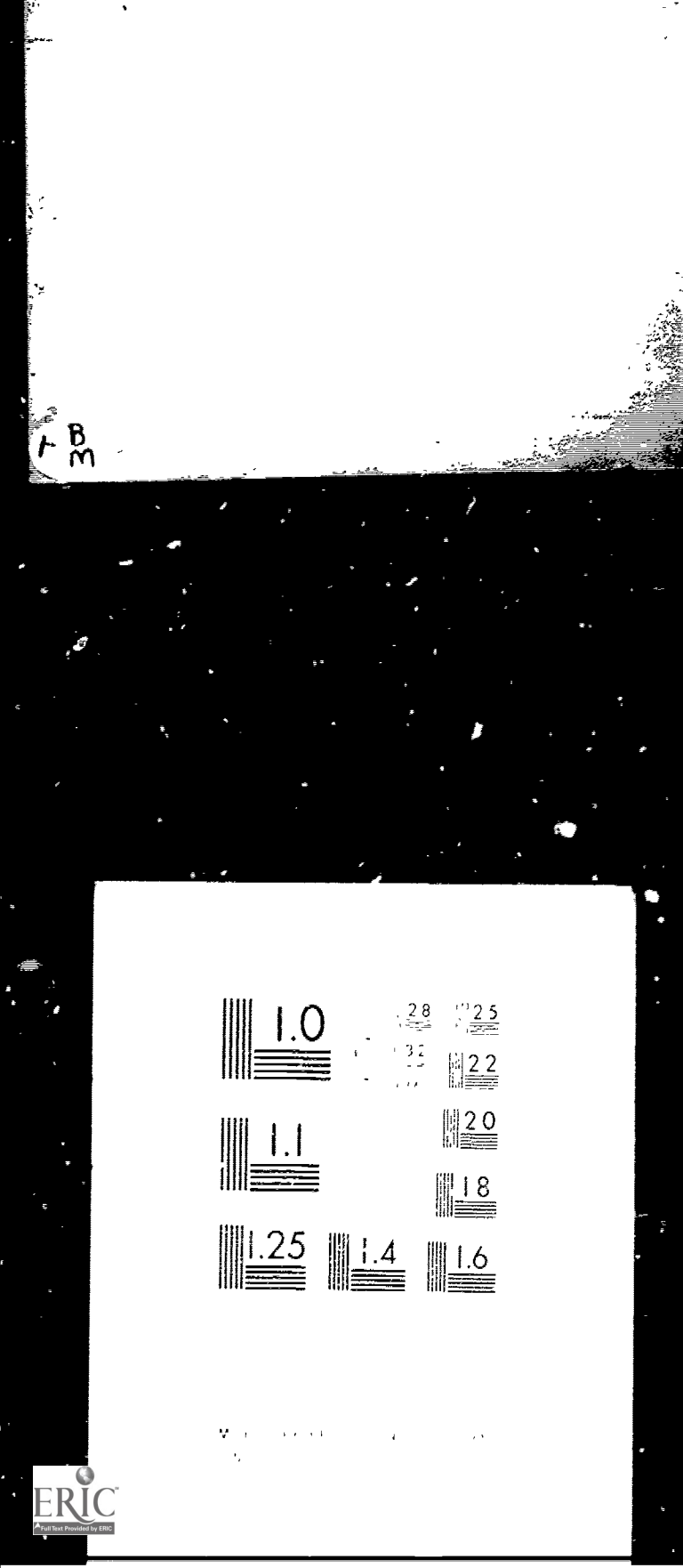
1.0
1.1
1.25
1.4
1.6
2.8
2.5
2.2
2.0
1.8

two methods were equally effective in the 106-113 group, and in the 114-129 group the thought method was superior. This claim raises a question. If the combined method was so generally inferior, how, in fact, did more combined method students better their thought method partners in the one school for which she gave figures? Either almost all of the students fell in the 106-113 range, which, if true, would raise questions about the representativeness of the sample; or the one school she cited could have been an exception where the combined method students did markedly better than those in other schools, which, if true, causes one to wonder why she cited that school. Her failure to give this kind of breakdown for all schools is a serious omission.

