DOCUMENT RESUME

ED 044 260 RE 003 051

AUTHOR Block, J. R.
TITLE A Critique of Research with the Initial Teaching Alphabet and Some Recommendations.
PUB DATE 66
NOTE 11p.
JOURNAL CIT i.t.a. Foundation Report; v1 n2, 3 p32-42 Summer-Fall 1966

EDRS PRICE EDRS Price MF-$0.25 HC-$0.65
DESCRIPTORS Beginning Reading, *Initial Teaching Alphabet, *Measurement Instruments, *Reading Research, Reading Tests, *Research Design, Teacher Attitudes

ABSTRACT

Research on the Initial Teaching Alphabet (i.t.a.) is viewed from two aspects--its limitations and how it can be improved. Thus far research has indicated that (1) transition from i.t.a. to traditional orthography is not the problem expected and (2) the children who are taught with i.t.a. read no better or worse than children taught with the conventional alphabet. Discussed are errors in research design related to (1) the use of control and experimental groups and materials used by each, the teachers, the investigator's attitude, sample size, and statistical analyses used; (2) the duration of the investigation, which in many studies is but one year; and (3) problems with criterion measures centering around such issues as final measures for evaluating i.t.a., the adequacy of reading measures, and correlated measures. Suggestions were made for improving research design which can be implemented as further studies replicate or build upon what has been done. Further research studies with i.t.a. could lead not only to its more effective evaluation but also to an examination of basic concepts relating to aspects of educational research. A bibliography is included. (DH)

A CRITIQUE OF RESEARCH WITH THE INITIAL TEACHING ALPHABET AND SOME RECOMMENDATIONS

J. R. Block Executive Director i.t.a. Foundation
Hofstra University, Hempstead, New York

This is a time of excitement in education. Innovations have come on the educational scene by the score and have had considerable impact. Perhaps it was Sputnik that made Americans aware that we must advance as rapidly in education as in science. Perhaps it is only an affluent society, concerned with becoming a "great" one, that we have both an awareness of the need for quality education and the capacity to provide it. The fact remains that we are making an enormous effort to improve our educational effectiveness. But, rapid and exciting progress is not without price. Each innovation intended to improve the effectiveness of education is met with great acceptance by some, with strong rejection by others, and probably with cautious optimism by most. Educational television, for example, has experienced only limited success. Programmed instruction, too, has its cultists and its detractors. With both techniques we anticipate some depersonalization of the student-teacher relationship, yet recognize that our growing population demands some such solution. Undoubtedly, the ultimate balance will involve intelligent and perhaps creative use of each. But programmed instruction and educational television are, in the long run, merely devices which an educator may use to increase and expand his effectiveness.

The introduction of the Initial Teaching Alphabet has created similar controversies. It, too, is a "device". It suggests that one of the major causes for reading failure lies in the idiosyncrasies of our alphabet and spellings. It suggests further that, by introducing temporary changes to make these more logical and consistent, we may develop not only better readers, but better writers as well. Perhaps most importantly, there is the hope that we may increase the *love* of reading and writing. The alphabet looks a little different. To some, this is exciting: to others, frightening. Although literate adults have no difficulty in reading in i.t.a., some are troubled by it. Commonly expressed fears are: "How will a child be able to switch to the conventional alphabet?" "How will he ever learn to spell?", not that we can be particularly proud of how well our children spell when they learn to read using our conventional alphabet.

Some of the resistance to i.t.a. stems from genuine educational concern. Here, the fear is that exposure of children to a temporary alphabet, no matter how carefully designed, may jeopardize their future reading ability, and since reading ability is so basic to the educational process, thus may ultimately undermine their educational future.

Some resistance stems simply from resistance to change of any kind. There are many whose personality is such that change of any sort is threatening in and of itself.

Finally, there is resistance that arises out of commercial or professional interests. For many years, our approach to the teaching of reading has been based largely on various "methods." The emphasis shifts from time to time from essentially a "look and say" approach to a "phonic" approach, with a number of other variations. There are experts and proponents of each with deep professional and commercial commitments. There are also business organizations which have invested large sums of money in traditional texts and basal series designed to teach beginning or remedial reading. Obviously, for them a shift to a new alphabet would impose a considerable financial burden.

While educators do not like to acknowledge the impact of the commercial organization upon educational policy, it is undeniably significant. A given approach to education may succeed or fail in part as a reflection of the extent to which usable materials are available. The Initial Teaching Alphabet has had its predecessors in the form of other transitional alphabets or spelling reform movements. Sir Isaac Pitman was instrumental in developing "Fonotypy" in the late 19th century. Later, in both the 1920's and 1940's, the simplified spelling movement was popular. All of these movements, while experiencing considerable success, were ultimately discarded. One of the major reasons appears to have been a lack of sufficient range and quality of commercial ma-

terials (Harrison, 1964). Perhaps i.t.a. has become as controversial as it has because we have a competitive commercial market today. Many corporations have become involved in this educational innovation. In Great Britain, over thirty-five commercial institutions are preparing materials in i.t.a., and at least twelve are known to be doing so in this country. In addition to many hundreds of children's books, major manufacturers have prepared typewriters, rubber stamps, flash cards, and decal-type material. Printers have indicated that i.t.a. fonts are available, and major test publishers have transliterated well-known and widely-used achievement tests into i.t.a. Some individuals have viewed the introduction of these commercial organizations as a threat to education. Others view it as providing the necessary materials for a comprehensive program so that i.t.a. may have a full and conclusive evaluation.

