

Fourth Grade Writing Achievement and Instruction, 1974-1984: NAEP's Report Card

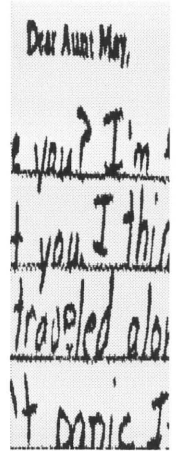
Anne Campbell

CN 6710, Princeton, NJ 08541-6710

Visible Language XXI, 2 (Spring 1987)

Anne Campbell, pp. 93-109

© Visible Language, Rhode Island School of Design
Providence, RI 02903



The National Assessment of Educational Progress reports its trend results for the writing achievement of 9-year-old students from 1974-1984. The results are based on writing exercises which were administered using identical administration procedures in at least two of three assessments. The nationally representative sample of responses was scored using two methods, primary trait scoring (task accomplishment) and holistic scoring (fluency). Three types of tasks were given to the students, persuasive, imaginative, and informative. Performance on the persuasive task was somewhat less successful in 1984 than in 1979. Nine-year-olds showed steady improvement from 1974 to 1984 in their ability to write in response to an imaginative task. Improvement on the informative task was very slight. The results for various subgroups reveal that Hispanics made the greatest improvement.

Adapted from *Writing Trends Across the Decade, 1974-84*
by Arthur N. Applebee, Judith A. Langer, and Ina V. S. Mullis

How well do 9-year-old students write as compared to ten years ago? Do they do better at one kind of writing than another? How well are minority students writing? These are some of the questions addressed in the writing report from the National Assessment of Educational Progress (NAEP).¹

NAEP is an ongoing, congressionally-mandated project established to conduct national surveys of the educational attainments of young Americans. Its primary goal is to determine and report the status and trends over time in educational achievement. NAEP was initiated in 1969 to obtain comprehensive and dependable national educational achievement data in a uniform, scientific manner. Today, NAEP remains the only regularly conducted national survey of educational achievement at the elementary, middle, and high school levels.

Since 1969, NAEP has assessed 9-year-olds, 13-year-olds, and 17-year-olds. In 1983, NAEP began sampling students by grade as well as by age. In addition, NAEP periodically samples young adults. The subject areas assessed have included reading, writing, mathematics, science, and social studies, as well as citizenship, literature, art, music, and career development. Assessments were conducted annually through 1980 and have been conducted biennially since then. All subjects except career development have been reassessed to determine trends in achievement over time. To date, NAEP has assessed approximately 1,300,000 young Americans. The students sampled are selected so that their assessment results may be generalized to the entire national population.

The trend results reported here are based on three assessments of writing achievement, from 1974 to 1984. These assessments took place during the 1973-74, 1978-79, and 1983-84 school years. (For convenience, each will be referred to by the last half of the school year in which it occurred.) The results are based on writing exercises which were administered using identical administration procedures in at least two of the three assessments. Thus, the discussion of trends in writing performance is limited to a rather small data set of three to five writing tasks.

The writing tasks were designed to reflect the differing purposes for which people write at home, at school, and in the community. These purposes have been categorized as informative, persuasive, and

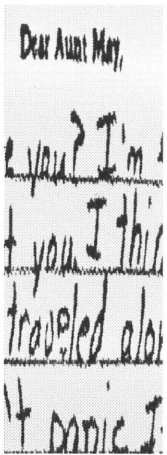
imaginative. Informational writing is used to share knowledge and convey messages, instructions, and ideas. It can involve reporting or retelling events as well as analyzing and examining concepts. Persuasive writing attempts to bring about some change or action by influencing others. It may entail advising, convincing, or refuting. Literary or imaginative writing provides a special way of sharing experiences and understanding the world. It finds expression in a variety of forms, such as stories, poems, plays, or song lyrics. The literary motive is also evident when language breaks conventional patterns.

Scoring **T**he exercises discussed in this report were evaluated using two procedures: primary trait scoring and holistic scoring. For each procedure, raters scored all 9-year-old papers collected from the two or three different assessments at the same time. Also each kind of scoring was done by a different group of raters.

Primary trait scoring (task accomplishment) focuses on how successfully each response accomplishes the rhetorical task specified by the writing prompt. It involves isolating particular features essential to accomplishing the task and then developing criteria for various levels of performance based on those features. Papers were rated against these performance criteria rather than in terms of relative quality within the population sampled.

