

A POST SECOND GRADE COMPARISON OF
ACHIEVEMENT BETWEEN STUDENTS RECEIVING
BEGINNING READING INSTRUCTION IN
INITIAL TEACHING ALPHABET AND TRADITIONAL ORTHOGRAPHY

A Master's Thesis
Presented to
The Faculty of the Department of Psychology
University of Wisconsin - La Crosse

Submitted in Partial Fulfillment
of the Requirements for the Degree
Master of Science in Education -- School Psychology

by
Randolph Jon Schenkat

August, 1972

WT
72
Sche
c.2

UNIVERSITY OF WISCONSIN - LA CROSSE

THE GRADUATE COLLEGE

Candidate: Randolph Jon Schenkat

We recommend acceptance of this thesis to the Graduate College in partial fulfillment of this candidate's requirements for the degree Master of Science in Education.
The candidate has completed his oral report.

James J. Ryan
Thesis Committee Member

Aug. 4 / 1972
Date

John Cleveland
Thesis Committee Member

Aug 4, 1972
Date

Bill E. Rossum
Thesis Committee Member

Aug 4, 1972
Date

This thesis is approved for the Graduate College:

James H. Erickson
Dean, Graduate College

Aug 17, 1972
Date

ABSTRACT

The purpose of this study was to make a post second grade comparison of achievement between students receiving beginning reading instruction in i.t.a. and T.O.

The experimental group was made up of 138 students who had received first grade instruction using the Early-to-Read i/t/a series during the two academic years 1968-69 and 1969-70. They had continued in the La Crosse Public Schools and had taken the Iowa Tests of Basic Skills (I.T.B.S.) at the beginning of third grade. The same number of students who received beginning reading instruction in the Ginn Series were selected for a control group. The two groups were matched by means of:

- 1) Ginn Pre-Reading Test given at the beginning of first grade,
- 2) sex
- and 3) socio-economic factors.

In comparing the achievement of the two groups, four measures of achievement were used: I.T.B.S. Reading, Spelling, Vocabulary, and Arithmetic scores. The scores of the two groups were classified according to three ability levels; high, medium, and low; by using the Ginn Pre-Reading Test results. An analysis of variance on the four measures of achievement was conducted. In addition, an analysis of variance was done using the means of the i.t.a. and T.O. groups on the Lorge-Thorndike I.Q. Test which was given concurrently with the I.T.B.S. Correlations were also obtained between Reading and I.Q., Reading and Pre-Reading level, Pre-Reading level and I.Q., and Vocabulary and I.Q.

No significant differences were found between the groups on the I.T.B.S. Reading, Spelling, or Arithmetic, however, the i.t.a. group was

significantly higher on the I.Q. test and the I.T.B.S. Vocabulary test. Regardless of method, girls performed at a significantly higher level on the I.Q. test and the I.T.B.S. Reading and Spelling tests. Neither method demonstrated more effectiveness with any particular ability level. A post hoc examination revealed a more perfect I.Q. matching might have resulted in lower results for i.t.a. students in this sample.

The findings indicate that i.t.a. reading instruction did not produce the superior results that its adherents purport. However, study limitations have to be considered: sensitivity of measuring instruments, difference in teaching materials, lack on control over second grade T.O. instruction, and teacher differences.

ACKNOWLEDGEMENTS

In the course of this study I have incurred a debt of gratitude to many individuals who have expended both time and effort assisting me. Lack of space prohibits me from thanking each individually, nevertheless, this paper would not be complete without special acknowledgement of several individuals' assistance.

My greatest debt is to Dr. James Ryan, chairman of my thesis committee for his patience and guidance. I also wish to thank the other members of the committee, Dr. John C. Cleveland and Dr. Richard Rasmussen for their helpful suggestions.

I am also indebted to the La Crosse Public Schools for the cooperation, information and suggestions that staff members provided.

My final debt is to my wife, Cyndy, for her encouragement and understanding.

TABLE OF CONTENTS

CHAPTER I	Page
INTRODUCTION	1
Statement of the Problem	5
CHAPTER II	
REVIEW OF THE LITERATURE	6
Characteristics and Purpose of the Initial Teaching Alphabet	6
Methodological Considerations	8
Early Experimentation with i.t.a.	11
United States Studies	14
Post First Year Studies	19
Sex Differences	21
Transition	22
Additional Effects of Reading Programs	23
CHAPTER III	
THE STUDY	25
Method	27
Instructional Program	29
Measuring Instruments	31
Results	33
Discussion	47
CHAPTER IV	
CONCLUSIONS AND IMPLICATIONS	54
REFERENCES	56

TABLE OF TABLES

	Page
TABLE 1	
Means for the i.t.a. and T.O. Groups on the Ginn Pre-Reading Test	30
TABLE 2	
Means I.Q.'s for the i.t.a. and T.O. Groups on the Lorge-Thorndike	35
TABLE 3	
Summary of Results of Analysis of Variance on the Lorge-Thorndike Intelligence Test	36
TABLE 4	
Means for the i.t.a. and T.O. Groups on the Iowa Reading Test	37
TABLE 5	
Summary of Results of Analysis of Variance on the Iowa Reading Test	39
TABLE 6	
Means for the i.t.a. and T.O. Groups on the Iowa Vocabulary Test	40
TABLE 7	
Summary of Results of Analysis of Variance on the Iowa Vocabulary Test	41

TABLE 8	Page
Means for the i.t.a. and T.O. Groups on the Iowa Spelling Test	43

TABLE 9	
Summary of Results of Analysis of Variance on the Iowa Spelling Test	44

TABLE 10	
Means for the i.t.a. and T.O. Groups on the Iowa Composite Math Tests.	45

TABLE 11	
Summary of Results of Analysis of Variance on the Iowa Composite Math Tests	46

TABLE 12	
Summary of Analysis of Variance Results of Each of the Measures in Terms of the Obtained Significance Levels . . .	48

CHAPTER I

Introduction

Reading is the basic skill on which success in life is predicated. According to Hilton (1972), "Reading is a versatile and dependable tool for learning, and without this tool a child is severely handicapped in the pace of today's life. One does not learn to read but instead learns to read to learn (p. 12)." Yet studies show many Americans are failing to acquire necessary reading skills for living.

The National Reading Council recently completed a research report titled, "Survival Literacy Sample" (Caughran, 1972). Its purpose was to determine the ability of adult Americans to read and complete the myriad of forms that are part of our daily living. The latter included: Social Security, driver's license, Medicaid, bank loans, and public Assistance forms. The results showed that 18.5 million Americans had marginal abilities, 7.1 million experienced considerable difficulties, and 4.3 million would be considered functionally illiterate.

It would seem fitting to question why reading failure is so prevalent in America. Lee's (1969) work sheds some light on this question.

In his study of languages in over thirty countries, he found:

there is cause to believe that it takes a British or American child from one to two years longer to learn to read than it takes children whose mother tongue is say, Italian, Castilian Spanish, Finnish or particularly Russian (p. 55).

Reading in the foreign languages mentioned is less difficult to learn because of the close matching of graphemes (the characters of a language) and the phonemes (the basic sound units of speech from which words are

built and meanings construed). In explaining why Russian is an easy medium for learning beginning reading Lee went on to state:

Russian has 32 character forms for its 32 phonemes, and in addition it makes use, with only a few exceptions of the same character forms for upper case (capital) and lower case (p.55).

English does not have a perfect matching of graphemes and phonemes. It has 26 graphemes in the alphabet while there are 40 frequently identified phonemes. In writing the sound symbols for the 40 phonemes, the 26 graphemes can be arranged in as many as 507 different ways, according to the calculations of Dr. Godfrey Dewey (1969). For example, there are 22 different grapheme combinations for representing the phoneme for the sound of "I".

One specialist, John Money (1966), indicated in the discussion of problems in beginning reading, "the real difficulty for the beginner comes in isolating the components of spoken words, the syllables and their sound units, and writing them, that is turning phonemes into graphemes (p. 26)." Thus the inconsistencies in the spelling of English because of the lack of systematic relationship between the letters of the alphabet and the phonemes of the spoken language appears to be one source of difficulty for the child in his beginning stages of reading.

Sir Cyril Burt in commenting on the importance of reading for young children stated the difficulty in teaching reading was in the English orthography which is more erratic and irregular than any other contemporary language as a result of its composite origin (Downing, 1964). The origins of the English language are in an unsettled Romanic Alphabet being applied to a Teutonic language with varying Saxon and Anglo pronunciations.

Since the time of the first Elizabethans, when our spelling began to solidify and came to be regarded as sacrosanct, teachers and men of letters

such as Sir Isaac Pitman, Bernard Shaw, and Benjamin Franklin have tried out numerous ways of overcoming the difficulties inherent in an inadequate alphabet and of rationalizing spelling (Pitman, 1969). While other countries, upon seeing difficulties in their languages, had Academies to resolve the problems, little change has been made to simplify English. As early as 1570, book titles such as this appeared: "A Methode or comfortable beginning for all unlearned, whereby they may bee taught to read English in a very short time with pleasure" (Downing, 1967a).

Following in line with the group of teachers and men of letters who had attempted changes in the English language is Sir James Pitman, a politician and publisher, whose grandfather invented a form of shorthand.

Because of societal resistance to a full scale change in English orthography, Pitman, instead, concentrated his efforts in a temporary simplification of the orthography for easier mastery in the beginning stages of reading. In the 1950's he began work on what was first called the Augmented Roman system at the University of London Institute of Education. The name, Augmented Roman, was an inappropriate choice, for it was thought the system involved archaic Latin backgrounds and certainly could not result in the simplification of English. Subsequently another title, the Initial Teaching Alphabet (i.t.a.), was adopted as a more appropriate descriptive name to replace the Augmented Roman alphabet.

