

DOCUMENT RESUME

ED 365 431

PS 021 845

AUTHOR Huitt, William G., Ed.
 TITLE Proceedings from the Annual Graduate Research Symposium (5th, Valdosta, Georgia, May 20, 1991).
 INSTITUTION Valdosta State Coll., GA. School of Education.
 PUB DATE 20 May 91
 NOTE 71p.; For 1990 Symposium Proceedings, see ED 333 334.
 PUB TYPE Collected Works - Conference Proceedings (021)

EDRS PRICE MF01/PC03 Plus Postage.
 DESCRIPTORS Academically Gifted; Comparative Analysis; Cooperative Learning; Elementary Secondary Education; Graduate Study; Group Dynamics; Higher Education; *Instructional Effectiveness; *Mathematics Instruction; *Music Education; Parody; Problem Solving; Self Concept; Teaching Methods; *Writing Instruction; Writing Research

IDENTIFIERS Explicit Instruction; *Project Follow Through; Writing Style

ABSTRACT

Covering a wide variety of issues of concern to educators, the papers compiled in this proceedings report present graduate student research on writing instruction, mathematics instruction, cooperative learning groups, a follow-up program to Project Head Start, relationships between language and music, and the self-concept of gifted children. The papers are as follows: (1) "Effect of Writing Parodies on the Writing Style of Advanced Ninth Graders" (J. Albert Rowell); (2) "An Evaluation of Explicit Instruction in Problem-Solving Heuristics in Mathematics" (Wanda L. Sumner); (3) "Effects of Group Processing Skills on the Social Interaction of Cooperative Learning Groups" (Emilie Allen); (4) "Have the Results of Project Follow Through Been Well Utilized?" (Janis N. Lee); (5) "Relationships between Language and Music" (Linda J. Star); and (6) "The Self-Concept of Gifted Children: A Developmental and Comparative Study" (Polyann Diamond). (MM)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

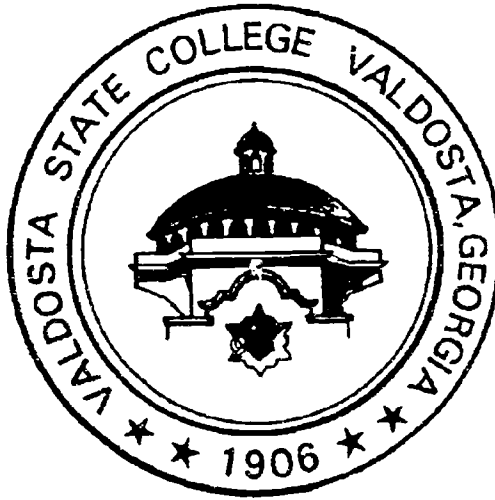
U. S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as
received from the person or organization
originating it.
Minor changes have been made to improve
reproduction quality.

Points of view or opinions stated in this docu-
ment do not necessarily represent official
OEI position or policy.

ED 365 431

PROCEEDINGS FROM THE FIFTH ANNUAL GRADUATE RESEARCH SYMPOSIUM



SCHOOL OF EDUCATION

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

*William G.
Hurt*

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

PS 45

Editorial Review Board

Glenda L. Akins, Early Childhood and Reading Education

C. Jay Hertzog, Middle Grades Education

William G. Huitt, Psychology, Counseling, & Guidance

Robert A. Hull, Special Education

Thomas G. Jackson, Health, Education, & Athletics

Robert O. Michael, Educational Administration & Supervision

Thomas M. Reed II, Special Education

Tonja Root, Early Childhood and Reading Education

Robbie J. Strickland, Secondary Education

Valdosta State College, a unit of the University System of Georgia, does not discriminate on the basis of race, religion, sex, color, national origin, age, disability, marital status, sexual orientation or as a disabled Vietnam-Era veteran in its education programs or employment.

Valdosta State College is an Affirmative Action/Equal Opportunity Employer

BEST COPY AVAILABLE

3

**PROCEEDINGS FROM THE FIFTH ANNUAL
GRADUATE RESEARCH SYMPOSIUM**

**SPONSORED BY
THE SCHOOL OF EDUCATION
VALDOSTA STATE COLLEGE
MAY 20, 1991**

WILLIAM G. HUITT, EDITOR

**The publication of these proceedings is funded by the Regional
Education Center, School of Education, Valdosta State College**

Table of Contents

Effect of Writing Parodies on the Writing Style of Advanced Ninth Graders J. Albert Rowell	1
An Evaluation of Explicit Instruction in Problem-Solving Heuristics in Mathematics Wanda Sumner	11
Effects of Group Processing Skills on the Social Interaction of Cooperative Learning Groups Emilie Allen	19
Have the Results of Project Follow Through Been Well Utilized? Janis Lee	29
Relationships Between Language and Music Linda Star	37
The Self-Concept of Gifted Children: A Developmental and Comparative Study Pollyann Diamond	45

Preface

The selections presented in this volume are edited versions of papers presented at the Fifth Annual Graduate Research Symposium sponsored by the School of Education at Valdosta State College on May 20, 1991. The purpose of the graduate research symposium is to acknowledge superior work of graduate students within the School of Education and to encourage students to engage in research activities.

The Symposium owes its beginnings to efforts made by the Writing Across the Curriculum Committee at Valdosta State College. This college-wide committee was established to increase awareness of the importance of writing as an academic and learning tool. The committee was appointed by the Vice-President of Academic Affairs and held a number of workshops on specific techniques that could be used by college faculty to impact students' writing skills. As more faculty encouraged their students to engage in writing at the graduate level, a means was needed to recognize the excellent student work being produced.

The first three papers in this volume describe research studies completed by authors in their classrooms. The first paper, "Effect of writing parodies on the writing style of advanced ninth graders," by J. Albert Rowell describes a research study whose focus is the improvement of writing. Wanda Sumner's paper "An evaluation of explicit instruction in problem-solving hueristics in mathematics"

describes using a variety of strategies to teach problem solving to fifth-grade students. In her paper entitled "Effects of group processing skills on the social interaction of cooperative learning groups," Emilie Allen reports on teaching social skills to third-grade students using cooperative learning. All authors found support for their treatments in improving student skills.

The use of research results is the focus of Janis Lee's paper entitled, "Have the results of Project Follow Through been well utilized?" Ms. Lee considers the results and controversy resulting from the Follow Through experiment of the 60s and 70s and suggests how dissemination of federally-funded research could be improved.

In "Relationships between language and music," Linda Star considers the relationship between the two symbol systems of language and music. Based on her review, she suggests that music should occupy a more prominent place in education curriculum.

In the last paper, Pollyann Diamond reports on her research, "The self-esteem of gifted children: A developmental and comparative study." The results suggest that students identified as gifted see themselves as more scholastically competent than their non-gifted peers. However, the self-esteem of students selected for a special gifted program remained high, while the self-esteem of students not selected for a special program declined.

Acknowledgements

As mentioned in the previous proceedings, these volumes owe a debt to many people. The members of the Writing Across the Curriculum Committee and the Vice-President for Academic Affairs are primarily responsible for providing the impetus to hold the original symposium. F. D. Toth, Dean of the School of Education, is primarily responsible for providing the resources to hold the symposium and publish the proceedings. The editorial review board, with the assistance of faculty readers Diane Andrew, Russ Deavours, John Hummel, Dan Kaeck, Ben McLain, and Marty Meyer, are primarily responsible for the selection of papers actually presented at the symposium. Ann Hilgert provided invaluable editorial assistance while Charlene McQuire provided typing assistance.

Biographical Sketches of Presenters

Emilie Allen holds a B.S. degree in Early Childhood Education from Auburn University and an M.Ed. in Early Childhood Education from Valdosta State College. She is currently a teacher at S. L. Mason Elementary School in Valdosta, Georgia.

J. Albert Rowell holds a B.A. degree in English and B.S. and M.Ed. degrees in Secondary Education, all from Valdosta State College. He is currently a teacher of English at Lowndes High School in Valdosta, Georgia.

Pollyann Diamond holds a B.S. degree in Music Education from Gettysburg College, an M.Ed. degree in Elementary Education from Armstrong State College, an M.Ed. degree in Educational and School Psychology from Mississippi State University, and an Ed.S. degree in School Psychology from Valdosta State College. She is currently a teacher in the Gifted Education Program for Valdosta City Schools, Valdosta, Georgia.

Linda Star holds B.M. and M.Ed. degrees in Music Education, both from Valdosta State College. She is currently a graduate assistant in Music Education at Valdosta State College.

Janis Lee holds a B.A. degree in history and an M.Ed. in Early Childhood Education from Valdosta State College. She currently teaches third grade at Hahira Elementary School in Hahira, Georgia.

Wanda Sumner holds an A.S. degree in mathematics from Abraham Baldwin Agriculture College, and B.S. and M.Ed. degrees in Middle Grades Education from Valdosta State College. She is currently a math/science teacher in Irwin County, Georgia and is enrolled in the Ed.S. program in Middle Grades Education at Valdosta State College.

Effect of Writing Parodies on the Writing Style Of Advanced Ninth Grade Students

J. Albert Rowell

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
J. Albert Rowell

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

This research project sought to determine if writing style and attitude toward writing may be improved through a guided process of writing parodies, exaggerated imitations of literary works. Two groups of advanced ninth-grade students at a rural Georgia high school participated in the study. The treatment group wrote four parodies over 8 weeks, while the control group wrote four traditional essays over the same period. The treatment group made statistically significant gains in writing style. Fluency and attitude about writing were not statistically significant between the treatment and control groups. The major conclusion was that writing parodies is an effective means of improving student writing style. Further research should attempt replication of this study and should determine if this process benefits students at other grade levels.

Prevailing methods of teaching composition generally fail to recognize and nurture student style. Emphasis upon correctness and the dominance of such artificial forms as the five-paragraph theme yield student writers who produce mechanically correct, stylistically deficient compositions. In Georgia, results of student writing assessments indicate significant deficiencies in writing style.

The state of Georgia has instituted a basic skills test to assess student writing competency. Students take the writing test in the fall of the 10th grade, and most individuals will meet or exceed the state-established passing score on the first at-

tempt. Trained raters evaluate student writing in five domains: content/organization, style, sentence formation, usage, and mechanics. Within each domain the student receives a score of inadequate, minimal, good, or very good. Among participants in the fall 1989 administration of the test, 70% received good or very good ratings for sentence formation; only 30% received inadequate or minimal (see Figure 1). Sixty-five percent received similarly high ratings for usage, and 60% scored at the good or very good level for mechanics. Sixty-six percent received good or very

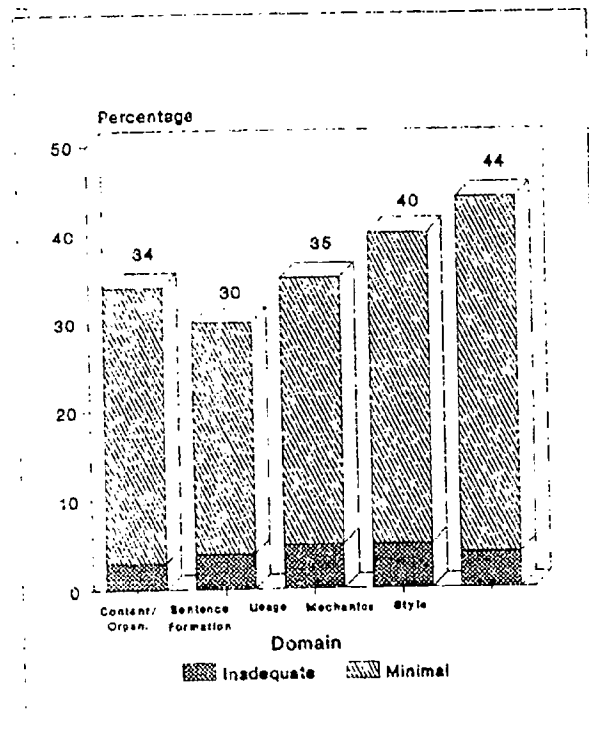


Figure 1. Deficiencies per domain in state writing test

good ratings for content/organization. Style, however, represented a weakness for many of the students; 44% were rated as inadequate or minimal on written composition and style (Georgia Department of Education 1989). Clearly, improved instruction is needed in the area of style.

To evaluate style in student compositions, raters judge a student's control of language by observing four criteria established by the Georgia Department of Education (1990): 1) concrete images and descriptive language; 2) readability; 3) varied sentence patterns; and 4) appropriate tone for topic, audience, purpose. These components of style are desirable outcomes of writing instruction, yet student writers in Georgia perform most poorly in this domain. One possible explanation is that teachers who are adept at teaching basic sentence structure, error detection, and outlining lack strategies for guiding students to an individual style. One teaching strategy that holds promise for encouraging student style is writing parodies of the works of established writers.

Imitation is a time-honored method of instruction across disciplines. When writing a parody, a student reads a work, identifies its fundamental structure and style, then writes an imitation which follows the original structure, but exaggerates key elements, including noticeable style traits. This research project investigates whether writing parodies significantly improves the writing style of ninth-grade students.

Review of the Literature

When teachers of English attempt to teach style, it is ironical they encounter difficulties, since the present generation of students seems to be particularly sensitive

to how things are done. In addition, students seem to want to be seen as unique, but in composition class students hesitate to reveal themselves as individuals. While precision in the mechanics of language is desirable, instructional focus should be revised to emphasize not only correctness but how students express themselves in writing. A survey of English teachers in the Southeastern United States found style to be the least emphasized aspect of composition (Graves, 1974). According to Corbett (1980), "The evidence supplied by annual bibliographies of style or of composition and by the articles that appear in other professional journals does not confirm that the teaching of style is currently flourishing in the writing classroom" (p. 80). The literature does describe some teaching approaches which aspire to substantially improve student writing through the use of models, imitation, and parody.

Teachers may avoid teaching the concept of style because the concept itself is not well defined. Jolly (1982) elaborates on the challenge to teachers of composition: "Teaching style, as opposed to dwelling on mechanical and grammatical conformity, is an elusive art, elusive because the word itself escapes exact definition. Perhaps Ella Fitzgerald's definition is adequate: 'It ain't what you do, it's the way howdya do it'" (p. 39). Narrow definitions limit style to choices of diction or sentence structure (Lindemann, 1987). Modern rhetoricians tend to reject the notion of style as ornamentation. Graves (1974) maintains, "Style is not mere decoration, nor is it an end to itself; it is rather a way of finding and explaining what is true. Its purpose is not to impress but express" (p. 189-190). Corbett (1980) offers a functional definition of good style as "not mere

ornament" but as a means of expression that "facilitates the conveyance of meaning to readers" (p. 85). Although individuals may disagree on its components, style is generally a distinctive, individual means of expression that clearly communicates with the reader.

How can teachers develop student writing style? Weathers (1970) suggests, "We can remind students of Aristotle's observation, 'character is the making of choices,' and point out that since style, by its very nature, is the art of selection, how we choose says something about who we are" (p. 187). Weathers proposes a specific exercise system: "We would advocate setting up recognized stylistic material as models; the models to be copied until the student can create similar but original versions of his own. It is a process of creative imitation . . ." (p. 189). In his review of research on* the use of models to teach writing, Hillocks (1986) found that studying models alone had a limited effect on writing quality. According to Hillocks, research indicates that using models is more productive than the study of grammar, and that models can help writers of various ages and abilities.

One of the oldest instructional methods is imitation. Its champions emphasize that this process leads ultimately to a distinctive, more individual style. Zinsser (1985) encourages the aspiring writer: "Don't ever hesitate to imitate another writer--every artist learning his craft needs some models. Eventually you'll find your own voice and shed the skin of the writer you imitated" (p.126). Jolly (1982) cites Lewis Mumford in support of imitation: "'Imitation then is but one device to hasten the acquisition of an individual's style, another tool to add to the writer's repertoire'" (p.

39). Gruber (1977) found that imitation produces measurably better writing with accurate diction, well-formed sentences, and overall improved organization. According to Gruber, creating imitative works allows students to acquire techniques to shape their own writing: "The act of imitation became a tool to achieve individual freedom; instead of stifling individual personalities, it liberated them" (p. 491). The process of imitation leads students to acquire individual style.

To further involve students in the active processing of a model, some individuals have introduced their students to parody, an exaggerated form of imitation, most often for humorous effect. By studying and writing parodies, students don masks, freeing themselves to experiment with new expression (Clark & Motto, 1986). Zahlan (1987) advocates parody as a means for students to develop individual style by "trying-on" the styles of established writers. According to Zahlan, "imitation to be effective must be informed, deliberate, creative" (p. 2). To facilitate development of student writing style, Zahlan advocates a program of assignments progressing from model analysis, through imitation and parody, to the realization of an individual style. Zahlan maintains that parody bridges the gap between acquisition of components of style and manifestation of individual style.

Style should be a vital concern for all teachers of composition. Parody holds particular promise as a means to develop style because the student must analyze and act on the original model. Parody involves students in the composing process and invites them, in a non-threatening way, to try on new styles. The literature indicates significant improvements in style are possi-

ble when parody is part of a systematic program of exposure and exercise, yet no research is available on the effectiveness of a parody-based program of intervention.

Method

This research project evaluated the effects of writing parodies on three dependent variables: student writing style, writing fluency, and student attitudes about writing. Two closely matched classes were selected for this research, and a previously developed writing scale measured writing style. Instruction was planned to provide the same experiences for treatment and control groups, except that treatment subjects would write parodies in lieu of the more traditional expository essays written by the control subjects. Pre- and posttreatment writing samples were evaluated to determine the effects of the treatment.

Subjects

Subjects for this study were ninth-grade students in two advanced literature classes at Lowndes High School, a large, comprehensive high school serving a predominantly rural population surrounding Valdosta, Georgia. A review of test scores--the language component of the Iowa Test of Basic Skills as well as gross scores and the style subscore for the eighth-grade Georgia Basic Skills Writing Test indicated the two groups of students were not significantly different. The classes selected for the study were taught by this researcher. The treatment group was selected by coin toss.

