

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.

**Criterion Variance in Signal Detection
Theory:
The Interactive Effect of Knowledge of
Results and Task Difficulty on Binary
Decision Tasks**

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

**Master of Arts
in
Psychology**

**at Massey University, Palmerston North,
New Zealand.**

Robert T. Taylor

2010

Acknowledgements

To my supervisor, Dr. John Podd, I would like to express my deepest gratitude for his tireless support and guidance, his endless supply of Signal Detection resources, and his belief in the research. His tutelage, and patience, in refining my Signal Detection knowledge has been invaluable.

Thank you to Malcolm Loudon for all his help in the programming of the computer application. His support and skills was an asset that this research could not have done without.

To my fellow postgraduate students, thank you for providing me a sounding board to bounce ideas off, for providing advice, and an avenue for venting frustration in general.

To my family and friends, thanks are due for their understanding and support, and for putting up with my continual absenteeism.

And finally, thanks are due to my partner Steph. Her relentless support and encouragement made the research process that much easier, and was pivotal in the completion of this research.

Abstract

Within traditional Signal Detection Theory (SDT) experiments decision noise is very rarely considered, with researchers clinging to the assumption that the decision criterion has no associated variability. This assumption is incorrect. Furthermore, two factors contribute to criterion fluctuation: task difficulty and the type of knowledge of results (KR) delivered to the observer. The accepted standard in SDT experiments is to provide veridical trial-by-trial feedback (TTKR_e). This type of KR may adversely affect observer performance when the decision task is difficult, as the KR may appear highly inconsistent to the observer. The present study hypothesised that providing KR relative to the optimal criterion location (TTKR_i) would minimise criterion fluctuation. The present Criterion Variance Model (CVM) assumes that the decision criterion in SDT is subject to fluctuation. Two hypotheses were derived to test the model: a) contrary to the assumption of SDT, the decision criterion in a signal detection task is a variable rather than a fixed value on the decision axis, and is present within binary discrimination tasks; and b) There will be an interaction effect between the type of TTKR provided and the difficulty level of the task. Specifically, TTKR_i will enable more accurate decision making than TTKR_e, but only for a difficult decision task. Forty-four observers took part in a simple binary decision task, discriminating whether a presented tone was high or low in frequency (Hz). All tones were easily discriminable from each other; thus, the experiment was free from sensory noise. Task difficulty was manipulated by varying the degree of overlap between the high and low distributions, from which the high and low tones were sampled. As predicted by the CVM, performance in a difficult decision task was affected by the type of KR provided. Observers who received TTKR_e performed less well than observers who received TTKR_i in the more difficult version of the task. Despite mean criterion location measures across groups approaching zero – the optimal location – criterion fluctuation was evident when observer error distributions were analysed. Furthermore, the degree of criterion fluctuation was large, and was associated with the level of task difficulty. A major caveat was the lack of a no KR condition. Consequently, the degree to which observers utilised the KR could not be fully assessed. Additionally, the number of tones may have been too small, possibly encouraging observers not to use the KR provided in a consistent manner. Further research should incorporate a no KR condition and increase the number of tonal stimuli while ensuring the tones are still separated by 3 or 4 JNDs. Despite these design issues, the results highlight the potential detrimental effects of veridical KR on performance, particularly under conditions of high uncertainty.

Table of Contents

Acknowledgements	iii
Abstract	v
Table of Contents	vii
Table of Appendices	ix
List of Figures	x
List of Tables	xiii
Introduction and Overview	1
Chapter I: An Overview of Signal Detection Theory	4
SDT Fundamentals.....	4
The Sensory Stage.....	6
Distinctions between Sensitivity Measures and ROC Functions	11
The Decision Stage	12
Chapter II: Criticisms of SDT and the Issue of Criterion Variance	17
The Evidence for Criterion Variance	17
Models of Criterion Variance	20
Chapter III: The Role of Knowledge of Results in Criterion Variance	28
Evidence for Knowledge of Results and the Effects on Criteria	28
The Interaction of KR and Task Difficulty, and the Introduction of Optimal KR.....	31
Chapter IV: The Present Research	34
Method	40
Pilot Investigation	40
Main Study.....	41
Observers	41
Apparatus and Stimuli.....	41

Design.....	42
Procedure.....	43
Chapter V: Results	45
Observer Performance	45
Hypothesis 1 (Criterion Fluctuation).....	48
Hypothesis 2 (Interaction)	55
Post Hoc Tests	56
Auto-Correlation Analysis.....	57
Chapter VI: Discussion	59
Supplementary Analysis.....	63
Limitations	64
Conclusions and Future Directions.....	65
References	68
Appendices	73

Table of Appendices

Appendix A:	Glossary of Signal Detection Equations	75
Appendix B:	Tonal Frequencies.....	77
Appendix C:	Information Sheet	79
Appendix D:	Consent Form.....	83
Appendix E:	Instructions.....	85
Appendix F:	ANOVA Tables and Calculation of Eta Squared.....	89
Appendix G:	Auto-Correlation Tables	91
Appendix H:	Response Distributions and Error Plots	95

