

Food Science Text Series

Harry T. Lawless
Hildegard Heymann

Sensory Evaluation of Food

Principles and Practices

Second Edition

 Springer

Contents

1	Introduction	1
1.1	Introduction and Overview	1
1.1.1	Definition	1
1.1.2	Measurement	3
1.2	Historical Landmarks and the Three Classes of Test Methods	4
1.2.1	Difference Testing	5
1.2.2	Descriptive Analyses	6
1.2.3	Affective Testing	7
1.2.4	The Central Dogma—Analytic Versus Hedonic Tests	8
1.3	Applications: Why Collect Sensory Data?	10
1.3.1	Differences from Marketing Research Methods	13
1.3.2	Differences from Traditional Product Grading Systems	15
1.4	Summary and Conclusions	16
	References	17
2	Physiological and Psychological Foundations of Sensory Function	19
2.1	Introduction	19
2.2	Classical Sensory Testing and Psychophysical Methods	20
2.2.1	Early Psychophysics	20
2.2.2	The Classical Psychophysical Methods	21
2.2.3	Scaling and Magnitude Estimation	23
2.2.4	Critiques of Stevens	25
2.2.5	Empirical Versus Theory-Driven Functions	25
2.2.6	Parallels of Psychophysics and Sensory Evaluation	26
2.3	Anatomy and Physiology and Functions of Taste	27
2.3.1	Anatomy and Physiology	27
2.3.2	Taste Perception: Qualities	30
2.3.3	Taste Perception: Adaptation and Mixture Interactions	30
2.3.4	Individual Differences and Taste Genetics	33
2.4	Anatomy and Physiology and Functions of Smell	34
2.4.1	Anatomy and Cellular Function	34

2.4.2	Retronasal Smell	36
2.4.3	Olfactory Sensitivity and Specific Anosmia	37
2.4.4	Odor Qualities: Practical Systems	38
2.4.5	Functional Properties: Adaptation, Mixture Suppression, and Release	39
2.5	Chemesthesis	41
2.5.1	Qualities of Chemesthetic Experience	41
2.5.2	Physiological Mechanisms of Chemesthesis	42
2.5.3	Chemical "Heat"	43
2.5.4	Other Irritative Sensations and Chemical Cooling	44
2.5.5	Astringency	45
2.5.6	Metallic Taste	46
2.6	Multi-modal Sensory Interactions	47
2.6.1	Taste and Odor Interactions	47
2.6.2	Irritation and Flavor	49
2.6.3	Color-Flavor Interactions	49
2.7	Conclusions	50
	References	50
3	Principles of Good Practice	57
3.1	Introduction	57
3.2	The Sensory Testing Environment	58
3.2.1	Evaluation Area	59
3.2.2	Climate Control	62
3.3	Test Protocol Considerations	63
3.3.1	Sample Serving Procedures	63
3.3.2	Sample Size	63
3.3.3	Sample Serving Temperatures	64
3.3.4	Serving Containers	64
3.3.5	Carriers	65
3.3.6	Palate Cleansing	65
3.3.7	Swallowing and Expectoration	66
3.3.8	Instructions to Panelists	66
3.3.9	Randomization and Blind Labeling	66
3.4	Experimental Design	66
3.4.1	Designing a Study	66
3.4.2	Design and Treatment Structures	69
3.5	Panelist Considerations	72
3.5.1	Incentives	72
3.5.2	Use of Human Subjects	73
3.5.3	Panelist Recruitment	74
3.5.4	Panelist Selection and Screening	74
3.5.5	Training of Panelists	75
3.5.6	Panelist Performance Assessment	75
3.6	Tabulation and Analysis	75
3.6.1	Data Entry Systems	75
3.7	Conclusion	76
	References	76