With all this activity, there has been great interest in i.t.a. Over two hundred articles have been published in the mass media and professional journals describing i.t.a. or reporting on its effectiveness. Vernon (1965) has suggested that this publicity has invalidated the i.t.a. studies in England. It is my opinion that publicity both in this country and in England was essential to gain sufficient cooperation from professionals and parents to permit evaluation of the effectiveness of this new medium.

With all this activity, what has the research indicated thus far? Viewed from the most conservative point of view, the research has shown that fears that there would be difficulty in transition and poor spelling have not been realized. While there is a great risk in overgeneralization, many of the studies (Alden and Maaning, 1965; Baker, 1966; Bosma and Farrow, 1965; Downing, 1965; Dunn, Muller and Neely, 1966; Jameson, 1965; Mazurkiewicz, 1964; Monteal, 1965; Myers, 1965; Regan, 1965; Shapiro, 1966; and Sloan, 1966) indicate significant differences in reading achievement in favor of i.t.a. groups when compared with controls, while a number of other studies (Downing and Jones, 1966; Hahn, 1966; Mountain 1966; Nemeth and Hayes, 1966; Tanyzer, 1966b; and Tanyzer, Alpert and Sandel, 1965) show no significant difference between i.t.a. taught students and control groups.

Even in the case of studies purporting to show no significant differences there is partial evidence for the success of i.t.a. Nemeth and Hayes (1966) for example, generally report that the i.t.a. group scored significantly higher than the basic T.O. controls but not always higher than a highly phonically-oriented T.O. approach. In addition they indicate that their evidence suggested that the greatest advantage of i.t.a. seemed to be for children in the lowest third of students ranked by I.Q. Similarly, McCracken (1966b) found that his i.t.a. group scored significantly higher than a T.O. control group on some of his measures but not on all. Both the i.t.a. and the T.O. control group scored significantly higher than a sub-control group unaware that they were part of an experiment.

Thus far, I am not aware of any study which has presented evidence that children taught with the Initial Teaching Alphabet do generally less well than those taught with the conventional alphabet. Thus, criticism of the alphabet has shifted from a focus upon potential difficulties in transition and spelling (which are presumably no longer a serious concern) to suggestions that i.t.a. either is not better than our conventional approach or that, given an early advantage to the i.t.a. group, the persistence of the advantage is questionable (Downing, Rose and Cutts, 1964; Southgate, 1965; and Vernon, 1965).

It is important to note that virtually every study investigating the effectiveness of the Initial Teaching Alphabet has contained one or more serious methodological errors. Of course, this is probably true of most studies in the field of education in general, and more broadly, in all areas of applied research. Investigators in educational research do not have the opportunity for rigid laboratory controls. They do not use lower animals and concern themselves with simple variables. They investigate inherently complex material in a complex applied setting. Although technically desirable, it is not always possible, and, indeed, at times would be unethical to present all combinations of one's experimental variations to persons who are learning to read. It would be difficult to obtain parental approval for many theoretically interesting combinations of variables. Thus, we must rely upon the ethics of the profession and permit the weight of number of studies to compensate for the lack of complete control.

Nonetheless, it might be useful to point out some errors in research design in studies of i.t.a. and to suggest what might be done to overcome or circumvent them. In some cases, the nature of the methodological error would be likely to reduce the apparent improvement achieved by the use of i.t.a. In other cases, the flaw in the design might be expected to increase the apparent effectiveness of i.t.a. In still other cases, it would be difficult to pre-

dict the net effect upon the dependent variable, but conclusions based on the studies would necessarily be extremely tenuous.

Although there have been a great number of specific methodological errors in the i.t.a. studies conducted thus far, they appear to fall into the three rather broad categories: errors in research design; the duration of investigations; and problems associated with the criterion measures. Each of these areas will be treated separately with a number of specific problems discussed under each. Obviously, the many problems discussed in this paper are interrelated and any independent discussion of categories must necessarily be somewhat arbitrary.

I. Errors in Research Design

Problems associated with research design involve such factors as the use of control groups and the ability to match them with the experimental groups for factors related to reading; the materials used by each; and the teachers involved. Additional problems include elimination or control of the Hawthorne Effect and the investigator's attitude, as well as considerations of sample size and the statistical analyses used.

Control Groups – Several studies have been conducted without benefit of control groups (Deverell, 1966; Georgiades and Downing, 1964; Hastings, 1966; Kemp, 1966; Mason, 1966; and Stevenson, 1966). Implicit in many of these is the suggestion that a control group does exist in the form of prior experience with many groups of students taught in a conventional manner. While these studies do not present statistical evidence for or against i.t.a., they may be fruitful as pilot studies. The opportunity they provide for direct personal experience may yield a greater sense of validity for many teachers than abstract sets of statistical tables.

Other studies have made efforts to match individuals or groups on such variables as age, sex, I.Q., and reading readiness[1] (Downing, 1965; Mazurkiewicz, 1966a; Myers, 1965; and Sloan, 1966). While these would seem to be relevant factors for matching, a number of the studies have been forced to resort to matching classes in two different schools. For example, Bosma and Farrow (1965) used students in a campus school matched with those in a neighboring public school district. While the authors state they have reason to believe the two schools are comparable and the subjects are matched on the demographic data noted above, the environment of a campus school might be expected to be quite different from a public school setting. As i.t.a. becomes increasingly accepted, it should be possible to replicate the research conducted thus far using more adequately matched groups.