The Initial Teaching Alphabet consists of 44 characters. This was based on the assumption that in the English language there are 40 frequently identified phonemes. To eliminate the many inconsistencies that he believed caused beginning readers problems such as the 18 separate ways of spelling the vowel sound in the word 'do', he developed an alphabet in which there was a close to perfect matching of

graphemes with the phonemes. As the word initial in the title "Initial Teaching Alphabet" implies, its purpose is as an aid in beginning reading. As soon as a certain level of reading fluency has been achieved with materials presented in the i.t.a. orthography, a transfer is made to materials printed in Traditional Orthography (T.O.). Pitman and other advocates of the i.t.a. orthography felt that several advantages should be gained from its use in this way. They believed that children should be able to progress through the beginning stages of reading more rapidly than the T.O. instructed student, that there should be less ambiguity present in spelling than is the case for T.O. students because of the close phoneme-grapheme correspondence and consequently that children should be able to express themselves in writing earlier. They also feel that pupils will not meet the confusion and failure often encountered in reading with T.O. and thus are more likely to develop positive attitudes toward reading and possibly toward school in general.

Since i.t.a.'s inception in Britain in 1961, this medium has become a common method of beginning reading instruction in the United States and much research has been done to empirically test its merits. Some of this work will be described below. In 1967 the La Crosse School System began instruction on an experimental basis using i.t.a. materials in a beginning reading program. During the initial year, two first grade classes in one school were given reading instruction using the i.t.a. materials. Since that time, four first grade classes in two different La Crosse schools have been taught with the i.t.a. method. This year approximately twelve per cent of the first grade student population of La Crosse is being taught by this method, using the Early-to-Read i/t/a series. Prior to this current study, no objective data has been gathered

to assess the effectiveness of the program in the La Crosse School System.

Statement of the Problem

The problem stated generally is whether beginning reading instruction with i.t.a. materials contributes to any measurable differences in the acquisition of reading proficiency than instruction using materials printed in traditional orthography.

More specifically the problem concerns the effectiveness of instruction with i.t.a. materials in first grade classes in the La Crosse School System. The main objective was to determine the long range effects of beginning reading instruction with i.t.a. relative to T.O. reading approach through a comparison of the achievement levels in reading and related skills attained by the pupils so instructed at the beginning of third grade. In addition to this general comparison, sex differences were also examined.

CHAPTER II

Review of the Literature

Characteristics and Purpose of the Initial Teaching Alphabet (i.t.a.)

Initial Teaching Alphabet consists of 44 characters. Each of the 44 characters of the alphabet represents one of the 40 frequently identified sounds in the English language. In commenting on the use of 44 characters for the 40 phonemes, Pitman explained:

It would have been possible to manage with fewer augmentations; on the other hand it was tempting to introduce even more. A balance had to be struck between simplicity and effectiveness. The fewer the sounds catered for and the fewer the characters, the smaller the burden of learning in the first instance, but too crude a classification of sounds makes the identification of words more difficult; a more precise classification of sounds, in excess of the basic 40 phonemes in English, may make the identification of words somewhat easier, but at the expense of increasing the burden of learning and the difficulty of eventually making the transition to orthodox print. My final choice of 40 phonemes and 44 characters represents a balance that seems to be reasonable (p. 120).

Twenty-four of the characters in the i.t.a. alphabet are identical with characters in the traditional alphabet while each of the remaining 20 characters represents a single phoneme, which in traditional orthography has an inconsistent letter-sound (grapheme-phoneme) correspondence. For example, the 18 lower case print signals for the phoneme "oo" in zoo, too, shoe, flew, do, through, etc. are reduced to only one in i.t.a.

In addition, the i.t.a. orthography does not distinguish between upper and lower case letter characters with capital letters representing only an enlarged lower case letter in i.t.a. Thus the twelve upper case T.O. print symbols for the phoneme "oo" are also reduced to only one in

i.t.a. In the case of the print symbol for "oo" then the 30 T.O. grapheme representations are reduced to one in i.t.a. Of the 20 new characters, 14 resemble the juxtaposition of two T.O. letters and the remaining six are new symbols.

Fewer whole word representations need to be learned because only one style of letter type is used in i.t.a., therefore, each whole word in i.t.a. has only one printed form rather than the variations in visual patterns in T.O. words resulting from the use of upper and lower case and various type styles.

Pitman (1969) made it quite clear that the new alphabet was not a perfect phonetic alphabet as some inconsistencies were retained in i.t.a. The letters C and K, double letters, Y as a vowel and consonant, and the unstressed neutral vowels (urd, uhr, uh) were retained as they appear in T.O. spelling. In defense of these inconsistencies, Pitman contended they remained for the sake of easing the final transition to T.O. Pitman stated:

There are indeed in its writing many compromises with the 'pure theoretical phonetic truth' because of the need to adhere as nearly as possible to orthodox spelling and the conventional appearance of words. Being a medium for teaching reading--or as some may prefer: a learning device--and not a spelling reform, the choice of i.t.a.'s spellings has been very flexible. Its purpose is limited as well as temporary. Phonetic considerations are relevant only in so far as they impede or foster the process of learning to read, the extraction of meaning from the silent page. Each decision has to be weighed and assessed in relation to the eventual objective: to provide an easier, surer way of teaching meaningful reading of texts printed in the roman alphabet and with orthodox spelling, and to infuse the process of learning with enjoyment from the very start. In designing the characters and in deciding how they can be best employed in the representation of each word, this objective has been paramount (p. 140).

As Pitman envisaged the use of i.t.a., once the student had acquired complete fluency in reading and comprehending the materials printed in i.t.a. transfer was to be made to T.O. As mentioned above, Pitman made concessions in the development of i.t.a. to promote an easy transfer to T.O. He anticipated no problems for students making the transfer. Part of his confidence was based on:

the well-known research finding that fluent readers use only minimal cues situated chiefly in the upper part of the line of print. Therefore, as far as possible, the upper part of the i.t.a. configuration of whole words are similar to the upper part of the T.O. configurations of the same words. (Downing, 1967b).

If words that are not recognizable by the use of the upper half of the line of print then Pitman felt that contextual clues would serve as the guide.

More recent research (Downing, 1967b) has shown that smaller units of processing than the top half of configurations of whole words should be considered as a factor in the transfer effect. Studies of errors i.t.a. students make in reading traditional orthography indicate that words which have highly similar configurations but some misleading individual letter or letters in the T.O. spelling (e.g., ch in school, s in island, c in ceiling) seemed to have caused errors in the post-transition T.O. tests.

Methodological Considerations

Before beginning to review the British and American research evaluating the i.t.a. instructional programs, two methodological concerns will be considered: (1) How success in reading is measured and (2) How teaching methods might differ among users of the i.t.a. orthography.

In the United States, evaluation of the i.t.a. reading programs are

most frequently carried out using the Stanford Achievement Series. The use of the Stanford series to evaluate its instructional programs has drawn criticism from Block (1966). He pointed out that most reading achievement tests including the Stanford series are constructed within the framework of assumptions made from the evaluation of the more traditional basal series. Block thought that the i.t.a. approach was based on widely different assumptions. He compared measuring i.t.a. reading achievement with the Stanford series to measuring success of modern math with the traditional math achievement tests. If this argument stood, it would be logical to question the merits of the studies which have investigated the effectiveness of i.t.a. through the use of achievement tests like the Stanford.

Contrary to Block's belief, empirical data has been gathered which seems to support the use of the Stanford series. Bond and Dykstra (1967), who received a federal grant to do a follow-up analysis of 27 United States Office of Education (U.S.O.E.) reading evaluation studies (five of these studies used i.t.a. materials), concluded that, "it is probably not necessary to use different tests to evaluate the reading ability of pupils who learn to read by means of linguistic programs from those used to test reading ability of pupils from the basal reading programs" (p.118). They went on to state in the conclusion of their study, "the high inter-correlations found in the variety of reading measures used in this study indicate that reading at the end of first grade is largely a unitary accomplishment depending upon the ability to recognize words accurately and to associate meaning with those words (p.119)".

Correlations were obtained by Bond and Dykstra to support the emphasis they placed upon the unitary aspects of word recognition as a

principle factor in reading proficiency. "It should be noted that the correlations between the two reading tests (paragraph meaning and word reading) was so high that little differential prediction could be expected (p. 118)." Similarly, at the end of second grade Bond and Dykstra (1968) stated: "the high intercorrelations between the word recognition and paragraph meaning tests indicate that reading ability at the end of the second grade is largely a unitary accomplishment. It is likely at this stage that word recognition is so demanding that comprehension is highly influenced by it (p. 60)."

It would seem that Block's beliefs thus were not borne out by the extensive evaluation of the empirical research by Bond and Dykstra which showed that reading in a unitary accomplishment that appears to be assessed just as adequately by conventional tests, such as the Stanford, without regard to the medium of instruction.

A further methodological consideration is the selection of the specific instructional approach for teaching reading using i.t.a. materials. Most typical is the use of the Early-to-Read i/t/a series which stresses a language-experience approach together with the teaching of the 44 characters of Pitman's alphabet. Pupils are taught to recognize and write each character of the new alphabet and to associate its visual form with its speech equivalent. Intensive training in visual and auditory discrimination accompanies the teaching of each character. After several characters have been introduced, pupils are taught to use an analytical procedure to "unlock" the pronunciation of new words. This approach emphasizes the analysis and synthesis of the characters of the word as the principle method of word identification.

In Britain and in other English speaking countries it is more

common when comparative evaluations are being made for a standard basal series to be reprinted using the i.t.a. alphabet but retaining the exact content and illustrations. A standard British series which is often transliterated is "Janet and John". Britain's foremost researcher on i.t.a., John Downing, feels that unless a transliterated series is used, all variables in the experiment are not comparable. Downing (1970) stated: "It is clearly illogical for i.t. a. invariably to be used with one set of reading materials and T.O. invariably to be used with another set, if we wish to isolate the separate effects of the two media (p.70)."