Frogner seems justified in claiming that her study shows that formal grammar instruction is not essential in developing an awareness of sentence structure, and her elaboration of the thought method is a major contribution, one which anticipated the sentence-combining activities that are currently being advocated. Her study does not, however, demonstrate the clear superiority of the thought method, much less the ineffectiveness of grammar instruction.

Harris (1962). Harris conducted a long-term (two year) study in which he compared an approach to sentence structure which emphasized formal grammar, to one which emphasized extensive writing and an intuitive, non-terminological approach to sentence errors. Both groups were taught the same curriculum

# Grammar Instruction and Writing: What Does the Research Really Prove?



Thomas Newkirk
Assistant Professor
University of New Hampshire

> In view of the widespread agreement of research studies based upon many types of students and teachers, the conclusion can be stated in strong and unqualified terms; the teaching of formal grammar has a negligible or, because it usually displaces instruction and practice in actual composition, even a harmful effect on the improvement of writing. (Braddock, Lloyd-Jones, and Schoer, 1963)

Writing teachers who look to research for guidance are likely to be disappointed, for they will meet with a mass of equivocation, contradiction, and "areas to be explored." The findings on the relationship of formal grammar instruction to writing seem to stand out, however, because the research, it is claimed, has been exhaustive, covering a period of seventy years, and the results have invariably shown that grammar instruction does not aid student writing. One noted British educator has put the matter even more strongly. Wilkinson (1971) has claimed that:

> ...by 1929, when R. L. Lyman brought out his Summary of Investigations, it was possible to assert a beneficial effect for formal grammar instruction through ignorance (or defiance) of the evidence of a large number of empirical studies. (p. 32)

It would seem that the matter is closed for all but the ignorant and the defiant.

Perhaps. But Braddock et al. themselves offer evidence that the research may not be as convincing as they claim. The authors noted that "Uncommon, however, is carefully conducted research which studies the effect of formal grammar on actual composition over an extended period of time. (p. 37)" There

would seem to be an inconsistency here. One the one hand the claim is made that the ineffectiveness of grammar instruction has been clearly established; on the other hand there is the admission that all but a few of these studies are seriously flawed.

The following is an examination of the research of the relationship of grammar instruction to writing. It is not an exhaustive survey of the research, much of which is available only in the form of heavily-abbreviated abstracts; rather it is an attempt to examine the methodology used to examine the question and to suggest even the few well-run studies do not offer clear-cut results.

## The Term "Formal Grammar Instruction"

The term "formal grammar instruction" has been used very loosely in many of the grammar studies. The most serious difficulty with the term is determining whether it refers to a) a body of grammatical principles and concepts, b) a method of teaching this content, or c) both. In the past most researchers have failed to distinguish between content and method. "Formal grammar instruction" was for them a composite of a content, which included a range of grammatical rules and terms, and a method, which included the memorization of these rules and terms, and the analyzing of sentences provided by the textbook. Far more time was spent on analysis than on the production and manipulation of sentence structures. While

this describes the way in which formal grammar was frequently taught, such a sweeping definition creates problems in interpretation.

Any experiment which attempts to examine one variable must be planned in such a way that groups vary in only one significant way. The problem with claiming that "formal grammar instruction" is a single variable is that even if the formal grammar group performs less well than some experimental group, it is impossible to determine what element of formal grammar instruction was responsible for the poorer performance. Was it the inherent uselessness of the concepts? Was it the deductive teaching approach? Was it the emphasis on analysis over production? Was it the failure to individualize? It is impossible to make these determinations because "formal grammar instruction" is not a single variable; it is a complex of variables. In effect, then, researchers have explored one method of teaching formal grammar as if it were the only method; alternatives have not been explored.