In addition, there is the question of the ways in which the control groups should differ from the experimental groups. Some investigators have suggested that only the media be changed and that the content of the material remain constant (Downing, 1965). There is a marked methodological advantage in this procedure since one changes only a single independent variable at a time and may more adequately identify the specific factor to which any change may be attributed. On the other hand, it can be argued that one virtue of i.t.a. is that it permits an author to relate his materials more directly to the interest and sophistication levels of his audience, since he is no longer concerned with a rigidly controlled vocabulary as a means of handling the inconsistencies of written English. A study presenting the vocabulary of an i.t.a. series with its less rigid vocabulary control to beginning readers in T.O. represents a combination of variables which one would hesitate to manipulate in fear of doing irreparable damage to the child's ability and attitude toward reading.

Nonetheless, if i.t.a. may be used to present more challenging materials, then its effectiveness in this approach must be evaluated. Further, one may argue that there are several ways to control the idiosyncracies of English. One such method is to present a highly controlled vocabulary so that these idiosyncrasies may be introduced gradually and systematically. Another method is to provide a transitional alphabet to eliminate problems initially and at the same time provide more satisfying reading material. The idiosyncrasies of written English may be gradually and systematically introduced as the transition takes place. If vocabulary control reduces the complexities of written Eng-

[1] I would raise the question of the utility of matching on the basis of reading readiness measures. First, we do not usually use measures of speaking, listening or writing readiness. Second, such measures are highly correlated with I.Q. Finally, there is an implicit assumption in the construction of the measure that the child will be reading in T.O. i.t.a. may involve substantially different skills and perceptual abilities.

lish, then there would be an advantage in superimposing a second control of a transitional alphabet only if vocabulary control is not completely effective.

Obviously, both kinds of studies are needed (i.e., those which change the alphabet only and those which change content as well). Since some series emphasize the stimulating nature of the materials as well as the alphabet, it is suggested that investigators consider the use of dependent variables other than reading achievement *per se*. For example, measures of personality, interests, or attitudes toward reading and toward school may be more crucial dependent variables in such studies if it could be shown that the mechanics of reading may be mastered as well in i.t.a. as in T.O.

A further problem in matching lies in the control of the teacher variable. In some cases, investigators have pointed out that requiring teachers to teach both traditional orthography (T.O.) and i.t.a. would not be possible since it is difficult to insist that a teacher shift completely from one medium to another. Morale would likely suffer and school administrators are seldom willing to antagonize their staff unnecessarily. Downing and Jones (1966) tried to control this variable by having each teacher spend half her time with the T.O. group and half with the i.t.a. group. Their preliminary results suggest that this procedure may have been unpopular enough to lower the reading achievement level of the children. Others have attempted to match teachers according to length of experience or other comparable measures (Sloan, 1966).

To the extent that teacher attitudes are important factors in their effectiveness with children (and there are few who would question this assumption), it is of vital importance to provide evidence of adequate control of this variable.

Once again, as i.t.a. becomes increasingly popular, strong positions both positive and negative, will continue to develop. It should be progressively easier to find groups of teachers who are deeply committed either for or against i.t.a. (It is rare to find among those who are familiar with i.t.a. persons lacking strong attitudes.) Teachers with opposing attitudes should be matched in terms of strength of feeling as well as amount of teaching experience and other variables associated with teaching competence. It is possible that there are important differences between teachers who support i.t.a. and those who do not. This problem would be interesting to study in itself.

The Hawthorne Effect – The well-known Hawthorne Effect (that is, the tendency for the results of an experiment to be largely determined by the fact that the subject knows he is participating in an experiment) has been cited by those critical of i.t.a. and by almost every investigator finding positive results, although Dunn, Muller and Neely (1966), for example, believe that the Hawthorne Effect was "not sufficient to explain the results" in their study. It is generally assumed that the Hawthorne Effect has operated in such a way as to motivate teachers of the i.t.a. groups to expect higher levels of achievement than they would normally expect. It should be pointed out that the Hawthorne Effect *could* have exactly the reverse effect in some studies, particularly since many reading experts have expressed strong negative attitudes toward i.t.a. The Hawthorne Effect is extremely difficult to control. If it has not been adequately controlled in the past, it should be increasingly easy to find groups equally committed for and against i.t.a. Further, as there is greater exposure of the public, teachers, and administrators to the Initial Teaching Alphabet, whatever novelty effect the alphabet may have had in the past should be a less potent factor in future studies.

Early studies are likely to have been conducted by those deeply concerned with the problem. Some investigators have been concerned with the preparation of i.t.a. materials while others have been interested in other systems.[2] Myers (1965) served as the teacher of the experimental group in her study with a colleague serving as the control group teacher. While these investigators are both competent and ethical, many recent studies have shown that studies using lower animals performing simple tasks (Kintz, Delprato, Mattee, Persons and Schappe, 1965; McGuigan, 1963; and Rosenthal, 1963) may unconsciously bias the results of a study. With time, a greater number of investigators undoubtedly will become interested in research with i.t.a. As a result, there

---

[2] Investigators associated with i.t.a.:
Baker – The Freedom to Read Series
Downing – The Downing Readers
Matuszkiewicz and Tanyzer – The Easy-to-Read Series

Investigators associated with other systems:
Fry – The Diacritical Marking System
McCracken – The Sheldon Basic Reading Series
Mountain – Challenge Readers

should be an increasing amount of evidence presented by those who are less psychically involved with the alphabet.