Mazurkiewicz (1966), author of the Early-to-Read i/t/a series takes a view opposite to Downing contending that the benefits of i.t.a. cannot be achieved to their fullest unless reading is freed from the traditional look-say approach of transliterated basals. He feels that i.t.a.'s merit is more than the use of new augmented alphabet with any reading series. Rather its effect comes from all the components of the Early-to-Read series: the language experience, the training in visual and auditory discrimination, and the word analysis that is employed.

In comparing the outcomes of research on i.t.a. it is important to know what type of approach to beginning reading instruction is involved. Knowing that the Early-to-Read i/t/a series is used gives the evaluator some indication that a consistent approach spelled out in the series' manual was used. With the use of a transliterated basal series, a greater knowledge of teacher individual differences is needed in order to have an adequate idea of what is incorporated in the reading program.

Early Experimentation with i.t.a.

The earliest evaluations of the i.t.a. orthography as a measure instruction in beginning reading were carried out in England by

John Downing, who is one of i.t.a.'s strongest advocates. Because the favorable outcome of these evaluations was the launching point for the rather extensive use of i.t.a. both in Britain and America, the results will be reviewed in some detail.

The first experiment with the i.t.a. method of reading instruction in real class situations was conducted in 1961 by John Downing after Pitman had convinced the University of London's Institute on Education and the National Foundation for Educational Research in England and Wales that his approach to reading instruction had merit and should be tested.

Four hundred thirteen pupils in the beginning classes of 20 infant schools which are equivalent to our primary level (K-2), were instructed with the i.t.a. materials while 687 students in 33 infant schools received beginning instruction in T.O. and served as the control group.

There was an attempt to match the i.t.a. and T.O. classes and schools on such variables as pupil-teacher ratios, urban-rural locations, size of schools, and the type of school organization. Also as a further control, pupils in both the i.t.a. and T.O. groups used the same reading series, "Janet and John" with the texts of the experimental group being printed (transliterated) in i.t.a. In addition, there was an attempt to control for the Hawthorne effect--the effect on teachers and classes produced by the knowledge that their work is the subject of research--by providing as nearly as possible equal treatment for both groups by those conducting research. For example, the control classes received the same number of class visits by researchers as the i.t.a. classes.

Comparisons were made between the two reading programs in terms of (a) the time required for pupils to successfully complete the basal readers (b) pupil performance on the Schonell Graded Word Reading Test (which

measures skills in reading single printed words) and (c) pupil performance on the Neale Analysis of Reading Ability Test which assess three areas of reading achievement: speed, comprehension, and accuracy.

The results according to Downing (1964b) indicated that at the end of the first school year, 62% of the i.t.a. group had passed the beginning of Book 2 in the Janet and John series. By that time only 23% of the control group had done so; while six months later, only 55% in the control group had reached this level of performance. In other words, after one and a half years nearly half of the control group were not yet beyond Book 1, while only 15% of the i.t.a. pupils remained at that stage. At the higher levels, 25% of the i.t.a. group were in Book 4 or beyond after only one year while 3.6% of the control group had gotten that far. Six months later 56.7% of the i.t.a. group were on or beyond Book 4 while only 12.6% of the control group was at that level.

On the Schonell test, which was administered in a transliterated form to i.t.a. students, at the end of the first school year 50% of the i.t.a. group scored 10 or more compared with 16% of the control group. Six months later 48.6% of the i.t.a. group scored at or beyond 35 words on the Schonell test while only 3.5% of the control group scored at that level.

The Neale Analysis of Reading Abilities was administered to a subsample of pupils in both groups (146 i.t.a. and 190 T.O. pupils) after one and a half years of instruction. A transliterated form was used for the i.t.a. group. In all three areas of reading performance measured by the test; speed, comprehension, and accuracy the i.t.a. group showed a significantly higher level of performance ($P < .01$).

In order to obtain an indication of the extent to which the i.t.a.

pupils could transfer their reading skills to the conventional alphabet and spelling a month later an alternate form of the Neale test printed in T.O. was administered to both groups. Although only 50% of the i.t.a. group was reported to have made the transition to T.O., their results were significantly better than those of the control group on the accuracy and comprehension measures ($P < .01$) and on the measure of reading speed ($P < .05$).

Downing (1964b) stated at this point in the research on the effectiveness of i.t.a., "We would only submit that the preliminary evidence indicates that Sir James Pitman's initial teaching alphabet, with its design for transfer to conventional print at the later stage, does appear to be a very promising solution to the problem of teaching beginners to read (p. 62)." Because the preliminary findings by Downing contained strong evidence that i.t.a. did add to the ease of beginning reading instruction, interest in the use of i.t.a. was generated in the United States.

United States Studies

The initial teaching alphabet appeared as a formal approach to reading in American public schools in the fall of 1963. Its first large-scale use and evaluation was in Bethlehem, Pennsylvania (Pitman, 1969). In the Bethlehem research, conducted by Dr. Rebecca Stewart (1965), 114 pairs of first grade students were matched in sex, chronological age, socio-economic status, and I.Q. within two points. The pairs differed on the form of reading instruction they received. The two approaches to reading in the Bethlehem district were in a language arts-centered co-basal reading program and the use of i.t.a. as a medium. The Lower Primary level of

the California Reading Tests which had been administered to all first graders in the city's system at the end of first year of reading instruction was used to evaluate reading performance.

The test results showed that 67% of the 114 students instructed in i.t.a. were reading at or above a 2.5 grade placement while only 39% of the T.O. group were above the 2.5 grade level.

The spelling section of the Stanford Achievement Test, which measured spelling proficiency in Traditional Orthography, was also administered at the end of the year to the i.t.a. and T.O. instructed groups. The results indicated that there was no significant difference between the two populations in the ability to spell as measured by the Stanford spelling test.

In another early study, Tanyzer and Alpert (1965) who compared i.t.a., the Lippincott, and the Scott, Foresman series found that in samples of approximately 200 first grade students for each of the three methodologies that the Lippincott and Early-to-Read i.t.a. groups obtained significantly higher than the Scott, Foresman group on a composite measure of reading proficiency and on all of the subtests of the Stanford Achievement Test. In addition, the Lippincott group showed a significantly higher level of achievement on the Vocabulary and Spelling subtests of the Stanford than the i.t.a. group.

In another study to evaluate i.t.a. in which there was a more complete effort to control for the Hawthorne effect, Fry (1966) compared another new reading system, the Discritical Marking System (D.M.S.) with i.t.a. materials. The D.M.S., which Fry developed, was also a means of increasing the grapheme-phoneme correspondence. Fry felt with two new methodologies, enthusiasm would be induced in teachers and thus compensate for what is usually a more energetic teaching in the first year of a newly

introduced experimental program. Along with the i.t.a., and D.M.S. a third group received beginning reading instruction in a traditional basal series, the Sheldon readers. Fry's method of evaluation involved randomly assigning 21 classes into one of the three methods. Pretests indicated there was essentially no difference among the seven D.M.S., the seven T.O. classes, and the seven i.t.a. classes. The Stanford achievement series showed no statistically significant differences in any of the reading subtests for the three methodologies after one year of reading instruction. In the spelling subtest on the Stanford, the i.t.a. classes were significantly inferior to the D.M.S. and T.O. classes when T.O. spelling standards were applied.

Hahn (1965) also believed the Hawthorne effect could be controlled by "employing relatively unfamiliar programs so that the spontaneity of interest and the excitement associated with using different materials would be a common experience for all (p. 591)." Hahn went on to state "that at the close of the school year most of the teachers were enthusiastic about the programs they had selected (p. 591)." The three instructional approaches used were a traditional basal reading series, i.t.a., using a transliterated British series (Janet and John), and language arts model.

In the study of reading achievement results of the 300 or more first grade students in each of the three matched treatment groups, Hahn found no significant differences on the San Diego County Reading Attitude Survey in the groups but did find the i.t.a. and language arts approach resulted in significantly higher scores than the basal reader approach on the Word Reading subtest of the Stanford.

Chasenoff (1968) took a different approach to control for the Hawthorne effect. One of the teachers of a team consisting of two teachers taught

the T. O. reading program to a self contained first grade class and the i.t.a. method to another class while the other teacher taught the remainder of the academic subjects. Thus an enthusiastic teacher should benefit the i.t.a. and T.O. programs equally according to Chasenoff. Seven two-teacher teams taught 14 classes in his study. In selecting an experimental and control group at the end of the academic year first grade pupils were closely matched on such parameters as: sex, chronological age, I.Q.'s, socio-economic status, reading readiness, sibling placement, and teacher judgement. One hundred sixty-nine students comprised the experimental group, and 150 pupils formed the control. Near the 140th day of instruction the groups were administered the Stanford Achievement Test, Form W. The experimental group received their test in a transliterated form. The total experimental group showed higher mean scores on the Word Reading and Word Study tests that were significant at the .01 level and a mean score on the Spelling test that was higher at the .05 level.

Near the 150th day of instruction another form of the Stanford Achievement Test presented in T.O. was given both groups. There was no significant difference between the mean score for i.t.a. group of 20.37 and the mean for the T.O. group of 20.72.

Krech (1969) "in order to control the differences in methodology which is created by using different reading series (p. 170)" introduced a i.t.a. transliterated form of Scott Foresman's, then experimental series, Dick and Jane, which was used by 397 students in the study. The control group of 165 pupils received beginning reading instruction in the T.O. Dick and Jane series. On the four subtests of the Stanford Achievement series (word meaning, paragraph meaning, vocabulary, and spelling), printed in i.t.a. for the i.t.a. group, which were administered

at the end of a year's reading instruction the i.t.a. group showed significantly superior levels of performance at the .001 level after adjustment for inequalities in the pretest measures.

