DOCUMENT RESUME

ED 287 458 IR 012 843

AUTHOR	Becker, Henry Jay
TITLE	The Impact of Computer Use on Children's Learning: What Research Has Shown and What It Has Not.
PUB DATE	23 Apr 87
NOTE	88p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, DC, April 20-24, 1987).
PUB TYPE	Information Analyses (070) -- Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Computer Assisted Instruction; Elementary Secondary Education; *Instructional Effectiveness; Literature Reviews; *Meta Analysis; Microcomputers; Models; National Surveys; *Research Design; *Research Needs; *Research Problems; Student Attitudes

ABSTRACT

This examination of the research on computer-based instruction for the major academic subjects in the elementary and secondary grades considers three varieties of empirical data--surveys of computer-using teachers, school district program evaluations, and experimental field studies. A discussion of the findings of two national surveys of computer-using educators conducted in 1983 and 1985 is followed by a review of research on the effectiveness of computer-assisted instruction. The methodologies used and the major conclusions reached by reviews and meta-analyses of research conducted since the mid-1960s are then discussed, followed by a review of more recent studies of the impact of computer-based approaches on children's learning. A best-evidence synthesis approach is used to provide an overview of the type of research being done and individual studies are discussed in two categories--those using randomized assignment of classes to treatment groups, and those using comparison groups but without randomized assignment. The implications of these studies for further research in this area are then considered, and a model for field experiments is proposed. The results of the 1985 national survey of computer-using educators are presented in 28 figures and tables, including a capsule summary. (MES)

The Impact of Computer Use on Children's Learning: What Research Has Shown and What It Has Not

Henry Jay Becker

Center for Research on Elementary and Middle Schools

The Johns Hopkins University

April, 1987

Paper presented at the Annual Meetings of the American Educational Research Association, Washington, D.C., April 23, 1987.

We have all read the ads for the products of educational software publishers:

"Announcing a breakthrough program
for helping students learn
to read and write...."

"I 'aced' the test!...Thanks to my teacher and Skills Bank."

"Science teachers: we have the Apple program for you."

Should schools use computers and computer programs like these to help students learn traditional academic content? For which students and subjects are computers a helpful learning medium, and which instructional approaches using what software provide appropriate learning activities and content to meet existing curricular goals? With schools spending perhaps one-fifth as much on computers and their associated software, training, and extra staffing as they do on all books and other instructional materials combined, these become important questions.

Research about the instructional effectiveness of computer-based programs has been going on for more than two decades. Yet, on what evidence do we base our recommendations to invest or not to invest in computer-related approaches to providing instruction? What can we say that "research has shown"? This paper will examine some of the research on computer-based instruction for the major academic subjects in the elementary and secondary grades. It considers three varieties of empirical data -- surveys of computer-using teachers, school district program evaluations, and experimental field studies. We review particular studies and their findings primarily to assess the value of research in the field in general. In the last section, we propose a model for expanding the empirical basis of our conclusions about the impact of computer approaches on academic achievement.

Survey Data

The least costly and least complex way to find out how computers have affected learning is to ask computer-using teachers what they believe the impact of computers has been in their school. On two occasions, in 1983 and 1985, I did that -- fielding two surveys in more than 4,000 schools nationwide. The surveys could not test improvements in students' actual achievements. Nor could we compare students' achievements in classes using computers with those in similar classes providing traditional instruction on the same content. But we could measure teachers' perceptions of the effects of the use of computers at their own school and in their own classes.

The validity of these teachers' judgments varies -- some erring toward over-skepticism; others, toward over-enthusiasm. But by looking carefully at the data (for example, by paying more attention to relative perceptions across different outcomes) we can extract clues about how learning in typical classroom settings has been most affected by computers and how it has been least affected.

In the most recent survey, done two years ago and involving roughly 8,000 teachers and principals, computer-using educators saw significant benefits occurring mainly in four areas: student motivation, student cooperation and independence, opportunities for high-ability students in programming activities and in other higher-order thinking and writing skills, and oppportunities for low-ability students to master basic math and language arts skills. Nevertheless, the promise of computers was not yet seen as being fulfilled for average students, for diagnosing specific instructional needs, for systematic individualization of instruction, nor for improving most learning of facts and concepts.

Specifically, the teachers saw computers as primarily helping students to enjoy their school experience more and motivating them to pay attention to academic work. Four out of ten believed that student enthusiasm in school subjects for which they used computers was "much improved" because of computers. Teachers and principals also saw computers giving opportunities to special populations -- gifted students and learning disabled. But they were much less sanguine about actual learning of traditional academic content. Only 7 percent of the schools' primary computer-using teachers believed that learning in regular school subjects by "average" students was "much improved" because of computers.

Not surprisingly, learning outcomes were viewed as having been more greatly affected in schools with more computers, although the correlations were quite modest, in the .10 to .20 range, probably because of the unreliability of the single-item measures of perceived outcome. In addition, there were correlations of similar magnitudes between how computers were used at the school and the teacher's perceptions of the computer's impact on different aspects of student learning. A higher proportion of school computer time spent in word-processing activities was associated with more perceived learning by above-average and average high school and middle school students. A higher proportion of time spent in drill-and-practice activities was associated with an increased perception that average students in the middle grades and below-average students at all grade levels were learning more. On the other hand, the proportion of school computer time being spent on computer programming was not positively associated with any perceived learning outcomes that were measured (Becker, 1986b).

But what are we to make of these teacher and principal perceptions? The large and representative sample sizes -- essentially hundreds of times the number of schools, teachers, and students on whom data is gathered in typical evaluation reports or field experiments -- lend plausibility to the findings. External validity is substantial. And even with imprecise measurement and individual differences in objectivity, comparisons among types of schools (e.g., according to how they use computers), as well as comparisons of different perceived outcomes, are likely to provide fair and valid data about relative differences in impact across schools or outcomes. But because surveys are such high-inference measures of behavior -- let alone of cause and effect relationships ("...how much impact have computers had...?") -- even unbiased perceptions at each site produce limited evidence about actual effects and effectiveness of instructional programs.

To more accurately assess these effects, we must turn to more rigorous methodologies. In particular, we need to examine studies that contrast the achievements of students who have been using computers with those of students who have not. We have dozens of such studies -- evaluations of specific instructional programs undertaken by program developers themselves, school district evaluations of instructional technology programs implemented on a grand scale throughout the district, and doctoral dissertations reporting specific experiments, generally in a few classrooms, but often in fairly typical school settings. The next two sections of this paper address the question, "What do these program evaluations, dissertations, and other studies tell us about how computer use is affecting children's learning?"

Effectiveness of Computer-Assisted Instruction: Research Reviews

Research about the effectiveness of using computers in instruction has gone on since the mid-1960's, and attempts to abstract the major conclusions from these studies began in the early 1970's (Visonhaler and Bass, 1972; Jamison, Suppes, and Wells, 1974; Edwards, et al., 1975; Hartley, 1977; Thomas, 1979; Burns and Bozeman, 1981). Lately, reviews have primarily been reviews of earlier reviews on the subject rather than reviews of more recent original investigations (e.g., Hasselbring, 1984; Okey, 1985; Stennett, 1985). However, two groups of researchers have conducted a series of meta-analyses, attempting to statistically summarize central tendencies by actually examining the original program evaluations and experiments that most reviewers have not looked at for years (Kulik, Bangert, and Williams, 1983; Kulik, J., Kulik, C-L, and Bangert-Drowns, 1985; Bangert-Drowns, Kulik, J., and Kulik, C-L, 1985; Niemiec and Walberg, 1985).

Most reviews and meta-analyses have concluded that computer-assisted instructional programs studied in the 1960s and 1970s were generally more effective in raising students' scores on standardized achievement tests than alternative approaches. Less systematic computer activities (e.g., simulations, problem-solving tasks) aimed at affecting higher order thinking skills have had a more mixed evaluation in these reviews.

However, an absolutely remarkable fact about this entire literature is that 99 percent of the studies that it draws upon employed computer systems, software, and social organization of computer implementations that characterize, at most, one or two percent of all of the instruction on traditional academic subject-matter using school computers today.

The most obvious distinction between the instructional environments in which this research was undertaken and the environments in which computers are used for school instruction today is in terms of the hardware in use. Microcomputers -- stand-alone desk-top computers --first began entering substantial numbers of schools in 1982. Between 1982 and 1983, the proportion of schools with microcomputers doubled from about 25% to slightly over 50% (Becker, 1983, calculated from Fig. 1). By 1985, the proportion of computers in use in schools that were NOT microcomputers was down to 4%, and two-thirds of those were at high schools where computer programming activities predominated. Yet all but two of the more than 200 studies contained in the Kulik, et al. and Niemiec and Walberg meta-analyses were published prior to 1983. Even more significantly, studies involving microcomputers constitute only one out of 64 studies in the two most recent Kulik, et al. reviews and only two out of 224 studies in the Niemiec and Walberg meta-analysis.

It is true, at a general level, that much of the instructional use on school microcomputers resembles the routine drill-and-practice that characterizes computer activity on the centralized, multi-terminal systems on which the early research had focused. As of 1985, roughly one-half of all activities on school computers, other than computer programming, involved drill-and-practice and tutorial computer-assisted instruction in math, language arts, and other subjects (Becker, 1986a). However, schools employ a much greater variety of computer activities than in the days before microcomputers --for example, word processing occupies at least 15 percent of all student computer time -- and the entire instructional climate is likely to be vastly different even for the routine C.A.I. undertaken on school microcomputers.

Schools' acquisition of microcomputers and the organization of their use are vastly different from the patterns by which centralized computer systems and complete packages of computer-assisted instruction were bought by school systems and evaluated in the studies which comprise the vast majority of those reviewed to date. With microcomputers, decisions about when students should use computers, how to link their use with other instructional activities, which software to use, and how to coordinate use of computers among different classes and among students within a class have been teacher decisions and school-level decisions, with the outside intervention of no more than a single district computer-coordinator or subject-matter supervisor. In contrast, the earlier implementations were central office administrative decisions, and the protocols for using software were part of the package that the school or school system bought into.

Features of the hardware also are likely to affect climate, and perhaps student learning as well, in different ways. For example, most microcomputer implementations involve the "floppy shuffle." In only a small minority of instructional use (under 20 percent) are microcomputers linked to a central source for supplying instructional programs -- a "network." In the vast majority of cases, each student's computer must be "loaded" with the appropriate software. And most comprehensive instructional packages (e.g., a mathematics series) involve many different diskettes, each of which has to be appropriately selected and loaded. The use of individually selected media for each student station is one example of how the mechanics of using computers for instruction is very different with microcomputers than with a single, centralized computing facility. In this one respect, microcomputers may be a step backwards from the older mini-computer-based systems, and may produce less effective results.

In other ways, though, microcomputer-based software and systems may result in more favorable outcomes than those observed in older studies. Schools may now have a much greater variety of software available than in cases where they bought into a single instructional package. They certainly have hundreds of times as many products to select from. Instructional quality may vary, but it is likely to have varied as well among the earlier large computer-based systems, and the wider market is likely to have a beneficial impact on instructional quality over the long run. In addition, the motivational characteristics of the microcomputers, with their much more sophisticated and colorful graphics, may have important consequences for student learning and performance.

In addition to being outdated, the existing literature about the effectiveness of computer-based instruction suffers from methodological problems. The earliest research, reviewed, for example, in Vinsonhaler and Bass (1972), explicitly dealt with instructional programs that were SUPPLEMENTARY to the instruction and practice afforded to students in the control treatments. Thus, allocated time was not held constant between students practicing skills on computers and students not using computers. One instance of unequal time allocation in these early studies cited by Clark (1985) was that of Suppes' experiments in the 1960's. Schools where students were provided 10 minutes additional practice time using computer-assisted drills scored significantly higher on post-tests than a control not given the extra time. However, in one school the principal saw to it that control classes had an additional 25 minutes of practice time using traditional workbooks, and in that experiment, the control students outscored the students in the computer-based treatment (Suppes and Morningstar, 1969).

On the other hand, Kulik, Kulik, and Bangert-Drowns (1985) found no difference in effectiveness between 16 studies of elementary school computer use where the computer treatment provided extra time and 12 studies where it did not. If computer-based programs are effective uses of time, it seems reasonable to expect that they would be even more effective when they represent an increase in the time devoted to a subject-matter over a traditional approach. Perhaps other factors were artificially raising the effects found for ALL of the studies, regardless of what specific implementation was occurring.

Clark, re-analyzing a sample of 42 of the studies analyzed by the Kulik group, found problems with three-quarters of the research designs of the studies included in their meta-analyses (Clark, 1985). In addition to unequal-time, Clark identified several other problems: -- disproportionate attrition from the experimental group, non-random assignment of students to treatment, absence of comparable instructional content supplied to the students in the control treatment, and possible teacher quality differences between computer and control treatments. The Kuliks/Bangert-Drowns 1985 papers, however, recognized problems of implementations such as those cited by Clark but approached the issue by calculating measures of effectiveness ("effect scores") under varying conditions -- for example, where the same instructor taught both the computer and the control classes, and where the instructors were different. Kuliks/Bangert-Drowns generally find small differences in effect sizes between studies conducted under different conditions, but Clark's re-analysis finds much more substantial differences using a sample from the same data base. Clark, for example, suggests that the effect of using computers is reduced to roughly 25% its original size when only studies using the same teacher for both computer and control treatment groups are considered.

Certainly, it is worth someone's time to clarify the results of these early C.A.I. studies done on terminal-using centralized computer systems. But, given the predominant attention that schools are giving to microcomputer-based approaches, it would seem more important to evaluate evidence from more recent studies, particularly those using microcomputers and microcomputer software. Have there been such studies, and what conclusions can we draw from them? The following section reports an initial effort to examine this emerging literature.

Effectiveness of Computer Approaches to Instruction: Recent Studies

Meta-analytic procedures have become the most common method for summarizing findings from existing research on the impact of computer-based approaches on student learning. In meta-analysis, all identified studies meeting some often vaguely described minimum criteria contribute equally to a common statistical measure of the differential effectiveness between two approaches to instruction (e.g., computer-based and 'traditional'). Meta-analysis has been criticised in general (Slavin, 1984) for providing a distorted synthesis of research findings because the least adequate research designs are often systematically biased in the direction of finding favorable effects for one of the contrasting treatments.