For the exercises reported here, five levels of proficiency were defined for each task: **unrateable**, **unsatisfactory**, **minimal**, **adequate**, and **elaborated**. **Unrateable** responses included those that were blank, off task, unreadable, or “I don’t know.” Across tasks, **unsatisfactory** responses were those that failed to reflect a basic understanding of the informative, persuasive, or imaginative purpose of the writing. **Minimal** responses recognized the elements needed to complete the task but were not managed well enough to ensure the intended effect of the writing that resulted. **Adequate** responses included those features critical to accomplishing the underlying purpose; responses scored as adequate are likely to have the intended effect. **Elaborated** responses went beyond the merely adequate, reflecting the higher level of coherence and elaboration that is highly desirable, if not absolutely necessary.

The trend items assessed in 1983-84 were also scored holistically to provide an estimate of the overall, relative fluency of the writing. When rating holistically, the readers concentrate on their general impression of a writing sample relative to the other papers they have read. Holistic scoring is sensitive to a range of different skills,



including organization, quality of content, grammar, usage, spelling, punctuation, and the choice of words, but its ratings are based on overall impressions.

Guidelines for the holistic scoring (fluency) were developed by the chief readers and table leaders — all of whom were experienced holistic readers — who began by surveying the pool of papers for each task and selecting examples representing six levels of proficiency for that task (a seventh level was routinely used for blank or unrateable papers). Levels 1-3 were used for bottom-half papers, and levels 4-6 for top-half papers. Chief readers used the sample papers to train readers first to decide whether papers were “top half” or “bottom half” and then to make finer distinctions. In general, holistic scoring produces a roughly normal distribution of scores for the total sample of papers, with scores equally distributed around the center of the scale. The results of this process indicate students’ levels of fluency.

The purpose of NAEP’s holistic evaluation was to detect changes in writing performance for each task. Thus, papers written in response to a particular writing task by 9-year-olds in the two or three assessments of writing including that task were randomly mixed together and rated relative to each other. The differences in performance reported between assessment years are a direct result of that comparative process.

Trends in Persuasive Writing, 1979-1984 **A**t age 9, students were asked to write a persuasive letter to their Aunt May to convince her that they were old enough to travel alone so that they could go to visit her. To accomplish this task effectively, students had to take a stand and support it with some appropriate reasons.

Unsatisfactory. Writers who performed unsatisfactorily on this task failed to take a stand or took a stand but did not support it with any concrete reasons. In the letter in figure 1, for example, the writer offers only the global appeal to “give kids a chance” without providing any reasons to believe that such a chance would be worth taking.

Figure 1. Unsatisfactory Response.

Dear Aunt May,

I know every body say
I to little to travel but
I'm not if you sometimes
give kids a chance to prove
something to you Maybe it
won't be so bad. Maybe
we can help you to
give kids a chance.

Yours truly,
your nice

Minimal. In order to achieve at least a minimal level, the writers needed to clearly take a stand and support it with at least one reason that was appropriate to their point of view. The result could be quite short, as in the letter in figure 2.

Figure 2. Minimal Response.

Dear Aunt May,

I'm old enough to travel alone. Please let me
come. The reason I'm old enough because
I've already traveled alone. Please excuse me
for not writing long
love,

Adequate. Performance at an adequate level required the writer to take a clear stand and support it with a brief argument or at least two appropriate reasons. Responses at this level, if not eloquent, seem to have at least a chance of persuading the reader to agree to the writer's

request, as in the response in figure 3.

Figure 3. Adequate Response.

Dear Aunt May,

I am old enough to travel alone. My mom said I could. So I hope you believe I can go travel alone. I did it before. Don't worry. I am very good when I am alone, I am.

Elaborated. Writers who went beyond the merely adequate offered an extended argument or an interrelated list of reasons to support their stand. The paper in figure 4 was rated as elaborated because of its interrelated list of reasons.

Figure 4. Elaborated Response.

Dear Aunt May,

How are you? I'm fine. I would really like to visit you. I think I'm old enough because I have traveled alone before and because I don't panic. If something bad would happen another reason is I know how to take care of my self. I have been on many other busses and planes alone and I have made it. I would really like to visit you.

