TECHNIQUES
IN
TESTING

· Harold S. Madsen ·

**Oxford
American
English**

TECHNIQUES
IN
TESTING

Harold S. Madsen

Illustrations by Steven Schindler.

·EDITORS' PREFACE·

It has been apparent for some time that little attention has been given to the needs of practicing and student teachers of English as a Second Language.* Although numerous inservice and pre-service teacher-training programs are offered throughout the world, these often suffer for lack of appropriate instructional materials. Seldom are books written that present practical information that relates directly to daily classroom instruction. What teachers want are useful ideas, suggestions, demonstrations, and examples of teaching techniques that have proven successful in the classroom—techniques that are consistent with established theoretical principles and that others in our profession have found to be expedient, practical, and relevant to the real-life circumstances in which most teachers work.

It was in recognition of this need that we began our search for scholars in our field who had distinguished themselves in particular instructional aspects of second language teaching. We sought out those who had been especially successful in communicating to their colleagues the characteristics of language teaching and testing techniques that have been found to be appropriate for students from elementary school through college and adult education programs. We also sought in those same

---

*In this volume, and in others in the series, we have chosen to use *English as a Second Language (ESL)* to refer to English teaching in the United States (as a second language) *as well as* English teaching in other countries (as a foreign language).

scholars evidence of an awareness and understanding of current theories of language learning, together with the ability to translate the essence of a theory into practical applications for the classroom.

Our search has been successful. For this volume, as well as for others in this series, we have chosen a colleague who is extraordinarily competent and exceedingly willing to share with practicing teachers the considerable knowledge that he has gained from many years of experience in many parts of the world.

Dr. Madsen's book is devoted entirely to the presentation and exemplification of practical testing techniques. Each chapter of his book contains, in addition to detailed consideration of a wide variety of techniques, a number of activities that teachers can perform that tie the content of the book directly to the teachers' responsibilities in their classes. With this volume then, a critical need in the language teaching field has been met.

We are extremely pleased to join with the authors in this series and with Oxford University Press in making these books available to our fellow teachers. We are confident that the books will enable language teachers around the world to increase their effectiveness while at the same time making their task an easier and more enjoyable one.

Russell N. Campbell
William E. Rutherford

Editors' Note: Apologies are made for the generalized use of the masculine pronoun. It is meant to be used for simplicity's sake, rather than to indicate a philosophical viewpoint. We feel that the *s/he, her/him, his/her* forms, while they may be philosophically appealing, are confusing.

·ACKNOWLEDGMENTS·

·CONTENTS·

TECHNIQUES IN TESTING

·CHAPTER ONE·
INTRODUCTION

Testing is an important part of every teaching and learning experience. This book on testing has been prepared for both experienced and inexperienced teachers of English as a Second Language (ESL) who feel a need to improve their skills in constructing and administering classroom tests at the middle and secondary school levels or in ESL courses for adult students.

We begin in this introductory chapter with a brief look at the history of language testing and then consider the current status of language testing.

In later chapters we provide specific explanations, descriptions, examples, and precautions for the preparation of tests that you and your colleagues might wish to construct as part of your teaching responsibility. In Part I of the book we examine tests of English language features including vocabulary, grammar, and pronunciation tests. In Part II we consider tests of language use including reading, writing, speaking, and listening. The Appendix lists some of the tests used around the world today to measure overall language proficiency such as the Test of English as a Foreign Language (TOEFL) and the Michigan Tests, as well as tests to measure language dominance.

THE IMPORTANCE OF TESTING

How Testing Helps Students Learning English

Well-made tests of English can help students in at least two ways.

First of all, *such tests can help create positive attitudes toward your class.* In the interest of motivation and efficient instruction, teachers almost universally aim at providing positive classroom experiences for their students. There are some important ways that testing can contribute to this aim. One that applies in nearly every class is a sense of accomplishment. In the early 1970s students in an intensive ESL program were being taught from an unstructured conversation-based text. These students complained that while they had ample opportunity to converse in English, they were "not learning anything." Soon afterwards, however, periodic evaluation provided them with a sense of accomplishment that ended their dissatisfaction. Tests of appropriate difficulty, announced well in advance and covering skills scheduled to be evaluated, can also contribute to a positive tone by demonstrating your spirit of fair play and consistency with course objectives.

