Contents

Acknowledgements 9

1. Brief Introduction to Measurement (in Kinanthropology) 11

2. Historical Paths to Modern Test Theory 14

3. Groundwork for Item Response Theory 17

4. Item Response Theory (IRT) 19
- 4.1 Introduction 19
- 4.2 Unidimensional dichotomous IRT models 20
- 4.3 Unidimensional polytomous IRT models 24
- 4.4 Assumptions required for unidimensional IRT models 30
- 4.5 Parameter estimation in IRT models 32
  - 4.5.1 Latent trait (θ) estimation 33
  - 4.5.2 Item parameters estimation 37
- 4.6 Information and standard error of the θ estimates 38

5. Computerized Adaptive Testing (CAT): Historical and Conceptual Origins 44

6. Testing Algorithms in Unidimensional IRT-based CAT 51
- 6.1 Starting 53
- 6.2 Continuing 54
- 6.3 Stopping 59
- 6.4 Practical issues related to item selection in CAT 59
  - 6.4.1 Item pool 60
  - 6.4.2 Content balancing 61
  - 6.4.3 Exposure control 62

6.5 Evaluation of item selection and trait estimation methods used in computerized adaptive testing algorithms 63

7. Empirical Part – Problem Statement 66

8. Aims and Hypotheses 68

9. Methods 70
9.1 Item pool, IRT model used for item calibration, dimensionality analysis 72
9.1.1 General description of the item pool 72
9.1.2 Item calibration 72
9.1.3 Dimensionality analysis 72
9.2 CAT simulation design and specifications 73
9.2.1 Step 1. Simulate latent trait values (true θ) 73
9.2.2 Step 2. Supply item parameters for the intended item pool 73
9.2.3 Step 3. Set CAT algorithm options 73
9.2.4 Step 4. Simulate CAT administration 75
9.3 Analysis of simulation results 76

10. Results 78
10.1 Dimensionality 78
10.2 Number of administered items in CAT simulation 79
10.3 Bias of the CAT latent trait estimates 87
10.4 Correlations 91

11. Discussion 95

12. Conclusions 102

Summary 103

References 105

Appendices 117
Appendix A – IRT parameters (a – discrimination and b's – thresholds) for the Physical Self-Description Questionnaire items (source: Flatcher & Hattie, 2004) 117
Appendix B – R code used for the simulation of the PSDQ CAT 122
Appendix C – Test information and corresponding standard error for the Physical Self-Description Questionnaire item pool 125

Appendix D – Example of R code used to create Figure 1 126
Appendix E – Online application for adaptive testing using the Physical Self-Description Questionnaire 127

List of Tables 130
List of Figures 131