An alternative approach, called the "best-evidence synthesis," uses as a part of its analysis the "effect size" statistic that meta-analysis uses to link studies otherwise reporting non-comparable statistical measures. But best-evidence synthesis differentiates among studies meeting common minimal "search" criteria -- and pays correspondingly more attention to studies with the highest quality research designs and most substantively relevant (Slavin, 1986).

This paper reports the beginnings of a best-evidence synthesis on the effects on children's learning of computer-based instructional programs. A search was begun for empirical research about the effects of computer-based approaches on basic curricular categories of learning (math, language arts, writing, science, etc.) in grades 1 through 12. We limited the search to reports produced since 1984 and including achievement measures as outcomes. Studies using mini-computers were included as well as those using microcomputers, mainly to see whether they still dominated research reports. Requests went out to school district research personnel in over 100 districts for studies done in their districts; the ERIC/Education Index on-line data base was searched as was Dissertation Abstracts. References and personal contacts led us to other sources. However, not all items meeting our initial criteria were obtained, as it was clear that more rigorous criteria were necessary.

The 51 reports that were obtained included 11 dissertations, 13 reports of school district evalutions, 15 published articles, and 12 unpublished papers. Forty-five were done using microcomputers, and six using mini-computer systems. Of the 51 studies, 11 had no comparison group, and measures of effects were limited to over-time gains on standardized achievement tests. Such reports are not necessarily of no value, and we have put these aside for further scrutiny, but for the present we narrowed our search to research designs that contained comparison groups.

Of the remaining 40 studies, 8 were excluded from consideration because the studies did not employ pre-test controls in their analysis AND neither classes nor students were randomly assigned to computer vs. traditional instruction. Lacking both pre-tests and random assignment, it was impossible to equate the two comparison groups in any way.

Seven studies were removed because the treatment period was shorter than eight weeks, and the study therefore measured the effectiveness of a very limited intervention. Finally, we chose not to consider 8 studies involving fewer than 40 children or where each treatment involved only a single class of students and the experimental and control classes were taught by different teachers. Extremely small samples require massive effects to be statistically significant, and, when significance limits are reached, there is a strong likelihood that factors other than the explicit treatment variable are responsible. For example, if two teachers with differentially effective teaching styles or teaching skills each teach only one class in a study, their "teacher effect" would be confounded with the method -- computer- or traditional-instruction -- that they used. Even having two or three teachers per method only allows us to draw tentative conclusions about method effects.

What about the major element of experimental design -- random assignment? Random assignment of individuals to treatments is of course difficult to accomplish in a school situation. Indeed, of the 17 remaining studies, only one (Hawley, Fletcher, and Piele, 1985, described in Table 1) accomplished this true experimental design. Six more (also described in Table 1) were able to randomly assign intact classes of students to alternative computer- and non-computer experiences. Even that level of randomization would be powerful if sample sizes were large -- say a dozen classes in each treatment. But the largest study of the seven that did any randomized assignment at all (Reid, 1986) involved only 6 classes per treatment, and those were spread among grades 1 to 6 and apparently involved a control group that may not have been appropriate for a comparison of "computer" vs. "non-computer" approaches. Only one study with a class-level randomized control group had as many as four classes at the same grade level using computers for subject-matter learning (Turner, 1985).

In the absence of random assignment, researchers employ procedures for identifying a "matched sample" of students or classes in the same or other schools. Some of these non-random comparison groups seemed more appropriate than others. For example, in four of the ten reports, matching (or luck) produced comparison groups whose pre-tests were quite close to those of the computer treatment group (within .25 standard deviations). But in other cases, either no data was presented by which one could judge the pre-test achievement differential or pre-test means differed by enough that even equating by covariance analysis may be insufficient. The "better" non-random pre- and post-test studies are listed in Table 2. Those with more problems in design or presentation are listed in Table 3.

Effect sizes shown in the tables and discussed in the text below were calculated, where possible, from pre- and post-test means and control group post-test standard deviations reported by the researchers. The effect size indicates the differential gain on the achievement criterion between the experimental group and the control group, expressed as a proportion of the control group's standard deviation. Thus, it is a measure of association between treatment and outcome. Only one study that was reviewed included calculations of effect size. Studies instead focused on tests of statistical significance, which are largely irrelevant for large studies and which can identify only huge treatment effects in small studies. Effect sizes of .4 or .5 are substantial, often indicating a doubling of achievement gains between the two groups. Even effect sizes of .2 or .3 are certainly valuable, if consistently realized. But to achieve statistical significance for an effect size of .2, one requires more than 150 students (5 to 7 classes) per treatment; for an effect size of .4, at least 50 students (2 classes) per treatment. And that ignores the implications of clustered sampling of intact classes, which would raise the minimum sample sizes still further. So it is unfortunate that most studies focus on statistical significance, and sometimes make it difficult to obtain accurate effect size estimates.

Studies with Randomized Assignment

As noted above, only one study, by Hawley, Fletcher, and Piele (1986), was identified where students were randomly assigned to classes which were in turn randomly assigned to treatments. In many other ways, this study is representative of the non-district wide, microcomputer-based studies in Tables 1 through 3. It focused on mathematics instruction in the elementary grades (grade 3 and 5). It used a popular off-the-shelf software package (Milliken Publishing's Math Sequences) but not an externally orchestrated instructional program. The period of instruction between pre- and post-tests occupied between one semester and one full school year (4 months, in this case). And it was quite a small study -- only one classroom per grade level per treatment and four classes in all, all at one school.

Interestingly, the study occurred in Canada, in a small Saskatchewan town, among a very modal student population. Students in the computer class used computer drills for roughly 50 minutes per week (in 12 minute sessions). They also did paper-and-pencil math drills, but only for two-thirds the time of the control group classes. Both computer and control classes had the same amount of math instruction and in this study, as in only one or two others, the same teacher taught mathematics to both the computer and control classes at the same grade level. Six computers were available for each computer-using class -- a 3:1 ratio of students to computers -- and the computers were located in the classroom. Students took turns using computers individually.

In all, the study represents what might be a fairly typical instance of using computers for elementary school mathematics, and because it is well-designed, except for its small size, its results could have generalizability to many other school situations. What did the study find? Both third and fifth grade computer classes made greater gains on a standardized achievement test (a Canadian adaptation of the Iowa Test of Basic Skills) than did control classes. The effects for third grade students were largest on the computation sub-scale and the effects for fifth grade students were largest on the problem-solving and math concepts sub-scales. But effect sizes were positive for all three sub-scales for both grade levels, although each of the three sub-scales had statistically significant differences in gain scores for one but not both grade levels. The overall effect size for the third grade was +.48 and for the fifth grade, +.31, indicating gains over a four-month period of roughly one-half and one-third of a standard deviation more than the control group's gains. This study also evaluated cost differentials and found the program to be cost-effective as well.

Of the six studies randomly assigning treatments to intact classes, the one involving the most students was a comprehensive typing, reading, and language arts program that has National Diffusion Network approval (Reid, 1986). This study, done after N.D.N. approval, was conducted by the developers of the instructional program in a high-achieving suburban Salt Lake City school district that is the State of Utah's model district for computer-based education. All elementary schools had computer laboratories of 20 or more Apple microcomputers linked in a local area network. The control class in this study "spent equally as much time on the computer as the treatment groups." However, it is not clear whether their reading and language arts instruction involved computers or not. As was typical in many of the research reports reviewed, there was no description of the instructional process given to the control classes.

The instructional program provided in the computer treatment, called "Basic Literacy Through Microcomputers," involved heavy emphasis on typing skills, teacher-led oral activities with computer programs, and rapid recognition of words flashed on the computer screen. The teacher and class apparently went to the school's computer lab together for the instruction. Each student spent 30 minutes per day (2 hours per week) using the programs. The evaluation covered a 4-month period.

The study reported on covered all six elementary grade levels (1-6) split among two elementary schools, and thus involved only a single classroom per treatment at each grade level. Control classes came from the same pair of schools. In a majority of cases, the classes at the same grade level had substantial differences in pre-test scores (over .5 standard deviations). To compensate, we report mean effect sizes for pairs of grades -- 1st and 2nd; 3rd and 4th; and 5th and 6th, and combine reading and language test scores (Metropolitan Achievement Tests).

The effect sizes for data reported from this study are huge. The overall average is +1.00. The effect is largest for the lowest grade levels, but all two-grade means are above +.8. Poorly performing control group classes contributed somewhat to this effect. They gained less than the norm population did. But the computer classes made nearly four times the mean gains of the norming population, adjusted for the time period between tests. This study certainly is worthy of a validation in a more typical setting, involving a clearly described control group experience and conducted by researchers not connected with the program developers.

The remaining studies in Table 1 had at least one problem that detracts from the strength of their conclusions. The dissertation by Turner (1985) focused on differences between paired and individual computer-assisted-instruction mathematics practice at the 3rd and 4th grade levels. Twelve classes were randomly assigned to paired use, individual use, or control -- no computer use; thus there were 2 classes per treatment per grade level. Teachers were rotated among the treatments so that all teachers taught classes of each type during the study. However, the total computer experience was rather brief. Although the study lasted 14 weeks, each student or student pair apparently had less than 10 hours' use of the computers (fifteen minutes every other day). Both paired and individual C.A.I. users gained marginally significantly more than the control classes (and not distinguishably from each other). An effect size could not be calculated, because standard deviations reported were standard deviations of gain scores rather than of the test score distribution.

The "computer immersion" project reported by Ferrell (1985) evaluated a method in which two classes of 6th grade students used microcomputers for mathematics drills for 40 minutes per day (roughly 3 hours per week) over a 6-month period. Two classes in the same middle school and taught by the same teacher served as controls. Classes were randomly assigned to these treatments. Twenty-six computers were housed in the teacher's classroom. Students used a variety of drill-and-practice software, primarily S.R.A. math drills and D.L.M. "arcademics" game drills. The researcher found a statistically significant difference in test scores (using the previous Spring's score as pre-test covariate) favoring the computer classes. But she rejected the idea that her results had practical significance saying that the difference amounted to only 1.9 answers on the post-test. However, the practical impact depends upon the standard deviation on that test. If it was as low as 5 questions, that difference is equivalent to an effect size of .4 -- comparable to impressive results from other "successful" C.A.I. evaluations. Standard deviation data were not reported by the investigator.

One other study in Table 1 dealt with mathematics. In this case, the focus was on logical problem-solving skills (Melnik, 1985). The computer programs used gave students practice with logical thinking and decision-making (Code Quest and The Factory, by Sunburst). This experience was contrasted to a direct instruction program that explicitly emphasized solving mathematics word problems and to an "untreated" control group. The outcome measure (pre- and post-) was the applications sub-scale of the Stanford Achievement test. Fifth grade students in an ethnically mixed suburban Chicago district participated -- two randomly assigned classes using computers and two in each of the other two treatments. Ten weeks of instruction separated pre- and post-tests, with the computer exposure amounting to one hour per week. As with other small studies, pre-test differences between groups can arise in spite of random assignment. In this case, the control group started out at least one-third a standard deviation higher than the other two. That was the only treatment that did not experience a substantial gain on the Stanford test. The effect size for the computer vs. control treatment comparison was +.45. There was no difference, however, between the computer approach and the word problem approach.

Two of the studies in Table 1 concerned the computer's effect on writing quality. One (Welzel, 1985) had some appropriate design elements: four randomly assigned classes per treatment (but spread among grades 3 through 5); 15 weeks in duration; a specifically planned 8 week sequence of writing activities following a 3-week keyboarding skills training program; and blind holistic scoring of writing samples (control classes' essays re-typed on a dot-matrix printer to resemble the computer treatment classes' work). However, writing practice time was not held constant. In one control class, the teacher emphasized writing across the curriculum, and the students did a substantial amount of writing; in the other three control classes, the language arts program involved writing no more than once every two weeks. Also, the computer-using classes only used the word processing program (Bank Street Writer) for 15 minutes, twice per week -- not a substantial writing experience. And the pre-test results of the computer classes may have been substantially affected by having been done on the computer, after only three weeks of typing experience. Effect sizes averaged -.10 for two different writing samples.

For the other writing study in Table 1 (MacGregor, 1986), we have only limited data, although we did not obtain the complete report of the study. Four sixth-grade classes were randomly assigned to one of three computer and one control condition. The computer groups varied according to paired or individual assignment and whether additional pre-writing software (Random House Story Starter) was used in addition to the word processor (Scripsit on the TRS-80). The study lasted 10 weeks, but it was unclear how much time per week the students wrote on the computer (30 minute sessions, using the two computers available in each class). The writing experience in the control group classes lasted equally as long as the computer groups'. The researcher reported only the result of some significance tests, which were generally reported to be positive and significant ($p<.05$). However, no details were given in this particular summary of the study.

Studies With Comparison Groups But Without Randomized Class Assignment

Several of the microcomputer-based studies in Tables 2 and 3 are worth commenting upon. Two of the studies in Table 2 used interesting designs. Ash (1985) employed a cohort design in which student records for non-Chapter I students in the fifth-grade of one elementary school were identified for four years running -- for the two years before math instruction employed computers and for two during the computer-based program.

Fourth and fifth grade CAT scores were used to compute gains in mathematics achievement. The C.A.I. was supplementary, though, representing an addition of 40 minutes per week on mathematics instruction. Chapter I students were removed from the analysis because during the pre-computer period they had received additional mathematics instruction.

Teachers and a lab aide were present in the computer lab while students worked at a variety of math software including some mathematics programs from Sunburst, S.R.A., Milliken, and C.D.C. Plato (programs converted to microcomputer use). Data for 51 students involved during the two-year computer program were compared with data for 71 students from the previous two years. Gains on the computation sub-test of the California Achievement Tests were slightly less for the cohorts using the computer (ES=-.16), but gains on the concepts and applications sub-test were substantially higher (ES=+.99). The cohort design is called into question only because a much larger fraction of the fifth grade cohort was excluded during the computer years than in the pre-computer years (50% vs. 33%). However the pre-tests for the two treatment populations were nearly the same.