A second way that English tests can benefit students is by helping them master the language. They are helped, of course, when they study for exams and again when exams are returned and discussed. Where several tests are given, learning can also be enhanced by students' growing awareness of your objectives and the areas of emphasis in the course. Tests can foster learning, too, by their diagnostic characteristics: They confirm what each person has mastered, and they point up those language items needing further attention. Naturally, a better awareness of course objectives and personal language needs can help your students adjust their personal goals. For example, one person might note your strong test emphasis on aural comprehension, and he might also find that he had missed several vocabulary items on a recent test. One logical step would be for him to concentrate on the *meaning* of troublesome words, especially in a spoken context. Learning to spell them or recognize them in a printed context would become a second priority.

In short, properly made English tests can help create *positive attitudes* toward instruction by giving students a sense of accomplishment and a feeling that the teacher's evaluation of them matches what he has taught them. Good English tests also

help students learn the language by requiring them to study hard, emphasizing course objectives, and showing them where they need to improve.

How Testing Helps Teachers of English

We who teach English as a Second or Foreign Language are generally expected to be accountable for the results of our instruction. Our tests can help us answer the important question "Have I been effective in my teaching?" In other words, we can use them to diagnose our own efforts as well as those of our students. As we record the test scores, we might well ask ourselves the following questions: "Are my lessons on the right level? Or am I aiming my instruction too low or too high?" "Am I teaching some skills effectively but others less effectively?" "What areas do we need more work on? Which points need reviewing?" "Should I spend more (or less) time on this material with next year's students?"

And tests can provide insights into ways that we can improve the evaluation process itself: "Were the test instructions clear?" "Was everyone able to finish in the allotted time?" "Did the test cause unnecessary anxiety or resentment?" "Did the test results reflect accurately how my students have been responding in class and in their assigned work?"

Tests, then, can benefit students, teachers, and even administrators by confirming progress that has been made and showing how we can best redirect our future efforts. In addition, good tests can sustain or enhance class morale and aid learning.

THE STATE OF THE ART IN LANGUAGE TESTING

Recent Historical Trends

Language testing today reflects current interest in teaching genuine communication, but it also reflects earlier concerns for scientifically sound tests. Testing during the last century and the early decades of this one was basically *intuitive,* or subjective and dependent on the personal impressions of teachers. After the

intuitive stage, testing entered a *scientific* stage, a time that stressed objective evaluation by language specialists. We are now in a *communicative* stage, a time when we emphasize evaluation of language use rather than language form.

During the long intuitive era, teachers, untrained in testing, evaluated students in a variety of ways. Facts about English often weighed as heavily as skill in using the language. As a result, students had to label parts of a sentence and memorize lists of language patterns (I am, we are, you are, he/she/it is, etc.). Another characteristic of these rather subjective tests was abundant writing in various forms including translation, essay, dictation, précis, and open-ended answers based on reading comprehension. Some of this evaluation was quite sound, especially for advanced students. There was also reliance upon knowledge of grammatical information exemplified by directions such as, "Rewrite the following sentences substituting the present perfect continuous form of the verb."[1]

The scientific era followed the intuitive stage in testing. During the scientific era, many changes occurred. Testing specialists with linguistic training entered the scene. Careful linguistic description suggested that language mastery could be evaluated "scientifically" bit by bit. Objective tests were devised that measured performance or recognition of separate sounds, specific grammatical features, or vocabulary items. These tests often used long lists of unrelated sentences that were incomplete or that contained errors in grammar or usage. Students completed or corrected those sentences by selecting appropriate multiple-choice items. Subjective written tests began to be replaced by objective tests because the latter could be scored consistently even by untrained people. Specialists started to evaluate tests statistically, looking at the effectiveness of each question and measuring the examination's *reliability* and *validity* (see page 178). The new tests and statistical procedures contributed to a growing body of language research. Among the interesting by-products were language aptitude tests, which were designed to predict success in learning a second language.

Tests today are mainly concerned with evaluating real com-

munication in the second language. In this communicative era of testing, we feel that the best exams are those that combine various subskills as we do when exchanging ideas orally or in writing. In particular, communicative tests need to measure more than isolated language skills: They should indicate how well a person can *function* in his second language. Fortunately, in constructing tests today, we don't need to turn our backs on the developments of the scientific period. Earlier, as we recall, there was a movement from rather subjective to highly objective testing. Today's tests tend to use the best features of these two extremes. A relatively new test type, the *cloze*, reflects this compromise: From a story or essay, words are removed at regular intervals (every seventh word, for example). After considering the context, the examinee has to fill in the missing words. Thus, the task is holistic—that is, grammar and vocabulary and overall meaning are tested simultaneously. But the scoring is quite objective.