Ryan (1967) found inconsistencies in the results between comparisons of first year instructed i.t.a. and T.O. students in two different Minnesota cities. In his analysis of the end of first grade results in twelve T.O. and four i.t.a. classes in the St. Cloud system it was found that the i.t.a. classes had developed to a greater degree a higher performance on those reading skills measured by the Stanford Word Reading, Paragraph Meaning, and Vocabulary Tests and by the Gates Reading Test than did the T.O. classes. There was no achievement difference between the T.O. and i.t.a. classes on the Word Study test of the Stanford. On the Stanford Spelling test, which measured spelling proficiency in T.O., the T.O. pupils showed a higher level of performance than the students in the i.t.a. classes.

In 13 i.t.a. and three T.O. first grade classes in Owatonna, a comparison of reading achievement revealed that the i.t.a. group scored higher on the Stanford Spelling test, given only in T.O. than did the T.O. instructed classes. There were no other indications of reading achievement differences between the i.t.a. and T.O. students in this sample.

In summary, an overview of American first year studies does not lead to the consistent pattern of results that Downing found in his British studies. The studies would suggest that at the end of first grade it would be probable to expect inferior spelling performance in the T.O. alphabet by i.t.a. students. In only one study was it found to be inferior to a T.O. based instruction in any area of the Stanford other than

spelling. This was in a comparison of the Vocabulary subtest between Lippincott and i.t.a. instructed pupils. It would be concluded that at the end of first grade instruction i.t.a. except in the area of spelling did not have any detrimental effect and showed some positive gains. However, there was little data to support the confidence that i.t.a. adherents had initially expressed with respect to this approach.

Post First Year Studies

Several of the above mentioned studies considered reading achievement beyond the first year of reading instruction. In decisions regarding effectiveness of an instructional program the long range achievement or performance of pupils could be considered of most significance. Any differences observed immediately following the instruction may be of little significance if they do not persist over time. On the other hand, differences in later reading performance might appear because of a delayed effect in one of the forms of reading instruction.

In the Stewart Study (1965), carried into the second year, objective information was available for measures of word recognition, comprehension, and spelling. At the end of the second year the group receiving its instruction during first grade performed significantly better ($P .01$) than the T.O. instructed students on the spelling subtest of the Stanford Achievement test. The California Reading Test, Upper Primary was administered to 196 pairs of students drawn from the i.t.a. and control populations in April of their second year in school. The results revealed a difference at the .01 level of significance in the word recognition subtest favoring the i.t.a. group over the T.O. students. There was a slight but not statistically significant difference in the comprehension subtest also favoring the i.t.a. instructed group.

Fry (1969) continued his study into the second and third year of reading instruction. He stated, "Silent reading as measured by the Stanford Achievement Test did not show any significant differences among any of the three methods of instruction on any subtest with several minor exceptions (p. 361)."

Ryan (1968), in a post second grade follow-up of his St. Cloud study compared the reading achievement of 76 pupils who had received beginning reading instruction in i.t.a. to the reaching achievement of 161 T.O. taught pupils by an analysis of the achievement data of the Iowa Test of Basic Skills (I.T.B.S.) which was administered early in third grade and the results of the Gates Reading Test. The results showed that the i.t.a. instructed students maintained a higher level of proficiency than the T.O. students along the dimensions of achievement measured by the Iowa Vocabulary test and the Gates Reading tests. Ryan stated, "Despite the rather high correlation between the Iowa Vocabulary and the Iowa Reading tests, a similar difference was not reflected in the latter measure (p. 8)." The results of the Iowa spelling test showed a similar pattern to that shown in first grade with the i.t.a. students' performance inferior to that of the T.O. students.

Correlations were obtained between the measures of reading of i.t.a. students at the end of first grade assessed by the Stanford Achievement Test series and the Reading Test of the I.T.B.S. by Ryan. The correlation of the Stanford Paragraph Meaning to the Iowa Reading Test was .73.

In summary, the limited second grade results presented here would tend to favor i.t.a. over T.O. in some indices of reading achievement such as word recognition in the California Test, Vocabulary in the I.T.B.S., and the Gates Reading Test. Other than the questionable area of spelling

achievement, it can be said in the least that i.t.a. does not produce negative results and there is some evidence that superior results occur in comparison to other reading methods on indications of reading achievement selected in the above studies.

Sex Differences

Studies on sex differences in reading readiness and achievement have consistently indicated that girls are superior to boys.

This is a common finding in all reading approaches. Although this agreement exists there are differing theories as to its causes. It is common that a strictly heredity view supports the girls' superiority. There are, however, studies which do not confirm this belief. An interesting project matched 1338 American children with 1015 German children on such factors as I.Q. and socio-economic status (Preston, 1962). This study found German boys performed better than girls while the opposite was true with American children in reading areas. An explanation of these results would be cultural, suggesting that perhaps in the United States reading and language are approached more as feminine activities.

Bond and Dykstra (1968) in their evaluation and analysis of the 27 U.S.O.E. funded reading studies found, "On the average, girls were superior in reading achievement at the beginning of first grade. Girls were also superior in reading achievement at the end of second grade. In most cases, however, differences in reading achievement disappeared when achievement scores were adjusted for differences in readiness. The superiority of girls in achievement at the end of the second grade was apparently largely a function of their greater readiness at the beginning of first grade. In addition, none of the treatments had a unique effect on the achievement

of boys or girls. The absences of significant sex by treatment interactions indicated that girls tended to be better readers in all programs (p. 63)."

Transition

Transition from the use of the i.t.a. alphabet to traditional orthography seems to be commonly accepted by the layman as a possible stumbling block which has kept i.t.a. from a more wide spread use. Although the research presented thus far cannot unequivocally show that transition is no problem, the following studies have not indicated any adverse effects from i.t.a. instruction.

Shapiro (1971) contended that, "transition was not an incidental or automatic process; rather it appeared to be a complex stage of learning that warranted both comprehensive instruction and more detailed research (p. 57)."

He developed an 87 unit instructional package (G.C.R.P.) which was designed to elicit specific responses identified as crucial to the demonstration of adequate transfer behavior. He assumed in developing this package that the Early-to-Read i/t/a transitional materials (E.T.R.) were not adequate to insure proper transition.

The study designed examined the transitional period of each of three sets of four classes receiving 1) no transitional materials 2) E.T.R. Method 3) G.C.R.P. materials. An analysis of covariance revealed the following results:

1. In the Word Meaning and Word Study Skills, the G.C.R.P. group mean was significantly greater than that of both the E.T.R. and control groups.
2. The mean Spelling scores of both the G.C.R.P. and the control groups were greater than the mean of the E.T.R. group.

3. The Language score of the G.C.R.P. group was significantly greater than that of the E.T.R. group which was significantly greater in turn than that of the control group.
4. There were no significant difference between the three populations on the Stanford Paragraph Meaning test.

It would seem that an important consideration which has not been found in this review of the literature is the extent to which achievement level of i.t.a. instructed children decreases during and following their transition experience. This may well be an indication of the quality of the transitional program rather than the inherent transfer problem.

Additional Effects of Reading Programs

A distinction should be made between the results of standardized spelling tests and spelling errors occurring in free writing. Studies such as Fry's, Tanyzer's, and Ryan's St. Cloud investigation most commonly show deficiencies in the spelling of i.t.a. students when spelling is measured in terms of correct spelling of words in traditional orthography at the end of the first grade on standardized tests. This is common because transfer has not been made by all i.t.a. students. I.t.a.'s merit presumably comes in words attempted in free writing because words are spelled like they sound. Therefore, children show larger writing vocabularies at an earlier age. This is reflected in the length of stories they write.

Chasenoff (1968) and Fry (1969) both report that i.t.a. children wrote longer stories than T.O. students in first grade although these differences were absent by the end of second grade. In Hahn's evaluation of i.t.a. he stated, "Mazurkiewicz and others have shown that children who are provided with writing skills early in the reading program become active independent learners who find delight in expressing their ideas

and in reading the writings of others (p. 590)." There is no objective evidence, however, to clearly prove what Mazurkiewicz claims in this respect.

Hahn's statement leads to a subjective area where perhaps i.t.a. could present its strongest case if reliable measures could be used. Bond and Dykstra (1968) state, "a limitation of most studies of this nature is that it is much more difficult to measure attitudinal aspects of reading than the more mechanical aspects of the reading act...a more important outcome of any reading program is to engender in pupils a desire to read (p.69)."

CHAPTER III

The Study

The purpose of this study was to make a comparison between students who had received their beginning reading instruction with i.t.a. and T.O. materials with respect to their reading achievement at the end of second grade. The main comparison was made using the scores on the Reading test of the I.T.B.S. obtained for all pupils at the beginning of third grade. In addition to the Reading test, comparisons were also made on the Spelling, Mathematics and Vocabulary tests included in the I.T.B.S. These four were selected for the following reasons: The Reading test of the I.T.B.S. according to Herrick (1959) is a:

test of reading comprehension designed to evaluate the specific comprehension skills involved in grasping details and purpose, analyzing organization, and evaluating a reading selection. Most of the items for grade 3, 4, and 5 deal with comprehension of details (Buros, p. 33)."

Thus it will be seen if the decoding skills that are purported to be a product of i.t.a. instruction can carry over to the area of comprehension.

The Vocabulary test was included because of its high correlation (.75) with the Reading test of the I.T.B.S. (Lindquist and Hieronymus, p. 48) and its overt resemblance to the Word Reading subtest of the Stanford Achievement Series, with the recognition, however, that the type of tasks involved in this test are somewhat similar to those typically used in general intelligence tests.