Instrumentation

Central to this project was the analytic style scale (see Appendix A) developed by the author (Rowell, 1990). An analytic design was selected to precisely measure four aspects of style and to reduce bias found in holistic ratings (Davis, Scriven, & Thomas, 1987). Patterned upon Diederich's (1974) five-point scale, the scale provided rating criteria for low, middle, and high ratings, which would receive 1, 3, and 5 points, respectively, for each of the four components of the scale. Raters could also award 2 and 4 points for a component, if the rater determined that an essay's quality fell between two rating criteria.

Three components of the scale--use of sentences, voice, and wording--were derived from previously existing scales. The criteria for use of sentences were modified from Cooper's (1977) criteria for syntax, and the criteria for wording were taken verbatim from Diederich (1974). The criteria for voice were originally a blend of descriptions for flavor in Diederich and voice in Cooper. After consultation with other teachers of English, this researcher added concreteness/imagery as a component of the scale. This component measures the degree of elaboration in discussion of a topic. The rationale for including this component was that students make qualitative writing choices which affect the clarity of their ideas; these choices are an indication of individual style.

Hypotheses

This research design tests three hypotheses:

1. Students who write parodies as their

experimental treatment will not achieve significantly better style scores from pretest to posttest than control students who write conventional narratives and expository essays.

2. Students who write parodies as their experimental treatment will not show significant gains in fluency from pretest to posttest when compared to control students who write conventional narratives and expository essays.

3. Students who write parodies as their experimental treatment will not demonstrate significantly improved opinions of the writing course or of their own abilities from pretreatment to posttreatment when compared to control students who write conventional narratives and expository essays.

Design and Procedure

This research employed a quasi-experimental pretest-posttest control group design. Control and treatment groups were not randomly-assigned but were preexisting literature classes taught by the researcher. Both classes followed the same program of instruction, reading, and testing. Writing assignments were the independent variable for this project. Subjects in the control group wrote a narrative composition and three expository essays typical of English composition assignments. Subjects in the treatment group wrote four parodies of specific works studied in the class. This research project considered three dependent variables: student writing style, as evaluated by the analytic style scale; fluency, as measured by total number of words written on a timed writing assignment; and students' opinions of the writing course and

of their own abilities, as measured by a researcher-developed questionnaire.

Pretreatment assessment for both groups involved a writing sample and a questionnaire. Pretreatment writing samples for this project were written in 52-minute class periods on the same day for both groups. These essays were then typed to eliminate the influence of handwriting on the raters' evaluations; the typed essays were identified only by subject-selected, six-digit numbers. On the day following the initial writing assessment, subjects completed a researcher-developed questionnaire of their opinions of writing, the course work, and their writing abilities. The preliminary questionnaire used a four-point Likert scale to measure subjects responses to eight statements.

During the treatment period of 8 weeks, the subjects in both groups completed four units of a course in literary genres: short story, drama, novel, and nonfiction. To conclude each unit, the subjects completed a writing assignment. Control subjects wrote a narrative and three expository essays during the research, while treatment subjects wrote a parody after each unit of study. The protocol for the treatment group included discussion to identify style characteristics for each writer studied. Subjects in both groups wrote on the same days; only writing topics differed.

The same procedures were followed for the posttest as for the pretest. The subjects wrote their essays in a 52-minute class period. At the beginning of the class, the subjects were given their topics and directions as before. The subjects identified their essays with a self-selected, six-digit number. The essays were typed and identified only by the six-digit number then

mixed with the pretest essays and placed in numeric order according to the six-digit identification number. With these numbers and the same four-point rating scale on pre- and posttest assessments, raters did not know if a paper originated from a pre- or posttest or from a control or treatment group.

On the day following the posttest, the control and treatment subjects completed a researcher-developed questionnaire. This questionnaire was identical to the one the subjects had previously completed, except for the addition of an open-ended response question. Again, the questionnaires were only identified by class.

Five experienced teachers of composition volunteered to rate the essays for this project. These same raters piloted the original analytic style scale and were familiar with its application. By the end of the evaluation sessions, each essay had received two evaluations. For this research a discrepant score was defined as a difference of five points or more between the total scores assigned by the two raters. In the event of discrepant scores, the essay was given to a third rater with the previously assigned scores omitted. The score farthest from the mean of the three scores was discarded.

Questionnaire results and essay ratings were then tabulated, recorded, and analyzed. Total word counts for each essay, necessary to evaluate treatment effect on fluency, were generated by the word processing software used during the typing of the essays prior to evaluation by the raters.

Results

This research was designed to determine if the writing of parodies by students

of English would have any significant effect on students' attitudes about writing, their fluency in controlled writing situations, and their writing style as evaluated by raters using an analytical style scale. Using the Statistical Package for the Social Sciences (SPSS/PC+), the researcher computed means for control and treatment groups. Gains in each area were determined by calculating the means of post-treatment scores less pre-treatment scores, and the significance of any gains were analyzed by using *t* tests.

The writing of parodies had no significant effect on student attitudes about writing or on their opinions of their skill as writers; probability levels were at .245 and above, too high to assume any significant effect of the treatment. The open-ended response question on the posttreatment questionnaire generated some response.

The control subjects were generally vague in their comments about completed writing assignments: a typical response read, "Students who take this course in [the] future should be required to complete these kinds of assignments because it helps to improve your writing abilities for the future." Another student commented, "They are helpful in their own way, but they are boring, not interesting." The students who wrote parodies appeared more emphatic in their opinions; most endorsed the use of parodies, although a few students suggested limiting the number. One subject responded, "Students should be required to complete these types of writing assignments because they need to learn different types of writing. Parodies are a lot more fun than regular assignments so long as you don't give too many parodies to the class." Another wrote, "They get boring after one or two. But they make you

think." One subject remarked, "It's very hard to think of something to write on. I think students should have a choose [sic] to write a parody or a regular essay."

A second aim of the research was to determine if the writing of parodies made students more fluent writers, as measured by the number of written words. Both the treatment group and control group improved in fluency from pretreatment to post-treatment; differences in gains between the groups were not statistically significant.

In the qualitative measure of student writing style, the treatment group did score several statistically significant gains. The reported means are for the sums of the scores assigned to the papers by two raters. Each component of the style scale could receive a score from 1 to 5; therefore, the sum of scores for each component would range from 2 to 10. The range for total scores was 8 to 40. A review of mean

scores for each component and total scores shows that on the initial assessment the treatment group performed more poorly than did the control group; however, the treatment group demonstrated significant gains in writing style following the treatment period (see Table 1). The group of subjects who wrote parodies marked measurable gains on all four components of style and actually exceeded the control group's performance on three of the four components on posttreatment assessment. The performance of the control group remained essentially the same from initial to final assessment. The treatment group demonstrated marked gains on all four components of the style scale; gains for voice and concreteness were significant ($p < .05$). Gains for the treatment group on total style scores were also statistically significant ($p < .05$) (see Table 2).

Table 1. Means for Scores on Student Writing Assessments

	Group	Pretreatment	Posttreatment	2-Tail p ^a
Sentences	Treatment	5.05	6.05	.06
	Control	5.52	5.52	
Voice	Treatment	4.36	5.77	.04
	Control	5.57	5.81	
Concreteness	Treatment	4.05	5.55	.04
	Control	4.95	5.24	
Wording	Treatment	4.91	5.91	.07
	Control	5.19	5.33	
Total	Treatment	18.36	23.27	.03
	Control	21.24	21.90	

^aProbability levels for *t* tests comparing mean gains of treatment vs. control.

Table 2. Pretreatment to Posttreatment Gains in Total Scores

Group	N	Mean Gain	Standard Deviation
Treatment	22	4.91	5.87
Control	21	0.67	6.04
Pooled Variance Estimate			
	t Value	Degrees of Freedom	2-tail probability
	2.33	41	.025

Results of this research disprove the null hypothesis regarding the influence of parody writing on the development of student writing style. Directed writing of parodies of works read in class lead to statistically significant improvements in student writing style. The null hypotheses regarding the influence of writing parodies on student fluency and on student attitudes about writing have been supported. The writing of parodies has not been shown to have a statistically significant effect on either student writing fluency or on student attitudes.

Conclusions and Implications

This research demonstrated that improvement in student writing style can be effected in a relatively short period of time by having students experiment with various styles through the writing of parodies. Several factors may have contributed to these improvements in writing style. First, the instructional method and requirements of the writing assignments may have com-

pelled the treatment group to consciously consider and to imitate writing styles of various writers studied during the treatment period. Second, the treatment group appeared to have been more personally engaged in writing parodies than was the control group in writing essays. Treatment students were able to express their senses of humor as they attempted to elicit laughter through their writing; control students, in general, reflected little of their personalities in their essays. This aspect of the parody exercises may explain the significant gains in voice scores by the treatment group.

Several aspects of this study may have influenced its results. First, the treatment period of this research was limited to 8 weeks, which permitted the writing of four parodies in the treatment group and four traditional writing assignments in the control group. During a longer time period, different results, especially pertaining to affect, may be possible.

Second, the researcher taught both groups and possibly allowed expectation

biases to intrude upon the experimental setting. Of course, any research situation is suspect when the researcher is directly involved in administration of the treatment. This researcher exercised care to provide nearly identical ministrations over the experiment period to both the treatment and control groups; however, researcher expectancy is a threat that can be eliminated only through replication by educators who do not have an interest in the experimenter's results. This researcher has somewhat successfully used parodies with other levels of students. Research could be undertaken to measure the effectiveness of writing parodies on the writing style of students other than those at the gifted or advanced level.

Third, subjects for the study were advanced ninth-grade students. Results may not generalize to the broader population.

In conclusion, writing parodies offers teachers an exciting opportunity to integrate composition instruction and the study of literature. Guiding students through the writing of parodies actively involves teachers and students in considering the style of various writers, and the inherent element of humor promotes student engagement in these exercises. This research suggests that writing parodies may be one means to develop student writing style.

References

- Clark, J., & Motto, A. (1986). The uses of parody and excess in composition. Exercise Exchange, 31(2), 11-13.
- Cooper, C. (1977). Holistic evaluation of writing. In C. Cooper & L. Odell (Eds.), Evaluating writing: Describing, measuring, judging (pp. 3-31). Urbana, IL: National Council of Teachers of English.
- Corbett, E. (1980). Ventures in style. In A. Freedman & I. Pringle (Eds.), Reinventing the rhetorical tradition (pp. 79-87). Ottawa: Canadian Council of Teachers of English.
- Davis, B., Scriven, M., & Thomas, S. (1987). The evaluation of composition instruction (2nd ed.). New York: Teachers College Press.
- Diederich, P. (1974). Measuring growth in English. Urbana, IL: National Council of Teachers of English.
- Georgia Department of Education (1989). Basic Skills Writing Test: System report for Lowndes County. Atlanta: Author.
- Georgia Department of Education (1990). Teacher's guide: Georgia Basic Skills Writing Test. Atlanta: Author.
- Graves, R. (1974). A primer for teaching style. College Composition and Communication, 25, 186-190.
- Gruber, W. (1977). "Servile copying" and the teaching of English composition. College English, 39, 491-497.
- Hillocks, G., Jr. (1986). Research on written composition: New directions for teaching. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills and National Conference on Research in English.
- Jolly, P. (1982). A question of style. Exercise Exchange, 26(2), 39-40. (ERIC Document Reproduction Service No. ED 236 601)
- Lindemann, E. (1987). A rhetoric for writing teachers (2nd ed.). New York: Oxford.
- Rowell, J. A. (1990). Development and validation of an analytic scale for eval-

uating writing style. Unpublished manuscript.

Weathers, W. (1970). Teaching style: A possible anatomy. College Composition and Communication, 21, 144-149. In G. Tate & E. Corbett (Eds.), The Writing Teacher's Sourcebook (pp. 187-192). New York: Oxford University Press.

Zahlan, A. (1987). Liberating parody: Strategies for teaching style. (ERIC Document Reproduction Service No. ED 285 191)

Zinsser, W. (1985). On writing well: An informal guide to writing non-fiction (3rd ed.). New York: Harper and Row.

Wanda L.
Sumner.

An Evaluation of Explicit Instruction in Problem-Solving Heuristics in Mathematics

Wanda L. Sumner

This study investigated the effectiveness of teaching problem-solving heuristics on a daily basis to 66 fifth graders in a small, rural school. The explicit teaching of problem-solving strategies was compared to the basal arithmetic program. The strategies included using logical and spatial reasoning, making lists or tables, working backwards using flow charts, finding patterns, and drawing pictures. Results indicated that explicit teaching of problem-solving strategies was more effective than using the basal arithmetic program to teach problem solving to fifth-grade students. Differences in attitudes were noted. While all students initially expressed fear and dislike of word problems, after treatment the explicit-treatment group reported a more favorable attitude toward problem solving than did the basal-textbook group.

George Polya (1957) stated that the first duty of all mathematics teachers is to develop students' abilities to solve problems. According to the National Council of Teachers of Mathematics (NCTM) (1989), problem solving, the focus of school mathematics for the 1980s, is not treated as a separate content area but is the process by which students learn and do mathematics.

The development of problem-solving abilities using higher order thinking skills is the major goal of mathematics, but many students have difficulty problem solving. Results of the Fourth National Assessment

of Educational Progress (NAEP) (cited in Kouba et al., 1988) indicated that students' difficulties with verbal problems are not caused by an inability to read the problem or by a lack of computational skills. Students could identify the correct operations for one-step problems but had difficulty analyzing nonroutine and multi-step problems.

Kouba et al. (1988) presented several reasons for students' difficulties in solving word problems. They suggested that students lose track of information and resort to guessing without considering the reasonableness of an answer or use inappropriate strategies, such as key words or number size, when solving multi-step or nonroutine word problems. Moreover, students find solving word problems and nonroutine problems perplexing because they are unsure of how to decide which strategy to use.

Even though developing skill in problem solving is a major goal in mathematics, instruction in this area is troublesome because problem solving is a complex process rather than a set of algorithmic skills (NCTM, 1989). Swing and Peterson (1988) noted that problem solving requires students to use mathematical concepts to establish links between problem information and everyday meaning. A problem is a dilemma in which a person wants something but does not immediately recognize what to do to get it. Routine story problems concern the translation and application of a simple arithmetic procedure; nonroutine problems demand more reason-

ing, since the choice of mathematical schema is not clear. Problem solving is then defined as the process of applying previously learned knowledge to new and unfamiliar settings.

Problem-solving heuristics are guidelines that can be applied across domains. Heuristics emphasize higher order thinking skills such as questioning, analyzing, translating and illustrating results, working backwards, drawing pictures and diagrams, and using trial and error strategies (Polya, 1957). Problem-solving strategies simplify the process of translating and solving problems.

Teachers of mathematics can no longer assume that the ability to solve problems will develop with time and maturity. Problem solving needs to be taught in a systematic rather than an incidental way. This study tested the effectiveness of explicitly teaching problem-solving heuristics to fifth graders on a daily basis with emphasis on higher order thinking skills using Polya's strategies. The explicit method of teaching problem solving was compared to the approach used in the basal arithmetic program.

Review of the Literature

The National Council of Supervisors of Mathematics (NCSM) (1989) stated that students who graduate in the year 2001 will frequently face a 19th century, computation-dominated mathematics curriculum. However, due to changing technology and essential competencies, students must develop mathematical skills and become expert problem solvers. Further, students need to ask questions, to become risk takers, and to develop cooperative learning skills in solving and posing problems.

Finally, students need to analyze and find alternative solutions for problems and to check for reasonableness of results.

To become expert problem solvers, students need to solve word problems from text and real life settings. Students also need to encounter questions with multiple answers. Suydam (1982) stated that "one major conclusion is warranted from the research: teach problem solving strategies!" (p. 59). Suydam indicated that problem-solving strategies teach flexibility and give students a repertoire from which to draw as they meet a wide variety of problems.

Heightened awareness of the importance and of the difficulty in teaching problem solving has resulted in increased efforts to identify specific instructional. Robert Gagne (cited in Darch, Carnine, & Gersten, 1984) indicated that attention is needed in translating "concretely stated problems into mathematical form. Emphasis needs to be given to this phase as a separate kind of capability to be learned by students" (p. 359). Students have difficulty choosing the correct operation to complete a mathematical equation or sentence (Fischbein, Deri, Nello, & Marino, 1985). These researchers believe students' concrete models of division and multiplication hinder their ability to choose the appropriate operation to solve word problems. Sellke, Behr, and Voelker (1991) concluded that data tables, used to represent relationships, helped overcome the constraints of intuitive models. Darch et al. (1984) found a significant difference in problem-solving ability among students taught to translate story problems into mathematical equations using in an explicit, step-by-step approach. The researchers inferred that with explicit teaching of problem-solving processes, students who understand the

structure of the operations no longer depend upon superficial clues and can use their knowledge flexibly to solve word problems.

Unlike translation problems, process or nonroutine problems require more general strategies. Difficulty depicting process problems may even cloud the planning of solution strategies. Several studies (Charles & Lester, 1984; Essen & Hamaker, 1990; Lee, 1982; Moyer, Sowder, Threadgill-Sowder, & Moyer, 1984) were conducted to test the effectiveness of using explicit instruction with strategies adapted from or similar to Polya's heuristics. Lee as well as Charles and Lester found that students were able to learn and effectively use heuristics to solve process problems. Students were taught to make a plan, draw pictures, make a table or chart, look for patterns, and to look back. Charles and Lester theorized that explicit teaching of heuristics facilitates student learning and enables students to solve related problems.