Love,

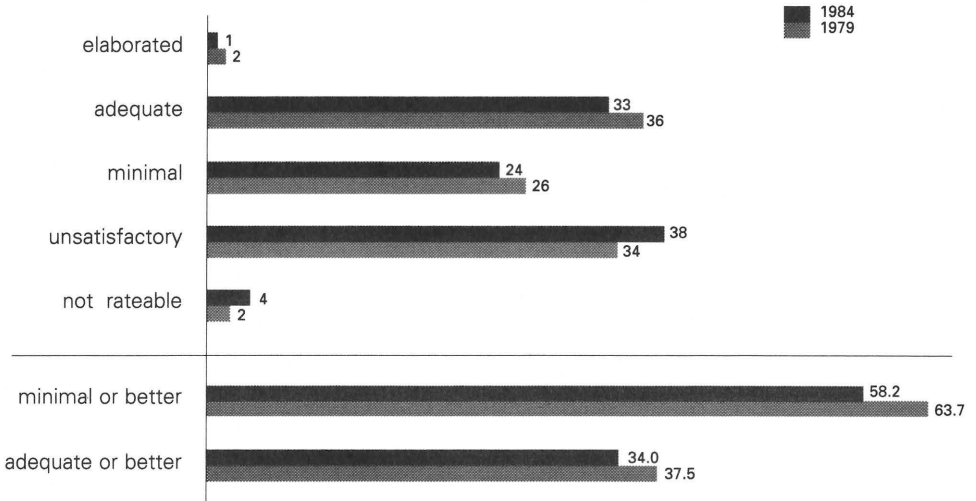
P.S Please tell me the answer

Changes in Persuasive Writing, 1979-1984 In 1984, one-third of the 9-year-olds wrote an adequately supported persuasive letter, and another one-fourth wrote a minimally supported letter that indicated at least an understanding of the task. However, as shown in figure 5 performance on this task was somewhat less successful in 1984 than in 1979: whereas sixty-four percent attained the minimal level or better in 1979, this percentage dropped to fifty-eight percent in 1984. Although these decreases in effectiveness were small, they were accompanied by parallel and significant drops in overall fluency

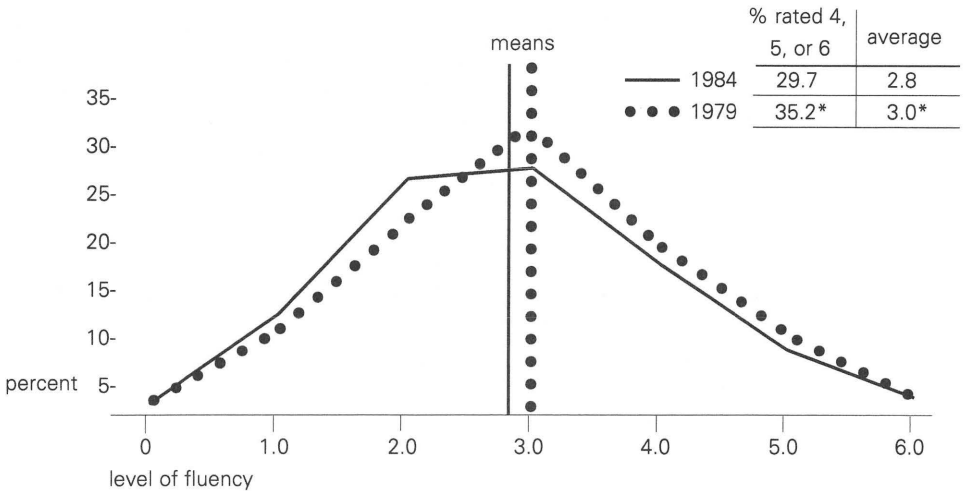
as judged by holistic ratings on this task, both in the percentage of better papers and in the average level of performance.

Figure 5.
Two Views of Persuasive Writing Achievement of 9-Year-Olds, 1979-84.
 (Aunt May: Persuasive Letter)

Percentage of Students at Each Level of Task Accomplishment



Percentage of Students at Each Level of Fluency



*Statistically significant differences from 1984 at the .05 level

Trends in Imaginative Writing, 1974-1984 **T**he exercise used to measure trends in performance in imaginative writing since 1974 was based on a picture of a box with a hole in it and an eye peeking through the opening. The prompt asked students to imagine themselves in the picture and then to describe the scene and how they felt about what was going on around them. They were encouraged to make their description “lively and interesting.”