Another supplementary C.A.I. program was reported by Todd (1985) in a study of reading and math programs for fourth graders. In this case, different students received C.A.I. reading than those who received C.A.I. math, and each group then served as the control group for the other. All fourth grade classes in four schools participated (N=241) making this one of the larger microcomputer-based studies in the set. Pre-tests showed that the groups were comparably matched. The students received 15 minutes per day additional time in the subject for which they used computers (one half-hour when pairs were used) -- or 33 hours total during the year. Educational Activities' Diascriptive

Reading program was used for the C.A.I. reading treatment and Milliken Math was used for mathematics. The investigator reported a number of logistical and technical problems with the reading software which may have been responsible for the negative effect produced on achievement gains in reading (ES=-.31). However a substantially greater attrition rate in the mathematics group than in the reading group (25% vs 13%) may have played some role. The mathematics achievement data was partly garbled in the report, so only effect sizes for problem-solving and computation could be computed. Those turn out to be +.08 and +.29 respectively.

The largest known study of the effects of microcomputer-based approaches to instruction is among those listed in Table 3. This is a statewide evaluation in Arkansas covering 22 school districts in grades 4 to 6 (McDermott, 1985). In this study, called "Project Impact," data are provided for literally hundreds of classes using one of several categories of computer-based instruction: C.A.I. using stand-alone Apple computers; C.A.I. using stand-alone Commodore 64 computers, C.A.I. programs on networked Apple computers and linked to a management system for individualizing instruction; and a locally developed management system for mathematics practice on Commodore computers. Actual computer experience varied from school to school, but most of the implementations followed patterns of 12 to 20 minute sessions held between 3 and 10 times every two weeks. Software varied widely from district to district, but was mainly products by small independent educational software producers (Micro-Ed, Microcomputer Workshops, Sterling Swift, Hartley, Edusoft, Educational Activities) as well as by national publishing firms (S.R.A., Zaner Bloser) and school district products (San Diego School District).

Effect sizes were calculated and reported by the investigator on a classroom by classroom basis. No data was given on the selection of comparison classrooms -- whether any were randomly chosen, how each was matched to a computer-using classroom or how the pair differed on pre-test achievement. Comparison classes may have been located in the same school building, but even this was not explicitly stated in the report. No analyses of teacher differences between computer-using and control classes were given. For all of these reasons, the study was included in Table 3 rather than Table 2.

The effect sizes were computed on annual achievement tests, although it was not clear whether the tests used were statewide tests or the S.R.A. series. In any event, reported effect sizes, when averaged across classrooms and schools were all positive. The highest ES was reported for the locally developed management package (ES=.62, based on eight pairs of classes). Other means (for math, reading, and language arts programs on Commodores, stand-alone Apples, and networked Apples) fell in the narrow range between +.14 and +.24 -- indicating small positive effects of the computer-based approaches. Individual classroom comparisons varied widely, however. Thirty percent of the 77 comparisons for Apple C.A.I. (across math, reading, and language arts outcomes) had NEGATIVE effect sizes as did 29% of the comparisons for Commodore C.A.I.

Implications of These Studies for Further Research on Effectiveness

The variability among classrooms in the Arkansas data is an important reminder that studies involving two or three classrooms per treatment can never be regarded as definitive, regardless of the quality of their research design. But neither are studies reporting data from hundreds of classrooms inevitably informative. To be most helpful, studies need to describe in detail how computers (and control treatments) were implemented at each site, and to do so in some kind of standardized way so that comparable aspects of the implementation can be measured for each study. What we need are studies with the size and range of the Arkansas Project Impact study and the detailed descriptions of the best of the reports in Tables 1 and 2.

It is also important to note that most of the data reported in studies cited here relate to math, reading, and language arts achievement in upper elementary grades. A few studies of writing and problem-solving were identified and met our minimal criteria for inclusion, but only one covered the high school grades and none involved science or social studies learning.

And even were we to consider all of the mathematics, reading, and language arts studies cited above to be without damning flaws, together they do not come close to providing prescriptive data for deciding whether and how to use computers as adjuncts for instruction in these subjects. Different hardware and software, different arrangements for using computers in the context of classroom and teacher-directed learning in school buildings, and different ways of integrating computer activities with the rest of the instruction in that subject may all profoundly affect the direction and size of effects on student achievement. In addition, the use of computers for one set of skills (e.g., automaticity of algorithmic procedures) or concepts (e.g., pre-algebra mathematics) or one group of students (e.g., below-grade level students) may say very little about their value in other situations.

Thus, the existing evidence of computer effectiveness is very scanty if we prudently refuse to collapse such disparate studies as a system-wide effort to raise Chapter I students' math and reading scores, a writing project of a single junior high English teacher, and an evaluation of computer programming's effects on general problem-solving abilities of above-average achieving 5th graders.

But does that mean that computer-based treatments are so "iffy" that we will never have data that can guide policy-makers and school administrators in a useful way? No, I think we're in a better position than that. But to systematically build a body of evidence about the conditions and circumstances in which computers are most effective we need a new model for conducting field research in schools.

Multiplying Knowledge of Computer Effects: A Model for Field Experiments

Traditionally, field experiments have been conducted in order to study the impact of a specific instructional program implemented in similar ways in a small number of closely situated schools. In order to insure faithful implementation of the instructional program, close personal supervision of the sites by the investigator has been the norm. As a result, studies have been conducted in close proximity to the investigator's place of work, and they have tended to involve few schools, teachers, and classes. Often those studies have been of short duration, further limiting one's ability to identify important effects. School district-wide evaluations, the other type of study predominant in the literature reviewed above, are generally longer in duration and involve larger samples of schools, teachers, and classes. But these often lack alternative control-group treatments, and are much less likely to randomly assign classes or schools to treatments, particularly when computers are involved.

As a person who has collected and analyzed survey research data for my professional career, I have envied the adequacy of outcome data -- particularly student achievement data -- that field experiments make possible. But I find it extremely limiting to think of measuring effectiveness of computer-based instruction by collecting data from a characteristically unrepresentative and small sample of specially designed programs -- even where alternative treatments and randomized assignment are used. Thousands of teachers and schools pick from the marketplace of computer-related products and ideas in order to accomplish instructional goals. And they produce a wide range of approaches for using computers with their students. To truly understand the effect that computers are having on students in schools, one needs to combine the large domain and representativeness of survey research with the most important characteristics of carefully designed field studies -- appropriate experimental design, over-time measurement of student achievement, and control and measurement of program implementation.

This Fall, our Center begins a research project that attempts to combine the features of survey research and field experiments in an effort to better understand the typical effects and the range of effects on student academic achievement of instructional programs that use computers. This project goes under the name of "The National Field Studies of Instructional Uses of School Computers." In this first year of what I expect will be a three year project, I am concentrating on one subject-matter and one grade-span -- mathematics instruction in grades five through eight. Subsequent years will focus on other specific subject-matters and grade-spans.

Each study will be one school year in duration and involve 80 classrooms recruited across the entire country. The 80 classrooms will be composed of 40 pairs of experimental and control classrooms, each pair located in the same school and matched in terms of grade level and ability level. We are recruiting schools that have a certain minimum of computer hardware -- specifically, enough for the computer-using class to have access to one computer for every four students in that class. (Most schools responding to our invitation to participate have substantially more than this number.) Also, because randomized designs are being employed -- at least at the classroom level -- all teachers recruited must be computer knowledgeable. We are asking that all teachers have at least two years of experience using computers for instruction, and we try to identify schools that have an instructional program in place, or are adjusting their program to make use of new software that has recently become available.

At each school, each class among a pair of participating classes will be randomly assigned to an instructional treatment -- computer-use or no computer use. Where possible, students will be randomly assigned to classes as well. Where schools are departmentalized and a single participating teacher teaches multiple sections at the same grade and ability level, the teacher will teach a pair of classes, using computers in one class and but not in the other -- at least not for the subject being studied. Where classes are self-contained, participating teachers and their classes will be randomly assigned to treatments.

The specific implementation -- the kinds of software used, the mathematics objectives taught, and the social organization of computer use -- will be decided by the school and the teacher, and will therefore vary substantially from site to site. Some teachers will emphasize programs that focus on basic skills or factual recall, while other teachers will mainly use programs that engage students in writing, thinking, or problem-solving activities. Some teachers will use computers to provide individualized instruction or enrichment activities for certain students; others will organize computers so that all students in their class use computers in the same ways. Each school is expected to attempt to optimize the amount of time that students in the computer-using class use computers for mathematics during the year, with a 30 hour minimum (more if students use computers in pairs). The only structural requirement for the schools is that a comparable curriculum and similar "opportunities to learn" be given to students in both the computer-using and the control classroom at the same school. The curriculum and emphases will vary from school to school and grade to grade.

Because an investigator will not be on-site in most cases, we will regularly collect specific information from each classroom about the activities and assignments for each class. In addition, we will systematically collect data about the way each classroom implements its computer-based approach -- for example, in-class versus lab arrangements, paired vs. individual use, networked micros vs. stand-alone micros, and so on. For the mathematics study, the same or equatable standardized achievement tests will be used for both pre- and post-testing at all sites throughout the country, and a curriculum-specific test will be developed to match the content of instruction at each site, and given as a post-test as well. Finally, teacher and student characteristics will be measured by comparable questionnaire data at all sites. Thus, regardless of how different schools implement instructional programs using computers, we will systematically collect data about those implementations and be able to describe and contrast each mini-experiment within the overall study. Among other consequences, this will enable us to correlate each mini-experiment's effect size with variations in conditions and characteristics of each implementation.

Selection of schools and classrooms will balance two major goals: (1) broad representation among regions of the country, student academic ability, approaches to using computers, types of computers and software used, and grade level; and (2) having a sufficient number of classes using computers in similar ways (e.g., using similar drill-and-practice computational packages from major publishers; or doing mathematics word-problems using a game-like format) so that different approaches can themselves be compared.

How willing and able are schools to meet the requirements of such a project? The requirements are severe. For example, a school must have enough computer experience to plausibly implement a respectable program of instruction, but computers cannot be so much a part of the instructional program that a school would be unwilling to contribute an untreated control group. Also, class assignment and staffing patterns must be amenable to our research design -- departmentalized teaching assignments with heterogeneous student assignment to classes is ideal; other patterns can also be identified so that computer and control classes have roughly equal pre-test abilities.

Certainly, most schools cannot meet the needs of the research design. But there are roughly 40,000 schools in the United States enrolling at least one of the grades five through eight that have a sufficient number of computers available and used for mathematics instruction. Our project seeks roughly 20 to 25 schools (80 classes) per study.

Knowing the requirements for participation -- including randomized assignment of classes to computer- and traditional-instruction approaches, many more schools have applied for the first year's mathematics study than we can accept. So far, we have selected sixty of the eighty classrooms that we need. The sixty come from 16 schools in 14 districts from Florida to Oregon. The remaining sites will be chosen based on the quality of their program and to assure greater socio-economic representativeness.

The data that will be produced by the first of the national field studies, on mathematics achievement in the middle grades, will greatly expand the amount of research data available to date on the effectiveness of microcomputer-based instructional programs in that area. Subsequent research planned for middle school science, writing, and language arts will similarly multiply what we know about the effectiveness of instruction emphasizing computers that is occuring in those subject-areas. But this series of studies may still be spread too thin to answer questions about the variability of effectiveness of computer-based instruction.

I encourage other researchers to propose initiatives of their own or to expand my own sample with additional schools participating in comparable experimental designs so that we can begin to get more definitive information about the impact of computer use on children's learning among the broad range of school and classroom settings in which computers are used today.

REFERENCES

Ash, Ruth. (1985). An analysis of the cost-effectiveness of a computer-assisted instructional program. (Doctoral dissertation, Auburn University).

Atlanta Public Schools. (1984). Evaluation of the Atlanta Public Schools Computer Curriculum Corporation labs, 1984-1985. Atlanta, Georgia.

Becker, Henry J. (1983). School uses of microcomputers: Reports from a national survey. Johns Hopkins University Center for Social Organization of Schools, Issue No. 1, April.

Becker, Henry J. (1986a). Instructional uses of school computers: Reports from the 1985 national survey. Johns Hopkins University Center for Social Organization of Schools, Issue No. 1, June.

Becker, Henry J. (1986b). Instructional uses of school computers: Reports from the 1985 national survey. Johns Hopkins University Center for Social Organization of Schools, Issue No. 3, November.

Bangert-Drowns, Robert L., Kulik, James A., and Kulik, Chen-Lin C. (1985). Effectiveness of computer-based education in secondary schools. Journal of Computer-based Instruction, 12, 59-68.

Burns, P.K., and Bozeman, W C. (1981). Computer-assisted instruction and mathematics achievement: Is there a relationship? Educational Technology, 21(10), 32-39.

Clark, Richard. (1985). Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses. Educational Communications and Technology Journal, 33, 249-262.

Edwards, Judith et al. (1975). How effective is CAI? A review of the research. Educational Leadership, 33(2), 147-153.

Ferrell, Barbara G. (1985). Computer immersion project: Evaluating the impact of computers on learning. Texas Women's University. Paper presented at the Annual Meeting of the American Educational Research Association.

Gourgey, Annette, Azumi, Jann, Madhere, Serge, and Walker, Elaine. (1984). Computer-assisted instruction evaluation report, 1983-1984. Newark School District, Newark, New Jersey.

Harris, Carolyn DeMeyer, Sadacca, Robert, and Hunter, Beverly. (1985). Evaluation of the TSC Dolphin. Computer-assisted instructional system in the Chapter 1 program of the District of Columbia public schools. Human Resources Research Organization, Alexandria, Virginia.

Hartley, S. (1977). Meta-analysis of the effects of individually paced instruction in mathematics. (Doctoral dissertation, University of Colorado).