4	Discrimination Testing	79
4.1	Discrimination Testing	79
4.2	Types of Discrimination Tests	80
4.2.1	Paired Comparison Tests	80
4.2.2	Triangle Tests	83
4.2.3	Duo-Trio Tests	84
4.2.4	n-Alternative Forced Choice (n-AFC) Methods	85
4.2.5	A-Not-A tests	85
4.2.6	Sorting Methods	87
4.2.7	The ABX Discrimination Task	88
4.2.8	Dual-Standard Test	88
4.3	Reputed Strengths and Weaknesses of Discrimination Tests	88
4.4	Data Analyses	89
4.4.1	Binomial Distributions and Tables	89
4.4.2	The Adjusted Chi-Square (χ^2) Test	90
4.4.3	The Normal Distribution and the Z-Test on Proportion	90
4.5	Issues	92
4.5.1	The Power of the Statistical Test	92
4.5.2	Replications	94
4.5.3	Warm-Up Effects	97
4.5.4	Common Mistakes Made in the Interpretation of Discrimination Tests	97
	Appendix: A Simple Approach to Handling the A, Not-A, and Same/Different Tests	98
	References	99
5	Similarity, Equivalence Testing, and Discrimination Theory	101
5.1	Introduction	101
5.2	Common Sense Approaches to Equivalence	103
5.3	Estimation of Sample Size and Test Power	104
5.4	How Big of a Difference Is Important?	
	Discriminator Theory	105
5.5	Tests for Significant Similarity	108
5.6	The Two One-Sided Test Approach (TOST) and Interval Testing	110
5.7	Claim Substantiation	111
5.8	Models for Discrimination: Signal Detection Theory	111
5.8.1	The Problem	112
5.8.2	Experimental Setup	112
5.8.3	Assumptions and Theory	113
5.8.4	An Example	114
5.8.5	A Connection to Paired Comparisons Results Through the ROC Curve	116
5.9	Thurstonian Scaling	116
5.9.1	The Theory and Formulae	116
5.9.2	Extending Thurstone's Model to Other Choice Tests	118

5.10	Extensions of the Thurstonian Methods, <i>R</i> -Index	119
5.10.1	Short Cut Signal Detection Methods	119
5.10.2	An Example	120
5.11	Conclusions	120
	Appendix: Non-Central <i>t</i> -Test for Equivalence of Scaled Data	122
	References	122
6	Measurement of Sensory Thresholds	125
6.1	Introduction: The Threshold Concept	125
6.2	Types of Thresholds: Definitions	127
6.3	Practical Methods: Ascending Forced Choice	128
6.4	Suggested Method for Taste/Odor/Flavor Detection Thresholds	129
6.4.1	Ascending Forced-Choice Method of Limits	129
6.4.2	Purpose of the Test	129
6.4.3	Preliminary Steps	130
6.4.4	Procedure	131
6.4.5	Data Analysis	131
6.4.6	Alternative Graphical Solution	131
6.4.7	Procedural Choices	133
6.5	Case Study/Worked Example	133
6.6	Other Forced Choice Methods	134
6.7	Probit Analysis	136
6.8	Sensory Adaptation, Sequential Effects, and Variability	136
6.9	Alternative Methods: Rated Difference, Adaptive Procedures, Scaling	137
6.9.1	Rated Difference from Control	137
6.9.2	Adaptive Procedures	138
6.9.3	Scaling as an Alternative Measure of Sensitivity	140
6.10	Dilution to Threshold Measures	140
6.10.1	Odor Units and Gas-Chromatography Olfactometry (GCO)	140
6.10.2	Scoville Units	142
6.11	Conclusions	142
	Appendix: MTBE Threshold Data for Worked Example	143
	References	145
7	Scaling	149
7.1	Introduction	149
7.2	Some Theory	151
7.3	Common Methods of Scaling	152
7.3.1	Category Scales	152
7.3.2	Line Scaling	155
7.3.3	Magnitude Estimation	156
7.4	Recommended Practice and Practical Guidelines	158
7.4.1	Rule 1: Provide Sufficient Alternatives	159
7.4.2	Rule 2: The Attribute Must Be Understood	159
7.4.3	Rule 3: The Anchor Words Should Make Sense	159
7.4.4	To Calibrate or Not to Calibrate	159
7.4.5	A Warning: Grading and Scoring are Not Scaling	160