Sample Size – Many studies, particularly those assessing the value of i.t.a. for remedial purposes, have used relatively small samples (Baker, 1966; Barclay, 1966; Bosma and Farrow, 1965; Georgiades and Downing, 1964; Myers, 1965; and Tanyzer, 1966a). Some of these studies indicate a superiority of the i.t.a. groups; others show no difference. In all cases, the sample sizes are too small to permit more than general suggestions about the relative effectiveness of i.t.a. Most of the authors of such studies are well aware that their sample is of limited size and caution their readers about interpretation and over-generalization. There is evidence to suggest that, as i.t.a. becomes increasingly accepted, investigators in future studies will have larger sample sizes available. Such studies should present a more adequate base for generalization than from any single set of results.

Statistical Analysis – An integral part of any research is the consideration of the statistical techniques used to evaluate the quantitative aspects of the study. Statistical evidence can be quite compelling. It is a rare situation, however, when one meets all of the assumptions which underlie the mathematical properties of the formulae to be used. As one illustration, Tanyzer, Alpert and Sandel (1965), in studying over four hundred children in i.t.a. classes as compared with four hundred in T.O. classes, used the classroom as the unit of analysis. They observe that, since the children in a given classroom are exposed to a common teacher and relatively common experiences, children's scores do not represent individual observations. They cite Campbell and Stanley (1963) for this rationale. While there is much to be said for this argument, there is no attempt to consider the effect of the relatively large number of children in Tanyzer *et al*'s (1965) analysis. They used a total of thirty-four classes (17 i.t.a. and 17 T.O. classes) with over twenty children per class. Thus, their sample size is apparently reduced from over eight hundred to thirty-four. Since each class had over twenty children in it, data from such a group might be considered more reliable than would be the case if the classes had contained only five or six children each.

This problem is not easy to resolve, since students grouped in classes are *not* independent. However, until there is a generally accepted unit of analysis for research in education, the present author would suggest presentation of data using *both* the individual and the class as the unit of analysis. This would permit comparison with earlier or subsequent investigations using either approach. Each reader might then make his own evaluation of the extent to which he accepts the underlying assumptions.

II. Duration of the Investigation

Another important consideration in the evaluation of any investigation is the point in time at which final evaluation takes place. Since reading is a skill one uses over a lifetime, it is theoretically possible to evaluate one or more aspects of reading resulting from a particular method or medium at almost any time. Cutts has argued, for example, that the evaluation of i.t.a. cannot be complete until children taught with the alphabet reach the fourth or fifth grade (Downing, Rose, and Cutts, 1964). In contrast, most studies have been of one year's duration. The point to be made here is that any investigator or sponsor of an investigation must consider carefully the implications and assumptions of whatever time period is studied. With i.t.a. for example, this includes such factors as the number of children having made the transition at one point, the amount of instruction and exposure to reading materials during the period; the other educational experiences the experimental and control groups have subsequent to the study (i.e., do they take the same curriculum, and what assumptions about prior achievement are built into this curriculum); and how experienced are the teachers in using this new medium.

One-Year Studies – Frequently, for administrative reasons, the time for the completion of an investigation is arbitrarily fixed at the end of the first year. This was true, for example, of all of the studies sponsored by the U.S. Office of Education (Hahn, 1966; Mazurkiewicz, 1966b; Mountain, 1966; Nemeth and Hayes, 1966; and Tanyzer, 1966b). It is widely accepted among both the proponents and opponents of i.t.a. that at the end of the one year a large percentage of children will not have formally made the transition to our traditional orthography (usually from twenty-five to fifty per cent). The minimum goal of all programs using i.t.a. is for the child to be able to read effectively in traditional orthography. As a result, dependent variable tests at the end of one year are invariably given in T.O.[3] Since many children have not made the transition to T.O., a test in what is for them a new alpha-

[3] When tests are given in the medium in which the child was learning to read, i.t.a. groups invariably perform at a higher level than T.O. groups.

bet puts the i.t.a. group at a serious disadvantage. Many authors have noted that, despite this disadvantage, i.t.a. groups will *do at least as well as* T.O.-trained children. In contrast, McCracken (1966b) shows that children taught in T.O. who took a reading test transliterated into i.t.a., apparently lose about one year in reading level. Thus, one might assume that a comparable loss could occur for i.t.a. children although they typically score higher or at the same level as the control group.

Behn (1963) and Fabino (1963) have suggested that there is insufficient overlap even between different basal series in T.O. to warrant ease in transfer from one to another. In their studies, they were concerned with different variations in controlling vocabulary using the same alphabet. It would be interesting to study the effect of having children take a test printed in cursive writing as compared with conventional print when they had had limited training in cursive writing.

Since none of the proponents of i.t.a. suggest that there is a specific time at which the child *must* make the transition, a number of alternatives suggest themselves. Mazurkiewicz (1966a) attempted to resolve this problem by comparing the performance of the children who have made the transition with a matched group sample from the control group. While these children were matched on I.Q., sex and socioeconomic status, it is possible that these variables do not fully reflect their reading ability. Those who perform best with i.t.a. materials and make the transition from i.t.a. to T.O. first are among the best readers. While matching on other variables represents an attempt to control this, those in the matched control group may not stand as high in reading ability within their own group. While there is no evidence on this point, a more defensible procedure might be to compare those who have completed the transition with an equal number of T.O. children representing the highest reading scores in their group. An alternate procedure, waiting until most children have made the transition, might provide more representative data. However, this procedure suffers from the disadvantage that the slowest children determine the point of administration of the dependent variable. Thus, any gains made by the first children to transfer may be obscured since the advanced i.t.a. group and the T.O. group will have been exposed to identical curriculum experiences for a substantial period of time after transition.