Hawley, David E., Fletcher, J. Dexter, and Piele, Philip K. (1986). Costs, effects, and utility of microcomputer assisted instruction. Center for Advanced Technology in Education, University of Oregon.

Hasselbring, Ted S. (1984). Research on the effectiveness of computer-based instruction: A review (Report #84.1.3). The Learning Technology Center, George Peabody College of Vanderbilt University.

Jamison, Dean, Suppes, Patrick, and Wells, Stuart. (1974). The effectiveness of alternative instructional media: A survey. Review of Educational Research, 44(1), 1-67.

Kulik, J.A., Bangert-Drowns, R.L., and Williams, G.W. Effectiveness of technology in precollege mathematics and science teaching. (1983). Journal of Educational Technology Systems, 12, 137-158.

Kulik, James A., Kulik, Chen-Lin C., and Bangert-Drowns, Robert L. Effectiveness of computer-based education in elementary schools. (1985). Computers in Human Behavior, 1, 59-74.

McDermott, Cecil. (1985). Affecting basic skills achievement through technology: A research report on IMPAC. Arkansas Department of Education, Little Rock, Arkansas.

MacGregor, S. Kim. (1986). Computer-assisted writing environments for elementary students. Louisiana State University. Paper presented at the National Educational Computing Conference.

Melnik, Leah. (1986). Investigation of two methods for improving problem-solving performance of fifth grade students. (Doctoral dissertation, Loyola University, Chicago).

Millar, Garnet, and MacLeod, Alan. (1984). Microcomputer learning project: Willow Creek School Division. Alberta (Canada) Education.

Miller, Suzanne. (1984). A comparison of computer-assisted instruction with prescription learning and the traditional "pull out" program used in supplementing instruction of mathematics basic skills to chapter 1 students. (Doctoral dissertation, Oregon State University).

Niemiec, Richard, and Walberg, Herbert. (1985). Computers and achievement in the elementary schools. Journal of Educational Computing Research, 14, 435-439.

Okey, James R. The effectiveness of computer-based education: A review. (1985). Paper presented at the 58th Annual Meeting of the National Association for Research in Science Teaching, French Lick Springs, Indiana.

Reid, Ethna R. (1986). Basic literacy through microcomputers (various materials). Reid Foundation, Salt Lake City.

Rodefer, Janet Criswell. (1985). Teaching higher-level thinking skills through Logo. (Doctoral dissertation, West Virginia University).

Shattuck, Joan. (1986). Computer-assisted instruction: Evaluation Summary. Rochester City School District.

Slavin, Robert E. (1984). Meta-analysis in education: How has it been used? Educational Researcher, 13(8), 6-15.

Slavin, Robert E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational Researcher, 15(9), 5-11.

Stennett, R. G. (1985). Computer-assisted instruction; A review of the reviews. London (Ontario) Board of Education, Research Report #85-01.

Suppes, Patrick, and Morningstar, Mona. (1969). Computer-assisted instruction. Science, 166(October 17), 343-350.

Thomas, David B. (1979). The effectiveness of computer-assisted instruction in secondary schools. AEDS Journal, 12(3), 103-115.

Todd, Wilma. (1985). Effects of computer-assisted instruction on attitudes and achievement of fourth grade students in reading and mathematics. (Doctoral dissertation, North Texas State University).

Turner, Larry G. (1985). An evaluation of the effects of paired learning in a mathematics computer-assisted instruction program. (Doctoral dissertation, Arizona State University).

Vinsonhaler, John F., and Bass, Ronald K. (1972). A summary of ten major studies on CAI drill and practice. Educational Technology, 12(7), 29-32.

Wetzel, Keith A. (1985). The effect of using the computer in a process writing program on the writing quality of third, fourth and fifth grade pupils. (Doctoral dissertation. University of Oregon).

Table 1: Studies with Randomized Assignment of Classes to Treatments.

[All are microcomputer-based using individual diskettes unless otherwise indicated. All used national standardized tests at pre- and post-test unless indicated.]

Reference	Explicit Program	Grade Levels	Classes of Computer Treatment	Duration (weeks)	Time Per Week, Location, # of Computers	Computer Materials	Design	Effect Size
Hawley, Fletcher & Piele 1986 (CATE)	no	3,5	2 (1/gd.)**	15	50 min., in classroom(?) (6 micros)	Milliken Math with automated management	Students randomized to classes. Same teacher taught both treatments. Computer substituted for 1/3 of students paper and pencil drills.	ES:+.40 computation +.25 concepts +.22 problem-solving Computers also shown to be cost-effective
Reid, 1986 (Exemplary Reading)	yes	1-6	6 (1/gd.)	17	120 min. in lab with teacher (20+ networked micros)	networked own reading, language software	Classes randomized. Experimental treatment emphasized keyboarding instruction, and oral reading. Control group used computers in unspecified ways.	ES:+1.0 mean across 6 grade levels in reading, language
Turner, 1985	no	3,4	8 (4/gd.)	14	38 min., in lab with teacher (15 networked micros)	Milliken Math and other drills	Classes randomized. Teachers rotated among treatments. Also compared paired and individual use.	Paired and individual better than control marginally significant. No effect size calcs.
Ferrell, 1985	no	6	2	25	200 min., in lab with teacher (26 micros)	math drills, games (SRA,CDI,DLM,etc,)	Classes randomized. Intensive computer treatment -- 40 min. of CAI daily.	Significant F-test(<.05). Could not calculate effect size.
Melnik, 1986	NC*	5	2	10	60 min., location (?) (micros (?))	logic games: The Factory, Code Quest	Classes randomized to three treatments: computer, math word problems, control.	ES:+.45 vs. control -.05 vs.word problem on Stanford test of math applications
Wetzel, 1985	NC*	3,4,5	4 (1-2/gd.)	13	30 min., in lab with teacher (8 to 10 micros)	Bank St. Writer	Classes randomized. Blind holistic scoring of writing samples. "Computers in Process Writing Program." Good, thorough instructional plan hurt by unfortunate occurrences. Eg., treatment group did pre-test using computers after 3 weeks of keyboarding.	ES:-.10 average of 2 writing samples Highly variable control group results
MacGregor, 1986	no	6	3	10	30 min/day (once/week?) in classroom (2 micros)	Scripsit, Story Starter (pre-writing)	Classes randomized. Weekly writing instruction.	Little data. Mostly favored computer and most p<.05

*Non-commercial systematic instructional program either put in place by school or researcher.

** One class per grade level.

Table 2: Studies with Comparison Population Only. Pre-tests Show Close Match.

[All are microcomputer-based using individual diskettes unless otherwise indicated. All used national standardized tests at pre- and post-test unless indicated.]

Reference	Explicit Program	Grade Levels	Classes of Computer Treatment	Duration (weeks)	Time Per Week Location, # of Computers	Computer Materials	Design	Effect Size
Ash, 1985	no	5	5	28	40 min., in lab with aide and teacher (15 micros)	Math CAI--SRA,Sunburst geometry,Plato, Milliken math	Successive cohorts: two cohorts pre-computer compared with two cohorts using computers. Only non-Chapter I student data is included. More kids excluded during computer years. Informal criteria to allocate available software for individualized tasks.	ES:-.16 computation +.99 concepts (spring to spring gains)
Todd, 1985	no	4	c.5 to 6 classes	24	75 min., location (?) # micros (?)	Diascriptive reading, Milliken math	Computer treatment in one subject was control for other subject. Controls at different schools. Supplementary time for experimental group only. Reading was not as tied to teacher instruction as was math.	ES:+.29 math computation +.08 math problem-solving -.31 reading
Shattuck, 1985; 1986 (Rochester)	CMI* -CAI	4-6	1000+ kids (c.15 to 20 classes/gd.)	until spring test (30 weeks?)	daily in lab with terminals	CCC math and reading software (non-micro)	Supplementary CAI program. Comparison group: students elsewhere in the district matched on grade, sex, ethnicity, SES, and previous spring achievement scores.	Diff. in NCE gains below +/-3 for all grades, both subjects. No ES. Computer worse for high stanines both years.
Miller, 1984	CMI* -CAI	5-8	3 schools, 80 kids (c.1 class per grade)	between fall-spring tests (28 weeks?)	50 min., pull-out in labs with terminals	CCC math software (non-micro)	Chapter I alternative. Compared 3 programs used in different schools in district. Very close pretest match.	ES:+.29 on district test

*Computer-managed instruction using CAI.

Table 3: Studies with Comparison Population Only. Unclear or Poor Match.

[All are microcomputer-based using individual diskettes unless otherwise indicated. All used national standardized tests at pre and post test unless indicated.]

Reference	Explicit Program	Grade Levels	Classes of Computer Treatment	Duration	Time Per Week, Location, # of Computers	Computer Materials	Design	Effect Size
McDermott, 1985 (IMPAC)	no*	3-6	22 to 37 per subject (5-15/gd.)	school year	avg. 40 min., 4 micros in classroom	Varied software, mainly from small independent producers	Partial standardization of existing programs in the state. Comparison at same school (?). No description of instructional programs. No data on pre-test equivalence. Unclear whether state or standardized tests used.	ES:(supplied in report) +.22 math +.22 reading +.15 language arts
McDermott, 1985 (IMPAC)	CMI-CAI	4-6	17 (6/gd.)	school year	avg. 40 min., 8 micros (networked) in classroom	Produced by another school district (San Diego)	Same as above.	ES:(supplied in report) +.24
McDermott, 1985 (IMPAC:C-AIM)	CMI-CAI	3-6	8 (3 for grade 3)	school year	avg. 40 min., 8 micros (networked) in classroom	Testing program & misc. CAI software	Same as above.	ES:(supplied in report) +.62
Millar & MacLeod 1984 (Alberta)	no	6-9	80 kids (c.1 cl./gd.)	school year(?)	80 min (?) (class time or computer time)	Mixed software: MECC, Sunburst Milliken--mainly math, some reading	Students with IQ>125 and reading test 90th percentile. Comparison group from other district. No explicit computer treatment. Not clear whether others used computer but not just tested. Rough comparability between groups.	ES:+.25 reading, vocabulary +.08 math concepts
Atlanta Schools 1986	yes	elem-h.s. (majority Chapt. 1)	1000 kids per subject (3+ cl./gd.)	school year (40% dropped at sem.)	avg. 50 min. in lab with aide, 8-16 terminals	CCC reading, math software (not micro)	Comparison group unclear -- other students in district, categorized by Chapter 1 participation. Spring to spring gains measured. Probably a supplementary program.	Signif. +: reading, elem. -: Ch.I reading,middle Could not calculate ES.
Harris, Sadaccayes, & Hunter 1985 (HUMRRO)		4-6 (Chapt.1)	1500 kids (15+ cl./gd.)	1.5 school years	75 min. in lab with terminals, aide,teacher	Dolphin math, reading CAI (mainframe)	Comparison schools compared on non-achievement measures -- same except professional activity of Chapter 1 and parent involvement was lower in controls. Data presentation very limited.	No difference over 3 year interval (computer used for 1.5 of those years)
Gourgey, et al. 1984 (Newark)	yes	2-12	440 kids (most grades: c. 1-2 classes)	school year	50 min. per subject in lab with 4-8 terminals and with aide (teacher in some)	CCC reading, math software (not micro)	Supplementary program at 8 schools of varying target student populations. Comparison group of non-participants elsewhere in district selected, but not uniformly successful. Analysis broke sample into too fine categories to measure effects. Standard deviations not reported.	Either too small N's or large pretest differences prevent drawing any conclusions.
Rodefer, 1985	no	4-5	5 (2-3/gd.)	8 weeks	20 min., computers not described	Logo	Convenience sample of teachers, some trained for Logo, others not. No data presented on pre-test equivalence (on Ross Test of Higher Cognitive Process).	ANCOVA controlling on pretest, CAT reading, IQ, age. Teacher effects but no differences between Logo and controls.

*Provided some standardization in software and organization of use among districts.

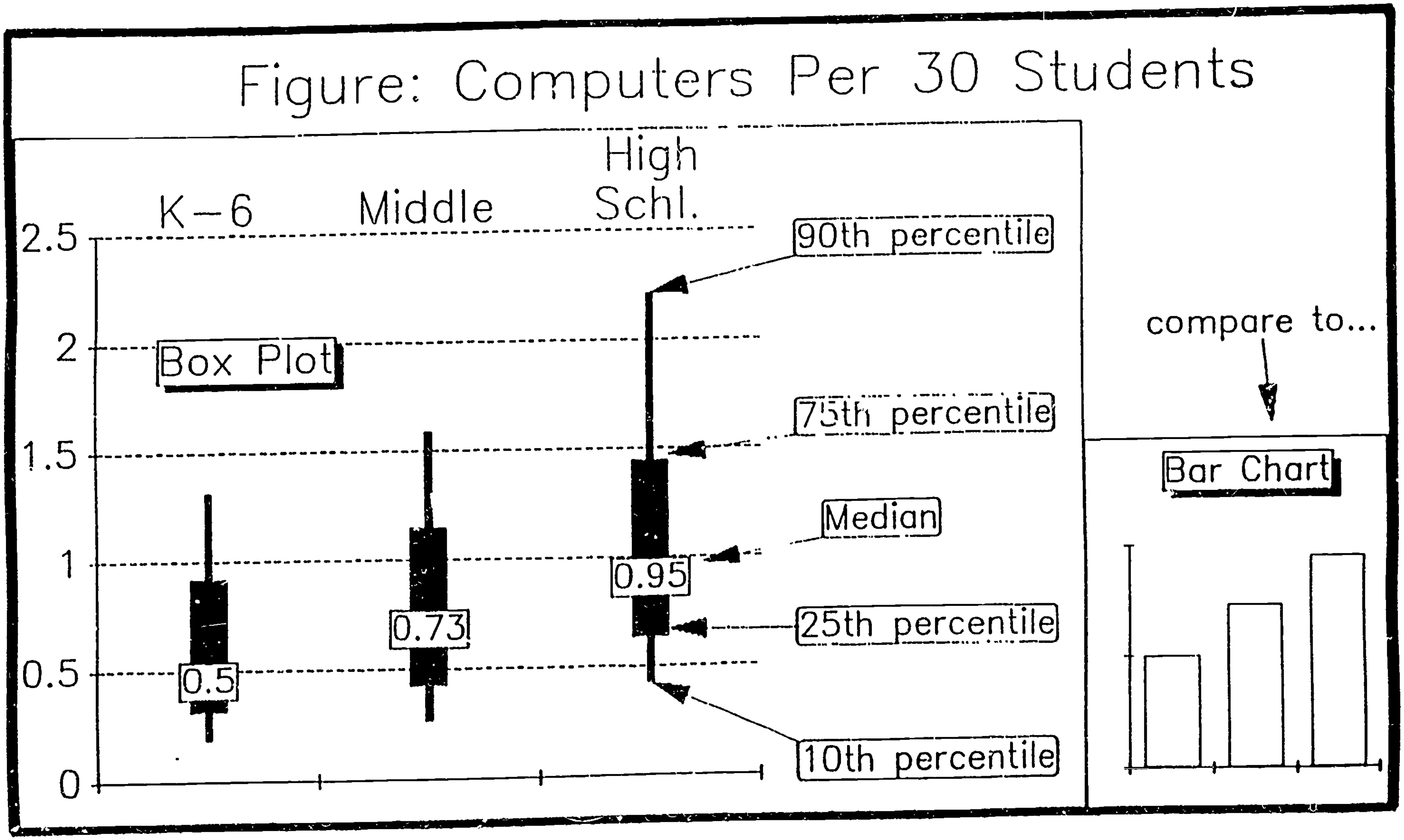


Figure: Computers Per 30 Students
K-6
Middle
High
Schl.
2.5
2
1.5
1
0.5
0
Box Plot
90th percentile
75th percentile
Median
25th percentile
10th percentile
0.5
0.73
0.95
compare to...
Bar Chart