7.5	Variations—Other Scaling Techniques	160
7.5.1	Cross-Modal Matches and Variations on Magnitude Estimation	160
7.5.2	Category–Ratio (Labeled Magnitude) Scales	162
7.5.3	Adjustable Rating Techniques: Relative Scaling	164
7.5.4	Ranking	165
7.5.5	Indirect Scales	166
7.6	Comparing Methods: What is a Good Scale?	167
7.7	Issues	168
7.7.1	“Do People Make Relative Judgments” Should They See Their Previous Ratings?	168
7.7.2	Should Category Rating Scales Be Assigned Integer Numbers in Data Tabulation? Are They Interval Scales?	169
7.7.3	Is Magnitude Estimation a Ratio Scale or Simply a Scale with Ratio Instructions?	169
7.7.4	What is a “Valid” Scale?	169
7.8	Conclusions	170
	Appendix 1: Derivation of Thurstonian-Scale Values for the 9-Point Scale	171
	Appendix 2: Construction of Labeled Magnitude Scales	172
	References	174
8	Time–Intensity Methods	179
8.1	Introduction	179
8.2	A Brief History	180
8.3	Variations on the Method	182
8.3.1	Discrete or Discontinuous Sampling	182
8.3.2	“Continuous” Tracking	183
8.3.3	Temporal Dominance Techniques	184
8.4	Recommended Procedures	185
8.4.1	Steps in Conducting a Time–intensity Study	185
8.4.2	Procedures	186
8.4.3	Recommended Analysis	186
8.5	Data Analysis Options	187
8.5.1	General Approaches	187
8.5.2	Methods to Construct or Describe Average Curves	188
8.5.3	Case Study: Simple Geometric Description	189
8.5.4	Analysis by Principal Components	192
8.6	Examples and Applications	193
8.6.1	Taste and Flavor Sensation Tracking	193
8.6.2	Trigeminal and Chemical/Tactile Sensations	194
8.6.3	Taste and Odor Adaptation	194
8.6.4	Texture and Phase Change	195
8.6.5	Flavor Release	195
8.6.6	Temporal Aspects of Hedonics	196
8.7	Issues	197
8.8	Conclusions	198
	References	198

9	Context Effects and Biases in Sensory Judgment	203
9.1	Introduction: The Relative Nature of Human Judgment	203
9.2	Simple Contrast Effects	206
9.2.1	A Little Theory: Adaptation Level	206
9.2.2	Intensity Shifts	207
9.2.3	Quality Shifts	207
9.2.4	Hedonic Shifts	208
9.2.5	Explanations for Contrast	209
9.3	Range and Frequency Effects	210
9.3.1	A Little More Theory: Parducci's Range and Frequency Principles	210
9.3.2	Range Effects	210
9.3.3	Frequency Effects	211
9.4	Biases	212
9.4.1	Idiosyncratic Scale Usage and Number Bias	212
9.4.2	Poulton's Classifications	213
9.4.3	Response Range Effects	214
9.4.4	The Centering Bias	215
9.5	Response Correlation and Response Restriction	216
9.5.1	Response Correlation	216
9.5.2	"Dumping" Effects: Inflation Due to Response Restriction in Profiling	217
9.5.3	Over-Partitioning	218
9.6	Classical Psychological Errors and Other Biases	218
9.6.1	Errors in Structured Sequences: Anticipation and Habituation	218
9.6.2	The Stimulus Error	219
9.6.3	Positional or Order Bias	219
9.7	Antidotes	219
9.7.1	Avoid or Minimize	219
9.7.2	Randomization and Counterbalancing	220
9.7.3	Stabilization and Calibration	221
9.7.4	Interpretation	222
9.8	Conclusions	222
	References	223
10	Descriptive Analysis	227
10.1	Introduction	227
10.2	Uses of Descriptive Analyses	228
10.3	Language and Descriptive Analysis	228
10.4	Descriptive Analysis Techniques	231
10.4.1	Flavor Profile®	231
10.4.2	Quantitative Descriptive Analysis®	234
10.4.3	Texture Profile®	237
10.4.4	Sensory Spectrum®	238
10.5	Generic Descriptive Analysis	240
10.5.1	How to Do Descriptive Analysis in Three Easy Steps	240