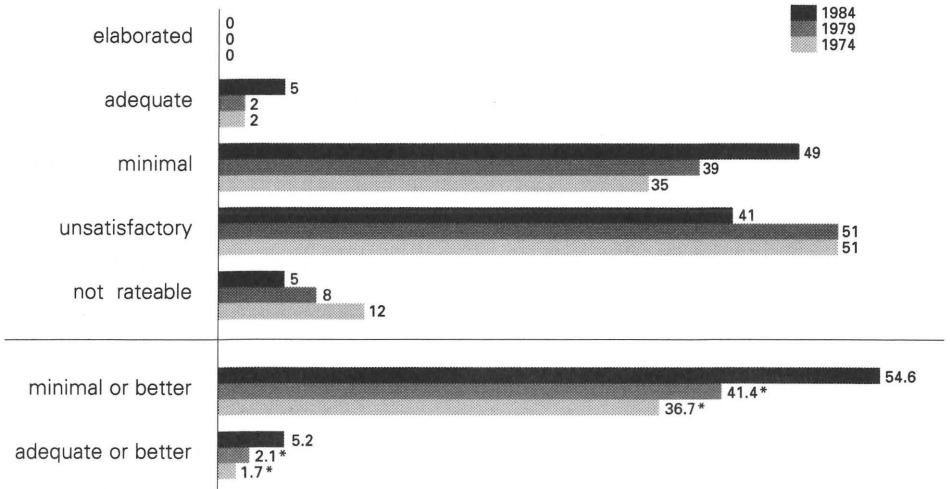
In order to accomplish their task successfully, students had to suggest the situation or imaginative world in which they found themselves and to reveal their attitudes through description of that situation. Writers who provided **unsatisfactory** papers left the situation undefined; if they mentioned attitudes toward this situation at all, these attitudes were unmotivated and unjustified. Writers who were rated **minimally** successful in accomplishing this task defined the situation or presented feelings and reactions but did not do both. They reflected a grasp of the imaginative character of the task but were unable to carry through with it. Writers whose performance was **adequate** defined a clear situation and provided a clear expression of attitudes and feelings. At this level of performance, their attitudes and feelings were presented in a fashion that was consistent and appropriate to the situation, although redundancy, vagueness, or abruptness may have been present. Writers who were most successful at this task developed a clear consistent situation and **elaborated** on the attitudes and feelings that were aroused within it.

The Hole in the Box task was administered in 1974, 1979, and 1984. Trends in performance for 9-year-olds are displayed in figure 6. Across that period, 9-year-olds showed a steady improvement in their ability to enter into and elaborate upon the imaginary situation. In 1974, only thirty-seven percent of the responses were rated as minimal or better; this rose to forty-one percent in 1979 and fifty-five percent by 1984. Even in 1984, however, only five percent of the responses were rated as adequate and essentially none as elaborated.

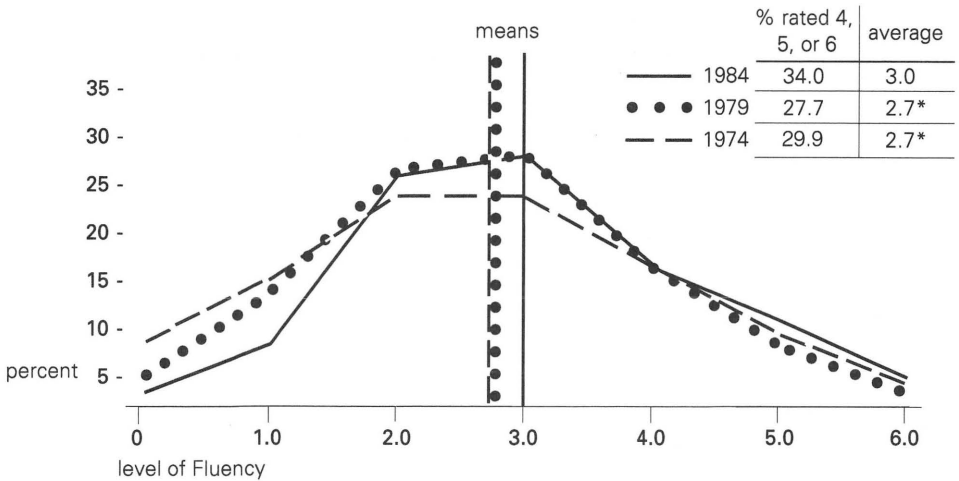
When measured by holistic ratings, overall fluency in response to this task also increased between 1974 and 1984, although the increases were less consistent and less dramatic. Mean scores (on a scale of 0 to 6) were 2.7 in 1974 and 1979, rising to 3.0 in 1984. Significant improvements were concentrated among the very poorest scores: the percentages of papers at the low end of the distribution fell from twenty-three percent in 1974 to twelve percent in 1984.