Figure: Computers Per 30 Students

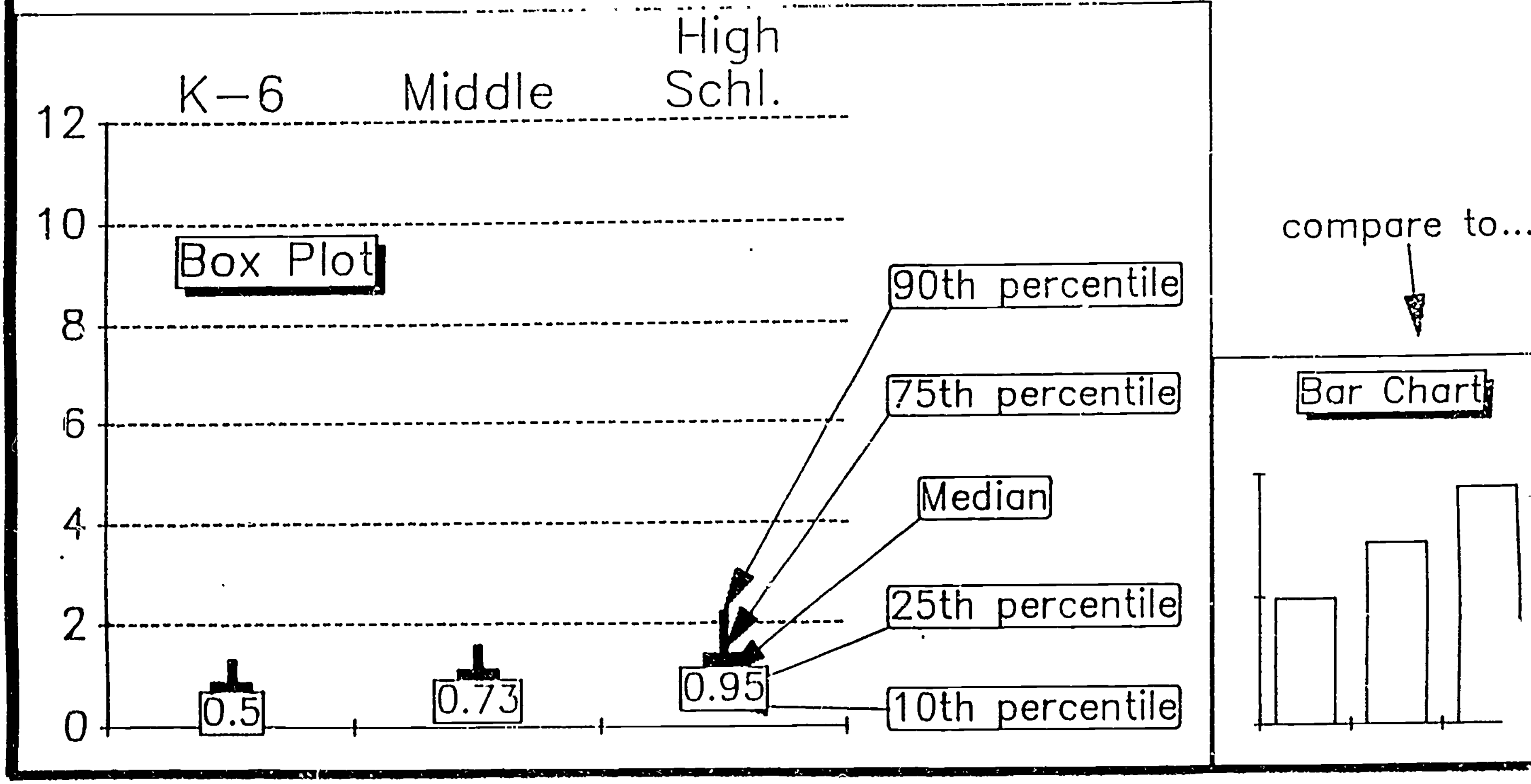


Table 4: Percent of Schools with a Substantial Number of Computers in a Given Location by Grade Span of School, Spring, 1985.

Grade Span of School	Percent of Computer Using Schools With: 7 or More Computers In Any One Room	15 or More Computers In Any One Room
K-6 Elem.	22%	6%
Middle/Jr.High	67%	35%
High School	85%	47%
K-8 Elem.	26%	4%
K-12 School	24%	4%
Jr./Sr. High	64%	26%
U.S. Total	39%	16%

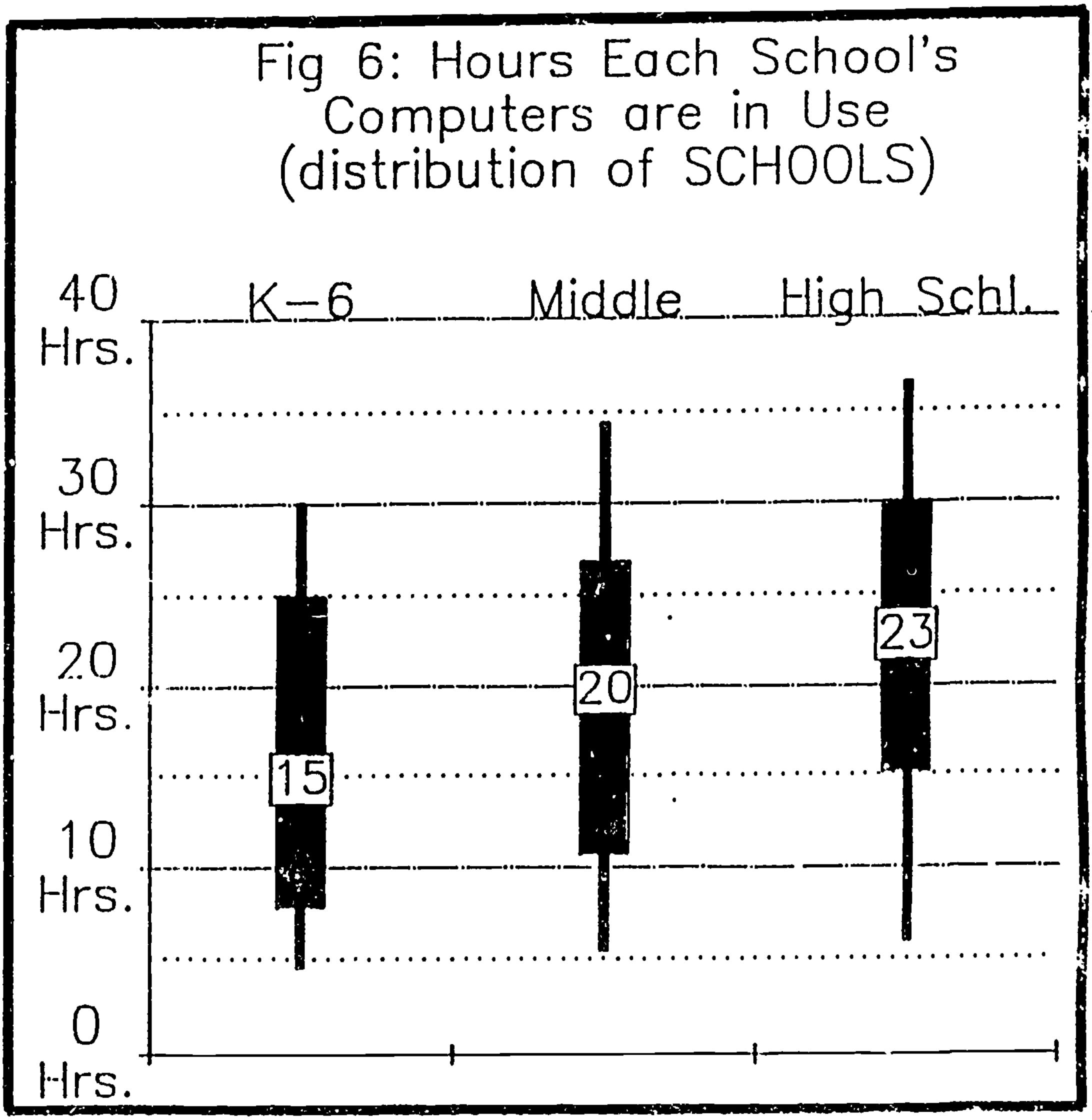


Fig 6: Hours Each School's
Computers are in Use
(distribution of SCHOOLS)
40 Hrs.
30 Hrs.
20 Hrs.
10 Hrs.
0 Hrs.
K–6
Middle
High Schl.
15
20
23

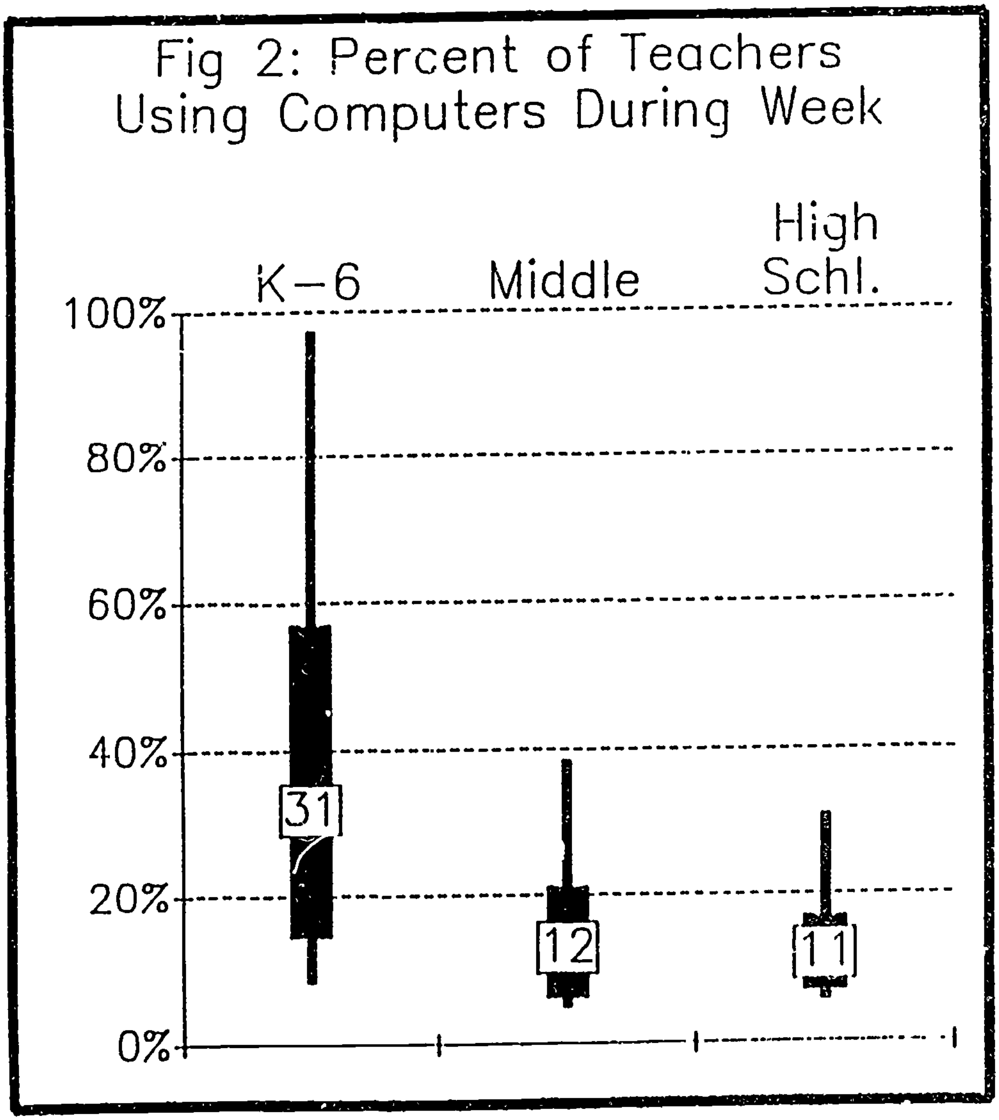


Fig 2: Percent of Teachers Using Computers During Week
K-6
Middle
High Schl.
100%
80%
60%
40%
20%
0%
31
12
11