The Spelling test was selected to assess one proficiency related to reading for which there is some evidence, as noted in the review of

the literature, that i.t.a. instruction may not be as effective as instruction initially with T.O.

The measure of mathematics achievement was included in the study because achievement in mathematics would not be expected to be highly related to achievement in reading. Adherents of the i.t.a. methodology emphasized that i.t.a. instruction should result in a general enthusiasm for learning and thereby a better performance in all academic areas because success was met early and consistently at school.

To have a basis for establishing equivalence with respect to existing pupils' abilities relevant to development of reading skills in the first grade, and to assess reading program differences for different readiness ability levels, scores on the Ginn Pre-Reading Test, routinely administered to all La Crosse public school pupils at the beginning of first grade, were also used. As a further check on the general academic ability of the pupils in the comparison groups at the time of the achievement test measures, scores on the Lorge-Thorndike Intelligence Test also administered at the beginning of third grade were obtained.

The general working hypothesis was that beginning reading instruction in i.t.a. would result in a higher level of reading achievement as measured by a standardized reading achievement test at the beginning of third grade than would T.O. instruction. It was also hypothesized that i.t.a. would be found to be a superior medium of instruction in any of the three pre-reading ability levels: high, medium, and low.

Based on the review of the literature, another hypothesis was there would be a difference in the achievement between boys and girls generally independent of type of beginning reading instruction with girls showing a superior performance to boys. It was further hypothesized that this

sex difference would not vary between the two instructional programs.

Also, contrary to some previous evidence, it was hypothesized that for i.t.a. instructed students, performance on the Spelling test would not be significantly inferior to that of T.O. instructed students.

Finally, it was hypothesized that i.t.a. students would obtain a higher level than the T.O. students on the Iowa Composite Arithmetic score.

Method

Subjects. The experimental (i.t.a.) group consisted of all children receiving beginning reading instruction with i.t.a. materials in the first grade in the La Crosse school system for the two academic years 1968-69 and 1969-70, who continued in the system into the third grade and for whom I.T.B.S. test data was available. Children who had been retained were excluded in this study in part because no test data was available for them. For both of the two academic years selected, i.t.a. instruction was provided in the three first grade classes at Josephine Hintgen Elementary School and one first grade class at Franklin Elementary School. There was a total of 80 boys and 58 girls in the experimental group.

The control (T.O.) group, which consisted of an equal number of students who had received first grade reading instruction in T.O. from other teachers during the same years as the experimental sample, was selected on the basis of matching scores on the Ginn Pre-Reading Test and availability of the I.T.B.S. test data.

In selecting the control group, the relative socio-economic status and the overall average reading readiness ability levels of the pupils

in the system's schools were considered. The matching procedure considered the difference that existed with respect to these factors between the two schools participating in the i.t.a. program (Hintgen and Franklin) by identifying two groups of schools that most clearly matched the different socio-economic characteristics and average reading readiness levels of those in each of the i.t.a. schools. Two schools were identified that matched Franklin in addition to its own other T.O. instructed first grade class and six schools were selected that matched Hintgen School. Thus for each i.t.a. school there was established a respective pool of first grade classes from which to select control group pupils. Separate pools of control group classes were identified for each of the two different years in which i.t.a. instruction was provided.

For each i.t.a. pupil in each school, a T.O. pupil of the same sex from the matching pool of comparable classes for the same year was selected whose Ginn Pre-Reading Test score was closest in value to that of the i.t.a. pupil. If there were two or more T.O. pupils who matched the respective i.t.a. pupil in all those respects then the T.O. pupil was randomly selected from among these subsets.

The result was that for each i.t.a. pupil one T.O. pupil of the same sex from a school of similar socio-economic level and average entering level who had an identical or very close Ginn Pre-Reading Test score was selected and included in the T.O. control group. These pupils were then classified into one of three categories (high, medium, low) with respect to their Ginn Pre-Reading Test scores to establish a basis for determining if there were differential reading program effects among those at the different initial reading readiness levels.

There was then exactly the same number of pupils of each sex with

matching pre-reading readiness abilities from comparable schools within each pre-reading readiness ability level who had been instructed with the i.t.a. and T.O. materials respectively in the first grade. The effectiveness of the matching procedures with respect to the pre-reading readiness measure and the numbers of pupils in each sex by ability level by program cross classification category is shown in Table 1.

Instructional Program

Four first grade teachers were involved in i.t.a. instruction. Two of the teachers had taught the only i.t.a. classes in the La Crosse system in the school year 1966-67. Thus for the two academic years covered in this study these two teachers had prior teaching experience in the i.t.a. media. They had volunteered to teach the new reading method and had attended a two day workshop presented by the publishers of the Early-to-Read i/t/a Series before their first year of teaching. The two teachers who began their instruction of the i.t.a. method in the 1967-68 year participated in a one day in-service conducted by the two experienced teachers and did independent study in i.t.a. instruction prior to meeting their first classes. For six of the eight classes considered, the teacher had at least one prior year of i.t.a. teaching experience.

The teachers involved in the instruction of T.O. students who served as the control group came from all but four of the city's schools so characterization regarding years of experience, background in teaching, and their methodology was not possible. There was one distinct common feature between the i.t.a. and T.O. classes in that all teachers instructed self-contained classes.

TABLE 1

Means for the i.t.a. and T.O. Groups on the
Ginn Pre Reading Test

Sex	Males				Females				M & F			
	High N=33	Med N=25	Low N=22	All Levels N=80	High N=29	Med N=20	Low N=9	All Levels N=58	High N=62	Med N=45	Low N=31	All Levels N=138
Pre Ability Level												
Treatment												
i.t.a.	68.48	64.08	55.32	62.63	68.14	64.60	55.55	62.76	68.31	64.34	55.44	62.70
T.O.	68.12	64.16	54.95	62.41	68.00	64.80	57.66	63.49	68.06	64.48	56.31	62.95
Both Programs	68.30	64.12	55.14	62.52	68.07	64.70	56.61	63.13	68.19	64.41	55.88	62.83

The i.t.a. materials used in the four classes were those included in the Early-to-Read i/t/a Series which consisted of seven readers, accompanying workbooks, teacher's manual, and certain instructional aids, such as the i.t.a. plastic character kit, flannel board characters, and flash cards. The T.O. classes were instructed with the Ginn Basic Readers by Russell. This series consisted of three pre-primers, one primer and a reader with the approach being characterized as look-say involving the presentation of words which are learned through use.

In the La Crosse School system the i.t.a. program was planned only for use in the first grade, there is no additional formal instruction with i.t.a. materials in second grade. Perhaps there were some differences in second grade reading instruction between the Franklin and Hintgen classes. There was no attempt at Franklin to keep the i.t.a. group together thus in the school's two second grade classes there were two mixed classes of i.t.a. and T.O. students. At Hintgen, on the other hand, with only three second grades in the i.t.a. students remained as a homogeneous group with the exception of children transferring to the school as second graders.

Measuring Instruments

The Ginn Pre-Reading Test was selected as the pre-instruction test because it is routinely administered to all first grade students upon their entrance into that grade. The scores of this test were used as an indication of potential reading abilities the pupils possessed. The test consists of five subtests: Vocabulary Readiness, Tactile-Visual Readiness, Visual Readiness, Auditory Readiness, and Comprehension Readiness. These five subtests yield a composite score which gives an idea

of overall readiness.

In dealing with the Ginn Pre-Reading Test, the distribution of scores was tabulated and cut off points were established to roughly break the distribution into thirds. The distribution of the Ginn Pre Reading Test (McCullough and Russell, 1961) is typically a J distribution being skewed to the upper end of the test consequently the one-third breakdown was 70-67, 66-61, and 60 and below. This coincides fairly well with the Ginn categories of Superior, Average, and Needs Help and is interpreted as giving an indication of the pre-reading skills a child possessed before starting instruction in reading.

The Lorge-Thorndike Intelligence Test which was given during the same period of time as the I.T.B.S., was selected as a post instructional test because it was considered to provide further information concerning the equivalence in academic ability at the time of the achievement testing between pupils in the experimental and control groups. The Lorge Thorndike Intelligence Test provides three scores: the Nonverbal score, Verbal score, and Full Scale score. For purposes of this study only the Full Scale score was used. In a technical review of this test, Freeman (1959) provided the following evaluation:

On the whole, the Lorge-Thorndike series is among the sounder group instruments available, from the point of view of psychological insights shown in selecting and developing the materials, and from the point of view of statistical analysis of the standardization data (Buros, p. 481).

The Iowa Tests of Basic Skills were used as the main post instructional test because they are used in the city system to provide a continuous evaluation of student achievement and are administered routinely at the beginning of third, fourth, fifth and sixth grade levels.

In commenting on the tests, Morgan (1959) stated:

Technically, the tests are beyond criticism. Items selected from three times as many tryout items were carefully edited for educational and statistical acceptability. Standardization of the test is excellent. It is based on a sample as near as humanly obtainable to a true random sample of U.S. schools (74,000 pupils in 213 school systems) and merits the claim that it is nationwide (Buros, p. 35).

At the third grade level for the composite tests' scores the split half reliability is .97 and for the four measures of achievement considered in this study split half reliability is around .90 (Lindquist and Hieronymus, 1964). Validity is referred to as curricular which means that identification and definition of the skill processes being tested was done before the test items were devised.

Herrick (1959) made the following comments:

Intercorrelations among the various subtests range from .37 to .83, with the average ranging from .60 to .70. As one would expect, the tests of vocabulary and reading comprehension have the highest intercorrelation with all other subtests, indicating a heavy loading of all subtests with vocabulary and reading skills. This suggests that a factor analysis of these data would reveal that 2 to 3 skills are really being tested rather than the 5 to 11. The intercorrelations suggest further that not much would be left after the effect of the vocabulary and reading skills were removed. (Buros, p. 32).

