
TABLE 18	
<i>Correlations among variables included in the standard regression analysis for the prediction of leniency in ratings</i>	141
TABLE 19	
<i>Results from a standard multiple regression using six personal and contextual variables to predict leniency in ratings</i>	143
TABLE 20	
<i>Results from the hierarchical multiple regression analysis for the prediction of leniency in ratings after controlling for the effects of rating form (Analysis 1) and for the effects of rating form and Likability (Analysis 2).....</i>	145
TABLE 21	
<i>Referees' access to information about task and interpersonal behaviours and results.....</i>	174
TABLE 22	
<i>Results of t-tests comparing mean accuracy values for ratings from three different forms using the original rating metric.....</i>	234

LIST OF FIGURES

Page

FIGURE 1

*Distribution of validity coefficients from studies in the
meta-analysis.....* 39

FIGURE 2

Rated usefulness of reference reports 124