Figure 6.
Two Views of Imaginative Writing Achievement of 9-Year-Olds, 1974-84.
 (Hole in the Box: Imaginative Description)

Percentage of Students at Each Level of Task Accomplishment



Percentage of Students at Each Level of Fluency



*Statistically significant differences from 1984 at the .05 level. No significance test is reported when the proportion of students is either >95% or <5%.

Trends in Informative Writing, 1979-1984 **T**he informative writing task presented students with a reprint of a surrealist painting by Salvadore Dali and asked them to describe it for a friend who had never seen it so that the friend could visualize the picture.

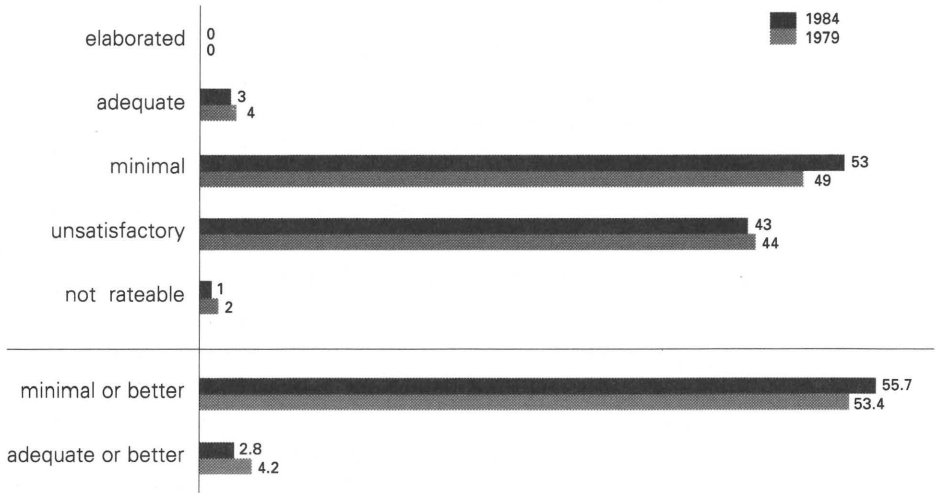
In order to accomplish this writing task successfully, students needed to select, organize, and present the details of the painting and to convey them in terms of the whole painting. Writers who responded **unsatisfactorily** to this task provided only the barest information, misinformation, or disjointed details so that the information did not fit. At the **minimal** level of performance, writers provided some details but in unrelated ways. They created no organizational framework for the reader to use to visualize how the various parts of the picture might fit together. To perform at the **adequate** level, writers needed to describe and interrelate most of the details in the picture and to present the details within an organizing framework. Writers presented **elaborated** papers when they wrote an extended description within a cohesive framework — spatial, formal, thematic, metaphorical, or narrative — to provide a context for the reader.

As figure 7 indicates, in 1984 over half of the 9-year-olds wrote minimal descriptions of the painting, with only three percent writing adequate papers. Although students wrote more papers that were at least at the minimal level in 1984 than in 1979, the changes in students' levels of performance were slight. No students in this younger age group presented an elaborated description.

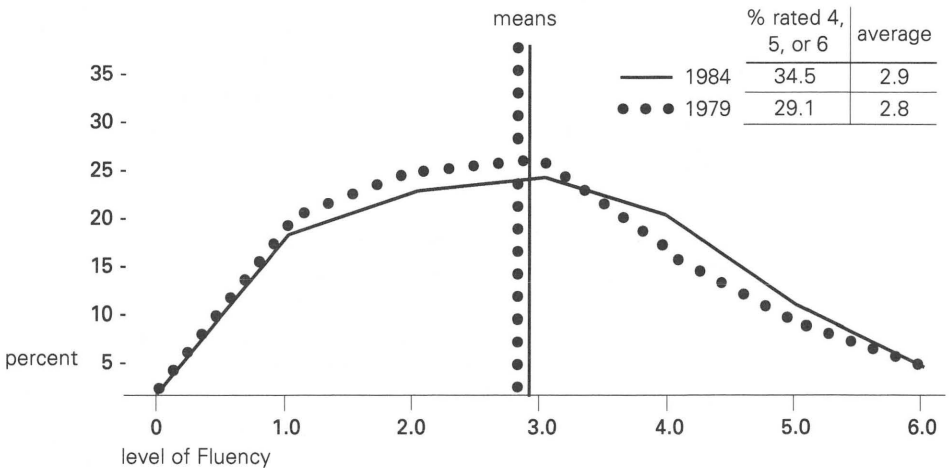
There was also a slight improvement in overall fluency as measured by holistic ratings on this task. The distribution shifted slightly up the scale, with five percent more students writing better papers (4, 5, or 6 on the fluency scale) in 1984 as compared with 1979. However, as with the results for level of task accomplishment, changes in fluency were relatively slight.