10.5.2	Studies Comparing Different Conventional Descriptive Analysis Techniques	246
10.6	Variations on the Theme	247
10.6.1	Using Attribute Citation Frequencies Instead of Attribute Intensities	247
10.6.2	Deviation from Reference Method	248
10.6.3	Intensity Variation Descriptive Method	249
10.6.4	Combination of Descriptive Analysis and Time-Related Intensity Methods	249
10.6.5	Free Choice Profiling	249
10.6.6	Flash Profiling	252
	References	253
11	Texture Evaluation	259
11.1	Texture Defined	259
11.2	Visual, Auditory, and Tactile Texture	262
11.2.1	Visual Texture	262
11.2.2	Auditory Texture	262
11.2.3	Tactile Texture	264
11.2.4	Tactile Hand Feel	268
11.3	Sensory Texture Measurements	270
11.3.1	Texture Profile Method	270
11.3.2	Other Sensory Texture Evaluation Techniques	272
11.3.3	Instrumental Texture Measurements and Sensory Correlations	274
11.4	Conclusions	276
	References	276
12	Color and Appearance	283
12.1	Color and Appearance	283
12.2	What Is Color?	284
12.3	Vision	285
12.3.1	Normal Human Color Vision Variations	286
12.3.2	Human Color Blindness	286
12.4	Measurement of Appearance and Color Attributes	286
12.4.1	Appearance	286
12.4.2	Visual Color Measurement	289
12.5	Instrumental Color Measurement	293
12.5.1	Munsell Color Solid	293
12.5.2	Mathematical Color Systems	294
12.6	Conclusions	299
	References	299
13	Preference Testing	303
13.1	Introduction—Consumer Sensory Evaluation	303
13.2	Preference Tests: Overview	305
13.2.1	The Basic Comparison	305
13.2.2	Variations	305
13.2.3	Some Cautions	306
13.3	Simple Paired Preference Testing	306

13.3.1	Recommended Procedure	306
13.3.2	Statistical Basis	307
13.3.3	Worked Example	308
13.3.4	Useful Statistical Approximations	309
13.3.5	The Special Case of Equivalence Testing	310
13.4	Non-forced Preference	311
13.5	Replicated Preference Tests	313
13.6	Replicated Non-forced Preference	313
13.7	Other Related Methods	315
13.7.1	Ranking	315
13.7.2	Analysis of Ranked Data	316
13.7.3	Best–Worst Scaling	317
13.7.4	Rated Degree of Preference and Other Options	318
13.8	Conclusions	320
<i>Appendix 1: Worked Example of the Ferris k-Visit Repeated</i>		
	Preference Test Including the No-Preference Option	320
<i>Appendix 2: The “Placebo” Preference Test</i>		321
<i>Appendix 3: Worked Example of Multinomial Approach</i>		
	to Analyzing Data with the No-Preference Option	322
References		323
14	Acceptance Testing	325
14.1	Introduction: Scaled Liking Versus Choice	325
14.2	Hedonic Scaling: Quantification of Acceptability	326
14.3	Recommended Procedure	327
14.3.1	Steps	327
14.3.2	Analysis	328
14.3.3	Replication	328
14.4	Other Acceptance Scales	328
14.4.1	Line Scales	328
14.4.2	Magnitude Estimation	330
14.4.3	Labeled Magnitude Scales	331
14.4.4	Pictorial Scales and Testing with Children	332
14.4.5	Adjustable Scales	333
14.5	Just-About-Right Scales	334
14.5.1	Description	334
14.5.2	Limitations	335
14.5.3	Variations on Relative-to-Ideal Scaling	336
14.5.4	Analysis of JAR Data	336
14.5.5	Penalty Analysis or “Mean Drop”	339
14.5.6	Other Problems and Issues with JAR Scales	340
14.6	Behavioral and Context-Related Approaches	340
14.6.1	Food Action Rating Scale (FACT)	341
14.6.2	Appropriateness Scales	341
14.6.3	Acceptor Set Size	342
14.6.4	Barter Scales	343
14.7	Conclusions	343
References		344