The major problem with definition in the history of the research has been the formal-functional distinction. In the the 1920's there was the widespread belief that much of the grammar taught in the schools had no practical effect on student writing or speaking performance. The thinking was that an effort should be made to investigate the oral and written errors of students, and, on the basis of this exam-

ination, to plan instruction to meet actual needs. By 1930 thirty-three studies of this type had been carried out ( Harap, 1930). Functional grammar was defined as:

> ...that application of the knowledge of grammatical items that will prevent the commission of an error in English or which will assist in the prevention of an error already made. ( Rivlin quoted in Leonard, 1933)

Meckel (1962) in his review of the research on grammar and writing, was especially critical of some of the advocates of functional grammar. He argued that they often viewed grammar as a kind of smorgasbord, from which items could be drawn without concern for orderly presentation. He also criticized these advocates for implying that formal grammar was essentially non-functional grammar:

> In some appraisals of the research there has been confusion between the term formal grammar as used to denote systematic study and mastery and the term used to mean grammar taught without application to writing and speaking. Systematic study does not preclude application. (p. 978)

It is fair to say that this confusion persists. Formal grammar, it often seems, cannot improve student writing because if grammar were taught in such a way that it did improve student writing, it would not be formal grammar. Formal grammar instruction, by definition, does not work.

## Correlation Studies

Early researchers used a correlation technique to determine the relationship between a knowledge of grammar and writing proficiency. The claim is often made that there is little

correlation between a knowledge of formal grammar and the ability to write, the implication being that such research demonstrates the futility of grammar instruction (Wilkinson, 1971). Researchers in the 1910's and 1920's, working with instruments unchecked for reliability found low correlations (.23-.30) between grammar knowledge and the ability to recognize errors. Later studies, however, found higher correlations. Segel and Barr (1926) found a correlation of .48 between grammar knowledge and the ability to select the appropriate grammatical alternative in a sentence. Robinson (1960) surveyed four British secondary schools and found high correlations in two (.69 and .64) and low correlations in two (.26 and .22) between knowledge of grammar and writing ability.

The largest correlation study was conducted by Wykoff (1945); it involved 5,125 Purdue freshmen and covered a period of six years. The study found that at all ability levels students with a knowledge of "usable grammatical terms" wrote better themes, at least 90% of the time, than students with a lesser amount of such knowledge. This would suggest a substantial correlation between grammatical knowledge and writing ability. In a more recent study O'Donnell (1963) attempted to determine the correlation between a knowledge of syntactic structure and writing ability. He concluded that in no case:

> ...(was) the degree of correlation...sufficiently high to provide a basis for the assumption that either knowledge of traditional grammar or awareness of the basic structural relationships of English will be regularly accompanied by excellence in written composition. (p.23)

In short, there is enough contradictory evidence to question the conclusion that there is a small correlation between a knowledge of grammar and writing ability.

The correlation technique is also highly suspect in itself. These studies are fundamentally limited because they can say nothing about causality. A high correlation in each of these studies would not prove that grammar instruction is beneficial because high correlations can exist between virtually unrelated phenomena. One of the standard examples is the high correlation between foot size and reading ability among children aged 6 to 11. This is so because older children generally read better than younger children and they have longer feet. It does not follow that elongating young children's feet will bring about any improvement in reading ability. Similarly, a low correlation does not necessarily indicate a lack of causality. For example, if, in some of the schools surveyed, grammar was not taught, or not taught well, it is likely that both good and poor writers would do poorly on grammar tests, resulting in a low correlation.

Another common claim is that correlations between grammar knowledge and writing ability are no greater than correlations between writing ability and proficiency in other subjects. (Wilkinson, 1971) Again the implication is that knowledge of grammar is of no more value to the writer than knowledge in such subject areas as mathematics. The danger in this line of reasoning can be illustrated by returning to the foot size-reading correlation. If it could be demonstrated

that the correlation between family income and reading ability were no higher than the correlation between foot size and reading ability, it would not be valid to conclude that because foot size was obviously not a factor, neither was family income. The point is that correlations can indicate many types of relationships, from causal to purely coincidental. The research itself can give no information about the type of relationship.