Figure 7.
Two Views of Informative Writing Achievement of 9-Year-Olds, 1979-84.
 (Dali: Descriptive Report)

Percentage of Students at Each Level of Task Accomplishment



Percentage of Students at Each Level of Fluency



No statistically significant difference from 1984 at the .05 level.

Summary of Trends in Writing Achievement Across the Decade, 1974-1984

To provide a sense of the kinds of changes that have taken place in 9-year-old students' writing achievement, the discussion will turn to the results for writing tasks that were included on more than one assessment. This includes results for tasks already discussed.

In order to report trends in writing achievement as accurately as possible, the data presented are confined to instances where: 1) the identical writing task was administered to the same age level in two or three assessments; 2) the task was administered in each assessment in the same way (using a paced audiotaped procedure in which each task was read to the students); and 3) responses collected in more than one assessment were evaluated at the same time by the same readers, using both the primary trait and holistic scoring methods.

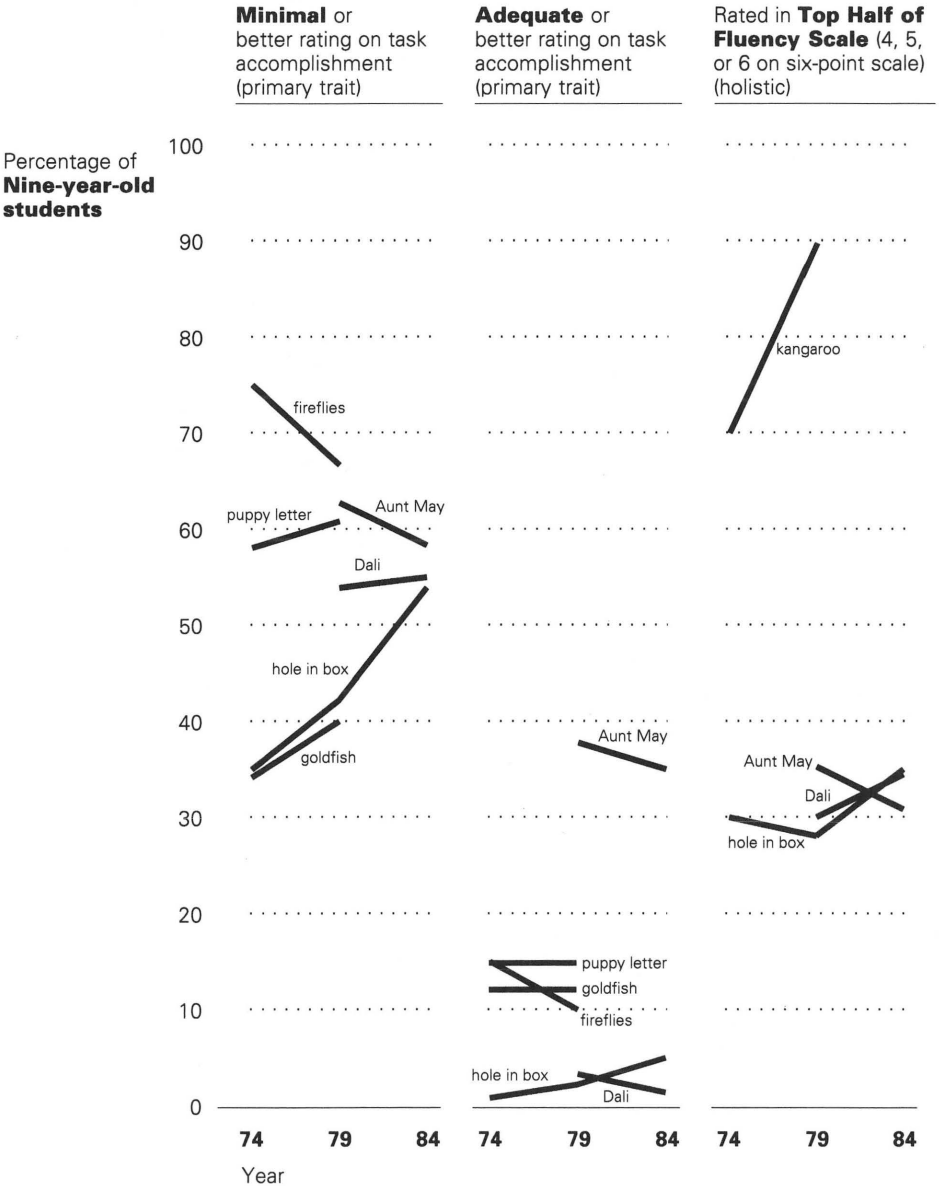
At age 9, four procedurally identical tasks were used to measure changes in writing achievement from 1974 to 1979; three procedurally identical tasks were used to measure changes from 1979 to 1984. One of these tasks was also included in the 1974 assessment and provides a direct link from assessment to assessment over the decade. These tasks represent the three types of writing assessed, informative, persuasive, and imaginative. For both figures 8 and 9, the informative task is "Dali." The persuasive tasks include "Puppy Letter," which asked students to write a letter convincing the landlord that they should get to keep a puppy, and "Aunt May." The imaginative tasks are "Fireflies," which directed students to write a story about a picture of a girl trying to catch fireflies; "Goldfish," which asked students to imagine what it would be like to be something besides a person, such as a goldfish, airplane, horse, or a tree; "Kangaroo," which had students write about a picture of a kangaroo jumping over a fence; and "Hole in the Box."