Essen and Hamaker (1990) and Moyer et al. (1984) studied the use of drawings to help students solve problems. Essen and Hamaker found that after instruction with self-generated drawings, fifth-grade students produced more drawings and improved their performance on word problems. According to Moyer et al., the drawn format produced superior performance across grades, especially for low readers. The researchers surmised that drawings provided students a concrete model and relieved working memory or memory overload.

Owen and Sweller (1989), however, question the teaching of heuristics because they doubt instruction in problem-solving strategies will overcome problems in the transfer of learning. Owen and Sweller

believe that failure in problem solving is likely to be from lack of specific subject knowledge or algorithmic skills. They conclude there is little evidence of successfully teaching general problem-solving techniques in mathematics education.

Word problems and nonroutine problems require students to use higher order thinking skills or more creative thinking processes. Gilbert-Macmillan and Leitz (1986) found that thinking and insight are stimulated when students explain their solutions to a group. The NCSM (1989) emphasized that in the future, creative problem solving will become more important since students must "know when and how to use computation; and they must develop proficiency in problem solving" (p. 44). According to Polya (1957), if students are to become expert problem solvers, they need to be actively involved in learning and decision making. Polya also stressed teacher/student interaction through imitation and practice. This paper examines the effectiveness of Polya's theory of teaching and learning mathematical problem solving through heuristics.

Method

Subjects

Participants in the study were 66 fifth-grade students, 10 to 12 years old, from two upper-level mathematics classes in the same rural elementary school ($N = 900$). Students were placed in the two classes based upon academic grades, scores on the Iowa Test of Basic Skills (ITBS), and teacher recommendation. The experimental group ($n = 33$) consisted of 17 males and 16 females, while 15 males and 18 females comprised the control group ($n = 33$).

Treatment

The regular classroom teacher, having 4 years' experience, taught both classes and administered the pretest and posttest. The control group worked only the problem-solving exercises taught in their basal textbooks. Students in the explicit-instruction group experienced the treatment in their daily mathematics class for 15 minutes each day for 10 days. During each 15-minute session, the teacher modeled specific problem-solving strategies developed by Polya. Students then solved a problem as a group under the teacher's guidance. Finally, students practiced the strategy alone then were encouraged to create and share problems using the strategy within cooperative groups.

Instruments

Students in the experimental and control group took a pretest to establish equivalence and then took the same test 11 days

later as a posttest to evaluate intervention effectiveness. The teacher-made pretest/posttest consisted of 29 problems derived from various fifth-grade basal mathematics books over the last 5 years. The test involved the operations of addition, subtraction, multiplication, and division. Problem-solving exercises covered one-step and multi-step word problems, logical reasoning, making list or tables, working backwards using flow charts, finding patterns, drawing pictures, and spatial reasoning. On each test, students were directed to find the correct answer using one of Polya's strategies for problem solving.

Results

The investigator scored both tests. Pretest and posttest scores (see Table 1) were obtained within a 12-day period and reported as mean problem-solving scores and standard deviations. Although there was no significant difference in the pretest

Table 1. Means and Standard Deviations for Problem Solving

	Mean	SD	t
Pretest			
Explicit Instruction	40.86	16.78	.62
Basal Textbook	38.09	14.67	
Posttest			
Explicit Instruction	65.00	17.00	5.03*
Basal textbook	46.49	12.56	
Gain			
Explicit Instruction	24.52	16.27	
Basal Textbook	8.39	9.02	

* $p < .01$

scores, there was a significant difference in the posttest scores ($p < .01$) with the explicit-instruction group gaining an average of 24.52 points while the basal-textbook group gained an average of 8.39 points.

Correlations of the ITBS reading, math, and problem-solving scores to pretest and posttest scores (see Table 2) show a significant correlation with both measures, with the correlation higher for posttest scores.

Table 2. Test Correlations: ITBS to Pretest and ITBS to Posttest

	ITBS to Pretest	ITBS to Posttest
Reading	.49*	.55*
Mathematics	.34*	.54*
Problem Solving	.35*	.58*

* $p < .01$

Test questions were grouped into related categories and the frequency of correct scores compared between groups (see Table 3). The explicit-instruction group showed greater increase in percent of correct responses among all five areas. The greatest increase was in the use of graphics to illustrate and solve word problems. The mean of the experimental group increased 34.8 points compared to a 9.6-

point increase for the control group for frequency of correct responses. Most students in the experimental group correctly answered questions which required constructing a drawing, making a list, completing a table, or using a chart and working backwards. Many students who missed the questions did not use the problem-solving strategies correctly or did not attempt to apply a definite strategy.

Table 3. Problem-Solving Heuristics, Pretest and Posttest (N = 29)

Category	Explicit Instruction			Basal Textbook		
	Pre	Post	Gain	Pre	Post	Gain
Patterns	38.9	53.5	14.6	20.2	24.3	4.1
Logic	30.3	64.7	34.4	29.3	30.3	1.0
Verbal	52.8	76.8	24.0	51.0	60.9	9.9
Graphics	29.3	64.1	34.8	36.4	46.0	9.6
Spatial	34.1	55.3	21.2	40.9	54.6	13.7

In the explicit-treatments group, 53.5% of the students correctly solved the number-patterns problems. Only 24.3% of the control group could do the same. After instruction, the explicit-treatment group increased their scores 14.6 points, while the basal-text group achieved a 4.1-point increase.

An increase of 34.4 points occurred in the area of logical reasoning for the experimental group compared to a 1.0-point increase for the control group. Most students in the experimental group attempted to answer these questions, and over 64% were successful. Results indicated that incorrect answers were possibly due to an incomplete understanding of place value and rounding.

With word problems and multi-step problems, the experimental group increased 24 points compared to a 9.9-point increase for the control group. During the pretest most students in both groups correctly analyzed and solved word problems. However, after instruction with Polya's method of problem analysis, the explicit-treatment group solved more of the multi-step problems correctly than did the basal-textbook group.

With respect to spatial reasoning, after instruction in drawing and visualizing cubes, the explicit-treatment group gained 21.2 points compared to a 13.7-point gain for the control group.

As an informal assessment of the problem-solving strategies, the research requested students write in their journals about their two weeks with our various problem-solving activities. Although most students thought problem solving was very difficult, they believed the topic was interesting and important. Students' comments ranged from, "This is hard," and "Mom

didn't find an answer," to "The problem was fun," or "My dad said that this is what he does." The explicit-treatment group liked the instructor and teaching method better than did the basal-textbook group. Generally, both groups still expressed the usual dislike and fear of word problems. However, after instruction, both groups stated they now felt more comfortable with word problems and believed they could solve problems similar to those taught in the unit. Students' comments ranged from, "It still doesn't tell me what to do," and "I think I can find the answer, but it is hard to find the right plan," to "I can see the pattern," or "Picking out data is easy," and "It's fun to trick my sister," or "These are fun, let's do some more." Many students in the explicit-treatment group expressed enjoyment of the new activities and strategies now associated with problem solving.

Discussion

The purpose of this study was to evaluate the effectiveness of explicit instruction in problem-solving heuristics. According to pretest and posttest results, the explicit-instruction group achieved a higher average answer than did the basal-textbook group. After 10 lessons in problem solving, most students in the explicit-instruction group could select and appropriately use a problem-solving strategy. However, no attempt was made to evaluate the impact of manipulatives, calculators, cooperative grouping, or problem-solving journals had on the results of the two groups. Although there were no significant differences between the two classes on the pretest, the results of the explicit-instruction method of teaching heuristics is beneficial in teaching students to solve problems.

Many basal textbooks do not teach problem solving as the focus of mathematics but as a topic to be covered at the end of each lesson or chapter. Scattered throughout the chapter are exercises that review daily work using word problems, but most of these problems do not challenge students to think about planning solution strategies. The results of this study indicate that explicit teaching of heuristics can improve a student's ability to understand the problem and to plan a solution. Heuristics provides an index of skills applicable to problem-solving situations. Heuristics improve a student's confidence in his/her ability to succeed in mathematics and to learn "how to think" about mathematics. Owen and Sweller (1989) warn there is no conclusive proof that instruction in general problem-solving strategies will overcome transfer of learning. They imply that instructors should teach domain-specific knowledge. However, according to Charles and Lester (1984), domain-specific strategies include heuristics which help the student organize and transform knowledge. Further research is needed to determine if teaching heuristics overcomes transfer of learning and if heuristics learned in mathematics can be applied to other domains.

Effort was made to minimize differences between the two classes by using cooperative groupings, calculators, and problem-solving journals in both classes. This study did not evaluate the effect of these aids and strategies on students' attitudes or on the problem-solving process. Future studies could evaluate the effectiveness of cooperative-learning groups and calculators in creative problem solving and creative problem formation.

Explicit teaching of problem-solving

strategies has definite implications for educators. Teachers should carefully evaluate the stage of learning for each student. Children at Piaget's concrete stage of learning need more manipulatives in learning problem solving than do children at the formal operational stage (e.g., Gage & Berliner, 1988). Both groups routinely used manipulatives to introduce and solve many of the problems presented in this study. However, the study did not evaluate the impact such manipulatives had on the success experienced by the explicit-instruction group. Future studies could examine the effect of extended use of manipulatives on the problem-solving process.

Summary

This paper has presented the findings and evaluation of a problem-solving unit in mathematics. Results are encouraging and support a mathematics program which emphasizes explicit teaching of problem solving and higher order thinking skills. Such explicit teaching helps students become more comfortable with problem solving as problem-solving abilities increase. By expanding their repertoire of skills, students made gains and developed critical attitudes toward solving problems which helped widen their views of mathematics.

References

- Charles, R., & Lester, F. (1984). An evaluation of a process-oriented instructional program in mathematical problem solving in grades 5 to 7. Journal for Research in Mathematics Education, 15(1), 15-34.

- Darch, C., Carnine, D., & Gersten, R. (1984). Explicit instruction in mathematics problem solving. Journal of Educational Research, 77, 331-358.
- Essen, G., van, & Hamaker, C. (1990). Using self-generated drawings to solve arithmetic word problems. Journal of Educational Research, 83, 301-312.
- Fischbein, E., Deri, M., Nello, M., & Marino, M. (1985). The role of implicit models in solving verbal problems in multiplication and division. Journal for Research in Mathematics Education, 16(1), 3-17.
- Gage, N., & Berliner, D. (1988). Educational psychology (4th ed.). Boston: Houghton Mifflin.
- Gilbert-Macmillan, K., & Leitz, S. (1986). Cooperative small groups: A method for teaching problem solving. Arithmetic Teacher, 33(3), 9-11.
- Kouba, V., Brown, C., Carpenter, T., Lindquist, M., Silver, E., & Swafford, J. (1988). Results of the fourth NAEP assessment of mathematics: Number, operations, and word problems. Arithmetic Teacher, 35(8), 14-19.
- Lee, K. (1982). Fourth graders' heuristic problem-solving behavior. Journal for Research in Mathematics Education, 13(2), 110-123.
- Moyer, J., Sowder, L., Threadgill-Sowder, J., & Moyer, M. (1984). Story problem formats: Drawn versus verbal versus telegraphic. Journal for Research in Mathematics Education, 15(5), 341-352.
- National Council of Supervisors of Mathematics. (1989). Essential mathematics for the twenty-first century: The position of the National Council of Supervisors of Mathematics. Arithmetic Teacher, 37(1), 44-46.
- National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: Author.
- Owen, E., & Sweller, J. (1989). Should problem solving be used as a learning device in mathematics? Journal for Research in Mathematics Education, 21(5), 322-328.
- Polya, G. (1957). How to solve it (2nd ed.). Garden City, NY: Doubleday Anchor Books.
- Sellke, D., Behr, M., & Voelker, A. (1991). Using data tables to represent and solve multiplicative story problems. Journal for Research in Mathematics Education, 22(1), 30-38.
- Suydam, M. (1982). Update on research on problem solving: Implications for classroom teaching. Arithmetic Teacher, 29(6), 56-60.
- Swing, S., & Peterson, P. (1988). Elaborative and integrative thought processes in mathematics learning. Journal of Educational Psychology, 1(1), 54-66.

Effects of Group Processing Skills on the Social Interaction of Cooperative Learning Groups

Emilie Allen

Cooperative learning is an instructional strategy designed to provide students with a positive learning environment and to address problems associated with teaching students of diverse backgrounds. In addition to enhancing student achievement, cooperative learning has shown positive impact on interpersonal relations, interaction with special education students, and self-esteem. This study was designed to teach necessary social skills to students in preparation for additional work with cooperative learning. Three measures of group interaction showed improvement at the end of the project. The results of this study suggest that social skills can be improved and that such improvement is a valued outcome in an educational setting.

The children who walk through America's classroom doors each fall present a unique and almost overwhelming challenge to their teachers. Each new group of children is a diverse mixture of academic abilities, emotional maturities, and motivational interests. They represent broad cultural and economic backgrounds. How does one teacher meet each child's individual differences? How is it possible to challenge the gifted and simultaneously incorporate the slower, less able student into the learning process?

A traditional classroom may not offer answers to these questions. In most schools, academically slower children compete with brighter classmates for teacher approval, for grades, and for positive

recognition. In addition, Slavin (1990) notes that low achieving students may lack the prerequisites to learn new material. They receive negative feedback daily on their academic efforts and learn that academic success is not within their grasp. Eventually, they may use antisocial or delinquent behaviors to gain a sense of self-worth.

Cooperative learning addresses problems associated with classroom competition and encourages students to help one another succeed. Dividing students into teams whose members are responsible for each other's learning as well as their own establishes a positive learning environment. Students encourage and help each other master skills and content presented by the teacher. In a cooperative setting, they convey to one another the idea that learning is important, valuable, and fun (Slavin, 1987a).

In the past, this researcher had much success with several short-term cooperative grouping projects. As a result of these positive experiences, cooperative learning was implemented in the classroom as an alternative to traditional instruction. As the year progressed, one problem continually threatened to undermine the project's success--too much time was required for teacher intervention to solve communication problems and to help resolve conflict among group members. Rules of group interaction obviously had to be mastered before successful group work could occur. The children had to understand the con-

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Emilie Allen

19

2.

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

cepts of cooperation and their application to group activities in order to achieve academic success.

Most literature on cooperative grouping appears to describe in detail how to organize the design material or how to implement grouping in a classroom. Very little information is available on teaching the necessary interpersonal and group skills essential for successful cooperation. In order to address this void in the literature, this study was designed to implement a unit on building interpersonal skills designed to create positive interdependence between group members and therefore, successfully impact the group's academic achievement.

Review of Literature

Cooperative learning refers to a set of instructional methods in which students of varying abilities work together in small groups toward a common goal (Slavin, 1987a). Based on 60 experimental studies, Slavin (1989) concluded that under certain circumstances the use of cooperative learning methods increased student achievement more than did traditional instruction. Two essential elements--group rewards and individual accountability--were necessary for this positive effect. The group must work together to earn recognition, grades, rewards, and other indicators of group success. The group's success must also depend on the individual learning of all group members. These methods consistently improved self-esteem and social relations among students (particularly race relations), and acceptance of mainstreamed students (Slavin, 1987a).

Comparisons to Traditional Instruction

According to McElroy (1989), education has an individualistic focus in a traditional classroom setting. Students work individually on academic tasks and compete with each other for grades and other rewards. One student's success may make it difficult for others to succeed. In these classrooms, students are usually grouped homogeneously for academic skill instruction. The teacher is the traditional classroom leader and controls academic content, learning environment, and instruction.

In contrast, in the cooperative classroom students are grouped heterogeneously and work together in these groups for grades and rewards. One student's success helps others achieve their goals. All group members are given the opportunity to lead and are encouraged to develop and use leadership skills. The students often participate in the decision-making process regarding curriculum and the learning atmosphere of the classroom.

Theoretical Models

Two theoretical models--developmental and motivational--explain the superiority of cooperative learning over traditional instruction. According to Slavin (1987c), the developmentalists believed that improved student achievement results from student interaction on learning tasks. As students discuss academic content, they resolve misunderstandings, thus higher quality understanding emerges.

Motivationalists agree that peer interaction can lead to cognitive growth, but interaction alone is not sufficient. Group

rewards and individual accountability are necessary to make peer interaction more effective. Slavin (1987c) states, "The cooperative reward structure is the critical element in increasing student achievement, because their use creates peer norms and sanctions supporting individual efforts" (p. 1163). Slavin (1983) determined that 25 of 28 studies on cooperative learning methods using group rewards based on the sum of individual learning performances showed significantly greater achievement in cooperative than in the control classes. More recently, 6 of 8 studies found similar positive results (Slavin, 1987a).

Academic Achievement

While there are many benefits of cooperative learning, enhanced achievement is the most important outcome and has the broadest research base. If properly structured, cooperative learning methods can significantly accelerate the learning of all children. Research on academic achievement overwhelmingly supports cooperative learning over traditional instruction. Of 35 nationwide studies conducted by Slavin (1987b) and stratified for age and ability, 29 (83%) found that students in various cooperative learning situations gained significantly more in achievement than did students in regular classrooms covering the same objectives.

Some of the greatest effects of cooperative learning on student achievement cited by Slavin (1987a) in the studies of Team-Assisted Individualization (TAI). This method combines cooperative groups with individualized instruction. In six studies, the TAI classes gained an average of twice as many grade equivalents on standardized mathematics computation measures than

did the traditionally taught control classes. Using another cooperative learning method--cooperative integrated reading and composition (CIRC)--classes gained 30%-70% of a grade equivalent more in reading comprehension, reading vocabulary, language expression, language mechanics and spelling than did the control groups. Numerous other studies consistently document academic gains for students.