## Experimental Studies

Since the 1930's it has been thought that the true test of an educational method was the experimental test that matched one method against another. In 1931 Symonds questioned the value of the correlation studies of the value of grammar instruction and urged a shift to experimental research. Since then there have been numerous studies in the area, but few meet the minimal requirements for an experimental study; most are short term studies involving younger children. Many fail to control for teacher bias, and many are inadequately reported. Rather than going through this body of research study by study, I will examine four studies that are consistently cited and that seem to avoid some of the major weakness of the majority of studies (Ash, 1935; Frogner, 1939; Harris, 1962; and Elley et al., 1976).

Ash (1935). Ash compared what he called a "stylistic approach" in the teaching of writing to the standard West Vir-

ginia curriculum of 1929-30, a curriculum which presumably stressed traditional grammar although Ash does not state this explicitly. The stylistic approach emphasized the sentence, the relationship of sentences, word choice and the relationship of paragraphs. His sample included 9 classes (three 7th grade, three 8th grade, and three 9th grade) in three West Virginia schools. Two of these schools taught the standard curriculum and one taught the experimental curriculum. The experiment lasted one semester after which students wrote a theme that was evaluated on a number of criteria, ranging from the number of words, to types of sentences, to grammatical correctness, to the quality of the compositions. Ash reported greater gains on most of these criteria for the experimental group, although he admitted that the experimental group started well behind the other groups.

There are three serious problems in interpreting the results of this study. First, Ash reported the changes in percentage gains, and large percentage gains can hide very small numerical gains. Since the experimental group started well behind the control groups, equal numerical gains by both experimental and control groups would result in greater percentage gains for the experimental group. Secondly, these percentage increases vary wildly from grade to grade in the experimental group. One two criteria of sentence complexity there was wild variation in the experimental group increases:

| | grade 7 | 8 | 9 |
|---|---|---|---|
| compound-complex sentences | +250% | +27% | -82% |
| complex sentences | +225% | +77% | +2% |

Finally, it is not clear how much grammar was taught in the experimental groups. Ash noted that two units of grammar were taught "incidentally," and one of the directly-taught units, that on the relationship of sentences, included a section emphasizing:

> Varying the form of the sentence as to length, position of the phrases and clauses, the voice of the verb, the type of discourse, and the order of the subject, predicate, and object.(55-56)

Ash did not explain whether these grammatical terms were taught, whether it was assumed that students already knew them, or whether the concepts wer taught without the use of grammatical terminology. At any rate, Ash himself did not conclude that the teaching of grammar has no value, only that schools often teach more formal elements of grammar than stusents actually need.

Frogner (1939). Frogner compared the effect of a "thought method" to a method which combined the thought method with drill in grammar. She defined the "thought method" in an earlier article (1933) as an approach which emphasized:

> ...the accurate expression of relationships in thought, the recognition of the value of dependent members as contributors to the main idea they define, qualify, and develop.

Her sample consisted of 107 matched pairs, 47 at the 9th grade level and 60 at the 11th grade level. It is not clear how the teachers for the two approaches were selected. The experimental group (thought method) was taught 7 units which included coordination and subordination of ideas in clauses,

subordination of ideas in phrases, and recognition of the sentence. These units were taught to the experimental group without reference to grammatical terminology.

Frogner's study is generally cited as evidence of the ineffectiveness of grammar instruction (Sherwin, 1969). Her results, however, are also difficult to interpret. At the end of each unit students were given a test on the material taught in that unit, and at the end of the semester-long experiment all students were given a general test on grammatical knowledge and on awareness of sentence structure. None of the evaluations required actual student writing. The tests at the end of individual units showed no difference between groups, but the general test on sentence structure at the end of the study showed the thought method was superior on all 18 criteria, three of these differences being statistically significant. Frogner does not state the level of significance she was referring to, nor did she state what these three criteria were. Later in the study she claimed that in one school where thirty-three 9th grade pairs (almost a third of her sample) were taught, the combined method students made superior gains in the sentence structure test in 17 cases, and the thought method students made superior gains in 16 cases although the margins of superiority were slightly larger for the thought method group.