15	Consumer Field Tests and Questionnaire Design	349
15.1	Sensory Testing Versus Concept Testing	349
15.2	Testing Scenarios: Central Location, Home Use	351
15.2.1	Purpose of the Tests	351
15.2.2	Consumer Models	352
15.2.3	Central Location Tests	353
15.2.4	Home Use Tests (HUT)	354
15.3	Practical Matters in Conducting Consumer Field Tests	355
15.3.1	Tasks and Test Design	355
15.3.2	Sample Size and Stratification	355
15.3.3	Test Designs	356
15.4	Interacting with Field Services	358
15.4.1	Choosing Agencies, Communication, and Test Specifications	358
15.4.2	Incidence, Cost, and Recruitment	359
15.4.3	Some Tips: Do's and Don'ts	360
15.4.4	Steps in Testing with Research Suppliers	360
15.5	Questionnaire Design	362
15.5.1	Types of Interviews	362
15.5.2	Questionnaire Flow: Order of Questions	362
15.5.3	Interviewing	363
15.6	Rules of Thumb for Constructing Questions	364
15.6.1	General Principles	364
15.6.2	Brevity	364
15.6.3	Use Plain Language	364
15.6.4	Accessibility of the Information	365
15.6.5	Avoid Vague Questions	365
15.6.6	Check for Overlap and Completeness	365
15.6.7	Do Not Lead the Respondent	365
15.6.8	Avoid Ambiguity and Double Questions	366
15.6.9	Be Careful in Wording: Present Both Alternatives	366
15.6.10	Beware of Halos and Horns	366
15.6.11	Pre-test	366
15.7	Other Useful Questions: Satisfaction, Agreement, and Open-Ended Questions	367
15.7.1	Satisfaction	367
15.7.2	Likert (Agree–Disagree) Scales	367
15.7.3	Open-Ended Questions	367
15.8	Conclusions	368
	Appendix 1: Sample Test Specification Sheet	370
	Appendix 2: Sample Screening Questionnaire	371
	Appendix 3: Sample Product Questionnaire	374
	References	378
16	Qualitative Consumer Research Methods	379
16.1	Introduction	380
16.1.1	Resources, Definitions, and Objectives	380
16.1.2	Styles of Qualitative Research	380
16.1.3	Other Qualitative Techniques	382

16.2	Characteristics of Focus Groups	383
16.2.1	Advantages	383
16.2.2	Key Requirements	384
16.2.3	Reliability and Validity	384
16.3	Using Focus Groups in Sensory Evaluation	385
16.4	Examples, Case Studies	386
16.4.1	Case Study 1: Qualitative Research Before Conjoint Measurement in New Product Development	387
16.4.2	Case Study 2: Nutritional and Health Beliefs About Salt	387
16.5	Conducting Focus Group Studies	388
16.5.1	A Quick Overview	388
16.5.2	A Key Requirement: Developing Good Questions . .	389
16.5.3	The Discussion Guide and Phases of the Group Interview	390
16.5.4	Participant Requirements, Timing, Recording	391
16.6	Issues in Moderating	392
16.6.1	Moderating Skills	392
16.6.2	Basic Principles: Nondirection, Full Participation, and Coverage of Issues	393
16.6.3	Assistant Moderators and Co-moderators	394
16.6.4	Debriefing: Avoiding Selective Listening and Premature Conclusions	395
16.7	Analysis and Reporting	395
16.7.1	General Principles	395
16.7.2	Suggested Method ("Sorting/Clustering Approach"), also Called Classical Transcript Analysis	396
16.7.3	Report Format	397
16.8	Alternative Procedures and Variations of the Group Interview	398
16.8.1	Groups of Children, Telephone Interviews, Internet-Based Groups	398
16.8.2	Alternatives to Traditional Questioning	399
16.9	Conclusions	400
	Appendix: Sample Report Group Report	402
	Boil-in-bag Pasta Project Followup Groups	402
	References	404
17	Quality Control and Shelf-Life (Stability) Testing	407
17.1	Introduction: Objectives and Challenges	408
17.2	A Quick Look at Traditional Quality Control	409
17.3	Methods for Sensory QC	409
17.3.1	Cuttings: A Bad Example	409
17.3.2	In-Out (Pass/Fail) System	410
17.3.3	Difference from Control Ratings	411
17.3.4	Quality Ratings with Diagnostics	412
17.3.5	Descriptive Analysis	413
17.3.6	A Hybrid Approach: Quality Ratings with Diagnostics	414