Interpersonal Relations

In addition to achievement gains, cooperative learning has positive effects on social, motivational, and attitudinal outcomes. According to Slavin (1990), cooperative learning, by nature, is a social method. Students engage in active rather than passive learning, which encourages them to work together toward common goals. The heterogeneous structure of this educational strategy groups children who in other circumstances may not choose to be together. By working in cooperative relationships, children gain respect for and understanding of others who may be very different from themselves. Ethnic, economic, and social barriers fall; and positive friendships can develop. Several studies by Slavin (1987b) support the positive effects of cooperative learning on close, reciprocated friendships - especially with cross-racial relationships.

In a study by Kagan (1989), race relations improved significantly. As cited by Kagan, when students in heterogeneous cooperative groups were asked to select close friends, another student's race was not a significant predictor of friendship choice. The highest level of intimacy choices among students were their teammates, regardless of race.

Mainstreaming

Public Law 94-142, which requires mainstreaming of the mildly mentally and physically handicapped into the regular classroom, has drastically changed the nature of teaching. In many classrooms, traditional instruction will not meet the individual needs of these students. According to Augustine, Gruber, and Hansen (1989), cooperative learning provides a positive way to include these special needs children in the classroom socialization process and simultaneously provide them with the extra academic help they need. Traditionally, handicapped children were taught in separate classes, thus increasing their feelings of isolation and difference. By being part of a group, their social skills and feelings of self-worth improved when accepted by regular students.

The benefits of cooperative learning for special education students are also supported by over 80 studies conducted by Johnson and Johnson (1989b). In addition to improved self-esteem, improved social skill, and improved psychological adjustment, cooperative experiences also promote academic gains, use of higher-level reasoning strategies, and greater critical thinking development for these special needs children.

Other Outcomes

Various studies of other outcomes of cooperative learning indicate positive gains in self-esteem, attitude toward school, feelings of success, and feelings of being accepted and valued by classmates (Slavin, 1990). Another study (Johnson, Johnson, Johnson, & Anderson, 1984) found that cooperation increases a student's ability to

make helpful decisions and altruistic choices.

As previously mentioned, cooperative learning has an extensive research base. Volumes of material are available to help teachers implement cooperative grouping in their classrooms. However, for these groups to be completely productive, it is vital to teach interpersonal and small-group skills essential for effective cooperation (Johnson & Johnson, 1989a). Other researchers state similar findings. This researcher agrees with the literature on the importance of teaching social skills, but found little information regarding the steps involved in teaching these skills. This study proposed to design a unit of material which outlines of the necessary steps for building cooperative social skills.

Methods

Subjects

This study was composed of 1 teacher and 21 public school third graders. The students were heterogeneously grouped, by race and ability, into 6 teams of 3 or 4 children. Classroom seating was arranged so desks of each team were together all day. There were 8 girls and 13 boys in the class with a ratio of 14 white to 7 black. The children all lived in a middle-size town in south central Georgia, and came from very diverse economic and social backgrounds.

Instruments

Three instruments were used to measure student interaction among group members. A teacher-made questionnaire of student attitudes on social interaction was

administered as a pretest and posttest. To ensure content validity, the questionnaire was patterned after team-building goals suggested in the literature on cooperative learning. The questionnaire was validated by a graduate research class.

A sociogram was also drawn to be used as a pre- and postmeasure of student interaction. At the beginning and again at the conclusion of the unit, each child listed three other students he/she would most like to work with in group activities. Their choices were diagrammed and compared to illustrate preferred interaction between group members.

Daily evaluations of group interaction represented the third method of measurement. For a 10-day period, each group participated in a daily fun activity. After each activity, the individual groups evaluated their ability to cooperate using a set of cooperation guidelines developed as a model at the beginning of the unit. (See Table 1 for a listing of the guidelines.) These evaluations were compiled on a chart for comparison.

Procedures

This study used a single-subject, single-group experimental design to measure the interaction of students working in cooperative learning groups. The study lasted 3 weeks and had two goals: (a) to build team spirit among group members and (b) to teach group processing skills that would enable the students to interact more efficiently and resolve conflict more effectively.

At the beginning of the unit, the students responded to an attitudinal questionnaire about the socialization skills of their group. They also listed friendship choices

that were later charted on a sociogram. This information was used as pretest data. The activities of the first week centered around building team spirit. Each group made up group names and designed posters to be displayed on the bulletin board. Fun, noncompetitive group activities were planned for each day.

Near the end of the week, a class discussion was held to "brainstorm." The children then listed different ways to demonstrate cooperative behavior in group activities. Following the discussion, the class outlined six "guidelines of cooperation," which were written on chart paper and posted with the team-name posters.

During the next 2 weeks the groups focused on building group interaction skills using the "guidelines" as a model. Fifteen to 20 minutes were spent each day in fun activities requiring the participation of all group members to be effective. At the conclusion of each activity, the groups reviewed the guidelines and evaluated their cooperative behavior. The evaluations consisted of listing guidelines each group performed well and areas needing improvement. To add encouragement, superstar awards were given to groups who demonstrated cooperative attitudes at times other than the scheduled group activity time. These stars were stapled around the team-name posters as visual incentives. Stars were also added for outstanding group performance on certain academic tasks such as spelling, math worksheets, and science and social studies projects.

At the conclusion of the unit, the students again responded to the attitudinal questionnaire and listed friendship choices. This information was used as posttest data and was charted for comparison with pretest results. The results of the daily evalu-

ations were also charted and included in the posttest data.

Results

This study included three measures of group interaction. Table 1 presents data on the attitudinal questionnaire. The responses

of the groups to each question were charted so that the pretest and posttest results could be compared. The total number of responses for each answer choice is recorded for each group. With the exception of question 2, the "never" responses decreased on every posttest question. For question 2, the "never" choice was not selected on

Table 1. Student responses for the attitudinal questionnaire

Question	Answer	Pretest Total	Posttest Total
1. We feel good about working together.	Always	10	8
	Sometimes	10	13
	Never	1	0
2. We help each other understand the assignment.	Always	10	13
	Sometimes	11	8
	Never	0	0
3. We help each other keep on task.	Always	6	13
	Sometimes	12	9
	Never	3	0
4. We disagree in a friendly way.	Always	3	5
	Sometimes	14	15
	Never	4	1
5. We accept each others' ideas.	Always	4	9
	Sometimes	15	12
	Never	2	0
6. We take turns.	Always	9	15
	Sometimes	10	6
	Never	2	0
7. We follow directions.	Always	3	15
	Sometimes	17	6
	Never	1	0
8. We think we are the best group of all.	Always	6	13
	Sometimes	9	4
	Never	6	4

either the pretest or the posttest. The number of "sometimes" and "always" responses increased on every question of the posttest. Some gains were quite dramatic. On questions 3, 7 and 8, the number of "always" choices more than doubled. The gains would have been even more positive for question 8, if Group 5 had answered the question correctly. In a class discussion of the results, students in Group 5 admitted they liked their group best and wanted to mark "always;" but since they had fewer superstar awards than did other groups, they felt they had to mark the "never" choice.

The students' daily evaluations of their group interaction are presented in Table 2. For 10 days, the individual groups listed the guidelines they performed well and the

ones they needed to improve. Each day's choices were charted for comparison. At the end of each week, the number of "improvements needed" was totaled. By comparing the totals for each week, an increase was demonstrated in the number of times each group recorded "0" for "improvement needed." Four of the five groups more than doubled their number of "0's" in the second week. Guideline 5 was selected seven times as "needing improvement" during the first week, Guideline 1 was selected six times, Guidelines 3 and 6 were selected five times, Guideline 2 was selected three times, and Guideline 4 was selected twice. During the second week no guideline was selected more than twice as "needing improvement."

Table 2. Students' daily evaluations of success with guidelines

Group	Rating	Pretest						Posttest					
		Days					Total 0's	Days					Total 0's
		1	2	3	4	5		1	2	3	4	5	
1	Success Improve	4*	5	1-6	4	4	1	1-6	1-6	2,3	2,4	1-6	5
		1	1,6	0	5	1		0	0	0	0	0	
2	Success Improve	1,6	1,3	1-6	1,4	1-6	4	1-6	1-6	1,2	1-6	1-6	5
		0	0	0	5	0		0	0	0	0	0	
3	Success Improve	2,3	1,3	5,6	0	4	2	1-6	2-6	2,3	1-6	1-6	4
		6	0	0	1,6	2,3- .5,6		0	1	0	0	0	
4	Success Improve	2,5	1	2,3	1-6	2-6	1	1-6	1-6	1,2	1-6	1-6	4
		3	5	1	0	1		0	0	3	0	0	
5	Success Improve	3,5	2,3	1,3	2,6	1-6	2	1,6	1-6	1-6	1-6	1-6	5
		2	5	0	5	0		0	0	0	0	0	
6	Success Improve	1,6	1,5	1,2	1,6	1	0	1-5	1-6	3,5	0	1-6	3
		4,3	3	3	5	5,6		6	0	0	1-6	0	

* Indicates number of guideline shown in Table 1

Table 3 represents the results of the third measure of interaction and shows data from the students' lists of friendship choices. On the pretest, Groups 1, 2 and 3 made no intragroup choices. However, in Group 6 only one child who made an intragroup choice (choosing two other members of the group); Groups 4 and 5 had no intragroup choices at all. Two children were not selected by any other children.

On the posttest, all groups, except Group 1, demonstrated increases in intragroup choices. Additionally, all children were selected at least once, leaving no one in isolation. The reasons for Group 1's decreased interaction are not evident from the data but may have partially resulted from a disagreement between two team members on the day of data collection.

Table 3. Students' intragroup friendship choices

Group	Pretest	Posttest
1	6	2
2	3	5
3	2	4
4	0	2
5	0	2
6	2	4

Discussion

All three measures of group interaction included in this study demonstrate positive gains. Evaluation of the tables and the sociogram show improvement in team members' acceptance of each other and in their ability to effectively interact. Each group also evidenced a tremendous increase in pride as a team. Based on teacher judgement, this researcher believes the improvement was significant. As the unit progressed, a positive increase occurred in each group's ability to resolve conflicts and resume work. Passive students become increasingly involved, and the leadership base of each group was more equally shared. Each group's ability to successfully complete assigned tasks improved tremendously, and less time was required for each task.

Personality disagreements and power struggles still exist within the groups. Realistically, since groups are made up of children with diverse personalities and abilities, not all difficulties can be completely resolved. Their continuing existence, however, may result from problems within the study.

The internal validity of the study was possibly threatened. The original study was designed to last 5 weeks; however, due to teacher absences for jury duty and illness and due to illnesses of several children, the study only lasted for 3 weeks. There was insufficient time to effectively teach all group processing skills needed for group interaction. Several times after the unit began, it had to be interrupted and restarted later. These interruptions greatly hindered the progress of the unit by necessitating repetition of review material, thus limiting time for work on new skills.

Other threats to internal validity may have affected the study's outcome. The actual structure of individual groups may have caused selection bias. Although the teacher assigned groups very carefully, certain groups may have had more compatible personalities than did others, since groups were not randomly assigned. Prior to the study, the children had been working in groups for several months. Since there had already been much teacher intervention due to cooperation problems, historical difficulties may have existed within some groups. From the beginning of the study, the children knew the purpose of the unit and what was expected of them. Their behaviors may have improved as a result of this knowledge.

Although problems existed within the study, this researcher believes the gains to be sufficiently positive to repeat the unit next year. The unit will be implemented at the beginning of the year for a longer period of time, and review sessions will be scheduled throughout the year. Hopefully, time spent teaching and practicing group processing skills will produce groups who interact less conflict, fewer disagreements, and fewer power struggles.

Schools should place greater emphasis on teaching group processing skills. As adults, individuals will be in group situations where interpersonal skills may be most important to employability, productivity and career success (Johnson & Johnson, 1989a). Social skills also directly relate to building and maintaining positive relationships as well as preserving emotional and mental health. Thus, social skills developed through cooperative efforts in school are important contributors to the happiness, wholeness, and career successes of our future generation.

References

- Augustine, D., Gruber, K., & Hansen, L. (1989). Cooperation works. Educational Leadership, 47(4), 4-7.
- Johnson, D., & Johnson, R. (1989a) Social skills for successful group work. Educational Leadership, 47(4), 29-33.
- Johnson, D., & Johnson, R. (1989b). Cooperative learning, what special education teachers need to know. The Pointer, 33(2), 5-10.
- Johnson, D., Johnson, R., Johnson, J., & Anderson, D. (1984). The effects of cooperative vs. individualized instruction on student prosocial behavior, attitudes toward learning, and achievement. Journal of Educational Psychology, 68, 446-452.
- Kagan, S. (1989). The structural approach to cooperative learning. Educational Leadership, 47(4), 12-15.
- McElroy, K. (1989). A taste of cooperativeness within an elementary school. The Pointer, 33(2), 34-38.
- Slavin, R. (1983). Using student team learning. Baltimore: John Hopkins University.
- Slavin, R. (1987a). Cooperative learning: Can students learn? Instructor, 96(7), 74-78.
- Slavin, R. (1987b). Cooperative learning: Where behavioral and humanistic approaches to classroom motivation meet. The Elementary School Journal, 88(1), 29-35.
- Slavin, R. (1987c). Developmental and motivational perspectives on cooperative learning: A reconciliation. Child Development, 58, 1161-1167.
- Slavin, R. (1989). Research on cooperative learning: Consensus and controversy. Educational Leadership, 47(4), 52-54.

Slavin, R. (1990). Cooperative learning: Theory, research, and practice. New Jersey: Prentice-Hall. —

Have the Results of Project Follow Through Been Well Utilized?

Janis N. Lee

Project Follow Through began in the late 60s as a follow-up to Head Start and was designed to serve needy children in grades K-3. This paper considers the results and controversy of the Follow Through experiment. Several longitudinal studies are discussed which confirmed that the Direct Instruction model was the most successful Follow Through program. This paper also examines federal dissemination and funding of the Follow Through programs and considers suggestions for improvement as well as future implications.

Project Follow Through (PFT) began in 1968 and, during the first 10 years, served 79,000 disadvantaged children in grades K-3. Between 1968 and 1978 the federal government spent \$500 million dollars on the project and \$50 million on its evaluation. Using a variety of learning and child-development theories, 22 teams created specific programs and teaching methods. Evaluation studies conducted since the inception of PFT have aroused considerable controversy, especially in 1977 and 1978.

Evaluation and Controversy

A 5-year study of PFT by Abt Associates (Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977) resulted in three major findings. The study discovered first that differences in effectiveness among sites within the same model were greater than differences in the effectiveness between PFT models. Other evaluators confirmed Abt's results (Anderson, St. Pierre, Proper,

& Stebbins, 1978; Rhine, 1983; Wisler, Burns, & Iwamoto, 1978). The second major finding was that no PFT model effectively raised test scores of disadvantaged children. One critique agreed that "if the concentrated effort of highly competent and well-funded sponsors working with a few sites cannot produce uniform results from locality to locality, it seems doubtful that any model program could do so" (House, Glass, McLean, & Walker, 1978, p. 154). The third finding by Abt was that disadvantaged children continued to perform poorly despite intervention programs.

Evaluations of PFT hoped to identify at least one or two models that would work in any school system. Such a strategy, while plausible 10 years previously, began to appear misguided (Anderson et al., 1978).

Abt Associates found, secondarily, that the Direct Instruction (DI) model, sponsored by the University of Oregon, achieved superior results overall, although the improvements were inconsistent across sites. Rhine (1983) found that two model programs were more successful than others: the DI model and the Behavior Analysis Model. However, House et al. (1978) claimed the Office of Education pressured Abt to use a method which erroneously maximized model differences in the models to make one model appear superior. House et al. further claimed that Abt's evaluation results were unfair to the various model sponsors. Abt Associates strongly refuted these criticisms (Anderson et al., 1978).

Four instruments were employed to evaluate the PFT models: (a) the Metropolitan Achievement Test (MAT), which

measures reading, language, and math; (b) Raven's Colored Progressive Matrices, which measures non-verbal problem solving ability; (c) Coopersmith's Self-Esteem Inventory; and (d) the Intellectual Achievement Responsibility Scale, which measures student attitudes concerning academic failure or success. The Abt study did not directly compare all 22 models but grouped results into three categories: basic skills models, cognitive-conceptual models, and affective-cognitive models (House et al., 1978).

House et al. (1978) questioned government involvement in the evaluation process and concluded that large scale evaluations are costly and unnecessary. Wisler et al. (1978) disagreed and stated the government should have provided stronger direction in the evaluation. These researchers concluded that success was sufficiently evident to indicate that compensatory education can work, although not on a widespread basis. Wisler et al. noted that "with a few exceptions, the models assessed in the national Follow Through evaluation did not overcome the educational disadvantages poor children have" (p. 179). The DI model showed the best pattern of success but was not successful in every location (Wisler et al.).

Hodges (1978) believed PFT to be successful, but mistakes on the part of the Abt evaluators produced the negative results. Hodges stated that unreported data would have shown overall positive results in reading and math achievement, parental involvement, and school morale. He also predicted that future examination of PFT would reveal more valuable information. "Just because Follow Through has not proved to be an easy, workable, inexpensive solution for all the educational prob-

lems of poor children does not mean it should be dismissed as just another failure in compensatory education" (Hodges, p. 191).

One Study of the Exception

Abt evaluators found DI, begun in P.S. 137 in the Ocean Hill-Brownsville section of Brooklyn, New York, to be consistently successful in all measured areas from 1968-1981. P.S. 137 is characterized as severely disadvantaged. Parental involvement was instrumental in maintaining PFT for 14 years despite budget cuts and excessively high turnover rates of teachers, principals, and PFT directors (Meyer, Gersten, & Gutkin, 1983).

DI employs the carefully sequenced Distar curriculum, ongoing teacher training, and constant monitoring of student progress. Since this model is highly structured, it is less dependent on specific people. New teachers can quickly learn to use it, which leads to greater potential for success in areas characterized by high turnover (Meyer et al., 1983).