Results

A Pearson Product Moment Correlation was run between the Ginn Pre-Reading Tests and the I.T.B.S. Reading scores to determine the extent that the control variable (pre-test scores) related to the main achievement variable. For the girls in both groups this correlation was .32 and for the boys it was .36.

Because of the relatively low correspondence between the pre-test

matching and this achievement measure, the Lorge-Thorndike Full Scale I.Q. data was examined to determine the comparability of the two groups. These results are shown in Table 2 which shows the mean I.Q. scores for i.t.a. and to students classified according to sex and reading readiness ability. To test the difference among these means with respect to each variable a three-way (2x2x3) factorial analysis of variance design was used. In this analysis, one dimension represented the two alternate programs, another dimension represented sex, and the third dimension represented the three abilities established by the pre-reading test. Since there were unequal numbers of subjects in the sub-groups the analysis was carried out using an unweighted mean solution as recommended by Winer (1962, pp. 242-243). This analysis was also used to test the difference among the means on each of the achievement measures.

Table 3 presents a summary of the results of the analysis of variance on the means of the Lorge-Thorndike Intelligence Test. The data revealed that there was a significant difference ($P < .05$) between the i.t.a. and T.O. instructed students on the Lorge-Thorndike Intelligence Test. There was also a significant difference ($P < .05$) between the I.Q. level of boys and girls in the study with girls having higher I.Q.'s. It further showed that the matching by ability level as determined by the Ginn Pre-Reading Test differentiated significantly between the three levels; high, medium, and low; at the $P < .01$ level. There were no significant differences indicated for any of the interactions among these variables. These results indicated that pupils in the two reading instruction groups were not adequately matched with respect to this measure of intelligence obtained at the beginning of third grade.

Table 4 shows the mean scores obtained in the I.T.B.S. Reading Test

TABLE 2

Means I.Q.'s for the i.t.a. and T.O. Groups on the Lorge-Thorndike

Sex	Males				Females				M & F			
	High N=33	Med N=25	Low N=22	All Levels N=80	High N=29	Med N=20	Low N=9	All Levels N=58	High N=62	Med N=45	Low N=31	All Levels N=138
Pre Ability Level												
Treatment												
i.t.a.	108.88	96.28	94.36	99.84	118.21	107.05	103.33	109.53	113.55	101.67	98.85	104.69
T.O.	101.90	96.72	94.59	97.74	111.00	95.33	92.50	99.61	106.45	96.10	93.55	98.67
Both Programs	105.39	96.50	94.48	98.79	114.61	101.19	97.94	104.57	110.00	98.85	96.20	101.68

TABLE 3

Summary of Results of Analysis of
Variance on the Lorge-Thorndike Intelligence Test

Source of Variation	SS	df	MS	F
A Reading Program	810.15	1	810.15	5.426*
B Sex	749.78	1	749.78	5.022*
C Pre Ability Level	3207.2	2	1603.6	10.741*
A x B	341.94	1	341.94	2.29
A x C	13.62	2	6.81	<1
B x C	137.5	2	68.8	<1
A x B x C	162.74	2	81.37	<1
Error (within cells)	39415.46	264	149.3	

*P < .05

TABLE 4

Means for the i.t.a. and T.O. Groups on the
Iowa Reading Test

Sex	Males				Females				M & F			
	High N=33	Med N=25	Low N=22	All Levels N=80	High N=29	Med N=20	Low N=9	All Levels N=58	High N=62	Med N=45	Low N=31	All Levels N=138
Pre Ability Level												
Treatment												
i.t.a.	3.81	2.69	2.67	3.05	4.40	3.71	3.32	3.81	4.10	3.20	3.00	3.43
T.O.	3.41	3.29	2.67	3.12	3.99	3.38	2.63	3.33	3.70	3.33	2.65	3.23
Both Programs	3.60	2.99	2.67	3.09	4.19	3.54	2.98	3.57	3.90	3.27	3.82	3.33

Note: Means expressed in Grade Equivalent Scores

for i.t.a. and T.O. Students classified according to sex and reading readiness ability. Table 5 gives a summary of the results of the analysis of variance on the means of the reading test of the I.T.B.S. It revealed that there was no significant difference in reading achievement between i.t.a. and T.O. instructed students as measured by the Iowa Reading Test. It further showed that there were significant differences in either methodology between the reading performance of girls and boys with the girls' performance reliably better ($P < .05$) than that of the boys. The matching by ability level as determined by the Ginn Pre-Reading Test differentiated significantly between the three levels at the $P < .01$ level. There were no significant differences indicated for any of the interactions among these variables. On the measure of reading achievement there were no statistically reliable differences between the students in the two reading programs.

Table 6 shows the mean scores obtained in the I.T.B.S. Vocabulary Test for i.t.a. and T.O. students classified according to sex and reading readiness ability. Table 7 presents a summary of the results of the analysis of variance on the means of the Vocabulary Test of the I.T.B.S. It revealed that there was a significant difference at the $P < .01$ level favoring the achievement of the i.t.a. instructed students over the T.O. instructed students on this measure. The analysis showed there was no significant difference between the performance of the boys and girls on the vocabulary test. It further showed that the matching by ability level as determined by the Ginn Pre-Reading Test differentiated significantly between the three levels at the $P < .01$ level. There were no significant differences indicated for any of the interactions among the variables. Thus on this alternate measure of reading achievement there were

TABLE 5

Summary of Results of Analysis of
Variance on the Iowa Reading Test

Source of Variation	SS	df	MS	F
A Reading Program	18.86	1	18.86	<1
B Sex	520.10	1	520.10	4.939*
C Pre Ability Level	1744.29	2	872.14	8.283**
A x B	168.30	1	168.30	1.598
A x C	130.20	2	65.10	<1
B x C	34.59	2	17.29	<1
A x B x C	86.41	2	43.22	<1
Error (within cells)	27796.8	264	105.29	

* P .05

** P .01

TABLE 6

Means for the i.t.a. and T.O. Groups on the
Iowa Vocabulary Test

Sex	Males				Females				M & F			
	High N=33	Med N=25	Low N=22	All Levels N=80	High N=29	Med N=20	Low N=9	All Levels N=58	High N=62	Med N=45	Low N=31	All Levels N=138
Pre Ability Level												
Treatment												
i.t.a.	3.97	3.13	2.71	3.27	4.37	3.84	3.23	3.81	4.17	3.48	2.47	3.54
T.O.	3.32	3.02	2.42	2.92	3.56	3.05	2.34	2.98	3.44	3.03	2.38	2.95
Both Programs	3.64	3.07	2.57	3.09	3.48	3.44	2.79	3.40	3.81	3.26	2.68	3.25

Note: Means expressed in Grade Equivalent Scores

TABLE 7

Summary of Results of Analysis of
Variance on the Iowa Vocabulary Test

Source of Variation	SS	df	MS	F
A Reading Program	773.74	1	773.74	8.378*
B Sex	207.85	1	207.85	2.251
C Pre Ability Level	1902.82	2	951.41	10.302*
A x B	131.46	1	131.46	1.423
A x C	24.58	2	14.79	<1
B x C	8.29	2	4.147	<1
A x B x C	29.40	2	14.70	<1
Error (within cells)	24381	264	92.352	<1

*P <.01.

statistically reliable differences favoring the i.t.a. instructed students.

Table 8 shows the mean scores obtained in the I.T.B.S. Spelling test for i.t.a. and T.O. students classified according to sex and reading ability. Table 9 presents a summary of the results of the analysis of variance on the means of the Spelling test of the I.T.B.S. It revealed that there was no significant difference in spelling achievement as measured by the I.T.B.S. between i.t.a. and T.O. instructed students. It further showed that there were significant differences in either methodology between the spelling performance of girls and boys with the girls' performance reliably better ($P < .05$) than that of the boys. The matching by ability level as determined by the Ginn Pre-Reading Test differentiated significantly between the three levels at the $P < .01$ level. There were no significant differences indicated for any of the interactions among the variables. On the measure of spelling achievement there were no statistically reliable differences between the students in the two reading programs.

Table 10 showed the mean scores obtained in the I.T.B.S. Composite Math Scores for i.t.a. and T.O. students classified according to sex and reading ability. Table 11 presents a summary of the results of the analysis of variance on the means of the Mathematics Composite scores on the I.T.B.S. The analysis revealed there was no significant difference in mathematics achievement as measured by the I.T.B.S. between the i.t.a. and T.O. instructed students. It also showed there was no significant difference between the performance of the boys and girls on the mathematics measures. It further showed the matching by ability level as determined by the Ginn Pre-Reading Test differentiated significantly between the three levels at the $P < .05$ level. There were no significant differences indicated for any of the interactions among the variables.