Frogner went on to claim that the thought method was superior in all cases where IQ was 105 or below, that the

two methods were equally effective in the 106-113 group, and in the 114-129 group the thought method was superior. This claim raises a question. If the combined method was so generally inferior, how, in fact, did more combined method students better their thought method partners in the one school for which she gave figures? Either almost all of the students fell in the 106-113 range, which, if true, would raise questions about the representativeness of the sample; or the one school she cited could have been an exception where the combined method students did markedly better than those in other schools, which, if true, causes one to wonder why she cited that school. Her failure to give this kind of breakdown for all schools is a serious omission.

Frogner seems justified in claiming that her study shows that formal grammar instruction is not essential in developing an awareness of sentence structure, and her elaboration of the thought method is a major contribution, one which anticipated the sentence-combining activities that are currently being advocated. Her study does not, however, demonstrate the clear superiority of the thought method, much less the ineffectiveness of grammar instruction.

Harris (1962). Harris conducted a long-term (two year) study in which he compared an approach to sentence structure which emphasized formal grammar, to one which emphasized extensive writing and an intuitive, non-terminological approach to sentence errors. Both groups were taught the same curriculum

four out of five periods each week, but instruction differed for the fifth period. His sample included 228 students drawn from five London secondary schools. At the end of the study the students were assigned to write a composition on a given topic, and the resulting compositions were evaluated on a number of criteria. Harris concluded thatof the 25 differences on the five most reliable criteria (words per common error, different sentence patterns, subordinate clauses, complex sentences, and number of correct sentences) six favored the grammar group, none at a statistically significant level, and 19 favored the the informal group, 10 at a statistically significant level. On the basis of these differences, it is generally claimed that the experimental group wrote with greater complexity, and with fewer errors than the control group (Petrosky, 1977).

There are serious weaknesses in the study, however. The concluding writing sample, one 80-minute essay on an assigned topic, is generally thought inadequate for measuring writing ability (Mellon, 1976). Harris failed to demonstrate that the teachers in the control group were as competent as those in the experimental group. There is also the question whether the Harris study actually demonstrated that the experimental group wrote with greater complexity. Mellon (1969) has made the astute observation that this claim rests on the fact that the experimental group wrote more complex sentences, although the study shows that there was no significant difference between the groups in the number of subordinate clauses written. Mel-

lon concluded that the two groups probably wrote at an equal level of complexity with the experimental group writing more complex sentences and the control group writing more subordinate clauses per complex sentence.

The argument for the superiority of the experimental method then comes down to the claim that the experimental group made fewer errors. While the tally of errors would seem to clearly show this, there is a question about Harris' concept of an "error." Braddock, Lloyd-Jones, and Schoer (1963) suggest that while Harris does not define "error," his examples indicate a rigid notion. For example, of the following two sentences:

> He found only three ducats.
> He only found three ducats.

the second would presumably be classified as "faulty positioning of adverb." Other categories of errors also look suspicious. Almost 50% of the control group was guilty of "misuse of prepositions," and one wonders if students did anything more serious than ending a sentence with a preposition. It is at least possible that many of the "errors" represent not so much grammatical errors, but the failure of students to write at the expected level of formality.

Elley et al. (1976). This study, carried out in one large New Zealand secondary school, attempted to compare three curricula, two of which included grammar (traditional in one and transformational in the other) and one which substituted

for grammar instruction additional practice in reading and writing. The Elley study avoids two of the major problems in the Harris study. First, the students were evaluated on a series of compositions at the end of each of the three years the study went on. Second, the three teachers in the study each taught all three approaches, thus minimizing the problem of controlling for the teacher variable. The researchers also seemed to exert more control over the actual instruction than did Harris.

The study involved 248 students in eight matched 3rd form (age 14-15) classes and followed them into their fifth year. The three curricula were as follows:

Group 1: Oregon Curriculum (transformational grammar strand, literature strand, and rhetoric strand).

Group 2: Oregon Curriculum (literature strand and rhetoric strand) free reading and creative writing.

