

Contents

EDITORS' PREFACE	ix
1 Introduction ALAN DAVIES	1
2 Basic Concepts in Testing ELISABETH INGRAM	11
1 Testing and language learning	11
2 Measurement	12
3 Test requirements	15
3.1 Reliability	15
3.1.1 Stability reliability	16
3.1.2 Equivalence reliability	16
3.2 Validity	18
3.3 Correlations	22
4 Norm-referenced versus criterion-referenced testing	26
5 Tests and experiments	32
6 Practical work	35
3 The Construction of Language Tests ALAN DAVIES	38
1 Introduction	38
2 Aims and purposes of testing	41
2.1 Research	41
2.2 Progress	42
2.3 Guide to curriculum	42
2.4 Representing terminal behaviour	43
3 The influence of programmed instruction	43
4 Types of test	44
4.1 Uses of tests	44
4.1.1 The four uses	45
4.1.2 Summary	48

4.2	Examinations and tests	49
4.2.1	Examination/test levels	49
4.2.2	Standardized tests	50
4.2.3	Ad hoc tests	51
4.2.4	Traditional examinations	52
4.2.5	New types of examination	52
4.3	Types of test levels and skills	54
5	Demands on language tests	55
5.1	Test criteria	55
5.2	Test virtues	57
5.2.1	Reliability	57
5.2.2	Validity	58
5.2.3	Types of validity	58
5.2.4	Schema for education	63
6	Language test analysis	65
7	Teacher made tests	69
7.1	Proficiency and achievement tests	69
7.2	Selection of language areas for testing	70
7.3	Language analysis	71
7.4	Work sample analysis	71
8	Item writing	76
8.1	The framework of a test	76
8.2	Examples of items	78
8.3	Conclusion	100
9	Practical work	100
10	Further reading	104
4	The Design and Interpretation of Experiments	
	RUTH CLARK	105
1	Introduction	105
2	Measurement	107
2.1	The choice of an adequate measure	107
2.2	Measurement: further sources of bias	109
3	Control	112
3.1	The natural experiment	112
3.2	The controlled experiment	113
4	Interpretation of results	116
4.1	Spanish-teaching experiment: nominal data	117
4.2	Training in the use of pronouns: ordinal data from independent groups	118

4.3	Understanding of logical connectives: ordinal data from related groups	122
4.4	Birth weight of normal and speech-retarded children: interval data from independent groups—subject variable	126
4.5	Summary of the four studies	130
4.6	Dispersion	131
5	The advantages and limitations of certain research designs	133
5.1	Related subject designs and the restriction of variation	133
5.2	Designs with a subject variable	137
6	Sampling	138
7	Randomization	138
8	Conclusion	139
9	Practical work	141
10	Acknowledgements	144
5	Procedures and Computations in the Analysis of Experiments	
	RUTH CLARK	146
1	Introduction	146
2	Analysing the contingency table	147
2.1	Calculating expected frequencies	147
2.2	The chi-squared test—calculating χ^2	148
3	Frequencies, the median and ranking procedures	149
3.1	Frequency tables	149
3.2	Cumulative frequency distributions and the median	150
3.3	Assigning a rank to each score	151
3.4	Sums of ranks and sums of ranks of differences	152
3.5	The calculation of U	155
4	Conclusion	156
5	Practical work	156
6	Statistical Inference RUTH CLARK	158
1	Introduction	158
2	The binomial distribution—probability and significance	159
3	The normal distribution and standard scores	167
4	The final steps in the analysis	172
4.1	Mann-Whitney U test—for comparing independent groups by means of the sum of ranks	173
4.2	Sign test—for comparing related groups using number of positive and negative differences	173
4.3	Wilcoxon matched pairs signed ranks—for comparing related groups using the sum of ranks of differences	174

4.4	Test of difference between means—for independent samples where the variable is on an interval scale and normally distributed in the population	175
4.5	Chi squared and median tests for analysing contingency tables	176
4.6	Summary	179
5	Increasing the sensitivity of an experiment	179
5.1	Sample size	180
5.2	The control of variability	180
6	One-tailed and two-tailed hypotheses	181
7	Factors affecting the choice of tests	182
7.1	Level of measurement	182
7.2	Parametric tests	184
7.3	Power	184
7.4	Normalization of scales	184
8	Conclusion	185
9	Practical work	185
10	Further reading	186
Appendix I: Some Statistics used in Language Tests		187
1	Method for working item analysis E_{1-3}	187
2	Method for working rank order (Spearman) correlations	189
3	Formula and method for working product moment correlations	190
4	Formula and method for establishing equivalence reliability Kuder-Richardson (21)	194
5	Method for finding square roots by using logarithms	198
Appendix II: Statistical Tables		200
A	Table of probabilities associated with values as extreme as observed values of z in the normal distribution	200
B	Table of critical values of chi square	202
C	Table of probabilities associated with values as small as observed values of x in the binomial test	203
ANSWERS TO PRACTICAL WORK		204
GLOSSARY OF SYMBOLS		222
REFERENCES		224
INDEX OF NAMES		227
SUBJECT INDEX		229