Figure 8 summarizes the trends of 9-year-old students from 1974-1984 on informative, persuasive, and imaginative tasks which were rated for both task accomplishment and writing fluency. The first set of results (Panel 1) shows trends in the percentage of papers at the minimal level or better in task accomplishment; this percentage includes all students who wrote minimal, adequate, and elaborated levels combined. The second representation (Panel 2) shows trends in the percentage of students achieving at the adequate level or better; it is the total percentage writing at the adequate and elaborated levels combined. Thus the first view depicts students' progress in moving from unsatisfactory performance to at least some minimal or basic level of performance. The second view depicts progress toward

responses rated at the adequate level or better.

The third view of trends in writing achievement presented for age 9 reflects the proportion of students in each assessment rated at the three highest levels on the fluency scale. Panel 3 shows global changes in writing performance from assessment year to assessment year.

Figure 8.



The 9-year-olds showed different trends in achievement on different writing tasks across the last decade. During the first five years of the decade (1974-79), the proportion of students reaching at least the minimal level showed slight increases on three of the writing tasks (two imaginative and one persuasive), but a decrease on a third imaginative task. Proportions of adequate or better performance on the same tasks remained relatively stable, but with a decline on the same imaginative task that showed the decrease at the minimal level. Fluency scores for the two tasks for which data are available for 1974 to 1979 moved in opposite directions.

During the second five-year period (1979 to 1984), there was a sharp increase in minimal performance on the imaginative task (describe an imaginary situation based on a picture), accompanied by a slight increase at the adequate level on the same task. The informative task (describe a painting) introduced in 1979 showed a slight increase at the minimal level accompanied by a very slight decrease at the adequate level. For persuasive writing, there was a slight increase at the minimal level from 1974 to 1979, followed by a decrease from 1979 to 1984 in performance on the persuasive task introduced in the 1979 assessment.