Further, with the rigid one-year duration, one may question how much formal instruction in reading had actually taken place at the end of one year. Much of this period is devoted to reading readiness preparation, which may or may not involve the use of a special alphabet. If the period of exposure to reading instruction is too short, it may be difficult to generate differences between any set of methods or media. In any event, it is my contention that *an i.t.a. program cannot be adequately evaluated in one year.*

Where significant differences have existed in favor of the i.t.a. groups, the question has been raised "How long will the advantage last?" (Downing, Rose and Cutts, 1964; and Vernon, 1965). It occurs to me as it has to others, that while some children may learn to read with i.t.a. who would not have learned to read at all, advantages accruing to the i.t.a. group may, in fact, disappear in the later grades. While only further research can verify this hypothesis, it seems quite reasonable. If children in the i.t.a. group are reading better, writing better, and have more positive attitudes toward school at the end of a study, unless there is a change in the traditional curriculum to permit them to make use of this new found skill, any differences between i.t.a. and T.O. groups must invariably disappear as they continue to be exposed to common curricula based upon what the T.O. group was assumed to have achieved. i.t.a. presents a challenge for rethinking the curricula beyond first grade. If there are advantages to i.t.a., we must take advantage of them.

Teacher Experience – It was noted earlier that teachers (and perhaps students and parents) in i.t.a. studies may be influenced by the Hawthorne Effect. It should also be pointed out that, in the first year of the study, a teacher may not feel completely at ease with the new alphabet, and, in fact, may not know how to deal with its full potential comfortably and easily in the classroom. Mazurkiewicz (1965) has suggested that his research is achieving better results in the second year of his study as teachers become more comfortable with the medium. In general, a classroom teacher has had many years of experience in teaching with the conventional alphabet. Most teacher-training courses in i.t.a. are limited to two or three days. While this may be a sufficient period in which to explain the principles of the alphabet and their use in the classroom, it may not be sufficient to develop a confidence and ease of working with this new medium which permits an adequate evaluation of its ultimate effectiveness in the hands of skilled teachers. Where possible, formal evaluation of an i.t.a. program should be delayed until after one year of experience with the alphabet.

III. Problems with Criterion Measures

Perhaps the most important consideration in any study is the criterion measure used. Such questions as, What do you want to measure? How reliable and valid are your measures? Are your different measures actually measuring different traits? Are your measures contaminated with some extraneous variable?, all must be considered. It is upon these considerations that one's final conclusions will be based.

What Measure Should Be Used? – A great many problems are associated with the final measures used to evaluate the Initial Teaching Alphabet in *all* of the i.t.a. studies conducted thus far. First, almost every study has focused largely, if not entirely, upon traditional measures of reading. Obviously, an alphabet may be used not only for reading, but for writing as well. Further, the printed word is widely accepted as affecting attitudes and emotions. Some evidence already exists that the writing of children taught with i.t.a. may be dramatically affected (Downing, 1964; Gardner, 1966; Montesi, 1965; and Reszek, 1966). Teachers have subjectively reported much greater interest in reading, more positive attitudes toward school in general, and greater independence on the part of the children. Stewart (1966) indicates changes not only in reading achievement, but in "independent learning, motivation, perseverance, the ability to observe, and the ability to write." Similar results are reported by Barclay (1966) working with emotionally disturbed children. Perhaps it would be profitable to include measures in other curriculum areas (e.g., arithmetic) to evaluate the role of i.t.a. in generating attitudes toward school. Before accepting or rejecting the present evidence, we clearly need better instruments than we now have to measure writing skills and attitudes. If i.t.a. accomplished nothing more than stimulating the development of more adequate measures of children's writing, attitudes toward school, attitudes toward reading, independent behavior, etc., it will have performed a valuable service to the field of education. It should make us increasingly aware that reading represents only one way in which we use our language. Our approach to reading will have an impact on a child's development in many spheres.

The Adequacy of Reading Measures – A crucial consideration of any measure is the degree of reliability and validity of the instruments. Most of the measures used in i.t.a. studies thus far have been reasonably reliable. In addition, most test publishers indicate that their tests have content validity. That is, most publishers indicate that the professional personnel who have constructed the tests have, in one way or another, surveyed a number of the most widely used texts and have tailored their tests to follow the curriculum covered in these materials. This raises the question, "To what extent are these tests valid for children who have learned to read with i.t.a. materials not using a controlled vocabulary?" To the extent that the curriculum as presented in such texts is different from the traditional curriculum, traditional reading tests will not provide valid measures.

An analogy may be drawn between the use of i.t.a. in teaching reading and the use of "Modern Math." In recognition of the change in approach to the teaching of mathematics, Harcourt, Brace and World, Inc. have recently published a new instrument in the Stanford Achievement Test series called the "Modern Mathematics Concepts Tests." In their promotional literature they note, "Mathematics is not new. The old rules still obtain; the relationships of quantities, figures, and processes have not changed. What has changed is the approach to teaching mathematics, the concepts taught as a basis for improved understanding and application of mathematics." In a covering letter with their brochure, the first two sentences read, "How is the Modern Math program succeeding? A major need at all stages of new curricula programs is objective evaluation of achievement in terms of new criteria, new content, new objectives and methods."

I raise the question as to whether or not tests developed to reflect curricula presented in traditional basal reading series are appropriate for use as dependent variable measures in i.t.a. studies. This is of particular importance in studies using i.t.a. books which depart from a controlled vocabulary. It is urged that traditional reading measures be carefully re-examined. Further, where necessary, new tests should be developed which reflect curricula goals existing between i.t.a. and T.O. programs and eliminate any variables which are unique to either program.