Group 3: Traditional grammar; class sets of plays, drama, and poetry; writing exercises from the text

In addition to the essays which were evaluated for mechanics, style, structure, and content (and a total of these), students were tested for a number of skills:

Reading vocabulary
Reading comprehension
Sentence combining
English usage
English literature

While there were isolated differences between groups which reached a level of significance, there was little difference between the three groups at any point of evaluation. It was, in effect, a dead heat. The researchers concluded that the results

demonstrate the ineffectiveness of grammar instruction:

> It is difficult to escape the conclusion that English grammer, whether traditional or transformational, has no beneficial influence on the language growth of typical secondary students. (p.18)

This conclusion goes well beyond the actual findings of the study. All that can be stated conclusively is that all three curricula worked equally well ( or equally poorly). One cannot pick one element out of the curriculum and say, in effect, that the growth that students made was not due to this element but was due to some other element. In other words, it would be as logical to state that extensive reading and writing have no effect on language growth because the group that read and wrote extensively made no more growth than groups which spent time on learning grammar.

Petrosky (1977) has claimed that the Elley study confirms the results of the Harris study, but in a way it contradicts the Harris findings. As noted above the Harris study rests its claim for the superiority of the experimental method on the fact that students made fewer errors. In the Elley study one of the categories where group two (extensive reading and writing) did significantly worse than the grammar groups was English usage. The researchers claimed that the differences were in relatively minor areas, but admitted that one of these "minor areas" was run-on sentences.

Those reviewing the research on grammar instruction have taken their conclusions a step farther. Petrosky, commenting on the Harris and Elley study, concluded:

> The study of grammar, while serving no ascertainable purpose, also exists at the expense of proficiency in reading and writing. (p. 88)

The surprising thing about the Elley study is that this is precisely what it did not prove. Those students who studied grammar performed just as well on the tests of reading and writing proficiency as did students who spent this time on extensive reading and writing.

## The Difficulty of Learning Grammar

There is considerable evidence that grammar, as it is traditionally taught, is difficult for students, particularly younger students, to master. In the Harris (1962) study, for example, only one of the five classes that were taught grammar averaged over 50% on the final grammar test. This result led him to conclude that it was inadvisable to teach grammar to early secondary students. In an earlier study carried out in Scotland, Macauley (1947) found that after several years of grammar instruction early secondary students had extraordinary difficulty recognizing the four basic parts of speech.

There is also evidence that students of average and below average intelligence have difficulty learning grammar as it is traditionally taught. Meade (1961) tested 104 high school seniors in a high school that stressed the teaching of grammar. He found that only 6% of the students with an IQ of 104 or below could answer 70% of the questions correctly, even though they had presumably been studying the material for

several years. Meade concluded that those students who can learn the principles of grammar have the opportunity of doing so without suffering the tiresome repetition of content which bright students often must face. For those students who have little chance of succeeding with the learning of grammar, Meade suggests that the content be eliminated from the curriculum.

## Conclusions and Discussion

Based on the research evidence, the following conclusions seem justified:

1. The term "formal grammar instruction" is vague, seeming to apply to a wide range of practices, many of which have not been experiemntally examined.

2. Correlation studies have come up with a wide range of correlations between grammar knowledge and writing ability.

3. The correlation technique, because it does not examine causality, can provide no information about the effects of grammar instruction.

4. There is no evidence that grammar, as it is traditionally taught, has a noticable effect on writing improvement. Yet it should be noted that few of the accepted practices in the teaching of writing have been experimentally validated.

5. Although there have been a great many studies of the relationship of grammar instruction to writing, few meet the minimal requirements for experimental research.

6. The few long term studies that do meet the criteria for an experimental study do not clearly establish the ineffectiveness of grammar instruction, nor do they clearly establish the superiority of alternate direct methods for teaching error avoidance.

7. There is evidence that younger secondary students and students of average and below average intelligence have difficulty learning grammar as it is traditionally taught.

8. Studies have concentrated on the relationship of grammar instruction to error-avoidance. Few have dealt with the possible relationship of grammar instruction to sentence construction and writing style.