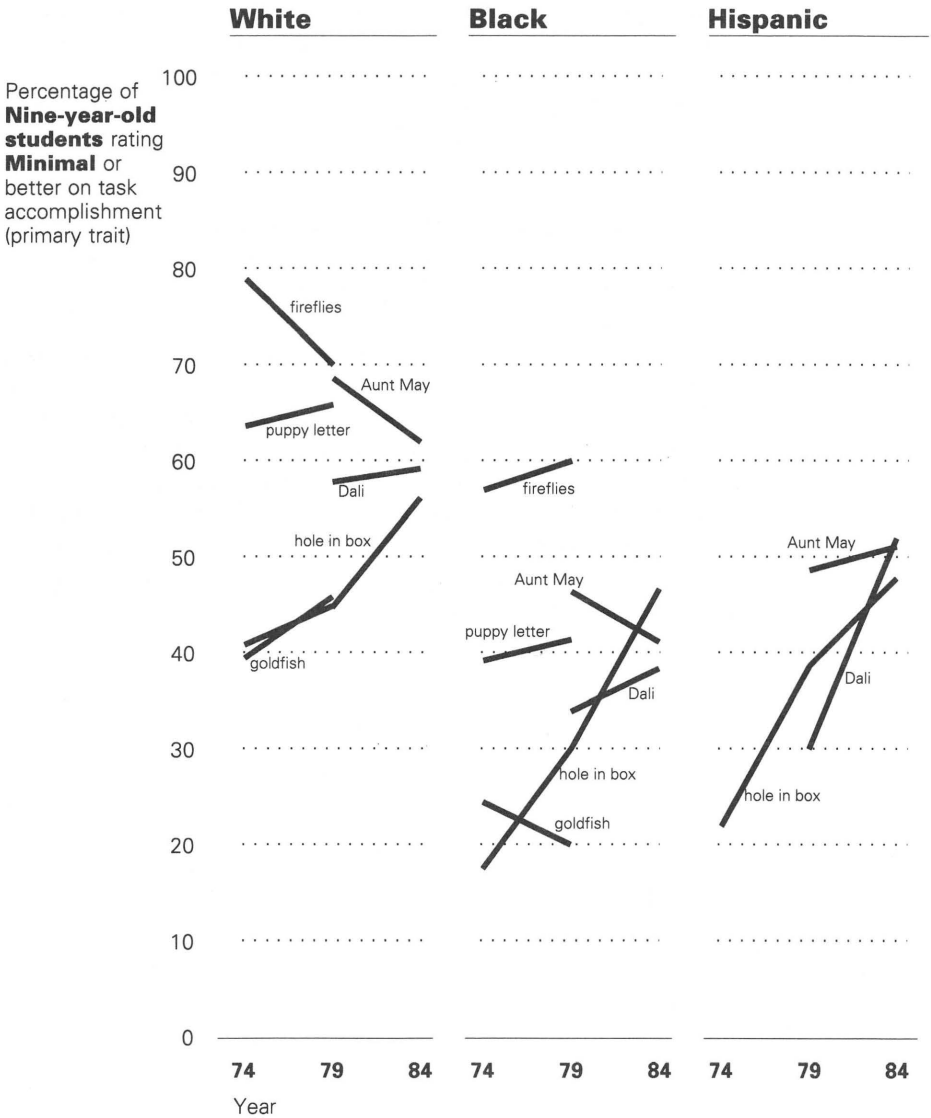
On balance, the writing performance of 9-year-olds was relatively stable from 1974 to 1979. Performance decreased on the persuasive task from 1979 to 1984, while informative writing skills remained about the same. Imaginative writing performance improved during that same period.

Subgroups **A**n understanding of the state of writing achievement of 9-year-olds is incomplete without attention to the diverse subgroups that comprise the nation. Do trends in the writing performance for particular subgroups parallel or help explain trends for the nation as a whole?

Performance of Black, Hispanic, and White Students All three writing assessments have examined the performance of Black students; for trend results on items included in the 1984 assessment, results are also available for Hispanic populations. For each assessment, the performance of the age group has been compared with that of their White age-mates.

Figure 9 depicts trends in writing performance for Black, Hispanic, and White students. The data shown are for percentages of students writing papers rated as minimal or better on task accomplishment for informative, persuasive, and imaginative tasks. As with the national

Figure 9.



results, the percentage of students writing papers rated as adequate or better was substantially lower for each of the three subgroups on each task.

Great differences appeared at age nine. White 9-year-old students showed mixed patterns of performance over time, whereas the per-

formance of Hispanic 9-year-olds improved on all three writing tasks analyzed in 1984. In fact, in 1984 only the Hispanic 9-year-olds showed improvement on the persuasive tasks. Black 9-year-olds also showed comparatively more improvement than their White age-mates from 1974 to 1979, but trends in the achievement of these two groups were very similar from 1979 to 1984.

Reflections on Trends in Writing Achievement

In summary, between 1974 and 1984, improvements at age 9 occurred on the imaginative writing task, although proficiency on informative and persuasive tasks remained relatively stable or even declined. This pattern may reflect the instructional emphasis in elementary schools: it may be that younger children are more likely to be asked to write creatively. However, this emphasis is too limited; elementary school children need to be engaged in informative and persuasive writing tasks appropriate to their levels of knowledge and interests. To move beyond the current levels of achievement, a more systematic program may be needed -- one focused more directly on the variety of different kinds of writing students need to learn to do and spanning a wider range of levels of complexity.