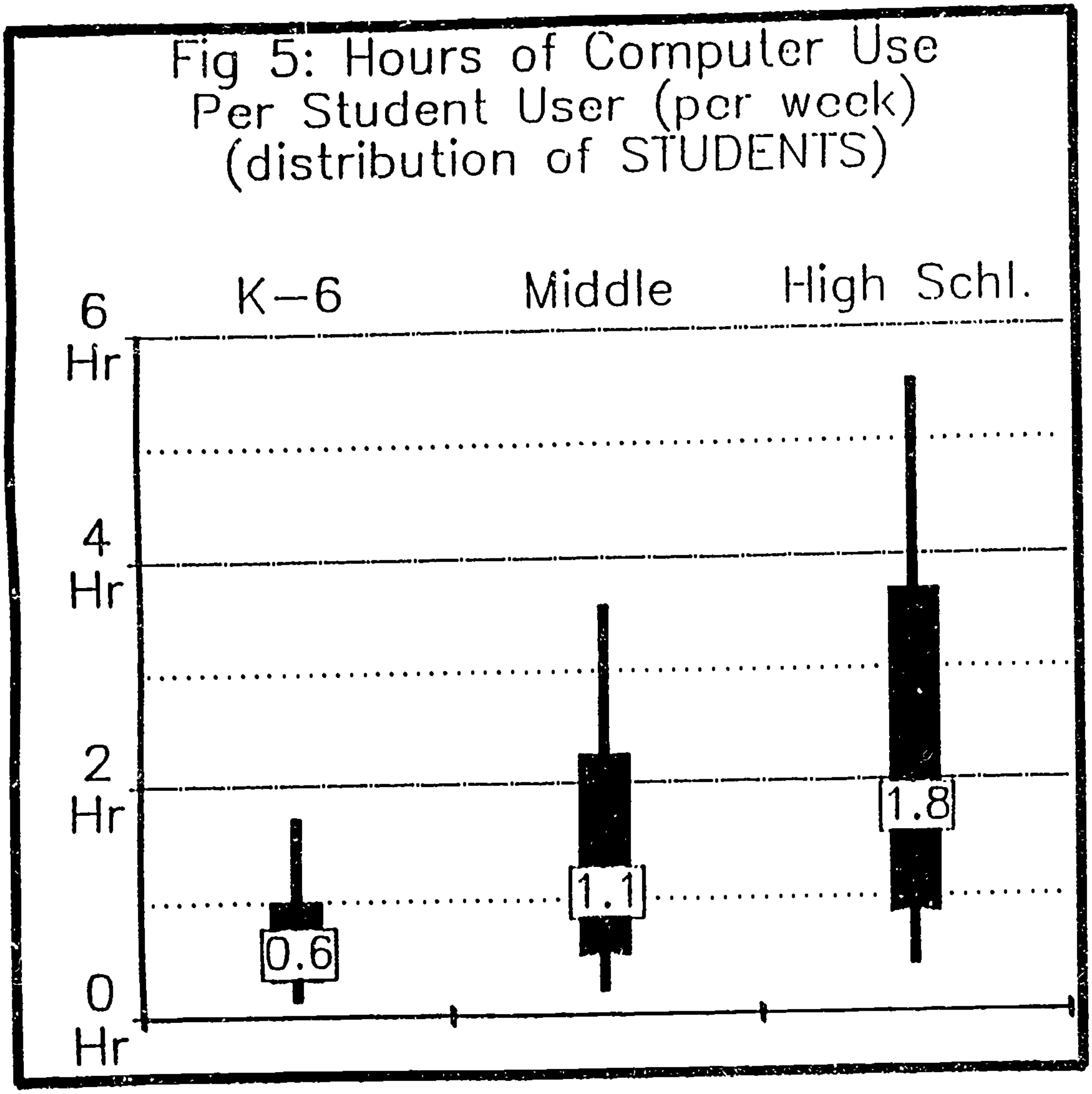


Fig 5: Hours of Computer Use
Per Student User (per week)
(distribution of STUDENTS)
K–6
Middle
High Schl.
6 Hr
4 Hr
2 Hr
0 Hr
0.6
1.1
1.8

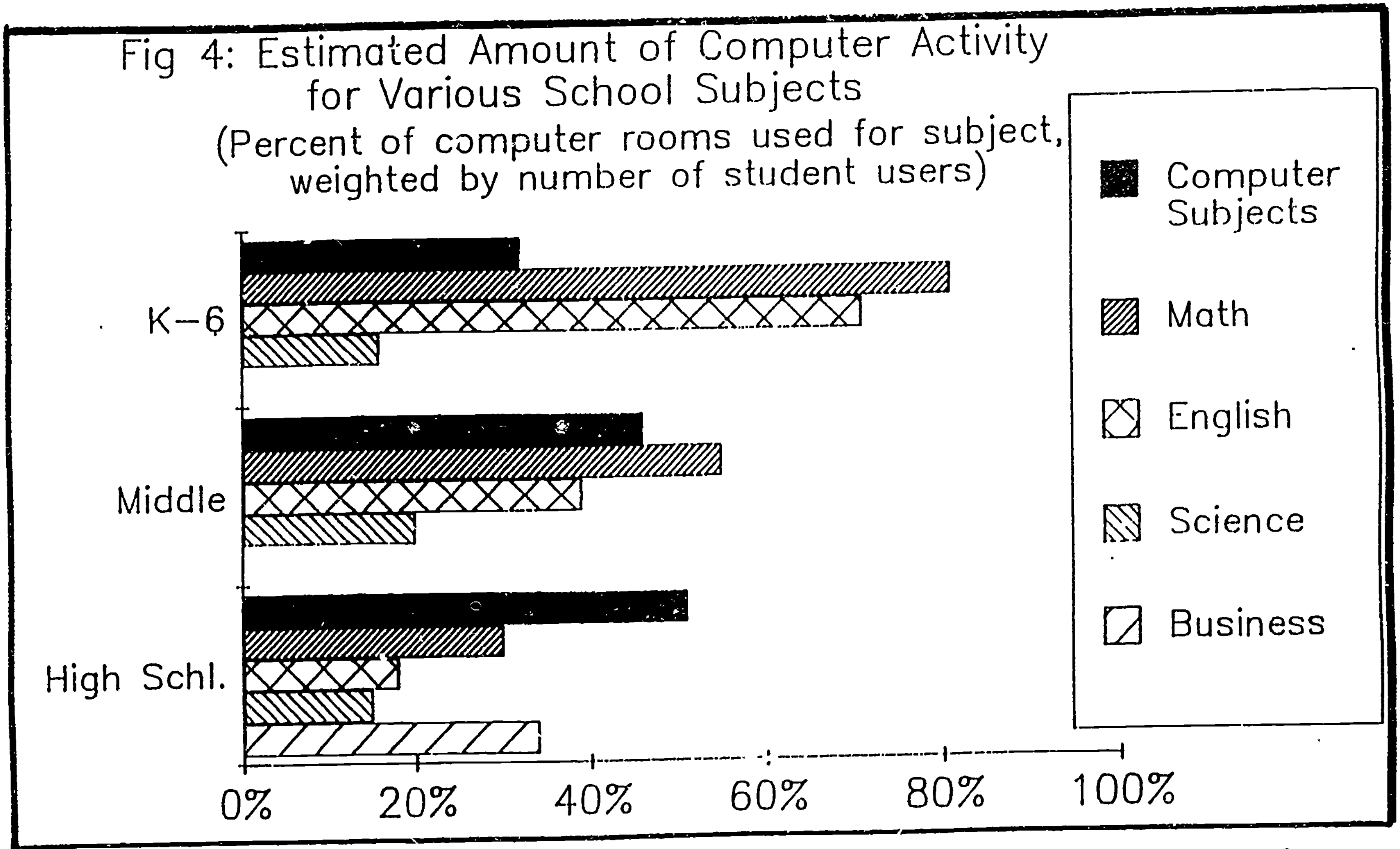


Fig 4: Estimated Amount of Computer Activity for Various School Subjects
(Percent of computer rooms used for subject, weighted by number of student users)
K–6
Middle
High Schl.
0%
20%
40%
60%
80%
100%
Computer Subjects
Math
English
Science
Business

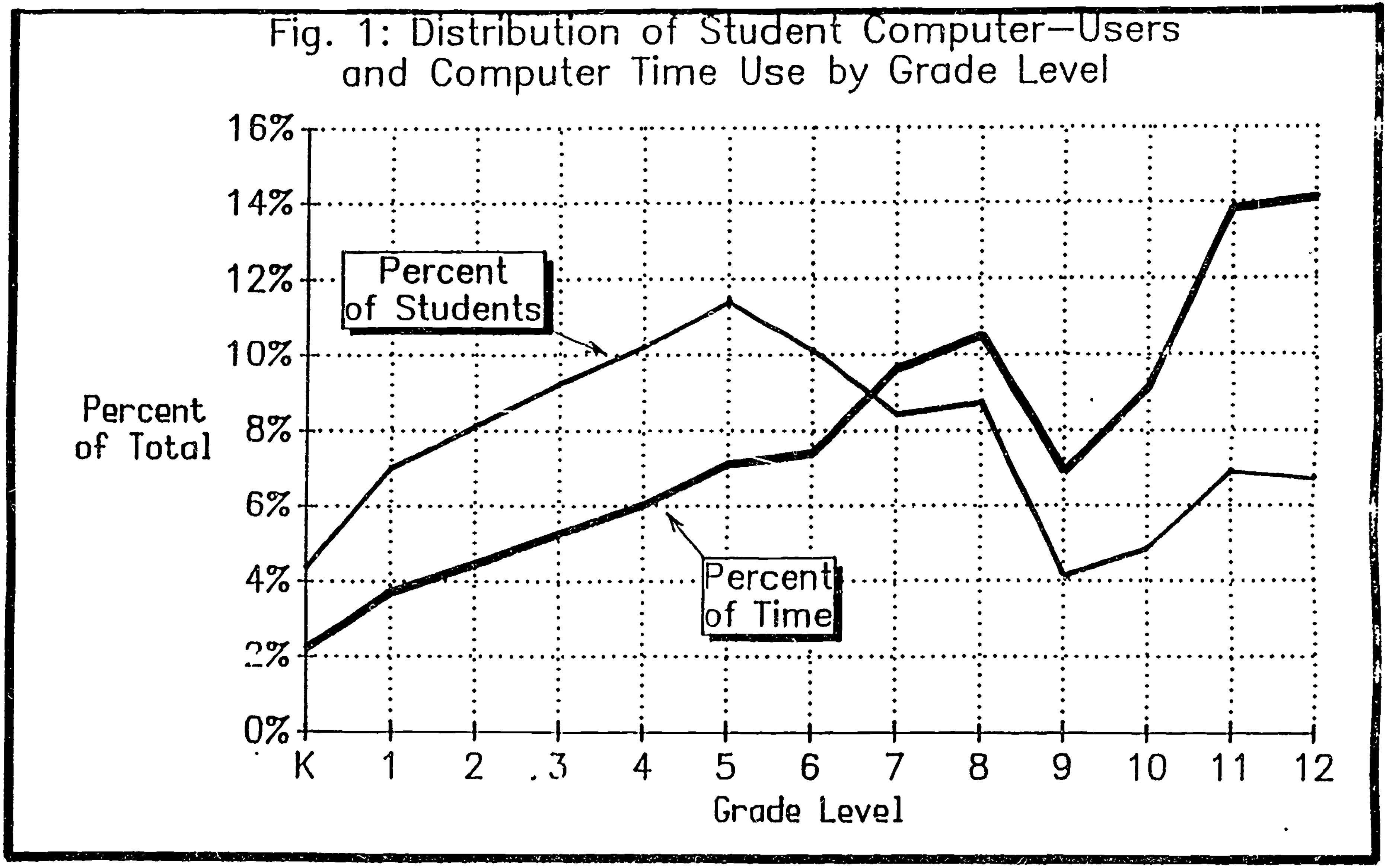
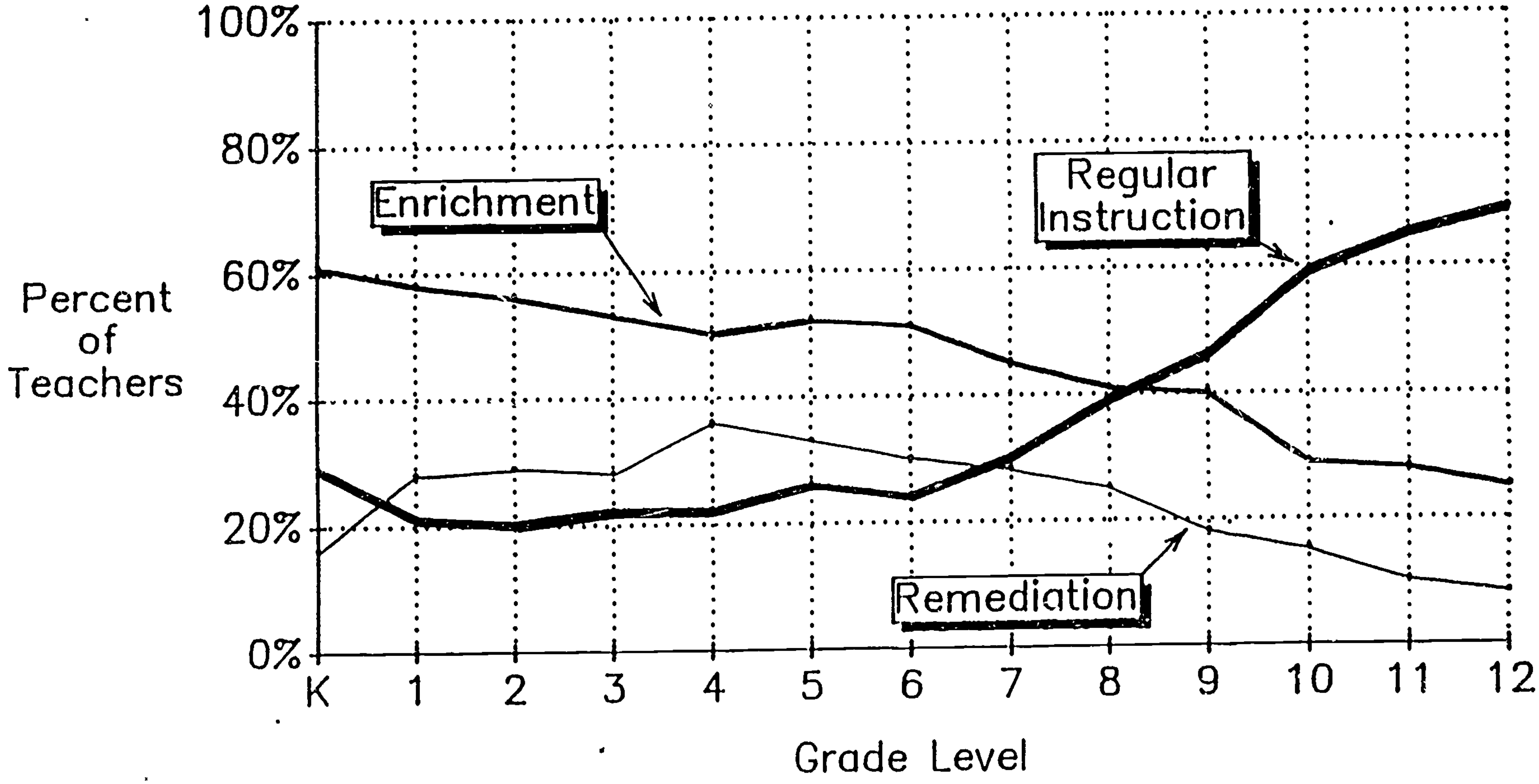


Fig. 1: Distribution of Student Computer–Users
and Computer Time Use by Grade Level
16%
14%
12%
10%
8%
6%
4%
2%
0%
Percent
of Total
Percent
of Students
Percent
of Time
K
1
2
3
4
5
6
7
8
9
10
11
12
Grade Level

Fig. 2: Function of Computer Activity for Teachers--
Enrichment, Remediation, or Regular Instruction--
by Grade Level*



*Percents May Add to More than 100% Because of Multiple Answers.