List of Figures

- Figure 1:** Varying frequencies of x produce two overlapping Gaussian distributions with equal variances. The area beneath the point where the two distributions cross reflects the area where x could be a noise + signal variables, or simply noise alone 6
- Figure 2:** (a) the degree of overlap is significant resulting in an attenuated distance between the means, and an increase in task difficulty; (b) the degree of overlap is reduced resulting in an increased distance between the means, and a decrease in task difficulty. 7
- Figure 3:** A cut point has been assumed along the decision axis at $z = 1$. For this value of x the HR is the area under the signal density marked with diagonal stripes, whereas the FAR is the area under the noise density shaded in grey. In this example $d' = 1$. .. 8
- Figure 4:** The ROC curve has been plotted for all values along the decision axis for both $d' = 1$ and $d' = 3$ conditions. As sensitivity decreases the bow becomes shallower and recedes toward the chance line – the positive diagonal..... 10
- Figure 5:** The z ROC curve has been plotted using the $z(\text{HR})$ and the $z(\text{FAR})$, producing a linear plot. The distance between the chance line and the plotted line for any value of z is equal to d' 10
- Figure 6:** The criterion splits the decision axis into ‘ S ’ and ‘ N ’ responses. Any stimulus magnitude equal to or falling to the left of k will produce an ‘ N ’ response, while any stimulus falling to the right of k will produce an ‘ S ’ response. 13
- Figure 7:** Varying the criterion location yields different β values. Criterion (a) assumes a lax position where $\beta = 0.63$; criterion (b) is optimal where $\beta = 1$; and criterion (c) assumes a strict position where $\beta = 1.58$. Bias can also be measured using c . Measures of c have also been provided (see Eq. 11). 14
- Figure 8:** Plotting the HR and FAR relative to the criterion location produces a point on the ROC curve that illustrates the degree of observer response bias. Criterion (a) reflects a lax position whereas criterion (c) reflects a strict position. Criterion (b) is

located at the optimal point. This ROC curve presented here is the result of normally distributed signal and noise densities with equal variances..... 15

- Figure 9:** Examples of binary (a) and rating (b) style tasks with regard to criterion and associated variability around the mean criterion position. Rating style tasks have multiple criteria, thus more variance is evident. 18
- Figure 10:** The interactive prediction for the current research. In a hard decision task $TTKR_i$ is expected to improve observer accuracy compared to that of $TTKR_e$. However, in an easy decision task the type of KR is expected to have little, or no, effect. 36
- Figure 11:** Probability distributions for the present research, showing degrees of overlap for both levels of difficulty. N = noise distribution; S1 = signal distribution $d' = 1$; S3 = signal distribution $d' = 3$ 38
- Figure 12:** Theoretical ROC functions for both levels of difficulty; a) $d'_{th} = 1$; b) $d'_{th} = 3$... 39
- Figure 13:** Trial sequence and corresponding times..... 42
- Figure 14:** a) Theoretical ROC for $d'_{th} = 1$ condition depicting events and ideal mean d'_{ob} values. Although the groups had a mean c close to optimal, there is a noticeable difference between the events and ideal KR group; b) Theoretical ROC for $d'_{th} = 3$ condition depicting events and ideal mean d'_{ob} values. In the easy version of the task, the criterion adopted was near optimal. As predicted by the CVM there is little difference in average performance for events and ideal KR..... 46
- Figure 15:** a) Accuracy measure, d'_{ob} , across KR groups for both $d'_{th} = 1$ (easy) and $d'_{th} = 3$ (hard) conditions. Performance was better in the hard ideal KR condition, compared to that of the events KR condition; b) Accuracy measures, A' , across KR groups for both $d'_{th} = 1$ (easy) and $d'_{th} = 3$ (hard) conditions. Performance again was better in the hard ideal KR condition, compared to that of the events KR condition 47
- Figure 16:** Example of an ideal error distribution, with the optimum criterion located between tones 7 and 8. If the criterion is fixed then high errors should only fall to the left of

the criterion, whereas low errors should only fall to the right. Errors should also reduce in frequency for tones further away from the optimum criterion 48

Figure 17: a) Distribution of errors for Observer 2; b) Distribution of errors for Observer 38. Both distributions reflect errors made using $TTKR_e$ under hard conditions. The optimum criterion is located between tones 8 and 9..... 50

Figure 18: a) Distribution of errors for Observer 27; b) Distribution of errors for Observer 24. Both distributions reflect errors made using $TTKR_i$ under hard conditions. The optimum criterion is located between tones 8 and 9..... 51

Figure 19: a) Distribution of errors for Observer 29; b) Distribution of errors for Observer 44. Both distributions reflect errors made using $TTKR_e$ under easy conditions. The optimum criterion is located between tones 10 and 11..... 52

Figure 20: a) Distribution of errors for Observer 23; b) Distribution of errors for Observer 42. Both distributions reflect errors made using $TTKR_i$ under easy conditions. The optimum criterion is located between tones 10 and 11..... 53

Figure 21: a) Distribution of errors for Observer 38. This error distribution was generated using $TTKR_e$ under hard conditions; b) Distribution of errors for Observer 27. This error distribution was generated using $TTKR_i$ under hard conditions. The optimum criterion is located between tones 8 and 9..... 55

List of Tables

Table 1:	Mean values for dependent measures across independent variables	45
Table 2:	ACF values for all observers across all conditions	58