TABLE 8

Means for the i.t.a. and T.O. Groups on the
Iowa Spelling Test

Sex	Males				Females				M & F			
	High N=33	Med N=25	Low N=22	All Levels N=80	High N=29	Med N=20	Low N=9	All Levels N=58	High N=62	Med N=45	Low N=31	All Levels N=138
Treatment	i.t.a.				T.O.				Both			
	3.49	2.75	2.61	2.96	4.44	3.49	3.47	3.80	3.97	3.12	3.04	3.37
Pre Ability Level	3.46				3.08				2.52			
	3.48	2.91	2.57	2.99	4.07	3.33	3.19	3.43	3.78	3.12	2.88	3.26

Note: Means expressed in Grade Equivalent Scores

TABLE 9

Summary of Results of Analysis of
Variance on the Iowa Spelling Test

Source of Variation	SS	df	MS	F
A Reading Program	118.93	1	118.93	1.091
B Sex	666.76	1	666.76	6.115*
C Pre Ability Level	1283.91	2	641.95	5.887**
A x B	202.18	1	202.18	1.854
A x C	66.68	2	33.34	<1
B x C	17.67	2	8.84	<1
A x B x C	5.66	2	2.83	<1
Error (within cells)	28787	264	104.04	

* P <.05

** P <.01

TABLE 10

Means for the i.t.a. and T.O. Groups on the
Iowa Composite Math Tests

Sex	Males				Females				M & F			
	High N=33	Med N=25	Low N=22	All Levels N=80	High N=29	Med N=20	Low N=9	All Levels N=58	High N=62	Med N=45	Low N=31	All Levels N=138
Pre Ability Level												
Treatment												
i.t.a.	3.70	3.13	2.94	3.26	3.95	3.42	2.93	3.43	3.82	3.27	2.94	3.34
T.O.	3.50	3.17	2.97	3.22	3.53	3.20	2.71	3.15	3.52	3.18	2.84	3.18
Both Programs	3.60	3.15	2.96	3.24	3.74	3.31	2.82	3.29	3.67	3.23	3.89	3.26

Note: Means expressed in Grade Equivalent Scores

TABLE 11

Summary of Results of Analysis of
Variance on the Iowa Composite Math Tests

Source of Variation	SS	df	MS	F
A Reading Program	60.41	1	60.41	1.160
B Sex	16.17	1	16.17	<1
C Pre Ability Level	915.74	2	457.87	8.789**
A x B	24.15	1	24.15	<1
A x C	22.45	2	11.23	<1
B x C	30.31	2	15.16	<1
A x B x C	9.88	2	4.94	<1
Error (within cells)	13753			

** P <.01

Table 12 is a summary of the analysis of variance results of each of the measures in terms of the obtained significance levels.

Discussion

The results of the study appear to offer mixed support for the general working hypothesis that beginning reading instruction in i.t.a. would result in a higher level of reading achievement as measured by a standardized reading achievement test at the beginning of third grade than would T.O. instruction. The results of the Iowa Reading Test revealed no significant difference between the i.t.a. and T. O. instructed students.

The Iowa Vocabulary Test, on the other hand, showed a significantly higher level of performance by the i.t.a. students. Since the Iowa Vocabulary Test was selected because of its correlation with the Reading Test on the I.T.B.S. and because of its overt similarity to the word reading task, it would seem logical to conclude that the i.t.a. students were superior to those in the T.O. group on reading proficiency as measured by this test.

However, the i.t.a. group was also found to have a significantly higher level of intelligence than the T.O. group of pupils as measured by the Lorge-Thorndike Intelligence Test. Consequently, because of this sample difference in intelligence, a correlation between the Lorge-Thorndike Intelligence Test and the Iowa Vocabulary Test for the study's population was obtained. The correlation between these two measures was .73. Such a high correlation suggests that the vocabulary test of the I.T.B.S. may be as much a measure of general intelligence as a measure of reading proficiency.

As shown, pupils in the experimental (i.t.a.) and control (T.O.) samples

TABLE 12

Summary of Analysis of Variance Results of Each
of the Measures in Terms of the Obtained Significance Levels

VARIABLES	IQ	Reading	Vocabulary	Spelling	Mathematics
Reading Program	.05*	-	.01*	-	-
Sex	.05**	.05**	-	.05**	-
Pre Ability Level	.01	.01	.01	.01	.01
x B	-	-	-	-	-
x C	-	-	-	-	-
x C	-	-	-	-	-
x B x C	-	-	-	-	-

ote: - no significant difference
* difference favoring i.t.a.
** difference favoring girls

were very closely matched on pre-reading ability as measured by the Ginn Pre-Reading Test. However, two years later in the first group administration of intelligence tests there was a disparity in intelligence between the two groups. Although reading readiness tests are not designed as intelligence measures per se, it seems reasonable to expect that pupils closely matched on a reading readiness measure should be at least roughly equivalent on general intelligence. The Lorge-Thorndike Intelligence Test results indicated they were not, but rather that the i.t.a. pupils were superior in that respect. There seems to be a fitting explanation for this discrepancy between the i.t.a. and T.O. groups in intelligence at the beginning of third grade that was not reflected in the first grade measure of readiness. A subsequent inquiry revealed that with one exception, kindergarten instruction in the La Crosse Schools involved the use of Ginn readiness materials. The one exception was Hintgen School, which contributed 108 of the 138 students in the i.t.a. sample. Kindergarten teachers at Hintgen, where no first grade classes were instructed in T.O., used more teacher made readiness materials and made more use of the Frostig program than they did the Ginn readiness program.

Consequently, it would seem that given a T.O. and i.t.a. student in this sample who were matched on the Ginn Pre-Reading Test, it would follow that the i.t.a. student had more school ability intelligence, as measured by the Lorge-Thorndike, because the T.O. student through his direct use of the Ginn pre-reading materials had been better prepared for the Ginn Pre-Reading Test. That is, it appears that because the i.t.a. pupils had not received the same type or amount of instruction on the reading readiness test related materials, their performance on the reading readiness

test was not as high as that of T.O. pupils of equal intelligence. They were consequently matched with lower intelligence T.O. pupils who were equivalent on the reading readiness measure.

It would seem advisable when conducting reading evaluation studies to carefully consider curricular differences in kindergarten that may exist between groups being compared and to select as pre ability measures instruments which take relevant differences into account. In retrospect, for this study an intelligence test and a pre-reading ability test less sensitive to the specific pre-reading program would have been more appropriate.

Thus the reading program difference favoring i.t.a. over T.O. instructed students on the Iowa Vocabulary Test seems from the post hoc examination more adequately explained by the mismatching of I.Q. level in the original pairing of the i.t.a. and T.O. students than by the evidence of a higher i.t.a. performance on a vocabulary test that has a resemblance to the Stanford Word Reading subtest. It is evident at the same time however, that even though they exhibited a higher average level of general intelligence, the i.t.a. pupils did not show a higher average level of reading proficiency as measured by the Reading Test of the I.T.B.S. which was initially assumed to be the principle index of reading achievement.

In fact, given the high correlation between reading and intelligence, if the present study had been undertaken with adequately matched, on the basis of I.Q., experimental and control groups it is quite likely that the T.O. group would have shown a superior achievement on the I.T.B.S. Reading Test. Whether this difference would have been statistically different can not be readily estimated.

It was hypothesized that the experimental (i.t.a.) group would not differ from the T.O. group on the I.T.B.S. spelling test. The results from the study confirm this hypothesis. This finding does not unequivocally counter the belief that i.t.a. produces inferior spellers because the spelling section of the I.T.B.S. consists only of finding the incorrectly spelled word in a series of five words which is different from the "spelling bee" approach of the Stanford series. It was in the spelling of orally presented words as measured by the latter test that critics of i.t.a. foresaw difficulties. The fact that the two groups differ in general intelligence must be recognized in interpreting the results as was noted in connection with the reading test.

As is generally found in instructional programs which compare boys and girls, girls in the present study showed superior achievement on the measures of reading and spelling regardless of beginning reading methodology. However, the analysis of variance on the means of the Lorge-Thorndike Intelligence test also indicated that the girls had a higher average intelligence than the boys in the sample. It is possible that some of the sex differences observed here were a consequence of the initial mismatching with respect to intelligence.

Although the correlations between pre ability levels and I.Q. and reading are not of a large magnitude (.32-.38), there were statistically significant differences on all of the I.T.B.S. tests among the high, medium, and low ability levels which were established by the Ginn Pre-Reading Test. However, this difference was observed across both instructional groups and since there was no significant interaction between instructional program and reading readiness level, the hypothesis that i.t.a. would be found to be a superior medium of instruction for

any of the three ability levels must be rejected. This indicates that i.t.a., which was designed to make beginning reading instruction easier and consequently provide greater help to those more likely to have difficulty learning to read (i.e. the lower ability students) did not have such an effect for the pupils in this study.

The hypothesis that i.t.a. students would show superior ability in mathematics because of the belief that i.t.a. effects school attitude and allows an early, consistent feeling of success which generalizes to other academic areas such as mathematics was not supported by the results. This result is not surprising for the hypothesis, as it was stated, assumed that before a significant difference for i.t.a. students in the Iowa Mathematics Achievement measure was shown there would have to be significant changes in other academic achievement areas more closely related to reading, which in fact were not observed. Here again the I.Q. differences in the two groups should be considered.

Assuming the Iowa Vocabulary Test is more a measure of intelligence than of reading proficiency, in general the study revealed no differences in the level of reading achievement between the i.t.a. and T.O. students on the third grade administration of the I.T.B.S.

At the same time it must be recognized that other factors in the instructional program could have contributed to these findings. For instance, because i.t.a. was only used on a trial basis the additional resources that the i.t.a. teachers had available were limited while the use of the T.O. method, which had been employed for several years, allowed for the accumulation of a much greater variety and quantity of resource materials.

If the study had been planned prior to the implementation of the

experimental (i.t.a.) reading program, a more careful selection of pre and post test measures might have shown differing results.

As has been discussed above the Ginn Pre-Reading Test did not prove to be an adequate instrument for initial matching. Regarding the use of the I.T.B.S. as a post measure, Ryan (1968) stated the tests, as a measure of reading achievement level tend to measure more general proficiencies rather than those specific to reading as such. A diagnostic reading test would possibly have been a more sensitive instrument to assess the specific kinds of reading skill differences between the two programs being examined here. This position, however, is not in line with the earlier statements by Bond and Dykstra which suggest that early reading skills are a single unitary trait, although they did not consider the I.T.B.S. as a post second grade measure in the studies they evaluated.

Whatever the measures selected, it is not likely that the i.t.a. group would perform at a superior level as the current results indicated the i.t.a. group with its higher level of intelligence did not achieve significantly higher scores on the measures of the Iowa Tests used in this study.