Table 2: Distribution of Computer Use, by Subject and Grade Level, Spring, 1985.*

SUBJECT (from course title, activity, and software in use)	GRADE LEVELS OF STUDENTS IN CLASS K-3	4-8	9-12
Mathematics**			
Topics below algebra or unspecified math	42%	27%	5%
Algebra, geometry	0%	1%	1%
Trig,. advanced math	0%	0%	1%
Subtotal: traditional math subjects	42%	28%	7%
English			
Language arts and spelling	18%	12%	4%
Reading	17%	8%	2%
Subtotal: language arts	35%	20%	6%
Writing	2%	2%	3%
Word processing in English class	3%	3%	2%
Subtotal: writing	5%	5%	5%
Computers and Problem-Solving			
Math topics: problem-solving, logo, programming activities	3%	6%	2%
Programming*** as specific topic or course	2%	13%	42%
Logo as specific topic****	4%	5%	1%
Computer literacy as specific topic or course	6%	10%	5%
Subtotal: computers & problem-solving	14%	34%	50%
Business and Word Processing			
Business, accounting, secretarial, other than word processing	0%	0%	6%
Word processing, other than in English	1%	2%	12%
Subtotal: business	1%	2%	18%
Other Subjects			
Science and nutrition	1%	3%	7%
Social studies	1%	4%	1%
Industrial arts and agriculture	0%	1%	5%
Others	2%	2%	2%
Subtotal: other subjects	4%	11%	15%
Total for All Subjects	100%	100%	100%

*Sum of individual entries may not equal subtotal entry because of rounding. Similarly, subtotals may not sum to grand total.

**See also first entry under 'Computers.'

***Excludes specific mentions of Logo.

****Includes some general problem-solving not classified elsewhere.

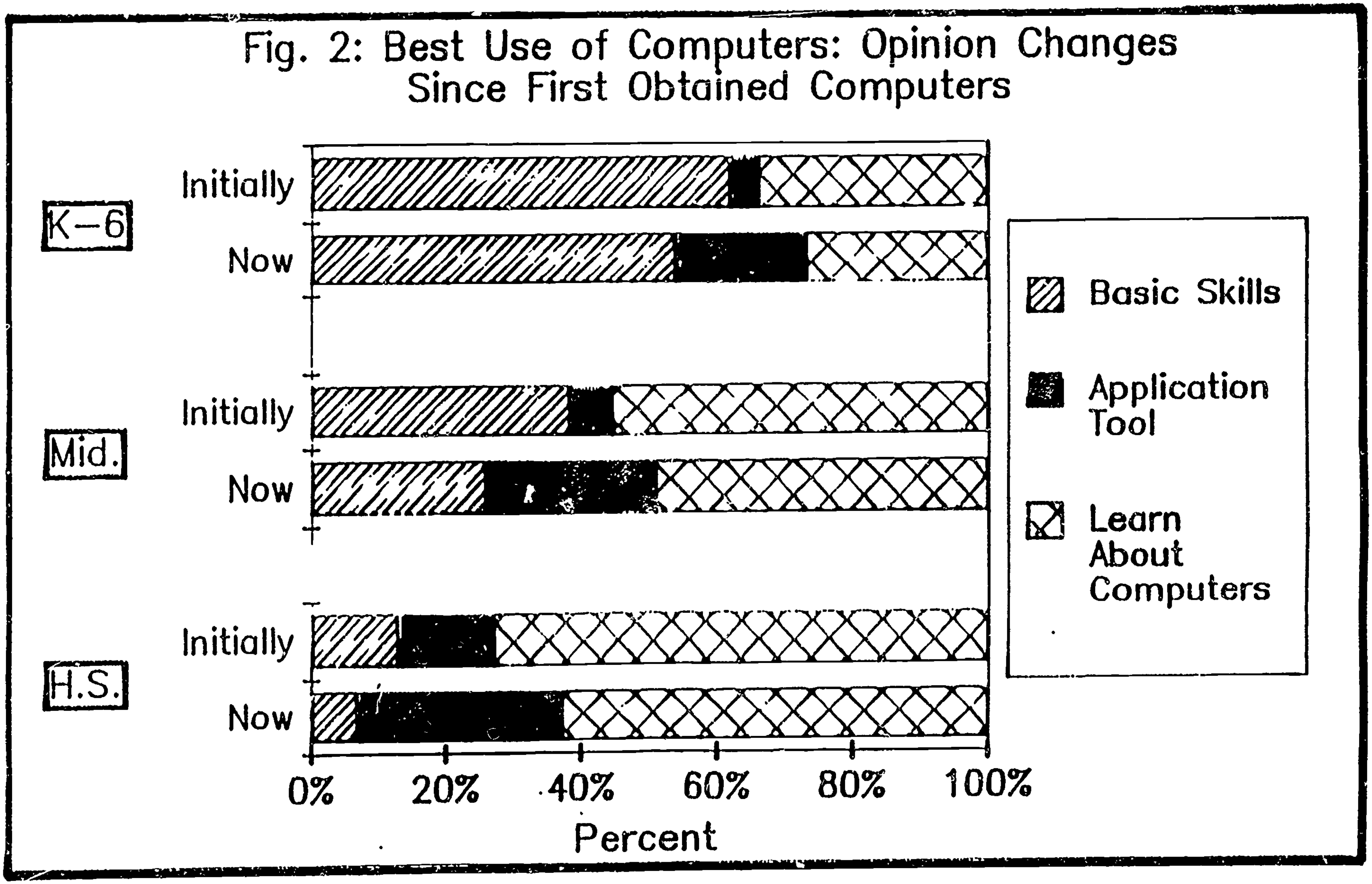
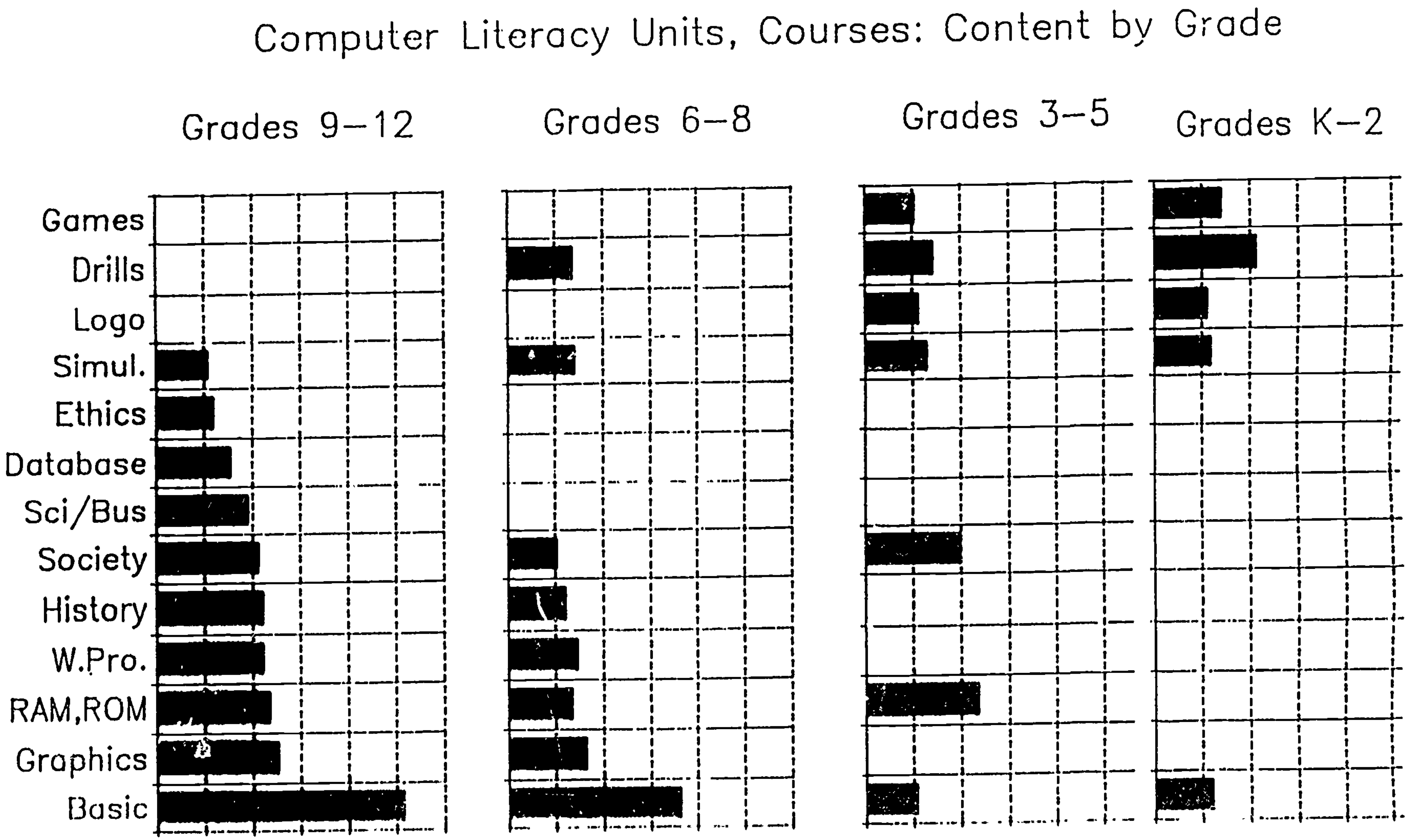


Fig. 2: Best Use of Computers: Opinion Changes Since First Obtained Computers
K–6
Initially
Now
Mid.
Initially
Now
H.S.
Initially
Now
0%
20%
40%
60%
80%
100%
Percent
Basic Skills
Application Tool
Learn About Computers


Computer Literacy Units, Courses: Content by Grade
Grades 9–12
Grades 6–8
Grades 3–5
Grades K–2
Games
Drills
Logo
Simul.
Ethics
Database
Sci/Bus
Society
History
W.Pro.
RAM,ROM
Graphics
Basic

Emphasis on Graphics in High School Programming Courses

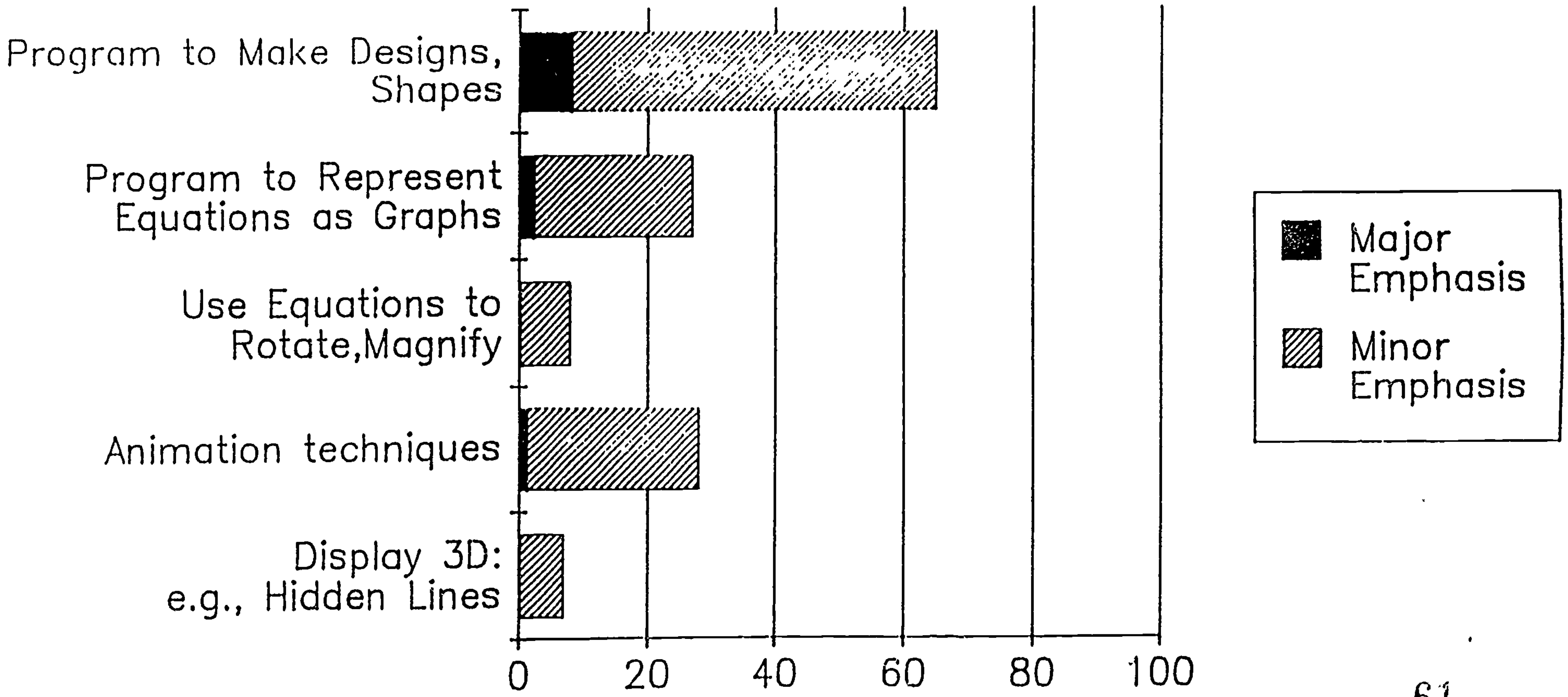


Table 1: Distribution of Computer Use by Girls for Elementary (K-6), Middle, and High Schools as Reported by Each School's Primary Computer-Using Teacher, Spring, 1985.