Since implementation of DI, students in P.S. 137 have surpassed students from comparably disadvantaged schools on the MAT in 1974, the Stanford Achievement Test (SAT) in 1975 and 1976, and the California Test of Basic Skills (CTBS) in 1977 (Meyer et al., 1983).

Continued evaluations of the students into the fourth and fifth grades demonstrated mean scores at grade level. Students receiving DI scored significantly higher on achievement tests when compared to all other students in that district not receiving DI. When compared to a group of 6,000 low socioeconomic, non-PFT students, the positive achievement test results generated

by students receiving DI were even more impressive. Among the seven PFT models sponsored in New York City, only the DI model demonstrated effective instruction (Meyer et al., 1983).

A Secondary Analysis

Gersten, Becker, Heiry, and White (1984) studied the relationship between entry IQ and yearly academic gains among 1,500 low-income students, from 20 communities, who had been through the DI program. Low IQs did not hinder academic progress in schools using DI, the only PFT model to bring low-income students near or above the national average performance on the MAT.

Reanalysis of Site Variability

Although the large number and variety of sites involved in the Abt study made precise comparisons difficult, the evaluators found the DI model to be generally effective in all three areas: basic skills, cognitive-conceptual, and affective-cognitive. Upon reexamination of the data, Gersten (1984) found cases of both good and poor matching of comparison groups. Positive effects of the DI model on math achievement were consistent, replicable, and significant among all urban sites.

"When the same analyses that Glass or the Abt Associates did are performed using samples of urban sites implementing the Direct Instruction model, much of the unexplained variation disappears" (Gersten, 1984, p. 421). Since Abt's 1977 study, sites which continued to use the DI model were examined. In later analysis, researchers gathered data exclusively from low-income students; only 80% of Abt subjects

were low-income. Data showed that increased performance levels had been maintained in four urban sites (Gersten).

Results of Several Longitudinal Studies

Meyer (1984) compared students receiving DI at the highly successful P.S. 137 in Brooklyn, New York to a control group of non-DI, similarly disadvantaged students in a nearby school. The students were then about 20 years old. Meyer compared them in three areas: a) ninth-grade reading and math achievement, b) high school graduation, and c) college application and acceptance.

In ninth grade, the DI students performed significantly higher than the control group in mathematics, yet were still approximately 6 months below grade level. Approximately 60% of the DI students graduated from high school, whereas only about 38% of the control students graduated. Thirty-four percent of the DI students applied to college and were accepted. Of the control group, 18.5% applied to college; 17% were accepted (Meyer, 1984).

Gersten and Keating (1987) investigated the first of the PFT students who completed high school in 1981 or 1982. They were compared to non-PFT students with the same ethnic background and family income who attended other schools. Students receiving DI attended six different schools. In all six schools, ninth-grade reading levels had come within 10 percentile points of the national average, well above the average reading levels of other low socioeconomic students. Fewer students dropped out of high school and more students applied to college. Students who participated in the 4-year (K-3) DI model realized greater benefits than did those

students in the 3-year program (1-3).

Engelmann, Becker, Carnine, and Gersten (1988) found that students in the DI model showed cognitive and affective gains as well. Students enrolled in DI performed better in reading comprehension, math problem solving, math concepts, and had a greater degree of self-esteem than did students taught with all other models.

Engelmann et al. (1988) found that teachers initially reacted negatively to highly structured lessons and in-class supervision, yet half of them changed their minds in view of classroom success. Engelmann's study demonstrated that money, good intentions, and social services do not guarantee a particular program's success. PFT sponsors had equal money and assistance, yet only a few models exhibited overall, though inconsistent, success.

While DI appeared most successful for all students, programs oriented toward individual learning styles provided the least academic achievement. Models using Piagetian concepts of self-directed learning were ineffective in increasing scores on standardized tests of basic skills in third grade. Only the carefully sequenced, highly organized DI model demonstrated gains in the cognitive areas of thinking and problem solving (Engelmann et al., 1988).

Some critics suggested there would be a dissipating or negative effect of DI programs in later years as a result of stifling students and evoking negativism toward education. To investigate these accusations Gersten, Keating and Becker (1988) compared fifth and sixth graders to students of similar demographics. The study utilized records from 1975 and 1976. For both years, students from the DI model outperformed the control group in most domains assessed by standardized achievement tests.

A second study followed junior and senior high DI graduates who had either attended a program from kindergarten through third grade or from first through third. Comparison groups were selected based on reception of Aid to Families of Dependent Children and on ethnicity. The study compared not only academic achievement, but graduation and drop out rates, college application and acceptance, and retention (Gersten et al., 1988).

The results of ninth-grade reading achievement favored the DI students as did high school reading achievement. By the ninth grade, students who had been enrolled in DI programs since kindergarten performed at or less than a year below grade level. Math achievement was similar but not as strong as reading achievement. A study of graduation and dropout rates showed inconsistent or insignificant results at two of the sites; results from a third site significantly favored DI. Although college acceptance was examined at only two sites, DI students demonstrated significantly better success. DI students were also retained significantly less often than were students in the control groups (Gersten et al., 1988).

A Preschool Model

Based on findings that students who attended a PFT program from kindergarten through third grade achieved greater academic success than did those students enrolled in a PFT program from first through third grades, a preschool for children living at the poverty level was begun in 1976. The school utilized the DI model and focused on 4- and 5-year-olds (Weisburg & Weisburg, 1988).

These students were compared to a

group of preschoolers in another intervention project. On the Wide Range Achievement Test measuring reading, spelling, and math, the Direct Instruction preschoolers scored between the 73rd and 90th percentiles. DI children who had been in both pre-kindergarten and kindergarten scored above the 90th percentile. The comparison group scored between the 30th and 45th percentiles. Academic success followed the DI students into regular school (Weisburg & Weisburg, 1988).

Dissemination

The agents charged with dissemination of PFT programs were the National Diffusion Network (NDN), and the Joint Dissemination Review Panel (JDRP). The NDN was composed of 21 resource centers funded by PFT. The 26 members of the JDRP represented groups in the Department of Education. They validated 48 exemplary Follow Through Programs (Rhine, 1983).

Among the programs validated and funded by JDRP and packaged and disseminated by NDN were PFT models which had not demonstrated improved academic instruction. Schools, therefore, could not discriminate between effective and ineffective programs (Watkins, 1988).

Funding

At the time of development, PFT was appropriated \$120 million. Before PFT began, funds were cut to \$15 million for 1968 (House et al., 1978). Between 1968 and 1983 almost \$1 billion were spent on PFT. At the height of the program in 1970, \$70 million were appropriated. In 1981 the number of funded programs was

reduced from 147 to 84, the number of sponsors reduced from 22 to 16, and the number of enrolled students reduced from 63,558 to 36,000. Funding continued for most of the validated programs (Rhine, 1983).

In 1982 PFT programs validated by JDRP received fewer funds than did non-validated programs. In 1983, no new programs were validated, and the highest funding again went to non-validated programs (Watkins, 1988). In its 1987 budget the Reagan administration called for termination of the Follow Through programs, which had last been reauthorized in 1984; however, Congress appropriated \$7.5 million for 1987, \$7.8 million for 1988, \$8.1 million for 1989, and \$8.4 million for 1990 (Cohn, 1987).

Conclusions

Since the federal government has spent considerable money on PFT, it seems unreasonable that results have not been better utilized. PFT tested many different models over a period of years. Reanalyses of original evaluative data as well as longitudinal studies of PFT graduates have repeatedly demonstrated that only the DI model obtained positive correlations across all measured areas. Longitudinal studies also demonstrated that DI produced a higher graduation rate, lower drop-out rate, and increased likelihood of college acceptance.

The results of program expansion to junior and senior high school must be considered. In regards to New York City which documented DI success, researchers stated, "Despite these important gains, however, 40 percent still dropped out of high school. This is significantly less than

the comparison group dropout rate of 58 percent, but it is far from ideal" (Gersten & Keating, 1987, p. 31).

Improvements could be made in PFT not only by expanding to the upper grades but also by considering sociocultural factors when educating disadvantaged children. "Schools are basically designed for white, middle-class children, and leave largely to parents the teaching of a most basic building block for intelligent behavior--namely, words and their referents" (Engelmann et al., 1988, p. 316).

We cannot depend on informal methods of disseminating information. The DI model was effective not only because of the material components but due to its implementation of specific teaching techniques, high teacher expectations, and parental support. "Learning these techniques is not easy. If teachers, even those who are experienced classroom teachers but new to Direct Instruction, do not receive intensive classroom supervision, the acquisition of these skills usually takes longer" (Carnine, 1988, p. 22).

Rather than improving or expanding the DI Follow Through program, the federal government has sought to reduce it. PFT results have not been carefully disseminated. Educational circles are replete with teaching ideas based on a variety of methods, few of which have been as thoroughly studied and successfully demonstrated as DI. The public should demand continued appropriations for only proven methods. Data relating to program effectiveness must be made available to school districts and colleges of education.

References

- Anderson, R., St. Pierre, R., Proper, E., & Stebbins, L. (1978). Pardon us, but what was that question again?: A response to the critique of the Follow Through evaluation. Harvard Educational Review, 48, 161-170.
- Carnine, D. (1988). Breaking the failure cycle in the elementary school. Youth Policy, 10, 22-25.
- Cohn, M. (Ed.). (1987). Congressional quarterly almanac, XLII. Washington, D. C.: Congressional Quarterly, Inc.
- Engelmann, S., Becker, W., Carnine, D., & Gersten, R. (1988). The Direct Instruction Follow Through Model: Design and outcomes. Education and Treatment of Children, 11(4), 303-317.
- Gersten, R. (1984). Follow Through revisited: Reflections on the site variability issue. Educational Evaluation and Policy Analysis, 6(4), 411-423.
- Gersten, R., Becker, W., Heiry, T., & White, W. (1984). Entry IQ and yearly academic growth of children in Direct Instruction programs: A longitudinal study of low SES children. Educational Evaluation and Policy Analysis, 6(2), 109-121.
- Gersten, R., & Keating, T. (1987). Long-term benefits from Direct Instruction. Educational Leadership, 44, 28-31.
- Gersten, R., Keating, T., & Becker, W. (1988). The continued impact of the Direct Instruction Model: Longitudinal studies of Follow Through students. Education and Treatment of Children, 11(4), 318-327.

- Hodges, W. (1978). The worth of the Follow Through experience. Harvard Educational Review, 48(2), 186-191.
- House, E., Glass, G., McLean, L., & Walker, D. (1978). No simple answer: Critique of the Follow Through evaluation. Harvard Educational Review, 48(2), 128-158.
- Meyer, L. (1984). Long-term academic effects of the Direct Instruction Project Follow Through. The Elementary School Journal, 84(4), 380-394.
- Meyer, L., Gersten, R., & Gutkin, J. (1983). Direct Instruction: A Project Follow Through success story in an inner-city school. The Elementary School Journal, 84(2), 241-252.
- Rhine, W. (1983). The role of psychologists in the national Follow Through Project. American Psychologist, 38(3), 288-297.
- Stebbins, L., St. Pierre, R., Proper, E., Anderson, R., & Cerva, T. (1977). Education as experimentation: A planned variation model (Volume IV-A): An evaluation of Follow Through. Cambridge, MA: Abt Associates.
- Watkins, C. (1988). Project Follow Through: A story of the identification and neglect of effective instruction. Youth Policy, 10, 7-11.
- Weisburg, R., & Weisburg, P. (1988). Effective preschools for at-risk children: Our hope for the future. Youth Policy, 10, 30-31.
- Wisler, C., Burns, G., & Iwamoto, D. (1978). Follow Through redux: A response to the critique by House, Glass, McLean, and Walker. Harvard Educational Review 48(2), 171-191.

Linda J.
Star

Relationships Between Language and Music

Linda J. Star

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

Since language and music are forms of communication, and both use symbol systems, other similarities may exist. This paper reviews the literature regarding structural, processing, and perceptive similarities between language and music. Also included is a review that indicates a high correlation between academic achievement and music. Recommendations and implications for educators and administrators are discussed.

Music education programs are frequently among the first casualties during budget cuts. Core subjects have obvious educational value, whereas music is often considered a frill or mere entertainment. Music education is neither frill nor solely entertainment; it is an activity that relates real life experience to the expression of experience. Music is a multi-modal art form that requires processing of complex temporal information, development of motor skills, assimilation and comprehension of coordinated sounds, and aesthetic appreciation.

Language and music share certain obvious traits. They both use symbol systems and sound, and both are forms of communication and expression. Musical development in humans has been compared to language development and is believed to be as universal to humans as is language. Sloboda (1985, p. 17) says "that humans have a general capacity to acquire linguistic and musical competence." Regarding music's cultural ubiquity, Serafine (1988, p. 1) states, "No culture exists without it

and all persons have a knowledge of it to a considerable degree." Citing anthropological studies, Dowling and Harwood (1986) argue that music does have adaptive value but at the societal rather than the individual level. They further suggest that musical abilities have evolved parallel to linguistic abilities, and that humans have enough musical ability to make music socially. If these assertions are true, then educators should reconsider music education in the schools.

This paper will explore relationships between music and language. Questions to be considered include: (a) Is music perceived according to structure? If so, is it similar to language perception? (b) Are processing strategies in music and language similar? and (c) Can predictive associations be made between music and academic achievement?

Hierarchical Structure

One area of study focuses on the hierarchical structure of language perception or music perception. That is, a sentence or piece of music contains primary structures upon which less important or ornamental structures are successively built. All the words in a sentence or sounds in a piece of music comprise the surface structure. By applying appropriate rules, deeper structures emerge. The final reduction is the most basic structure around which the sentence or music composition is built. In graphic form, a reduction resembles a family tree.

Sloboda (1985) compares the work of

linguist Noam Chomsky to that of the music theorist Heinrich Schenker. Almost simultaneously, Chomsky and Schenker developed similar theories of hierarchic structure in their respective fields. Chomsky described language as having surface structures which allow one to arrive at a deeper structure when a set of transformational rules are applied. He believed all natural languages have the same deep structure which makes language a universal intellectual skill. Schenker's theory describes a structure based on harmonic and contrapuntal movement. A limited number of deep structures can generate an unlimited number of surface structures. Chomsky's rules, however, are applied sentence by sentence, producing an unlimited number of deep structures.

Lerdahl and Jackendoff (1983) have developed a generative theory of tonal music which details rules used to derive a deep structure from the surface structure. They describe deep structure as analogous to phrase structure in language (p. 283). Their theory is restricted to the "formal characterization of the listener's intuitions about musical structure" (p. 332) and relies on the belief that humans have an innate cognitive musical capacity.

Although many theories of hierarchical structure in music exist, Serafine, Glasman, and Overbeeke (1989) could find no studies that tested the reality of its perception. Therefore, they performed a series of six experiments using six different groups of college students from either Yale University or Vassar College. The number of subjects and degree of musical experience varied in each group.

In order to test music perception according to hierarchical structure, an expert reduced several Bach excerpts to fore-

ground and middleground structures consistent with Schenker's theory. The first reduction from the surface is known as the foreground reduction, which can be further reduced to the middleground reduction. Each reduction contains fewer notes than the previous one. The expert then provided false reductions for each excerpt at both levels. Experiments 1-3 required subjects to match the correct reduction to the appropriate excerpt. Results showed a general tendency for subjects to correctly match a surface with its correct reduction. Metric placement of structural pitches seemed important in correct matching; however, the roles of aesthetic preferences and repeated hearings were unclear. Upon first hearing, subjects appeared to have more correct matches than after repeated hearings.

The experiments were revised, and newly composed pieces replaced the Bach excerpts. These compositions were generated from the structures in four forms: (a) two models having the same structure (M1 & M2), (b) a harmonically different piece having the same structure as the models (harmonic foil [HF]), and (c) a similar piece on the surface having different structure (counterpoint foil [CF]).

In experiments 4-6, subjects matched musical examples with a model and rated them according to similarity. Serafine et al. (1989) found that subjects were more likely to rate the HFs as being more similar to the model than were the CFs. Since the HFs differed more on the surface but had the same deep structure while the CFs had a different deep structure, Serafine et al. concluded that subjects were matching compositions according to perception of deep structure. Repeated hearings strengthened a subject's tendency to regard the HF

as similar to the model.

Experiment 6 provided no evidence that music is stored according to surface or deep structure. Repeated hearings may affect encoding. Nevertheless, the experimenters found sufficient evidence to support the theory of perception according to structure.

Processing Strategies

Levinowitz (1989) explored the relationship of language development to song production in 4- and 5-year-old children. The study examined how the imposition of words affected the learning of a song and how a child's language proficiency helped or hindered this learning process.

Subjects were thirty-five 4- and 5-year-old children from two nursery school classes in the same school. Levinowitz taught both classes for 5 months, one day per week for 30 minutes. Rhythm, movement, and rote-singing were included; both classes received the same instruction. Half the rote songs contained words, while the other half were sung with neutral syllables. Two criterion songs were taught the last month and were rated by two judges. One song contained words; the other was sung on "bum." The songs were somewhat harmonically and rhythmically similar.

Language development was assessed by the Peabody Picture Vocabulary Test-Revised (PPVT-R). Rhythm achievement and tonal achievement were assessed and their resulting scores correlated. No differences were found in rhythm performance between songs sung with or without words. A partial correlational analysis revealed a relatively strong relationship between rhythm performance without words and the following three indicators: (a) song with

words (.60), (b) song without words (.64), and (c) rhythm performance with words (.64). Results may indicate that underlying rhythms are involved in perception and processing of language and music.

A significant difference ($p < .01$) was found between tonal performances of the songs with or without words. The song without words was performed better than the song with words. Levinowitz hypothesized that a young child's attempt to learn a song with words may involve two different processes. Words may interfere with an ability to learn the melody. Due to a very low correlation between the PPVT-R and all other indicators, Levinowitz concluded that language development may have little to do with the ability to perform rote songs.