How exactly do these trends relate to our preparation of classroom tests today? For one thing, we are concerned almost exclusively with measuring skill. *Knowledge* of specific linguistic, literary, or cultural facts may be required in advanced ESL courses, but such information appears only indirectly if at all in our tests of language performance. Good tests can be written without requiring students to use any of the linguistic terminology so common in earlier evaluation. Secondly, certain exam types are now in rather restricted use—for example, translation tests are seldom used nowadays, although they occasionally show up on tests for advanced students, or on specialized tests for translators and on overseas leaving exams where translation is still part of the curriculum. Similarly, précis writing seems generally more appropriate for native speakers or advanced ESL students in an academic setting where the ability to summarize may be important.

Dictation tests, which were criticized during the scientific period, again have rather widespread acceptance. The essay is likewise widely used once more for classroom testing at intermediate and advanced levels. And objective tests are common-

place. Over the years, they have been used to measure all of the language skills. Now, a principal classroom application of objective exams is to evaluate progress in areas such as vocabulary and grammar. Cloze, a relative newcomer in language testing, is already being used around the globe.

Test Classification

We will outline here rather briefly some of the ways tests can be classified. Understanding contrasting exam types can be helpful to teachers since tests of one kind may not always be successfully substituted for those of another kind.

Table 1 · CONTRASTING CATEGORIES OF ESL TESTS

Knowledge Tests	Performance (or Skills) Tests
Subjective Tests	Objective Tests
Productive Tests	Receptive Tests
Language Subskill Tests	Communication Skills Tests
Norm-referenced Tests	Criterion-referenced Tests
Discrete-point Tests	Integrative Tests
Proficiency Tests	Achievement Tests

Let's review the contrasts shown in Table 1. First, tests of *knowledge* are used in various school subjects, from math and geography to literature and language. While ESL knowledge exams show how well students know *facts* about the language, ESL *performance* exams show how well a student can *use* the language. Because today's ESL teachers are concerned with teaching and measuring language skills, this textbook will be concerned only with performance tests.

The second contrast shown in Table 1 is that of *subjective* and *objective* examinations. Subjective tests, like translation and essay, have the advantage of measuring language skill naturally, almost the way English is used in real life. However, many teachers are not able to score such tests quickly and consistently. By contrast, objective exams, such as multiple-choice or matching tests, *can* be scored quickly and consistently.

To continue down the list in Table 1:

Productive measures, like speaking exams, require active or creative answers, while *receptive* measures, like multiple-choice

reading tests, tend to rely on recognition, with students simply choosing the letter of the best answer.

Tests of *language subskills* measure the separate components of English, such as vocabulary, grammar, and pronunciation (Chapters 2, 3, and 4 of this book). *Communication skills* tests, on the other hand, show how well students can use the language in actually exchanging ideas and information (Chapters 5 to 8 of this book).

Another set of contrasting tests is that of *norm-referenced* and *criterion-referenced* exams. Norm-referenced tests compare each student with his classmates (most classroom tests are like this). But criterion-referenced exams rate students against certain standards, regardless of how other students do.

Still another pair of categories is that of *discrete-point* and *integrative* tests. In discrete-point exams, each item tests something very specific such as a preposition or a vocabulary item. Integrative tests are those like dictation that combine various language subskills much the way we do when we communicate in real life.

A final classification is *proficiency* and *achievement* tests. Proficiency tests can measure overall mastery of English or how well prepared one is to use English in a particular setting such as an auto mechanics course or a university. Achievement tests, on the other hand, simply measure progress—gains for example in mastery of count-noncount noun use or mastery of the skills presented in an entire language text or course. This book concentrates on how to prepare achievement tests.

It should be apparent from this discussion that several labels can be applied to any one test. But we normally apply only one pair of labels at a time, just as we do in reference to an individual who might simultaneously be a friend, lawyer, wife, mother, and mayor.

Chapters 2, 3, and 4 of this book deal with measuring performance in language subskills. Chapters 5 through 8 cover integrative and communicative skills. The concluding chapter provides ways of evaluating your tests. Techniques for beginners and children are presented at the beginning of each chapter.