It should be noted here that in two surveys of the literature to find methods of appraising reading abilities and determining which reading abilities should be measured, Gillmore (1963) concluded that there was no single device adequate to measure all reading skills. She concluded that much of the evaluation is dependent upon the teacher who is in a position to observe the child daily. Future i.t.a. studies might include carefully constructed rating scales completed by teachers to evaluate the child's progress. Thoburn (1966) reports a study of the subjective evaluations of

teachers and administrators after using i.t.a. She indicates that they feel that substantial gains are possible with this medium. Georgiades and Downing (1964) and others have reported similar results.

Correlated Measures – One of the common methodological problems in research in education and in virtually all of the studies involved with the Initial Teaching Alphabet relates to the fact that a large number of tests or sub-tests are used. Very few of these studies indicate the degree of intercorrelation between these sub-tests which are treated as "independent" measures. It seems highly reasonable that spelling ability would be highly correlated with word recognition, reading comprehension and perhaps a number of other reading measures. While test manuals generally fail to report sub-test intercorrelations, Mitchell (1962) reports correlations between .51-.63 for total score on the Metropolitan Achievement Readiness Test and its sub-scales. She reports average sub-test intercorrelations of about .55. Thus, these measures are not independent of each other. Furthermore, children in educational research studies are taught in the same classroom, using the same books, by the same teacher. All of these factors tend to produce a high correlation between measures which are treated as though they were independent. Thus, in any given study, if one measure shows a significant difference in favor of the i.t.a. group, then it is highly probable that all measures will indicate this. Contrariwise, if one measure shows no significant difference, then it is highly probable that no significant differences would be reflected on the other measures in the study. In evaluating the results of each study, one is tempted to conclude that it has shown either that i.t.a. was more effective, or that it was generally comparable to T.O. in all measures.

It is strongly recommended that future investigations of the effectiveness of this alphabet (or for that matter, any method or medium) attempt to use measures which are as independent of one another as possible. An absolute minimum requirement should be that the investigator indicate the degree of inter-relationship between the various dependent variable measures used to permit his readers to evaluate his results.

Contaminated Measures – Another question which occurs to me is "What do our conventional reading tests measure?" Generally, test manuals present correlations between their instrument and other publishers' reading tests. They seldom indicate the correlation between the test and other measures such as intelligence.

Strang, McCullough, and Traxler (1955) however, have cited three studies suggesting that the correlation between measures of reading and measures of intelligence is between .50 and .80. Further, Austin, Bush and Huebner (1961) show graphically the marked role of intelligence in accounting for reading achievement, particularly in the first and second grades. To the extent that reading achievement tests are correlated with I.Q., it is difficult, if not impossible, for any method or medium for teaching reading to show any real superiority over any other method. It is generally assumed that I.Q. remains relatively constant. Obviously, an adequate study of reading ability must partial out the factor of intellectual ability and in fact, some investigators have attempted to do this (Shapiro, 1966). Yet, if I.Q. accounts for a large proportion of the variance of reading scores, one is left with the question of what is left in traditional reading measures when I.Q. is removed? What evidence of validity do the tests indicate under these conditions? Once again, if raising these questions in the evaluation of the Initial Teaching Alphabet serves to promote an interest and effort toward developing more adequate measures of a child's reading ability independent of his intellectual level, then i.t.a. will have performed an important service to the field of education.

Summary and Conclusions

Despite the research errors (and there have been few educational methods, media or materials which have been as intensively studied as i.t.a.) some general conclusions may be reached even now. One source of resistance to i.t.a. has been with regard to the potential "risks" involved. These are expressed largely in terms of the belief that children will not be able to make the transition. That they will not be able to spell in T.O. That in the long run they will be less well off than T.O.-taught children. The evidence suggests clearly that this is not the case. Children taught with i.t.a. do *at least as well* as T.O.-taught children – using T.O. standards for evaluation. This despite the fact that relatively large percentages of children are tested with an alphabet which they have not been reading. Downing and Jones (1965) for example, report that only 17% of their four- and five-year old children had made the transition at the time of testing, yet the total group did not score significantly different from children in the control group.

Expressed from a more positive point of view, one can point to a number of studies conducted by different investigators which present quite positive results using both objective and

subjective data. The question for the administrator should become "If there seems to be little risk, can I afford to ignore the potential gain suggested by i.t.a. research?"

There have been many methodological errors in studies attempting to evaluate the effectiveness of i.t.a., but to no greater degree than we usually find in educational research. Some of these errors may have had the effect of making i.t.a. appear more effective than it is, while it is obvious that many of these errors have served to reduce the apparent effectiveness of the alphabet. Some of the problems associated with the early studies will be easily resolved in future studies. With greater interest and experience with i.t.a., the Hawthorne Effect should be reduced to a negligible degree by now. The influence of publicity, if it has had any adverse effects, should be less potent. It should also be possible to obtain larger and more adequately equated groups. Further, with greater interest among diverse groups, control of the teacher variable should be easier, and with an increasing number of studies, the role of the attitudes of the investigator should come under control.

It should be emphasized that these controls will not come automatically. It is simply that we should be better able to control them once we are aware of their influence and have a wider range of school systems, teachers, and investigators interested in determining the effectiveness of i.t.a.

One obvious conclusion is that studies may be of varying duration, but in any event, greater than one year. Further, investigators must consider the effects of other aspects of the curriculum on the time period they select, and select measures appropriate to the duration selected (e.g., measures of simple reading skills, attitudes toward reading, etc.).