Young children may employ two different processes in language and music development. Hahn's (1987) study considered whole-to-part language processing by beginning string students, ages 9-12. Students tended to drop out of music with the introduction of traditional note reading, taught part-to-whole. Hahn's hypothesis represents Chomsky's (1969) and Goodman's (1984) psycholinguistic view of language learning, which maintains that language is learned whole-to-part, from general to specific rather than vice-versa. Hahn's research is also based on music perception theories developed by Longuet-Higgins (1978), Lerdahl and Jackendoff (1983), and Sloboda (1978), which explain that music is learned whole-to-part, similar to processes employed in language learning.

Subjects consisted of 18 students from two beginning public school string classes. Hahn worked closely with the two music teachers to coordinate the study. Both

classes received the same prereading instruction for the first 2 months. Following this instruction, the control group received traditional note reading instruction, while the experimental group received whole-to-part note reading instruction. A posttest was administered after the 11th music reading lesson.

On the sight-reading and music reading portions of the posttest, students in the experimental group performed significantly better than students in the control group. They performed similarly on the recognition task. Hahn suggested that the control group performed the recognition task using short-term memory, producing attention to surface elements. Due to the experimental group's superior performance on the sight-reading and music reading portions of the tests, Hahn concluded that whole-to-part note reading strategies provide a viable method to teach music reading.

In an attempt to discover the influences of learning style and language dependency on aural discrimination tasks, Schmidt (1984) studied 75 freshman music students at Indiana University. Previous studies determined two types of learning styles: (a) field dependence/field independence and (b) reflection/impulsivity. Field dependence/field independence describes whether an object is perceived in context or out of context. Reflection/impulsivity describes the speed and deliberateness of decision making and the accuracy of the response. The Group Embedded Figures Test (GEFT) (Witkin, Oltman, Raskin, & Karp, 1971) tested for field dependency, and the Matching Familiar Figures Test (MFF) (Kagan, 1964) tested for reflection/impulsivity. Language relationships in aural skills were tested based on a construct developed by Day (1981). This construct describes

language-bound individuals as those who process aural information by using language rules. Language-optional individuals use language rules or set them aside, depending on the situation. A test of temporal order discrimination (TOD) (Keele & Lyon, 1982) tested language-optional/language-bound perception. Scores on these three tests served as the independent variables. An entrance/placement examination which tested several types of aural skills served as the dependent variable.

A step-wise multiple regression determined which indicators were more important in predicting aural skills. The TOD test (16.1% of the variance) and the GEFT (4% of the variance) proved to be the best indicators of aural skills achievement. Language-optional subjects performed significantly better than language-bound subjects ($p < .05$). Field-independent subjects also performed significantly better than field-dependent subjects ($p < .05$). Schmidt (1984) suggested that the success of the language-optional students is probably due to reliance on short-term memory for each task (p. 166).

In the previous studies, subjects were from middle- to upper-middle-class socioeconomic backgrounds. The following two studies involved at-risk students.

In a study designed to help students with language problems access music, Tanner (1990) paired the Initial Teaching Alphabet (i.t.a.) with music. According to Tanner, "music will strengthen reading skills which, in turn, will allow the student greater access in learning musical skills" (p. 7). I.t.a. uses one symbol for each of the 44 sounds of the English language and is easily transferred to the 26-character alphabet. This system was chosen to eliminate confusion arising from the many

inconsistencies between spoken and written English.

Twelve elementary students enrolled in the resource program were tested during the first 4 weeks of classes and again in the spring during the last 4 weeks of classes. The experimental group was taught reading by the Early-to-Read i/t/a/ Program; music was taught using this system in conjunction with the Silver Burdett Music Series and supplemental folk songs. The control group was taught reading through a pull-out resource program and music was taught conventionally. They received reading instruction 5 days per week, 30 minutes each day, and music instruction 3 days per week, 30 minutes each day. Reading and music classes were taught to the experimental group using a multisensory approach: visual, auditory, and kinesthetic. Computers and a MIDI sequencer/synthesizer were used in music composition and playback. Computer programs using i.t.a. were written for both the music and reading classes. Students were encouraged to create their own songs and supply correct rhythms. Phrasing in music and language received rhythmic emphasis.

ANOVA results indicated significant improvement in all areas of the reading and music tests ($p < .01$). The control group improved, but not significantly. Tanner (1990) suggested that the interaction between the i.t.a. reading system and the multisensory teaching methods used with the experimental group served as mutual reinforcement.

A similar study involving at-risk students was undertaken by Everett (1990) to determine if music would improve reading and comprehension skills in third-grade students. Students in two classes from low

socio-economic backgrounds were selected and then subdivided into experimental and control groups.

The control group received normal and remedial reading instruction, while the experimental group received three extra 30-minute music activities per week in lieu of remedial reading. The study took place over an 18-week period. Activities in the experimental group remedial reading (music) class utilized visual, aural, and kinesthetic modes and incorporated language into musical elements by means of rap, games, and dances.

The results of a *t* test indicated that the experimental group scored significantly better than the control group in reading achievement ($p < .05$). A comparison of grade level changes using the Botel Reading Inventory revealed that all students in the experimental group attained a normal 1/2 year gain, while only three students in the control group demonstrated any gain.

Relationship between Music and Academic Achievement

In an extensive correlational study, Asmus (1990) collected and analyzed national, regional, and local statistics to determine the effects of music on academic achievement and concluded it has a positive effect. Asmus' literature review includes a study which positively correlated music electives in junior high and high school with higher scores on the Iowa Test of Basic Skills taken in the 6th grade. In another study, Kvet (1985) found that excusing 6th grade students from regular class time for music did not have a negative effect on academic achievement.

After collection and analysis of national data on test scores, monetary expendi-

tures, and student/teacher ratios, Asmus (1990) found student/ music teacher ratio to be the strongest indicator of academic achievement as measured by the Scholastic Aptitude Test or the American College Test.

In a related study, Hedden (1982) studied the effects of academic achievement on music achievement. After considering other factors (music background, attitude toward music, self-concept, and gender), Hedden found that academic achievement most strongly predicted music achievement. Inversely, Asmus (1990), using Hedden's data, found that music achievement was positively correlated with academic achievement as measured by the Iowa Test of Basic Skills.

Summary and Conclusions

Rhythm and meter seem to greatly influence structural perception (Serafine et al, 1989), song production (Levinowitz, 1989), and music and language achievement (Tanner, 1990). Levinowitz's study revealed a relatively strong relationship between rhythm and all other indicators of song production. However, language and music, when paired, may not contain congruous rhythmic or metrical structures. Perhaps for this reason, subjects learned rote songs without words more easily than songs with words. In Tanner's study, students composed texts and supplied correct rhythms. Perhaps, they intuitively supplied matching structures which produced success. Original creations might have made the subject matter more relevant, therefore more meaningful.

Short-term memory was cited in recognition tasks (Serafine et al., 1989; Hahn, 1987) and learning style (Schmidt, 1984).

Tasks calling for quick judgments produced attention to underlying structure, while recognition tasks seemed to produce attention to surface structures. Part-to-whole teaching methods focus attention to surface which may account for the control group's success on the recognition task in Hahn's study. Language-optional and field-independent perception, also attributed to short-term memory (Schmidt), may require another cognitive process such as one for discrimination.

Studies by Tanner (1990) and Everett (1990) correlated language achievement with music education. Everett's study, however, may have more practical implications for schools lacking elaborate equipment. Both studies used multisensory approaches with high-level student involvement. By its very nature, music is multisensory, which readily lends itself to this approach. Asmus (1990) found that music does strongly influence academic achievement.

These findings suggest other research: (a) What is the interrelatedness between underlying structure in music and language? (b) What place might rhythm and meter have in this relationship? and (3) Can the ability to discriminate be developed through music and transferred to language?

In conclusion, these studies suggest that music should occupy a more prominent place in education, particularly with at-risk students. This issue requires cooperation among administrators, music specialists, and special education personnel. Reading skills are very important, providing students access to other academic areas. If music helps students unlock language barriers, educators should pursue the relationship between language and music.

References

- Asmus, E. (1990). The influence of music education on non-musical indicators of educational attainment. Paper presented at the National In-Service Conference on the Value of Music Learning. Winston-Salem, NC.
- Chomsky, N. (1969). Syntactic structures. Paris: Mouton.
- Day, R. (1981). Music ability and patterns of cognition. In R. Day (Ed.), Documentary report of the Ann Arbor Symposium: Applications of psychology to the teaching and learning of music (pp. 270-283). Reston, VA: Music Educators National Conference.
- Dowling, W., & Harwood, D. (1986). Music cognition. Orlando: Academic Press.
- Everett, P. (1990). The effect of music on the development of basic reading skills of third-grade students. Georgia Music News, 51(1), 55-57.
- Goodman, K. (1984). Unity in reading. In A. Purves and O. Niles (Eds.), Eighty-third Yearbook of the National Society for the Study of Education: Becoming readers in a complex society Part 1. (pp. 79-114). Chicago: National Society for the Study of Education.
- Hahn, L. (1987). Music reading and language reading: Correlations in processes and instruction. Council of Research in Music Education Bulletin, 93, 41-48.
- Kagan, J. (1964). Matching Familiar Figures. Cambridge, MA: J. Kagan.
- Keele, S., & Lyon, D. (1982). Individual differences in speech fusion: Methodological and theoretical explorations. Perception and Psychophysics, 32(5), 433-442.
- Kvet, E. (1985). Excusing elementary school students from regular classroom activities for the study of instrumental music: The effect on sixth-grade reading, language, and mathematics achievement. Journal of Research in Music Education, 33(1), 45-54.
- Lerdahl, F., & Jackendoff, R. (1983). A generative theory of tonal music. Cambridge, MA: MIT Press.
- Levinowitz, L. (1989). An investigation of preschool children's comparative capability to sing songs with and without words. Council of Research in Music Education Bulletin, 100, 14-19.
- Longuet-Higgins, H. (1978). The perception of music. Interdisciplinary Science Reviews, 3(2), 148-156.
- Schmidt, C. (1984). The relationship among aspects of cognitive style and language-bound/language-optional perception to musicians' performance in aural discrimination tasks. Journal of Research in Music Education, 32(3), 159-168.
- Serafine, M. (1988). Music as cognition. New York: Columbia University Press.
- Serafine, M., Glassman, N., & Overbeeke, C. (1989). The cognitive reality of hierarchic structure in music. Music Perception, 6(4), 397-430.
- Sloboda, J. (1978). The psychology of music reading. Psychology of Music, 6(2), 3-20.
- Sloboda, J. (1985). The musical mind. Oxford, Great Britain: Clarendon Press.
- Tanner, D. (1990, November). Initial teaching alphabet, language experience and music. Paper presented at the National In-Service Conference on the Value of Music Learning. Winston-Salem, NC.

Witkin, H., Oltman, P., Raskin, E., &
Karp, S. (1971). A manual for the
Embedded Figures Test. Palo Alto:
Consulting Psychologists Press.

The Self-Esteem of Gifted Children: A Developmental and Comparative Study

Pollyann Diamond

This study compared aspects of the self-esteem of gifted and non-gifted, high achieving children in grades 3, 4, 5, and 6 using Harter's Self-Perception Profile for Children. Gifted children perceived themselves as more competent than their non-gifted peers only in the scholastic area ($p < .01$). Differences were found for both groups in the physical appearance area by grade ($p < .001$) and in the athletic ($p < .01$) and behavioral conduct ($p < .001$) areas by gender. Consistent with previous findings, boys were found to have higher self-esteem in the athletic realm, while girls expressed higher self-esteem in conduct. Also, compatible with other research was the finding that self-esteem associated with appearance declines over the elementary school years. Gifted children appear to be quite similar to their high achieving, non-gifted peers in all but scholastic areas. The difference in scholastic self-esteem favored the gifted group after only 1 year in a special program and continued to remain relatively constant over time, while the scores of the non-gifted peers declined. It is indeterminate if the difference in the scholastic area is due to the first year of the special program or to the effect of labeling a child as gifted.

Self-concept is a construct comprised of an individual's beliefs about various aspects of the self derived from interaction with significant others and self-appraisal of competencies. Self-esteem, on the other hand, is the value or worth an individual assigns to various dimensions within the construct of self-concept (Harter, 1983). In

the literature, the terms "self-concept" and "self-esteem" have been used interchangeably; self-esteem is the referent in formal measures of self-concept. The term "self-concept" will refer, in this paper, to the global construct; "self-esteem" will apply to particular aspects of the general construct (e.g., academic or social self-esteem). There is less agreement on the nature and measurement of the global construct but considerable agreement on the multidimensional quality of its constituents based on factor analytic studies (Byrne, 1984).

Historically, the construct of self-concept emerged in the works of William James and, notably, in the writings of Cooley around the turn of the century (Harter, 1983). Many schools of psychology have undertaken studies of the self-concept and have generated a variety of theories to account for the importance of this construct within the personality. Cooley suggested that self-image is a reflection of how others regard us. Piaget (1981), a representative of the organismic school, indicated that the self-concept is not merely a reflection of how others view us but stressed the additional importance of self-evaluation throughout childhood. Theorists of the humanistic perspective link self-esteem to both motivation and achievement citing the educational significance of deficiency needs, self-fulfilling prophecy, and the central role of positive self-concept development for students (Biehler & Snowman, 1986, p. 496). Bandura (1977), a neo-behaviorist, reinforced the connection between self-esteem and motivation, contending that successful achievement enhances

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Pollyann J.
Diamond

45

52

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

self-efficacy which, in turn, increases motivation and partially determines confidence, persistence, and goal-setting. Thus, organismic, humanistic, and neo-behavioristic schools acknowledge the importance of self-concept, but differ on the relationship between self-esteem and achievement.

Self-Esteem and Achievement

According to traditionalists, success enhances self-esteem (Harter, 1989). In the 1960s, the popular view emerged that self-esteem enhancement improved academic performance. However, Bachman and O'Malley (1977) found no evidence that high levels of school self-esteem were related to subsequent educational or occupational attainment. Rather, academic ability and achievement influenced self-esteem. On the other hand, studies do support that a student's perception of academic self-esteem may strongly impact achievement (Caslyn & Kenny, 1977; Harter & Connell, 1982). Perceived academic competence (i.e., high academic self-esteem) predicted motivational orientation; a greater sense of cognitive competence intrinsically motivates the student to perform academically (Harter & Connell). Thus, it appears that academic self-esteem, rather than global self-concept, is importantly related to achievement. Additionally, mood or affect may be a moderating variable in this relationship (Harter, 1983). Low self-esteem can be associated with diminished affect and energy to perform academically.

There is little agreement in the field of education that enhancement of self-esteem impacts achievement, or that achievement enhances self-esteem. A reciprocal relationship between achievement and dimen-

sions of self-esteem appears to be the emerging conceptualization. The importance of achievement to self-esteem seems to relate to the student's perception of academic competence (i.e., academic self-esteem) and the value the student attaches to that aspect of the self-concept.

Self-Esteem and Gifted Children

Gifted and high-achieving children would be expected to exhibit relatively high self-esteem regardless of the directionality of the relationship between self-esteem and achievement. Students are said to respond according to their self-perceptions, and the way teachers react to students influences the students' self-perceptions (Biehler & Snowman, 1986, p. 405). Huryn (1986) reported that gifted students generally believed their teachers viewed them as good students to whom they could give more work and send on errands since they were more reliable. However, 11 of the 60 students interviewed had encountered teachers who viewed them negatively; these students felt they had to hide their giftedness for the teacher to like them.

Several studies have examined teachers' reactions to and perceptions of the gifted. Findings indicate that two criteria have greatest impact on teacher perceptions: (1) years of teaching experience and (2) knowledge about the gifted. The more experienced teachers and those with greater knowledge of the characteristics of gifted demonstrate the greatest acceptance and positive regard for gifted students (Bran-sky, 1987; Cramond & Martin, 1987; Morris, 1987). Further, teachers may respond to gifted students based, in part, on their perceptions of gifted programs (Cologanello & Kelly, 1983; Schneider, 1987)

which may not always be favorable.

Historically, the self-concept of gifted children has been a topic of considerable discussion. In the last century, the gifted were stereotyped as puny, insane, and weird (Maker, 1977). Khatena (1982) cited Terman's longitudinal studies on gifted individuals which indicated that the gifted were generally superior on most measures, although considerable variability was found. Zaffrann and Colangelo (1979) questioned whether these qualities necessarily dispose the gifted to better acceptance by others, better personal adjustment, and greater self-esteem. Such children may find themselves isolated from peers. For example, Ross and Parker (1980) reported that gifted students exhibit lower expectations for social success than academic success.

Other investigators confirm that gifted and high-achieving students may lack social confidence and perceive themselves as less popular (Brody & Benbow, 1986; Whalen & Csikszentmihalyi, 1989). Khatena (1982) also expressed concern for affective development of the gifted and talented, viewing adequacy in this domain, including self-concept, as a prerequisite to the risk-taking behavior necessary for new discoveries. Thus, impairment in this realm may result in loss of productivity and other social consequences. Maker (1982) echoed these concerns by including development of a positive and realistic self-concept and positive attitudes toward learning as important goals for gifted programs.

The self-concept of gifted children concerns experts in the field. The literature does not support that the gifted automatically have positive self-concepts and adequate affective adjustment. McDowell (1984) summarized the literature by stating

that "feelings of being 'different' can have a deleterious effect on both the (gifted) child's emotional and social adjustment" (p. 26). He further commissioned administrators, teachers, and counselors to attend to this aspect of gifted children.

The consensus appears to be that the self-concept of gifted individuals is not uniformly positive, and that programs should make enhancements of self-esteem a major goal. The task is to determine what is being done to develop a positive self-concept in gifted children. What developmental trends are noted by the researchers? Has there been a systematic study to determine the effect of special programming on the self-concept of the gifted population?