17.3.7	The Multiple Standards Difference Test	414
17.4	Recommended Procedure: Difference Scoring with Key Attribute Scales	415
17.5	The Importance of Good Practice	417
17.6	Historical Footnote: Expert Judges and Quality Scoring	419
17.6.1	Standardized Commodities	419
17.6.2	Example 1: Dairy Product Judging	419
17.6.3	Example 2: Wine Scoring	420
17.7	Program Requirements and Program Development	422
17.7.1	Desired Features of a Sensory QC System	422
17.7.2	Program Development and Management Issues	423
17.7.3	The Problem of Low Incidence	424
17.8	Shelf-Life Testing	424
17.8.1	Basic Considerations	424
17.8.2	Cutoff Point	426
17.8.3	Test Designs	426
17.8.4	Survival Analysis and Hazard Functions	427
17.8.5	Accelerated Storage	428
17.9	Summary and Conclusions	428
	Appendix 1: Sample Screening Tests for Sensory Quality Judges	429
	Appendix 2: Survival/Failure Estimates from a Series of Batches with Known Failure Times	429
	Appendix 3: Arrhenius Equation and Q_{10} Modeling	430
	References	431
18	Data Relationships and Multivariate Applications	433
18.1	Introduction	433
18.2	Overview of Multivariate Statistical Techniques	434
18.2.1	Principal Component Analysis	434
18.2.2	Multivariate Analysis of Variance	437
18.2.3	Discriminant Analysis (Also Known as Canonical Variate Analysis)	438
18.2.4	Generalized Procrustes Analysis	439
18.3	Relating Consumer and Descriptive Data Through Preference Mapping	440
18.3.1	Internal Preference Mapping	442
18.3.2	External Preference Mapping	442
18.4	Conclusions	445
	References	446
19	Strategic Research	451
19.1	Introduction	451
19.1.1	Avenues for Strategic Research	451
19.1.2	Consumer Contact	453
19.2	Competitive Surveillance	453
19.2.1	The Category Review	453
19.2.2	Perceptual Mapping	455
19.2.3	Multivariate Methods: PCA	456
19.2.4	Multi-dimensional Scaling	458

19.2.5	Cost-Efficient Methods for Data Collection: Sorting	459
19.2.6	Vector Projection	460
19.2.7	Cost-Efficient Methods for Data Collection: Projective Mapping, aka Napping	461
19.3	Attribute Identification and Classification	462
19.3.1	Drivers of Liking	462
19.3.2	The Kano Model	463
19.4	Preference Mapping Revisited	464
19.4.1	Types of Preference Maps	464
19.4.2	Preference Models: Vectors Versus Ideal Points	464
19.5	Consumer Segmentation	465
19.6	Claim Substantiation Revisited	467
19.7	Conclusions	468
19.7.1	Blind Testing, New Coke, and the Vienna Philharmonic	468
19.7.2	The Sensory Contribution	469
	References	469
Appendix A Basic Statistical Concepts for Sensory Evaluation		473
A.1	Introduction	473
A.2	Basic Statistical Concepts	474
A.2.1	Data Description	475
A.2.2	Population Statistics	476
A.3	Hypothesis Testing and Statistical Inference	478
A.3.1	The Confidence Interval	478
A.3.2	Hypothesis Testing	478
A.3.3	A Worked Example	479
A.3.4	A Few More Important Concepts	480
A.3.5	Decision Errors	482
A.4	Variations of the <i>t</i> -Test	482
A.4.1	The Sensitivity of the Dependent <i>t</i> -Test for Sensory Data	484
A.5	Summary: Statistical Hypothesis Testing	485
A.6	Postscript: What <i>p</i> -Values Signify and What They Do Not	485
A.7	Statistical Glossary	486
	References	487
Appendix B Nonparametric and Binomial-Based Statistical Methods		489
B.1	Introduction to Nonparametric Tests	489
B.2	Binomial-Based Tests on Proportions	490
B.3	Chi-Square	493
B.3.1	A Measure of Relatedness of Two Variables	493
B.3.2	Calculations	494
B.3.3	Related Samples: The McNemar Test	494
B.3.4	The Stuart–Maxwell Test	495
B.3.5	Beta-Binomial, Chance-Corrected Beta-Binomial, and Dirichlet Multinomial Analyses	496

B.4	Useful Rank Order Tests	499
B.4.1	The Sign Test	499
B.4.2	The Mann–Whitney <i>U</i> -Test	500
B.4.3	Ranked Data with More Than Two Samples, Friedman and Kramer Tests	501
B.4.4	Rank Order Correlation	502
B.5	Conclusions	503
B.6	Postscript	503
B.6.1	Proof showing equivalence of binomial approximation Z-test and χ^2 test for difference of proportions	503
	References	504
Appendix C	Analysis of Variance	507
C.1	Introduction	507
C.1.1	Overview	507
C.1.2	Basic Analysis of Variance	508
C.1.3	Rationale	508
C.1.4	Calculations	509
C.1.5	A Worked Example	509
C.2	Analysis of Variance from Complete Block Designs	510
C.2.1	Concepts and Partitioning Panelist Variance from Error	510
C.2.2	The Value of Using Panelists As Their Own Controls	512
C.3	Planned Comparisons Between Means Following ANOVA	513
C.4	Multiple Factor Analysis of Variance	514
C.4.1	An Example	514
C.4.2	Concept: A Linear Model	515
C.4.3	A Note About Interactions	516
C.5	Panelist by Product by Replicate Designs	516
C.6	Issues and Concerns	519
C.6.1	Sensory Panelists: Fixed or Random Effects?	519
C.6.2	A Note on Blocking	520
C.6.3	Split-Plot or Between-Groups (Nested) Designs	520
C.6.4	Statistical Assumptions and the Repeated Measures ANOVA	521
C.6.5	Other Options	522
	References	522
Appendix D	Correlation, Regression, and Measures of Association	525
D.1	Introduction	525
D.2	Correlation	527
D.2.1	Pearson's Correlation Coefficient Example	528
D.2.2	Coefficient of Determination	529
D.3	Linear Regression	529
D.3.1	Analysis of Variance	530
D.3.2	Analysis of Variance for Linear Regression	530
D.3.3	Prediction of the Regression Line	530
D.3.4	Linear Regression Example	531