CHAPTER IV

Conclusions and Implications

The present study seems to offer little evidence to support the use of the Initial Teaching Alphabet. This finding is similar to the findings of the studies discussed in the review of the literature which for the most part indicate no conclusive results regarding the question of the effectiveness of i.t.a. as a beginning medium of reading instruction.

Adherents of i.t.a. suggest that instruction in i.t.a. develops larger writing vocabularies, allows a child more ability to express himself in writing, and provides a more positive attitude toward reading and school in general than does T.O. instruction. This study, with the limited post measures that were employed, did not offer an insight into these areas.

One factor to be considered in the post second grade comparisons of i.t.a. and T.O. achievement results is how much of the i.t.a. pupils' achievement at the end of second grade was a function of their initial instruction in i.t.a. and how much is a function of their later instruction in some other type of program.

As was discussed in the review of the literature, Pitman's original conception of how transfer was to take place was shown in later research not to be a full explanation of the process. The initial teaching alphabet was never purported to be the ultimate in the ease of teaching beginning reading and perhaps modifications can be made in the system that will

facilitate further use.

In considering the American studies that showed the success of i.t.a., it is important to recognize the type of i.t.a. reading program used. With the use of the Early-to-Read i/t/a series it is fitting to question whether the results achieved are accounted for in the use of the i.t.a. orthography or by the more structured use of a synthetic approach which are both incorporated in the design of the Early-to-Read i/t/a series. The Tanyzer and Alpert study showed Lippincott and i.t.a. superior to the basal series used in their study on all subtests of the Stanford while the Lippincott series was superior to i.t.a. on the vocabulary and spelling subtests. This might serve as an indication that the method of teaching (both used a synthetic approach) is more important than the i.t.a. orthography.

Perhaps a more basic question to consider in evaluations of reading programs is teacher differences. It is a quite common finding, for example Fry's study, that reading program comparisons show greater variations between classes in reading results that are shown in reading treatment differences. In looking for the sources of these variations, teacher differences play a large factor yet it is very difficult to objectively weigh these considerations in reading program evaluations.

Of final consideration in reading evaluations is, even though statistically significant differences may or may not be found, what are the practical implications of the selection of one reading program over another in various areas such as teacher training and skills and costs involved.

BIBLIOGRAPHY

- Asher, J.W. Comment on spelling achievement following i.t.a. instruction. Reading Teacher, 1968, 21, 153-6.
- ✓ Bleismer, E.P. and Yorborough, B.H. A comparison of ten different reading programs in first grade. Phi Delta Kappa, 1965, 46, 500-4.
- Block, J.R. A critique of research with the initial teaching alphabet and some recommendations. The i.t.a. Foundation Report, Vol. 1, nos. 2 and 3 (Combined Issue) Summer-Fall 1966, 32-42.
- Bond, G.L. and Dykstra, R. The cooperative research program in first grade reading instruction. Reading Research Quarterly, 1967, 2, 5-141.
- Bond, G.L. and Dykstra, R. Research program in first grade reading instruction: a follow-up. Reading Research Quarterly, 1968, 4, 49-70.
- Caughran, A.M. Should the survival literacy sample survive. Journal of Reading, 1972, 15, 429-35.
- Chasenoff, R.E. Two alphabets: a follow-up. Elementary School Journal, 1968, 68, 251-7.
- Chasenoff, R.E. Two alphabets. Elementary School Journal, 1967, 67, 257-64.
- Dapper, G. It's ten years later. Instructor, 1971, 80, 90-1.
- Dewey, G. English heterography or How we spel! Lake Placid Club Education Foundation, New York Revised edition, 1963. Ed. by J. Pitman, Alphabets and Reading. New York: Pitman, 1969, p.44.
- ✓ Downing, J. The Initial Teaching Alphabet. New York: Macmillan, 1964a.
- Downing, J. Teaching Reading with i.t.a. in Britain. Phi Delta Kappa, 1964b, 45, 322-29.
- Downing, J. How i.t.a. began. Elementary English, 1967a, 44, 40-6.
- Downing, J. What's wrong with i.t.a.? Phi Delta Kappa, 1967, 48, 262-66.
- Downing, J. British i.t.a. research. Reading Teacher, 1968, 21, 640-6.
- Downing, J. Alternative Teaching Methods in i.t.a. Elementary English, 1968, 68, 942-51.

- Downing, J. Cautionary comments on some American i.t.a. reports. Educational Research, 1970, 73, 70-2.
- Downing, J. A closer scrutiny of research data on i.t.a. Education, 1970, 90, 308-12.
- Downing, J. A psycholinguistic theory for i.t.a. Elementary English, 1970, 47, 953-61.
- Downing, J. I.t.a. American vs. British experience. Phi Delta Kappa 1971, 52, 416-17.
- Downing, J. Is i.t.a. atomistic? Reading Teacher, 1971, 25, 56-7
- Enstrom, E.A. Wanted: unbiased answers. Elementary English, 1967, 44, 748.
- Freeman, F.S. Review of Lorge-Thorndike Intelligence Test in O.K. Buros (Ed.) The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1957.
- Fry, E.B. Comparing the diacritical marking system, i.t.a., and a basal reading series. Elementary English, 1966, 43, 607-11.
- Fry, E.B. Comparison of beginning reading with i.t.a. DMS and T.O. after three years. Reading Teacher, 1969, 22, 357-62.
- Hahn, H.T. Three approaches to beginning reading instruction i.t.a., language experience, and basic readers. Reading Teacher, 1966, 19, 590-94.
- Hayes, R.B. I.t.a. and three other approaches in first grade extended into second grade. Reading Teacher, 1967, 20, 694-7.
- Hayes, R. B. Three year look at i.t.a., Lippincott, phonics and word power, and Scott, Foresman. Reading Teacher, 1969, 22, 363-70.
- Hayes, R.B. I.t.a. and three other approaches to reading in first grade. Reading Teacher, 1966, 19, 627-30.
- Herrick, V.E. Review of Iowa Tests of Basic Skills in O.K. Buros (Ed) The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1959.
- Hilton, E. Let's give reading top priority. Instructor, 1972, 81, 12.
- Holmes, J.A. Disadvantaged children and the effectiveness of i.t.a. The Reading Teacher, 1967, 22, 350-6.
- Kresh, E. Comparative effectiveness of i.t.a. vs. T.O. in first grade in the Pittsburgh public schools. Journal of Educational Research, 1969, 63, 170-2.

- Lee, W.R. Spelling Irregularity and Reading Difficulty. National Foundation for Educational Research in England and Wales, London, 1960. Cited by J. Pitman, Alphabets and Reading. New York: Pitman, 1969, 55.
- Leigh, T. I.t.a. and progressive teaching methods. The Reading Teacher, 1968, 22, 148-52.
- ✓ Lindquist, E.F. and Hieronymus, A.N. Manual for Administrators, Supervisors, and Counselors Iowa Test of Basic Skills. Boston: Houghton Mifflin, 1964.
- Mazurkiewicz, A. J. I.t.a. and T.O. reading achievement when methodology is controlled: extended into second grade. The Reading Teacher, 1967, 20, 726-27.
- Mazurkiewicz, A. H. I.t.a. and T.O. reading achievement when methodology is controlled. The Reading Teacher, 1966, 19, 606-10.
- McCullough, C. M. and Russell, D. H. Manual and Answer Key for Teachers: Ginn Pre-Reading Test. Boston: Ginn, 1961.
- Money, J. The Disabled Reader. Baltimore: John Hopkins Press, 1966.
- Morgan, G.A.V. Review of Iowa Tests of Basic Skills in O.K. Buros (Ed.) The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1959.
- Ohanian, V. Control populations in i.t.a. experience. Elementary English, 1966, 43, 373-80.
- Paxson, W. Wanted: unbiased questions. Elementary English, 1967, 44, 748.
- Peters, M.L. Influence of reading method on spelling. British Journal of Educational Psychology, 1967, 37, 47-53.
- Pitman, J. Alphabets and Reading, the Initial Teaching Alphabet, New York: Pitman, 1969.
- Preston, C. in Bond, G.L. and Dykstra, R. Research program in first grade reading instruction. Reading Research Quarterly, 1967, 2, 5-141.
- Ryan, J.J. A comparison of beginning reading achievement for pupils instructed with materials using i.t.a. and traditional orthography. Report No. 67-2. Minnesota National Laboratory, Minnesota State Department of Education, March, 1967.
- Ryan, J.J. A follow-up study of the effectiveness of the i.t.a. beginning reading program. Minnesota National Laboratory, Minnesota State Department of Education, December, 1968.

- Sabaroff, R.E. Comparative investigation of six reading programs. Education, 1971, 91, 303-14.
- Seymour, D.Z. Three linguistic problems using i.t.a. The Reading Teacher, 1968, 21, 422-6.
- Shapiro, B.J. The i.t.a. transition problem: a comparative study. Journal of Educational Research, 1971, 65, 57-60.
- Shapiro, B.J. I.t.a.--kindergarten for first grade. The Reading Teacher, 1967, 22, 307-311.
- Stewart, R.W. I.t.a. after two years. Elementary English, 1965, 42, 660-5.
- ✓ Strong, R. The Improvement of Reading. New York: McGraw Hill, 1967.
- Tanyzer, H. J. and Alpert, H. Effectiveness of three different reading systems on first grade reading achievement. Final report Cooperative Research Project No. 2720, Hofstra University, Homstead, New York, 1965.
- Tanyzer, H.J. and Alpert, H. Three different basal reading systems and first grade reading achievement. The Reading Teacher, 1966, 19, 636-42.
- Wijk, A. I.t.a. vs. regularized English. Education, 1968, 88, 300-7.
- ✓ Winer, B.J. Statistical Principles in Experimental Design. New York: McGraw Hill, 1962.