This last point raises one of the most pressing issues in educational research – the need for more adequate measures. We need measures of other factors related directly or indirectly to reading including, attitudes toward reading and school in general, writing ability, etc. Further, we must take a hard look at our reading measures. Are the sub-tests measuring different characteristics? What are we measuring besides I.Q.? Can we control for I.Q. and still have meaningful measures? Better instrumentation lies at the heart of more adequate research in education.

Future i.t.a. Studies – This paper has suggested that it would be profitable to replicate many of the early studies using more adequate controls. Future studies must obviously be longer than most of the early ones and investigate a wider range of relatively independent phenomena. Yet, there are some special considerations future studies must deal with. We need much more research in the area of special education and in the remedial use of i.t.a. We need to know more about its effectiveness in teaching English as a second language. We need to know, not only what happens when the transition has taken place, but what, if any, advantage there might be in prolonging exposure to i.t.a. in an effort to build increased confidence. We need to know what method or methods work most effectively with the Initial Teaching Alphabet. While the alphabet does not demand that a teacher utilize any particular method, it is possible that one approach is more effective than another. We have spent a great many years attempting to learn what teaching methods may be used to greatest advantage with our conventional alphabet. We should undertake similar investigations with i.t.a. We should learn how to use it; who should use it; when it should be used; how long it should be used, etc.

It seems reasonable to assume that i.t.a. (or for that matter, any educational innovation) will never be universally used no matter how effective it may be. Many minds are already closed to the issue. Some are convinced that the Initial Teaching Alphabet is the solution for which they have been waiting. Others are equally convinced that it does not work. Some argue that it obviously works and should be adopted. Some argue that it should not be accepted too quickly – others are arguing it should be rejected immediately.

This paper has attempted to suggest some improvements in design. It has suggested not only that we can do a more effective job in evaluating i.t.a., but that attempts to provide proper experimental controls inevitably will force investigators to re-examine many basic concepts in the field of educational research and evaluation.

REFERENCES

Alden, C.E., & Manning, R. The use of the initial teaching alphabet in reading instruction – a report on the pilot study conducted at Broadway School during the school year 1964-1965. Unpublished manuscript. Maple Heights, Ohio, 1965.

Austin, Mary C., Bush, C.L., & Huebner, M.H. *Reading Evaluation*. New York: Ronald, 1961.

Baker, M. Remedial Reading with i.t.a. In A. J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*. Hempstead, N.Y.: i.t.a. Foundation, 1966, pp. 158-165.

Barclay, G.L. i.t.a. and the Emotionally Disturbed. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 177-182.

Behn, Mary. Words Appearing in the First, Second and Third Pre-primers of Eight Reader Series. Cited by Doris V. Gunderson, *Research in Reading at the Primary Level: An Annotated Bibliography* OE-30008, Washington, U.S. Government Printing Office, 1963.

Bosma, R.L., & Farrow, V.L. Teaching Reading with i.t.a.: A Research Report, *Reading Horizons*, 6, 1, 6-19, Fall 1965.

Campbell, D.T., & Stanley, J.C. Experimental and Quasi-Experimental Designs for Research on Teaching. In N.L. Gates (Ed.) *Handbook of Research on Teaching*. Chicago: Rand-McNally, 1963, pp. 171-246.

University of Detroit. *Research Report on Basic Adult Education Program*, University of Detroit – Center for Continuing Education, Detroit, Michigan, October, 1965.

Deverell, A.F. The Canadian Scene. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*. Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 128-132.

Downing, J.A. *The Initial Teaching Alphabet Reading Experiment*. Chicago: Scott, Foresman, 1965.

Downing, J.A. *Examples of Children's Creative Writing from Schools using i.t.a.* London: London Institute of Education, 1964.

Downing, J., & Jones, Barbara. Some Problems of Evaluating i.t.a. A Second Experiment. *Educational Research*, February 1966, pp. 100-114.

Downing, J., Rose, I., & Cutts, W.G. Opinions Differ on the Value of i.t.a. *National Education Association Journal*, September, 1964, pp. 20-22.

Dunn, L.M., Muller, M.W., & Neely, M.D. An Interim Report on the Efficacy of the Initial Teaching Alphabet and the Peabody Language Development Kit with Grade One Disadvantaged Children. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 91-100.

Fabiano, F. R. A Study for the Franklin School Reading Plan. Cited by Doris V. Gunderson, *Research in Reading at the Primary Level: An Annotated Bibliography*, OE-30008, Washington, U.S. Government Printing Office, 1963.

Gardner, K. Remedial Reading with i.t.a. – In England. In A.J. Mazurkiewicz (Ed.) i.t.a. *and The World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 170-173.

Georgiades, N.J., & Downing, J.A. *Report on the uses of the Initial Teaching Alphabet in Remedial Reading Classes in Primary, Secondary, and Special Schools*. London: University of London Institute of Education, 1964.

Gillmore, Zelda C. Suggested Criteria for the Identification of Superior Readers in Grade One. Cited by Doris V. Gunderson, *Research in Reading at the Primary Level: An Annotated Bibliography*, OE-30008, U.S. Government Printing Office, 1963.

Hahn, H.T. Three Approaches to Teaching Reading in Grade One. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 268-272.

Harrison, M. *The Story of the Initial Teaching Alphabet*. New York: Initial Teaching Alphabet Publications, Inc., 1964.

Hastings, H.J. i.t.a. with Prison Populations. In A.J. Mazurkiewicz (Ed.), *i.t.a. and the World of English*, Hempstead, N.Y.; i.t.a. Foundation, 1966, pp. 213-216.

Herman, F.D. The Clearfield Area Schools i.t.a. Study 1964-1965. Unpublished manuscript. Clearfield Area Schools, Clearfield, Penn., July 15, 1965.