Current findings have resulted in inconsistent conclusions. Some researchers used global scores of self-esteem and reported no difference between gifted and non-gifted; others, using the same or similar measures, found that gifted children have significantly higher or significantly lower self-esteem than their non-gifted peers.

Bracken (1980) studied gifted elementary school students and their non-gifted peers employing self-esteem scales designed specifically for that study. No significant difference was found in the global self-esteem scores of the two groups. In a similar study, Ross and Parker (1980) administered the Sears Self-Concept Inventory to gifted students in grades 5 through 8. No sex differences were reported on either the academic or social subscale, but academic self-esteem of the gifted population was found to be significantly higher than social self-esteem. No significant difference was reported across grade levels. Kelly and Colangelo (1984)

reported similar findings using the Academic Self-Concept Scale and the Tennessee Self-Concept Scale (social self-concept) with students in grades 7 to 9. Thus, in the Ross and Parker study, academic self-concept was higher than social self-esteem among gifted students; in the Kelly and Colangelo study, social and academic self-esteem among the gifted were higher than age-mate controls.

Lehman and Erdwins (1981) attempted to further define the social incongruity experienced by gifted students as they relate to their non-gifted peers. Third-grade gifted students were compared to age peers and to sixth-grade students described as their mental-age peers. On social skills and sense of personal worth, the researchers found gifted students differed more significantly from peers of the same chronological age than from peers of the same mental age; generally, the gifted students were well-adjusted personally and socially and had good self-concepts. In families having a gifted child and a non-gifted sibling, parents may "idealize" the gifted child resulting in lower self-esteem for the non-gifted sibling (Colangelo & Brower, 1987; Cornell, 1983).

Contrast effects and social comparisons appear to be factors impacting the development of self-esteem. According to Bandura's social comparison theory (1977), gifted children may or may not increase self-esteem over the school years depending on the educational and social group with whom they associate. This theory may account for differences in self-esteem among gifted resulting from different service methods reviewed subsequently.

Some researchers have found significant differences between the self-esteem of gifted children and their non-gifted peers;

others have not. Results are inconsistent, probably due to the use of different instruments, different ages and abilities of children, and divergent control-comparison groups as well as comparison of the self-esteem of gifted and controls following diverse lengths of placement and types of special programming delivery models. Further comparative studies of global and unique dimensions of the self-concept of gifted children with both age peers and mental age peers is indicated.

Gender Differences in Self-Esteem and Peer Acceptance

Gender comparisons have been presented in the literature on how gifted children are perceived/accepted by peers and whether gender is a factor in self-esteem among the gifted or in relation to controls. Gender differences in peer acceptance of the gifted did not obtain, while attitudes toward non-gifted females were more favorable than those for non-gifted males (Killian, 1981). Using a peer nomination technique to assess perceived social competence of 10th-grade gifted and non-gifted males and females, Schneider (1987) reported a pattern similar to Killian's noting that methodological problems in previous studies may have obscured gender differences in ratings of social competence and acceptance since the rater's gender was not taken into account.

When comparing the self-esteem of gifted to non-gifted peers, gender differences were reported by several researchers using a variety of instruments. Colangelo, Kelly, and Schrepfer (1987), using the Tennessee Self-Concept Scale, reported that gifted girls, grades 7 through 9, had higher academic self-esteem than did non-gifted

girls. Gifted boys had higher general self-esteem than did non-gifted boys. No differences were noted regarding gender; however, other studies (e.g., Bartell & Reynolds, 1986) did find differences.

Bartell and Reynolds (1986), using the Coopersmith scale, reported that gifted girls in grades 4 and 5 had higher self-esteem than did gifted boys; non-gifted boys' self-esteem was higher than that of non-gifted girls. They commented that the difference may have been due to the reinforcement gifted girls received from teachers for academic performance compared with the tendency to reinforce boys for achievement in other areas.

Loeb and Jay (1987) reported gender differences in global self-concept for gifted students ages 9 through 12 on the Piers-Harris scale. Gifted girls had significantly higher self-esteem than did non-gifted girls; no effects were noted for gifted versus non-gifted boys. No gender differences were noted within the gifted group.

With regard to global self-concept, reports of gender differences are inconsistent. Two studies reported that gifted girls' self-concepts were significantly higher than those of non-gifted girls; however, differences within the gifted samples were not consistently significant. Additional study is needed to clarify this point. Gifted girls tend to have higher self-esteem than do their non-gifted female peers, while the same is not evident for boys.

Grade Effects and Self-Esteem

Several researchers have used the Piers-Harris Self-Concept Scale to assess gifted students, using the norm sample for comparison. The results, with regard to grade effects, have been inconsistent.

Karnes and Wherry (1981) found that, as a group, gifted fourth through seventh graders had significantly higher self-esteem than the norm sample for this instrument. No significance was noted for gender or grade, or between gifted enrolled in gifted programs and those not enrolled. In a study of gifted students in grades 5 and 6, Maddux, Scheibner, and Bass (1982) found significant grade effects using the Piers-Harris scale. Sixth graders had significantly higher self-esteem than did those subjects in the norm sample, while no difference was found between fifth graders in this study and those in a previous study. The norm sample could not be used for the fifth grade comparison since no fifth graders were included in the norm sample.

Placement Effects

Using the Piers-Harris scale, Coleman and Fults (1982) studied the global self-concept of fourth- through sixth-grade students enrolled in a gifted program and the global self-concept of high achieving students who remained in regular classes. No significant effects for sex, grade, or time period of testing (prior to enrollment or after enrollment) were reported. A surprising result was that gifted students enrolled in special programs had significantly higher self-esteem than did the norm sample but had a significantly lower self-esteem than did high achievers in regular classes. When returned to regular classes, the self-esteem scores of the gifted returned to preplacement levels within 8 months. This effect was especially significant in the area of academic self-esteem. The researchers hypothesized the difference was due to the reference groups of each population. The researchers hypothesized

that the results reflected the influence of the reference groups of the gifted high-achieving students. Gifted students in special programs compared themselves to gifted students, while high achievers in regular classes compared themselves to less gifted peers. As a result, the gifted enrolled in special programs had higher self-esteem but a more realistic perception of their ability than did those high achievers in regular classes. Chovan and Morrison (1984) reported similar results with students ages 9 through 12. Schneider (1987) observed that such findings would be predicted on the basis of Bandura's (1977) social comparison theory.

In an attempt to clarify this relationship between placement and self-esteem, Coleman and Fults (1985) assessed fourth-grade students before and after placement in gifted programs. They hypothesized that the children's self-esteem prior to placement would be higher than their self-esteem after placement; and those "just eligible," or borderline, would have a lower self-esteem than those with higher IQs. All gifted students in the study had high self-esteem. Preplacement self-esteem scores were higher than postplacement scores. High IQ gifted had higher self-esteem scores than lower IQ gifted, irrespective of time of assessment. Before placement, scores of both groups were approximately equal; after placement, those with high IQs had significantly higher scores than the lower IQ group. The researchers hypothesized that high IQ students did not show significant changes after placement because they were superior to both referent groups (lower IQ gifted and non-gifted students); lower IQ gifted students reflected the change in perception due to changes in referent group. Again, evidence was found

to support the social comparison hypothesis.

In a related study, Chapman and McAlpine (1987) completed a 2-year investigation of gifted and average-achieving children, grade 6, using the Student's Perception of Ability Scale. The gifted group had significantly higher self-esteem than the average group. The researchers found no significant group or time effect, indicating relative stability over the 2-year period in the self-perception of ability for both groups. This may indicate that self-perception of ability becomes stable as patterns of performance become stable; therefore, high levels of achievement do not necessarily produce increasingly high levels of academic self-esteem. Since the gifted sample in this study was not enrolled in special programs, the stability of their perceptions of ability may be reflective of the stability of the reference group. The researchers suggest that a study of gifted children enrolled in special programs may show that the academic self-esteem of these students is lower than that of similar students enrolled in regular classes; they also conjectured that this lower self-esteem is due to the character of the referent group for the special classes.

In three recent studies, the Harter scale (Harter, 1979), the Perceived Competence Scale for Children (Harter, 1982), or the Self-Perception Profile for Children (Harter, 1985) were used. These studies highlight the influence of reference groups on the gifted students' perception of their cognitive functioning. Chan (1988), using the original scale, compared three groups of Australian children: (1) gifted children, grade 7, enrolled full-time in enrichment classes; (2) gifted children, grades 5, 6, and 7, enrolled part-time in enrichment classes;

and (3) students enrolled in regular classes, grades 5, 6, and 7, in the same schools as those students in the enrichment classes were housed. The findings were similar to those of Coleman and Fults (1982, 1985). Gifted students had higher perceived competence on cognitive and general self-worth scales, regardless of sex or type of program, than did the controls. Gifted students in full-time enrichment classes had lower scores on cognitive and physical scales than those enrolled in part-time enrichment classes. The researcher noted that in cognitive functioning the former group compared themselves exclusively to similarly able peers, thus developing a more realistic and accurate perception of cognitive competence while experiencing adequate success to allow high perceptions of general self-worth.

Similar results were reported by Li (1988) using the revised Harter scale. The subjects were gifted fourth and seventh graders and non-gifted controls from urban Canadian schools. The gifted group had generally positive self-perception across the domains with significant group effects in scholastic competence and behavior. A significant group effect was also found for athletic competence, in favor of the control group. The gifted girls had significantly higher scholastic competence scores than either gifted boys or controls. This finding supports that of Coleman and Fults (1982) regarding the academic self-esteem of gifted girls. The findings also support those of Chan (1988) in which the gifted group had significantly higher scholastic competence scores than did the controls.

The third study also used Canadian students, grades 5, 8, and 10 (Schneider, Clegg, Byrne, Ledingham, & Crombie, 1989). The gifted students participated in

either self-contained gifted classes or regular classes (integrated gifted); the control group consisted of classmates of the integrated gifted. Some controls were matched for age, sex, and number of years in school; others were randomly matched. The randomly matched sample was used to complete the statistical comparisons. At all three grade levels, significant effects were noted for scholastic competence, favoring the integrated gifted over controls. With regard to school placement of gifted (self-contained versus integrated), a significant difference was found in grades 5 and 8 where the integrated gifted had higher perceived scholastic competence scores than did self-contained gifted; while not significant at grade 10, a difference favored the integrated group. These results are consistent with those of Coleman and Fults (1982, 1985), Chan (1988), and Li (1988).

Examination of current literature supports the contention that scholastic competence (i.e., academic self-esteem) is a distinguishing factor in the self-esteem of gifted children. A related variable in the formation of scholastic competence is the referent group. Initial studies found no difference between gifted students in special programs and gifted students not placed. Other studies found part-time special programs to be the most favorable for the self-concept of gifted children. Sex differences do not occur consistently but tend to favor female gifted over non-gifted controls. The developmental issue has yet to be systematically addressed regarding number of years in special programs and ages of the children. At present, little research is available addressing the effect of systematic programming on the self-concept of the gifted.

The purpose of the present study was

to assess developmental and comparative aspects of self-esteem with gifted children. Information is needed to describe the gifted and to address the effects of special programs. Does part-time placement (enrichment) of the gifted result in higher scholastic competence (Chan, 1988), and is it likely to continue after several years in gifted programs; or is this effect temporary, with scholastic competence stabilizing or even declining with increasing number of years in special programs (Chapman & McAlpine, 1987)? Also, do findings by Chan (1988), Li (1988), and Schneider et al. (1989) using Australian and Canadian children generalize to American children? The present study proposed to examine these questions and to specifically address the following hypotheses: (1) gifted children will score significantly higher than their high-achieving peers on a global measure of self-esteem; (2) gifted children will score significantly higher on a global measure of self-esteem with each successive year that they are enrolled in a special enrichment program; (3) gifted children will score significantly higher than high-achieving children not eligible for special programs on a measure of academic self-esteem; (4) gifted children will score significantly higher than their high-achieving peers on other areas of self-esteem; and (5) gender differences in self-esteem should favor female gifted over non-gifted female controls.

Method

Subjects

Subjects for this study were selected from students in grades 3, 4, 5, and 6 enrolled in three public elementary schools

in a small urban area. Two of the three schools house grades K-4 and serve as "feeders" for the third school which contains grades 5 and 6. Students enrolled in these schools were primarily from middle and upper middle income families.

Eligibility for participation first involved examination of second grade achievement test results from the years 1986 - 1989. All students defined as eligible to participate met state criteria for gifted programs regarding scores on a standardized achievement test, the Iowa Test of Basic Skills (ITBS). A student's Total Reading or Total Math must be at or above the 90th percentile or the Complete Composite at or above the 85th percentile. Participants were further delineated as gifted or non-gifted on the basis of Cognitive Ability Test scores from the same second-grade testing. Students scoring at or above the 96th percentile on two of the three areas (Verbal, Nonverbal, or Quantitative) or at or above the 96th percentile by averaging standard scores on two of the three areas were eligible to participate in the gifted program. Non-gifted students who became eligible for the gifted program after grade 3 were considered ineligible for inclusion in the study.

Control subjects were defined as those children who initially qualified on the achievement (ITBS) measure but failed to qualify on the aptitude test at the end of second grade. Thus, they were selected from grades 3, 4, and 5 from the same schools and regular classes as their gifted peers. They had nearly identical educational experiences in terms of teachers, curriculum, and peers apart from the enrichment placement.

All gifted students in the study were placed in the gifted program at the begin-

ning of third grade. Therefore, the sixth-grade group participated for 4 years, the fifth-grade group for 3 years, the fourth-grade group for 2 years, and the third-grade group for 1 year. With few exceptions, students in both programs had been enrolled in one of the K-4 schools and/or the 5-6 school within this school system for the duration of their education. Exact figures reflecting participation in the study are included in Tables 1 and 2. The gifted program was delivered on a resource basis. Students attended gifted classes 5 hours per week. The remaining 25 hours per week were spent in a regular classroom with the control group members as classmates.

Materials

The Self-Perception Profile for Children (Harter, 1985) was used to assess self-esteem. This scale yields scores on five specific subscales (Scholastic Competence, Social Acceptance, Athletic Competence, Physical Appearance, and Behavioral Conduct). Also included is a separate scale measuring Global Self-Worth. This scale

is not a composite of the five subscales but is computed from responses to items not included on any other subscale. The total scale is designed for use with children ages 8 to 13, approximately grades 3 to 6 or 7. Reliabilities reported in the manual for various subscales range from .71 to .86. The reliability reported for the Global Self-Worth scale ranges from .78 to .84. Reported means are approximately 3.0 with standard deviations falling between .50 and .85.

The total scale includes 36 items, six items per subscale, and is completed in approximately 30 minutes. Items are worded in a structured alternative format in which either choice is perceived to be equally attractive. The author of the instrument granted permission to duplicate it for use (Harter, 1985, p. 7).

Procedure

Permission from parents, teachers, and administrators was obtained for each participant. Parents and teachers evidenced a high level of cooperation. Most permission

Table 1. Number of Subjects by Sex, Grade, and Program

Subjects	Grade								Total
	3		4		5		6		
	G	C	G	C	G	C	G	C	
Male	15	15	9	15	14	17	11	12	108
Female	13	16	13	15	7	13	10	17	104
Total	28	22	30	21	30	21	21	29	212

Note. G = Gifted (N=92); C = Control (N=120)

Table 2. Mean Age for Gifted and Controls

Subjects	Grade								Total
	3		4		5		6		
	G	C	G	C	G	C	G	C	
Male	9-3	9-4	10-1	10-4	11-2	11-3	12-2	12-4	10-9
Female	9-1	9-2	9-11	10-5	11-0	11-5	12-1	11-8	10-7
Total	9-2	9-3	10-0	10-4	11-1	11-4	12-2	12-0	10-8

Note. Mean Age Gifted (G) = 10-7; Controls (C) = 10-8

slips were returned within the 10-day period established for collection; a total of 320 students (80% return) had permission to participate. Administration dates and locations were established with building principals and communicated to the teachers involved.

The scale was administered to groups of students, by grade, within a 1-week period in May. Each administration followed standard procedures as outlined in the Harter manual. Items were read aloud to third and fourth grades; fifth and sixth graders worked at their own pace.

After protocols were scored, an average was determined on each of the six scales. Results were recorded by code number on group tally sheets. Since the number of students in regular education and gifted programs was disproportionate (2:1 in some grades), random selection of approximately 30 regular education, high-achieving students in each grade was completed. Grade 6 consisted of only 29 high-achieving, control students; random selection was not necessary for this grade.

Results

Data were examined using a 2 x 2 x 4 ANOVA (Group Status x Gender x Grade in School). Significant main effects were found on the Scholastic Competence subscale for program and grade as well as on the Physical Appearance subscale for grade. Complete results are shown in Tables 3 and 4.

Scholastic Competence

Gifted students had significantly higher self-perceptions of their scholastic competence than did the control group, $F(1, 196) = 10.76, p < .01$ (see Table 3). Grade was also significant, $F(3, 196) = 4.97, p < .05$. Subsequent one-way ANOVAs were completed across grade levels for both instructional groups. Significant differences were obtained for grade within the control group $F(3, 119) = 3.88, p < .05$. Both groups recorded high mean scholastic competence scores compared with scale norms. The gifted group's mean scores remained high

Table 3. ANOVA Summary for Scholastic Competence Subscale

Source of Variation	SS	f	MS	F
Program	3.49	1	3.49	10.76**
Grade	4.84	3	1.61	4.97*
Gender	0.03	1	0.03	0.10
Program X Grade	0.50	3	0.17	0.52
Program X Gender	0.37	1	0.37	1.17
Grade X Gender	1.67	3	0.56	1.71
Program X Grade X Gender	0.30	3	0.10	0.31
Residual	63.55	196	0.32	
Total	74.74	211		

* $p < .05$, ** $p < .01$

over the four grades (no significant difference), while the control group's mean scores declined significantly over time. Post hoc comparisons using the Scheffe technique indicated that the sixth-grade, high achieving, non-gifted students had significantly lower scores on this subscale than third, fourth, and fifth graders within the same group.

