

# Contents

Foreword	ix
<i>C. Victor Bunderson</i>	
Preface	xiii
<i>Howard Wainer</i>	
<b>1</b> Introduction and History	<b>1</b>
<i>Howard Wainer</i>	
Prologue	1
The First Four Millennia of Mental Testing	2
The Origins of Mental Testing in the U.S. Military	4
The Origins of Admissions Testing for American Universities	6
Computerized Adaptive Testing	9
Important Issues in CAT	11
Challenges for the Future	16
The Structure and Use of a Gedanken Computerized Adaptive Test (The GCAT)	17
<b>2</b> System Design and Operations	<b>23</b>
<i>Bert F. Green</i>	
The Test Scenario	23
System Issues	26

vi CONTENTS

Hardware	28	
Human Factors and the Testing Environment		35
Software Issues	36	
References	38	
Exercises/Study Questions	40	
<b>3</b>	<b>Item Pools</b>	<b>41</b>
	<i>Ronald Flaugher</i>	
Introduction	41	
Steps in the Development of an Item Pool		43
Illustration of GCAT Item Pool Construction		49
On the Topic of Dimensionality	60	
Acknowledgments	61	
References	61	
Exercises/Study Questions	63	
<b>4</b>	<b>Item Response Theory, Item Calibration and Proficiency Estimation</b>	<b>65</b>
	<i>Howard Wainer and Robert J. Mislevy</i>	
Introduction	65	
Item Response Theory	67	
Estimating Proficiency	72	
Estimating Item Parameters	79	
Other Topics	91	
Technical Appendix: Linking Procedures		97
References	100	
Exercises/Study Questions	102	
<b>5</b>	<b>Testing Algorithms</b>	<b>103</b>
	<i>David Thissen and Robert J. Mislevy</i>	
Introduction and Background	103	
Item Selection CATs	107	
An Illustration from GCAT	117	
Topics Requiring Special Consideration	121	
Two-Stage Testing and Testlets	125	
Technical Appendix: Statistical Inference in CAT		130
References	132	
Exercises/Study Questions	132	

<b>6</b>	Scaling and Equating <i>Neil J. Dorans</i>	137
	Scores and Scales      138	
	Scaling and CAT: A Perspective      142	
	The GCAT to GP&P Equating: An Illustration      155	
	References      158	
	Exercises/Study Questions      160	
<b>7</b>	Reliability and Measurement Precision <i>David Thissen</i>	161
	Introduction and Background      162	
	Measurement Error      167	
	Other Sources of Error      173	
	Composite Scores      176	
	Comparison with Paper-and-Pencil Batteries      177	
	Illustrations of GCAT Measurement Precision      179	
	References      185	
	Exercises/Study Questions      186	
<b>8</b>	Validity <i>Lynne Steinberg, David Thissen, and Howard Wainer</i>	187
	Construct-Related Validity      190	
	Criterion-Related Validity      196	
	Threats to Validity      211	
	References      227	
	Exercises/Study Questions      231	
<b>9</b>	Future Challenges <i>Howard Wainer, Neil J. Dorans, Bert F. Green, Robert J. Mislevy, Lynne Steinberg, and David Thissen</i>	233
	Introduction      233	
	Practical Issues      234	
	What if the Model is Wrong?      237	
	Model Elaboration      240	
	Testlets      247	
	Policy Issues      256	
	New Possibilities      261	

viii CONTENTS

References	266
Exercises/Study Questions	272
CAT Primer References	273
Annotations and Acronyms Used	287
Author Index	289
Subject Index	295