Jameson, M.C. An experimental study of the effects of the initial teaching alphabet on reading achievement in the primary grades, with emphasis on grade one. A report on a primary reading experiment. N.D.E.A. Title III No. S7591X. Unpublished manuscript, Waterford Township School System. Pontiac, Mich., June 7, 1965.

Kemp, Valerie. i.t.a. and Multi-lingual Classes. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 150-154.

Kintz, B.L., Delprato, D.J., Mattee, D.R., Persons, C.E., & Schappe, R.H. The experimenter effect. *Psychological Bulletin*, 1965, 63, 4, 223-232.

Mason, F. Mary. i.t.a. at Miss Mason's School – Nursery Level. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 146-149.

Mazurkiewicz, A.J. A Comparison of i.t.a. and t.o. Reading, Writing and Spelling Achievement When Methodology Is Controlled. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966a, pp. 59-65.

Mazurkiewicz, A.J. A Comparison of i.t.a. and t.o. Reading Achievement When Methodology is Controlled. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966b, pp. 289-292.

Mazurkiewicz, A.J. Interim Report – Lehigh-Bethlehem i.t.a. Study. *Journal of the Reading Specialist*, March 1965.

Mazurkiewicz, A.J. A Tiger by the Tail, *Proceedings of the 29th Educational Conference of the Educational Records Bureau*. New York: Educational Records Bureau, 1964.

McCracken, R.A. A Two-Year Longitudinal Study to Determine the Ability of First-Grade Children to Learn to Read Using the Early-to-Read i.t.a. Program (An Interim Report of the First Year). In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966a, pp. 79-90.

McCracken, R.A. A Cross-Sectional Study to Determine the Ability of Children to Transfer in Reading from Traditional Orthography to the New Augmented Roman Alphabet. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966b, pp. 101-108.

McGuigan, F.J. The Experimenter: A Neglected Stimulus Object. *Psychological Bulletin*, 1963, 60, 4, 421-428.

Mitchell, Blythe C. The Metropolitan Achievement Readiness Tests as Predictors of First Grade Achievement. *Educational and Psychological Measurement*, 22, 4, 1962, pp. 765-772.

Montesi, Richard L. *The Initial Teaching Alphabet Experiment* September 1964-June 1965. Unpublished manuscript, Simsbury Public Schools, Simsbury, Connecticut, August, 1965.

Mountain, Lee H. A Comparison of i.t.a., Diacritical Marking System, and Basal Reader Programs. In A.J. Mazurkiewicz (Ed.), *i.t.a. and the World of English*, Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 280-288.

Myers, Nancy Joanne. A comparison of the achievement of children using traditional orthography with that of children using the Initial Teaching Alphabet in the Greater Johnstown School District, Grade One. Unpublished masters thesis, Indiana State College, 1965.

Nemeth, J.S. & Hayes, R.B. The New Castle Beginning Reading Study – A Preliminary Report. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English, Hempstead, N.Y., i.t.a. Foundation,* 1966, pp. 258-267.

Regan, J.F. An experimental study of the effectiveness of the i.t.a. method of teaching reading in year one. Unpublished manuscript. North Syracuse Central Schools, North Syracuse, N.Y., 1965.

Rosenthal, R. On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. *American Scientist,* 1963, 51, 268-283.

Shapiro, B.J. *A comparison of the Reading Achievement of i.t.a. and T.O. Groups in the First Grade, 1964-65.* Cleveland, Educational Research Council of Greater Cleveland, February 1966.

Sloan, C.A. *A Comparative Study of Matched Pairs i.t.a. – T.O. in Beginning Reading Instruction.* Unpublished manuscript, Board of Education of the Owatonna, Minnesota Public School System, 1966.

Southgate, Vera. Approaching i.t.a. Results with Caution. *Educational Research* 7, February 1965, pp. 83-96.

Stevenson, C. i.t.a. with Army Service Populations. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English,* Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 217-224.

Stewart, Rebecca. Two Years with i.t.a. – An Interim Report of the Bethlehem-Lehigh Research Project. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English,* Hempstead, N.Y., i.t.a. Foundation, 1966, pp. 121-124.

Strang, Ruth, McCullough, Constance M., & Traxler, A.E. *Problems in the Improvement of Reading,* 2nd ed., McGraw-Hill: New York, 1955.

Tanyzer, H.J. Using i.t.a. with Disabled Readers at the Secondary Level. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English,* Hempstead, N.Y., i.t.a. Foundation, 1966a, pp. 166-169.

Tanyzer, H.J. Effectiveness of Three Different Basal Reading Systems on First-Grade Reading Achievement. In A.J. Mazurkiewicz (Ed.) *i.t.a. and the World of English,* Hempstead, N.Y., i.t.a. Foundation, 1966b, pp. 273-279.

Tanyzer, H.J., Alpert, H., & Sandel, Lenore. *Beginning Reading – The Effectiveness of Different Media,* Mineola, N.Y., Nassau School Development Council, 1965.

Thoburn, Tina. *What do they think of i.t.a?* Cleveland, Educational Research Council of Greater Cleveland, January, 1966.

Vernon, M.D. i.t.a. in America. London Times Educational Supplement, December 31, 1965.

The Fourth International i.t.a. Conference will be held at McGill University, Montreal, Canada August 1967

The winter issue of the i.t.a. Foundation Report will contain a general outline of the program and forms for registration.

Individuals interested in presenting papers at the Conference should contact the Foundation as early as possible indicating the nature of the proposed presentation.

See p. 27