D.4	Multiple Linear Regression	531
D.5	Other Measures of Association	531
D.5.1	Spearman Rank Correlation	531
D.5.2	Spearman Correlation Coefficient Example	532
D.5.3	Cramér's V Measure	532
D.5.4	Cramér Coefficient Example	533
	References	533
Appendix E	Statistical Power and Test Sensitivity	535
E.1	Introduction	535
E.2	Factors Affecting the Power of Statistical Tests	537
E.2.1	Sample Size and Alpha Level	537
E.2.2	Effect Size	538
E.2.3	How Alpha, Beta, Effect Size, and N Interact	539
E.3	Worked Examples	541
E.3.1	The t -Test	541
E.3.2	An Equivalence Issue with Scaled Data	542
E.3.3	Sample Size for a Difference Test	544
E.4	Power in Simple Difference and Preference Tests	545
E.5	Summary and Conclusions	548
	References	549
Appendix F	Statistical Tables	551
Table F.A	Cumulative probabilities of the standard normal distribution. Entry area $1-\alpha$ under the standard normal curve from $-\infty$ to $z(1-\alpha)$	552
Table F.B	Table of critical values for the t -distribution	553
Table F.C	Table of critical values of the chi-square (χ^2) distribution	554
Table F.D1	Critical values of the F -distribution at $\alpha = 0.05$	555
Table F.D2	Critical values of the F -distribution at $\alpha = 0.01$	556
Table F.E	Critical values of U for a one-tailed alpha at 0.025 or a two-tailed alpha at 0.05	556
Table F.F1	Table of critical values of ρ (Spearman Rank correlation coefficient)	557
Table F.F2	Table of critical values of r (Pearson's correlation coefficient)	558
Table F.G	Critical values for Duncan's multiple range test (p , df, $\alpha = 0.05$)	559
Table F.H1	Critical values of the triangle test for similarity (maximum number correct as a function of the number of observations (N), beta, and proportion discriminating)	560
Table F.H2	Critical values of the duo-trio and paired comparison tests for similarity (maximum number correct as a function of the number of observations (N), beta, and proportion discriminating)	561
Table F.I	Table of probabilities for values as small as observed values of x associated with the binomial test ($p=0.50$)	562

Table F.J	Critical values for the differences between rank sums ($\alpha = 0.05$)	563
Table F.K	Critical values of the beta binomial distribution	564
Table F.L	Minimum numbers of correct judgments to establish significance at probability levels of 5 and 1% for paired difference and duo-trio tests (one tailed, $p = 1/2$) and the triangle test (one tailed, $p = 1/3$)	565
Table F.M	Minimum numbers of correct judgments to establish significance at probability levels of 5 and 1% for paired preference test (two tailed, $p = 1/2$)	566
Table F.N1	Minimum number of responses (n) and correct responses (x) to obtain a level of Type I and Type II risks in the triangle test. P_d is the chance-adjusted percent correct or proportion of discriminators	567
Table F.N2	Minimum number of responses (n) and correct responses (x) to obtain a level of Type I and Type II risks in the duo-trio test. P_c is the chance-adjusted percent correct or proportion of discriminators	567
Table F.O1	d' and B (variance factor) values for the duo-trio and 2-AFC (paired comparison) difference tests	568
Table F.O2	d' and B (variance factor) values for the triangle and 3-AFC difference tests	569
Table F.P	Random permutations of nine	571
Table F.Q	Random numbers	572
Author Index		573
Subject Index		587