Physical Appearance

Main effects for grade were noted on the Physical Appearance subscale, $F(3, 196) = 7.35, p < .01$ (see Table 4). Subsequent onr-way ANOVAs completed across the grades within both educational groups indicated a significant difference for both groups; for gifted $F(3, 88) = 2.71, p < .05$, and non-gifted $F(3, 116) = 4.63, p < .01$. Post hoc comparisons using the Scheffe technique showed significant differences

between grades 3 and 6 for the gifted group. Similar comparisons revealed significant differences between grades 4 and 6 for the control group. A comparison of grades 3 and 4 with grades 5 and 6 was also significant for controls. For both groups, perception of competence on the Physical Appearance sub-scale was significantly lower for students in grade 6.

Other Subscales

No main effects were noted on Social Acceptance, Athletic Competence, Behavioral Conduct, or Global Self-Worth with regard to program or grade. Mean scores for both groups in all areas were within the 2.75 to 3.52 range. Most scores were near 3.0 (reported mean for the norm group). Table A-1 provides a complete listing of mean scores for the six subscales for both groups and all grades.

Table 4. ANOVA Summary for Physical Appearance Subscale

Source of Variation	SS	f	MS	F
Program	0.01	1	0.01	0.02
Grade	10.81	3	3.60	7.35*
Gender	1.06	1	1.06	2.17
Program X Grade	0.18	3	0.06	0.12
Program X Gender	0.25	1	0.25	0.51
Grade X Gender	2.92	3	0.97	1.98
Program X Grade X Gender	2.98	3	0.99	2.02
Residual	96.06	196	0.49	
Total	114.24	211		

* $p < .01$

In addition to main effects already mentioned on the Scholastic Competence and Physical Appearance scales, main effects were noted for gender on the Athletic Competence and Behavioral Conduct subscales. Both effects are consistent with

Table 5. ANOVA summary for Athletic Competence subscale

Source of Variation	SS	f	MS	F
Program	0.07	1	0.07	0.14
Grade	0.06	3	0.02	0.04
Gender	6.06	1	6.06	11.61*
Program X Grade	1.05	3	0.35	0.67
Program X Gender	0.81	1	0.81	1.55
Grade X Gender	0.50	3	0.17	0.32
Program X Grade X Gender	1.45	3	0.48	0.93
Residual	102.27	196	0.52	
Total	112.22	211		

* $p < .01$.

Table 6. ANOVA Summary for Behavioral Conduct Subscale

Source of Variation	SS	f	MS	F
Program	0.00	1	0.00	0.01
Grade	1.88	3	0.63	1.57
Gender	12.83	1	12.83	32.13*
Program X Grade	2.09	3	0.70	1.74
Program X Gender	0.06	1	0.06	0.14
Grade X Gender	1.30	3	0.43	1.04
Program X Grade X Gender	1.06	3	0.35	0.88
Residual	78.27	196	0.40	
Total	97.79	211		

* $p < .001$

results of the norm samples in the manual. No main effects were noted on the Social Acceptance or Global Self-Worth scales. With regard to Athletic Competence, males consistently rated themselves as more competent than females $F(1, 196) = 11.61$, $p < .01$ (see Table 5). On the Behavioral Conduct subscale, females consistently rated themselves higher than males without regard to grade or program, $F(1, 196) = 32.13$, $p < .01$ (see Table 6). No significant interactions were noted in either analysis. A summary of ANOVA results by subscale is included in Table A-2.

Discussion

This study hypothesized that gifted children and high achieving, non-gifted peers (control group) would show a difference with respect to self-esteem. The Global Self-Worth subscale best charac-

terizes what is commonly termed global self-esteem. No main effects by program were noted on this subscale; thus, the hypothesis of a difference between the two groups on global self-esteem was rejected. The mean scores on this scale ranged from 3.10 to 3.38. The mean for the national sample of age-appropriate children reported in the manual was 3.0; therefore, both groups in this study exhibited ratings within a standard deviation of the normative mean.

The second hypothesis stated there would be a difference in the self-esteem of gifted students as a result of the number of years in a special program. There were no main effects for grade on the Global Self-Worth subscale for either group; this hypothesis, with regard to Global Self-Worth, was rejected.

The third hypothesis stated there would be a difference in the academic self-

esteem of gifted children and high achieving children not eligible for special programs. Significant main effects were noted by program and grade with no significant interaction on the Scholastic Competence subscale. Gifted students' scores were significantly higher than those of the controls, and the scores remained high across the grades. Control group means declined across grades with the sixth-grade mean significantly lower than the third-, fourth-, and fifth-grade means. The hypothesis was supported; students who participated in special programs appeared to have higher and more stable academic self-esteem than did their high achieving, non-gifted peers. This difference appeared after the first year of the program. Whether the program produced such differences or whether they were present prior to initial placement is indeterminate.

The fourth hypothesis stated that a difference was expected between the two groups in other specific areas of self-esteem. With the exception of the Scholastic Competence subscale, as noted above, no significant difference existed between the two groups on any subscale. The gifted students appeared much like their high achieving, non-gifted peers. This finding appears intuitively valid since academic placement constituted the unique difference between the two groups. Finally, with respect to the gender difference hypothesis, males (irrespective of program) produced higher self-esteem ratings on Athletic Competence, while females (irrespective of placement) showed higher esteem on Behavioral Conduct.

With regard to recent research using the Harter scales (Chan, 1988; Li, 1988; Schneider et al., 1989), results of this study supported the finding that the gifted have

significantly higher Scholastic Competence scores than do the controls. Unlike the earlier studies, Global and Behavioral subscales did not manifest main effects for program, nor was there a main effect for gender within the gifted sample. This difference in Global and Behavioral subscale results may have resulted from differences in control groups used in the various studies. Prior studies used grade peers with heterogeneous achievement levels; this study used gifted and control groups whose achievement levels were homogeneous and who shared the same teachers for the majority of the school day. Scores of these groups were expected to be more similar than those of groups with fewer commonalities. Unlike the Li (1988) study, the current study found no gender difference on the Scholastic Competence subscale for the gifted or controls. The lower Athletic Competence subscale scores for the gifted group, noted in the Li study, were likewise not replicated. These differences in results may have been due to either cultural differences between the Canadian students used in Li's study and the group used in this study or perhaps due to other factors.

Other recent research (e.g., Chapman & McAlpine, 1987), using a variety of instruments, also found that the gifted sample had a higher global self-esteem than did both the control and norm groups. Since few instruments isolate academic self-esteem, the global score is a composite of the subscales (i.e., an amalgam rather than a separate scale); thus, the higher self-esteem may actually have resulted from the higher academic self-esteem usually noted among gifted individuals rather than a global difference between the gifted and controls. This effect may also have been enhanced through use of a broader, more

heterogeneous control group. More variance in self-concept scores would have been expected within such a group, perhaps resulting in lower group means compared to the means from a more homogeneous control group.

The current study supported Bracken's (1980) results. In the former study, gifted and non-gifted age peers differed less in areas unrelated to academic success (e.g., athletic, social). In the present study, gifted children and their non-gifted, high-achieving peers differed only in the area of scholastic competence, an area relating directly to academic success. The current study replicated Bracken's findings, despite the use of a more stringent test using controls matched for achievement in the present study. Whether the difference in areas relating to academic success was due to placement in a gifted program, to a subtle difference related to intelligence, or to other factors is yet undetermined.

This study confirmed previous research in terms of the effect of identification and placement of the gifted on a specific dimension of self-esteem, i.e., academic, particularly where placement is on a part-time basis. The present investigation is particularly compelling since closely matched high-achieving controls were utilized. Selection and placement occurred early in the educational careers of the gifted students and remained stable across the remaining 3 years. Additionally, the global Self-Worth scale did not produce significant comparisons for gifted versus controls by gender or grade. Previous reports of gender differences on self-esteem measures may have been influenced by behavioral conduct and/or athletic competence self-esteem factors contained, to various degrees, within diverse generic

self-esteem measures and/or by comparisons between gifted and heterogeneous controls which had greater initial aptitude and achievement differences. Differences in control group selection and/or self-esteem measures could also account for the failure to find gender differences in academic or global self-esteem. Such effects may be weak and result when heterogeneous (lower ability) controls are used. Within the gifted group, gender differences have not been uniformly reported. Gender differences on the Behavioral Conduct subscale (females > males) and Athletic Competence subscale (males > females) were expected for all students based on previous findings.

Future research needs to address whether academic self-esteem is boosted as a result of selection for or academic experience in the gifted placement. Self-esteem assessment prior to determination of gifted versus high-achievement status is a necessary additional procedure. The roles of expectancy and labeling may very well continue to be paramount in a student's academic self-perception.

References

- Bachman, J., & O'Malley, P. (1977). Self-esteem in young men: A longitudinal analysis of the impact of educational and occupational attainment. *Journal of Personality and Social Psychology*, *35*, 365-380.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bartell, N., & Reynolds, W. (1986). Depression and self-esteem in aca-

- demically gifted and non-gifted children: A comparison study. Journal of School Psychology, 24, 55-61.
- Biehler, R., & Snowman, J. (1986). Psychology applied to teaching (5th ed.). Boston: Houghton Mifflin.
- Bracken, B. (1980). Comparison of self-attitudes of gifted children and children in a non-gifted normative group. Psychological Reports, 47, 715-718.
- Bransky, T. (1987). Specific program information: A key to attitudes about the gifted education program. Gifted Child Quarterly, 31(1), 20-24.
- Brody, L., & Benbow, C. (1986). Social and emotional adjustment of adolescents extremely talented in verbal or mathematical reasoning. Journal of Youth and Adolescence, 15(1), 1-18.
- Byrne, B. (1984). The general/academic self-concept nomological network: A review of construct validation research. Review of Educational Research, 54, 427-456.
- Caslyn, R., & Kenny, D. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? Journal of Educational Psychology, 69, 136-145.
- Chan, L. (1988). The perceived competence of intellectually talented students. Gifted Child Quarterly, 32(3), 310-314.
- Chapman, J., & McAlpine, D. (1987, April). Academic self-concept of gifted and average achieving students: A two-year study. Paper presented at annual meeting of the American Educational Research Association, Washington, D. C.
- Reston, VA: ERIC Clearinghouse on Handicapped and Gifted Children. (ERIC Document Reproduction Service No. ED 286 300)
- Chovan, W., & Morrison, E. (1984). Correlates of self-concept among variant children. Psychological Reports, 54, 536-538.
- Colangelo, N., & Brower, P. (1987). Labeling gifted youngsters: Long-term impact on families. Gifted Child Quarterly, 31(2), 75-78.
- Colangelo, N., & Kelly, K. (1983). A study of student, parent, and teacher attitudes toward gifted programs and students. Gifted Child Quarterly, 27, 107-110.
- Colangelo, N., Kelly, K., & Schrepfer, R. (1987). A comparison of gifted, general, and special learning needs students on academic and social self-concept. Journal of Counseling and Development, 66, 73-77.
- Coleman, J., & Fults, B. (1982). Self-concept and the gifted classroom: The role of social comparison. Gifted Child Quarterly, 26(3), 116-120.
- Coleman, J., & Fults, B. (1985). Special class placement, level of intelligence, and the self-concepts of gifted children: A social comparison perspective. Remedial and Special Education, 6(1), 7-11.
- Cornell, D. (1983). Gifted children: The impact of positive labeling on the family system. American Journal of Orthopsychiatry, 53(2), 322-333.
- Cramond, B., & Martin, C. (1987). Inservice and preservice teachers' attitudes toward the academically brilliant. Gifted Child Quarterly, 31(1), 15-19.

- Harter, S. (1979). Manual: Perceived Competence Scale for Children. Denver, CO: University of Denver.
- Harter, S. (1982). The Perceived Competence Scale for Children. Child Development, 53, 87-97.
- Harter, S. (1983). Developmental perspectives on the self-system. In E. Hetherington (Ed.), Handbook of child psychology: Vol. 4 (4th Ed., pp. 275-385). New York: John Wiley.
- Harter, S. (1985). Manual for the Self-Perception Profile for Children. Denver, CO: University of Denver.
- Harter, S. (1989). Causes, correlates, and the functional role of global self-worth: A life-span perspective. In J. Kalligian & R. Sternberg (Eds.), Perceptions of competence and incompetence across the life-span (pp. 67-100). New Haven, CT: Yale University Press.
- Harter, S., & Connell, J. (1982). A comparison of alternative models of the relationships between academic achievement and children's perceptions of competence, control, and motivational orientation. In J. Nicholls (Ed.), The development of achievement-related cognitions and behaviors. Greenwich, CT: J. A. I. Press.
- Huryn, J. (1986). Giftedness as deviance: A test of interaction theories. Deviant Behavior, 7, 175-186.
- Karnes, F., & Wherry, J. (1981). Self-concepts of gifted students as measured by the Piers-Harris Children's Self-Concept Scale. Psychological Reports, 49, 903-906.
- Kelly, K., & Colangelo, N. (1984). Academic and social self-concepts of gifted, general, and special students. Exceptional Children, 50(6), 551-554.
- Khatena, J. (1982). Educational psychology of the gifted. New York: John Wiley & Sons.
- Killian, J. (1981). Personality characteristics and attitudes of intellectually gifted secondary students. Dissertation Abstracts International, 42, 2060A. (University Microfilms No. 81-22,024).
- Lehman, E., & Erdwins, C. (1981). The social and emotional adjustment of young, intellectually-gifted children. Gifted Child Quarterly, 25(3), 134-137.
- Li, A. (1988). Self-perception and motivational orientation in gifted children. Roeper Review, 10(3), 175-180.
- Loeb, R., & Jay, G. (1987). Self-concept in gifted children: Differential impact in boys and girls. Gifted Child Quarterly, 31(1), 9-14.
- Maddux, C., Scheibner, L., & Bass, J. (1982). Self-concept and social distance in gifted children. Gifted Child Quarterly, 26(2), 77-81.
- Maker, C. (1977). Providing programs for the gifted handicapped. Reston, VA: Council for Exceptional Children.
- Maker, C. (1982). Curriculum development for the gifted. Rockville, MD: Aspen Systems Corporation.
- McDowell, J. (1984). Coping with social and emotional factors through various strategies: Help for the gifted student. The Creative Child and Adult Quarterly, 9(1), 18-27.

- Morris, S. (1987). Student teachers' attitudes toward gifted students. The Creative Child and Adult Quarterly, 12(2), 112-114.
- Piaget, J. (1981). Intelligence and affectivity: Their relationship during child development. (T. Brown and C. Kaegi, Eds. and Trans.). Palo Alto, CA: Annual Reviews (Original work published 1954).
- Ross, A., & Parker, M. (1980). Academic and social self-concepts of the academically gifted. Exceptional Children, 47(2), 6-10.
- Schneider, B. (1987). The gifted child in peer group perspective. New York: Springer-Verlag.
- Schneider, B., Clegg, M., Byrne, B., Ledingham, J., & Crombie, G. (1989). Social relations of gifted children as a function of age and school program. Journal of Educational Psychology, 81(1), 48-56.
- Whalen, S., & Csikszentmihalyi, M. (1989). A comparison of the self-image of talented teenagers within a normal adolescent population. Journal of Youth and Adolescence, 18(2), 131-146.
- Zaffrann, R., & Colangelo, N. (1979). Counseling with gifted and talented students. In J. Gowan, J. Khatena, & E. P. Torrance (Eds.), Educating the ablest: A book of readings on the education of gifted children (pp. 167-181). New York: F. E. Peacock Publishers, Inc.

Table A-1. Subscale Means and Standard Deviations

Subscale	Grade								
	3		4		5		6		
	G	C	G	C	G	C	G	C	
Scholastic									
Mean	3.40	3.12	3.47	3.36	3.52	3.23	3.21	2.85	
SD	0.62	0.64	0.53	0.60	0.44	0.63	0.45	0.52	
Social									
Mean	2.93	2.99	2.96	3.09	3.09	3.09	2.84	2.79	
SD	0.67	0.64	0.76	0.62	0.65	0.75	0.68	0.70	
Athletic									
Mean	2.89	2.97	2.83	2.97	2.93	3.02	3.02	2.79	
SD	0.68	0.68	0.69	0.69	0.80	0.83	0.75	0.79	
Physical									
Mean	3.18	3.15	2.98	3.06	2.85	2.89	2.59	2.53	
SD	0.65	0.76	0.70	0.65	0.86	0.69	0.75	0.66	
Behavioral									
Mean	3.19	3.22	3.27	2.95	3.00	2.97	2.75	3.13	
SD	0.60	0.62	0.55	0.76	0.77	0.64	0.82	0.62	
Global									
Mean	3.38	3.33	3.35	3.28	3.33	3.24	3.10	3.16	
SD	0.57	0.65	0.63	0.64	0.74	0.74	0.81	0.55	

Note. Subscale scores for: G = Gifted; C = Control

Table A-2. ANOVA Significance by Subscale

Subscale	F	Significance	Subscale	F	Significance
Social Competence			Physical Appearance		
Program	10.76	0.001	Program	0.02	0.899
Grade	4.97	0.002	Grade	7.35	0.000
Gender	0.04	0.757	Gender	2.17	0.143
Social Acceptance			Behavioral Conduct		
Program	0.12	0.734	Program	0.01	0.937
Grade	1.70	0.169	Grade	1.57	0.198
Gender	1.95	0.165	Gender	32.13	0.000
Athletic Competence			Global Self-Worth		
Program	0.14	0.710	Program	0.18	0.673
Grade	0.04	0.990	Grade	1.00	0.392
Gender	11.61	0.001	Gender	.04	0.836