9. Shaughnessy (1977) has described a method for teaching basic writing students to eliminate errors, and the teaching of formal grammar is a part of this method. It would be helpful if a case study of this method were conducted. Such a study could illustrate the precise way in which grammar instruction is incorporated into a teaching approach, and it would illustrate the results of the instruction.

The questions concerning the teaching of grammar are as much political questions as they are research questions. Those who advocate a return to the teaching of "basics" frequent- mean a return to workbook grammar. Those who oppose this "movement" are fond of citing research claims that declare the teaching of grammar does not help the student writer. In effect, the debate has become polarized, the question either/or. The research itself has become polarized, a weapon in the arsenal of the forces of light.

The intent of this review has not been to give solace to those who want to return to the workbook, and to the easy assumption that a knowledge of grammar will magically transfer to the student's writing without a need for the teacher to demonstrate its usefulness. Rather it has been to suggest that the question of the relationship of grammar instruction

to writing has not been conclusively resolved, and that the strident debate over the issue has perhaps done more to obscure than to clarify it. There may be a middle ground. There may be a way of teaching grammar so that it shows the student what he can do rather than continually reminding him what he cannot do. There may be a way of avoiding what has been called "the religion of nomenclature" in the teaching of grammar. The possibilities exist. The question is not closed.

References

Ash, I. C. An experimental evaluation of the stylistic approach in teaching written composition in the junior high school. Journal of Experimental Education, 1935, 4 (1), 54-62.

Braddock, R., R. Lloyd-Jones, and L. Schoer. Research in Written Composition. Urbana, Ill.: National Council of Teachers of English, 1963.

Elley, W.B. et al. The role of grammar in a secondary school English curriculum. Research in the Teaching of English, 1976, 10 (1), 5-21.

Frogner, E. Grammar approach versus thought approach in teaching sentence structure. English Journal, 1939, 28 (7), 518-526.

_____. Problems of sentence structure in pupils' themes. English Journal, 1933, 22 (9), 742-749.

Harrap, E. The most common grammatical errors. English Journal, 1930, 19 (6), 440-446.

Harris, R . An experimental inquiry into the functions and value of formal grammar in the teaching of English, with special reference to the teaching of correct written English to children aged 12-14. Unpublished Ph. D. dissertation, University of London, 1962. Summarized in Research in Written Composition, 70-83.

Leonard, J. Functional grammar-- what and where? English Journal, 1933, 22 (9), 729-735.

Macauley, W. The difficulty of grammar. The British Journal of Educational Psychology, 17 (3), 153-162.

Meade, R. Who can learn grammar? English Journal, 1961, 50 (2), 87-92.

Meckel, H. Research on teaching composition and literature. In N. Gage (Ed.), Handbook of Research on Teaching. New York: Rand McNally, 1963.

Mellon, J. Transformational Sentence Combining (Research Report # 10). Urbana, Ill.: National Council of Teachers of English, 1969.

_____. Round two of the national writing assessment-- interpreting the apparent decline in writing ability: a review. Research in the Teaching of English, 1976, 10 (1), 66-75.

O'Donnell, R. The Correlation of Awareness of Structural Relationships in English and Ability in Written Composition, United States Office of Education Cooperative Project No. 1524. Mount Olive, N. C. : Mount Olive College, 1963.

Petrosky, A. Grammar instruction-- what we know. English Journal, 1977, 66 (9), 86-88.

Robinson, N. The relationship between knowledge of grammar and ability in composition. British Journal of Educational Psychology, 1960, 30 (part 2), 184-186.

Segel, D. and N. Barr. Relation of achievement in formal grammar to achievement in applied grammar. Journal of Educational Research, 1926, 14 ( ), 401-402.

Shaughnessy, M. Errors and Expectations. New York: Oxford University Press, 1977.

Sherwin, J. Four Problems in Teaching English: A Critique of the Research. Urbana, Ill.: National Council of Teachers of English, 1969.

Wilkinson, Andrew. The Foundations of Language. London: Oxford University Press, 1971.

Wykoff, G. The relation of a knowledge of grammar and punctuation to writing. Educational Administration and Supervision, 1945, 31 (1), 54-62.