	Median Percent Use by Girls (50th Percentile)			Low Percent Use by Girls (20th Percentile)			High Percent Use by Girls (80th Percentile)		
	Elem. (K-6)	Middle School	High School	Elem. (K-6)	Middle School	High School	Elem. (K-6)	Middle School	High School
Overall Use	50%	50%	50%	50%	40%	35%	50%	50%	50%
Before/After School	30%	15%	20%	0%	1%	3%	50%	50%	50%
Word Processing	50%	50%	60%	0%	6%	45%	50%	50%	80%
Elective Programming	9%	45%	44%	0%	20%	30%	50%	50%	50%
Game Playing	50%	30%	10%	20%	5%	2%	50%	50%	40%
Use Index*	45%	38%	40%	30%	26%	29%	50%	50%	51%

*Additive index of non-missing entries on previous five aspects of use.

Student–Computer Ratios & Computer–Using Teachers
At Schools Attended by Median Black and White Students

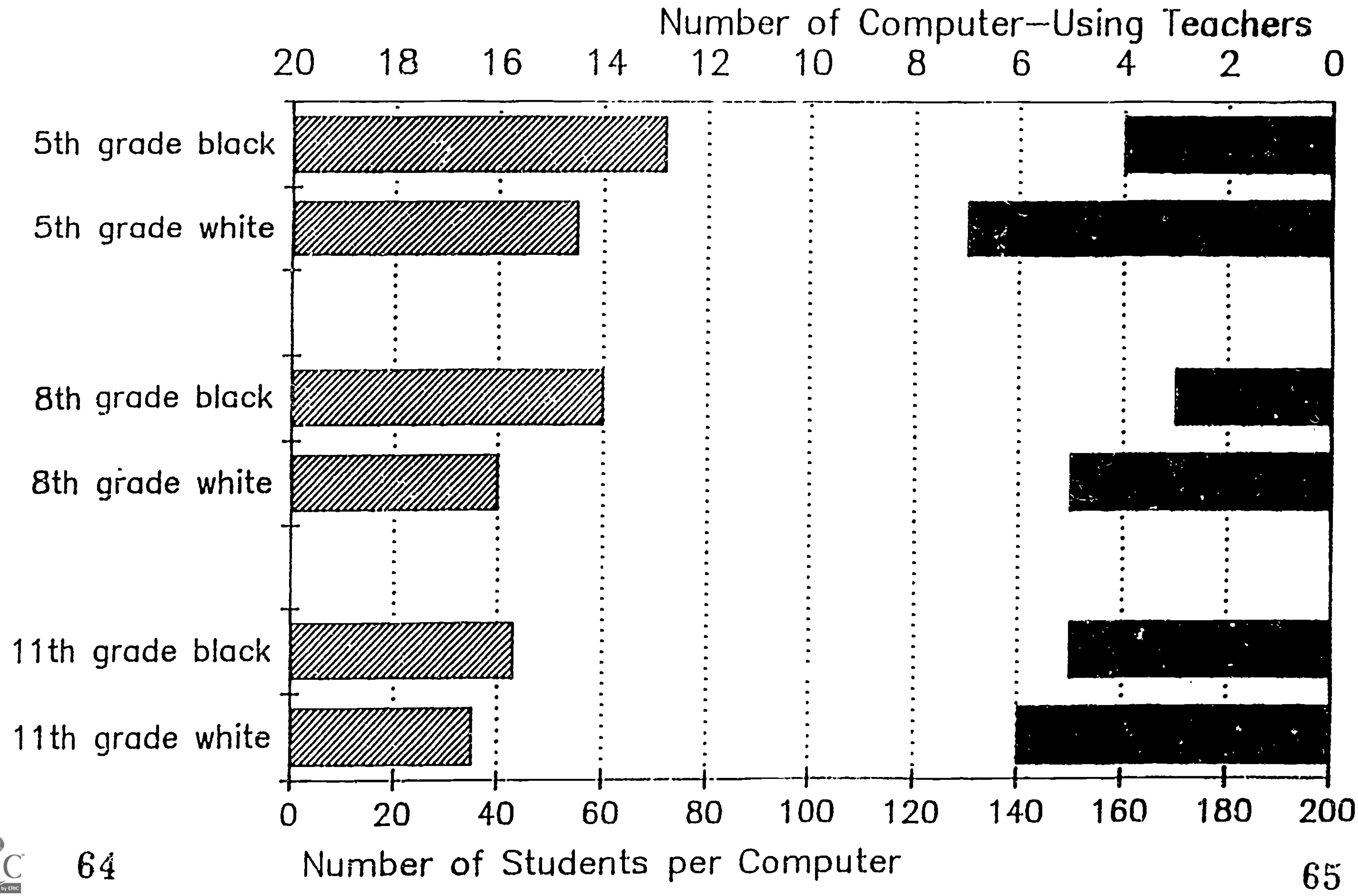


Fig. 3: The Most Active Users, by Sex and Academic Ability

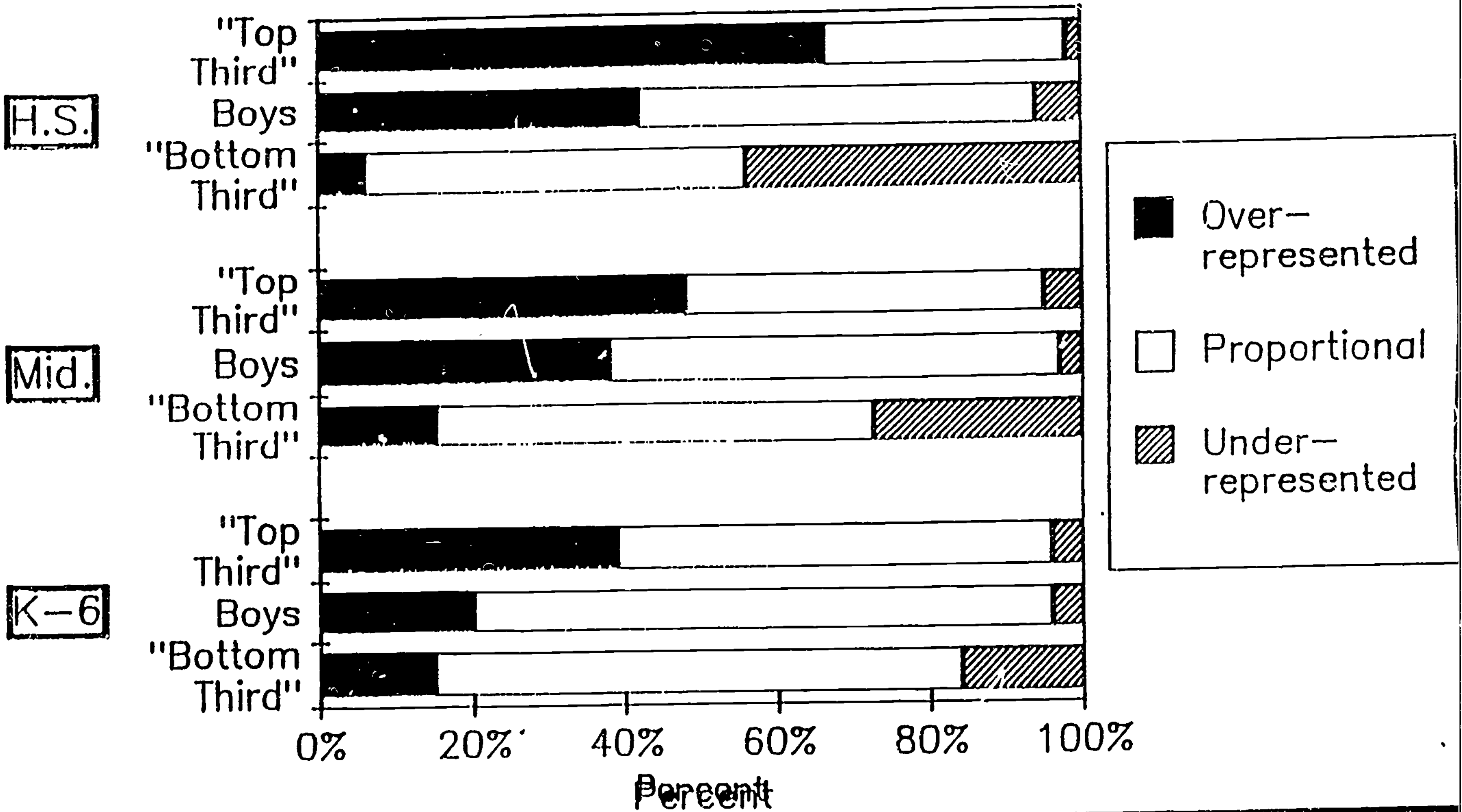
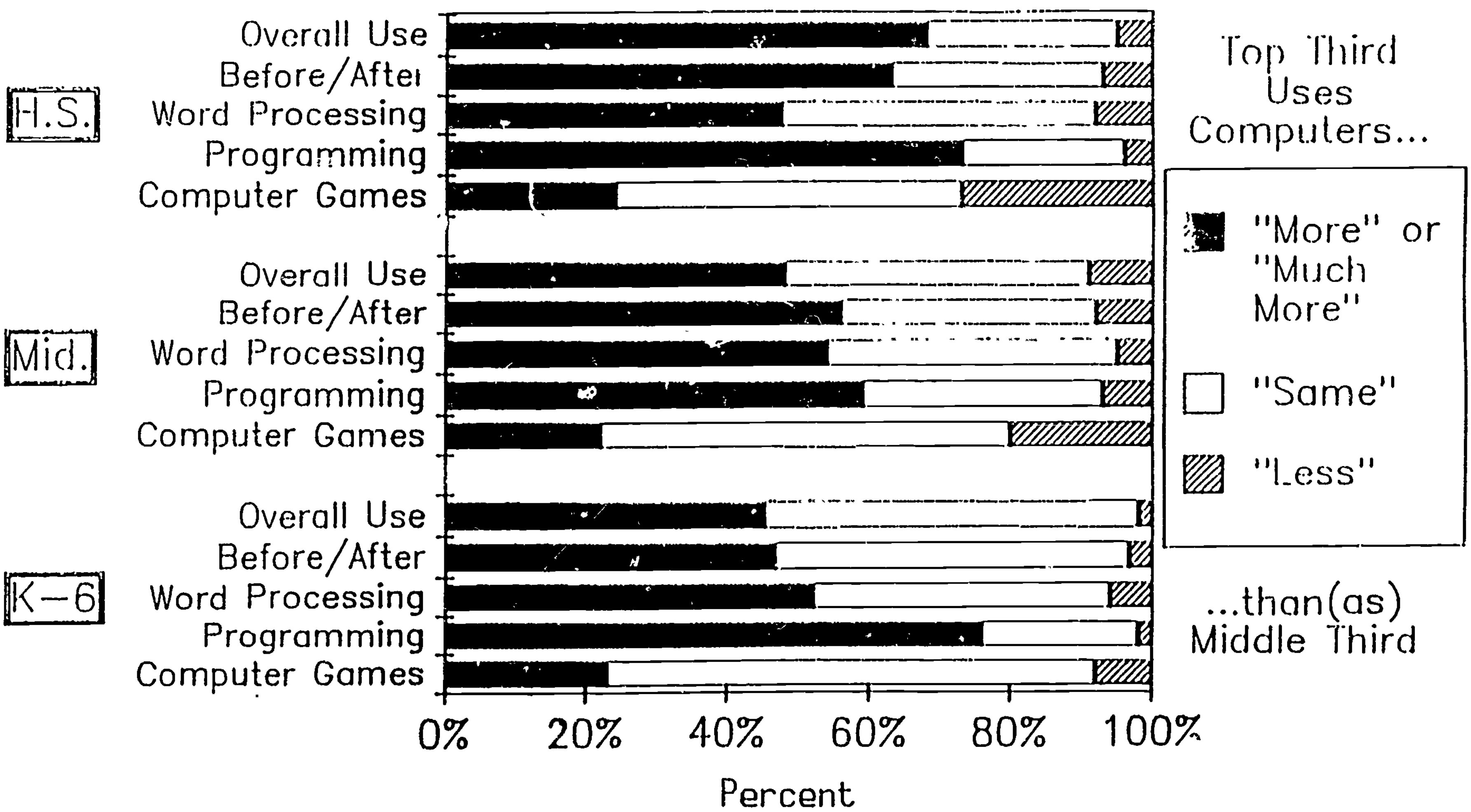


Fig. 4: Computer Use: Top Third vs. Average Students



Students Most Affected by Using Computers, by Grade Level and Ability

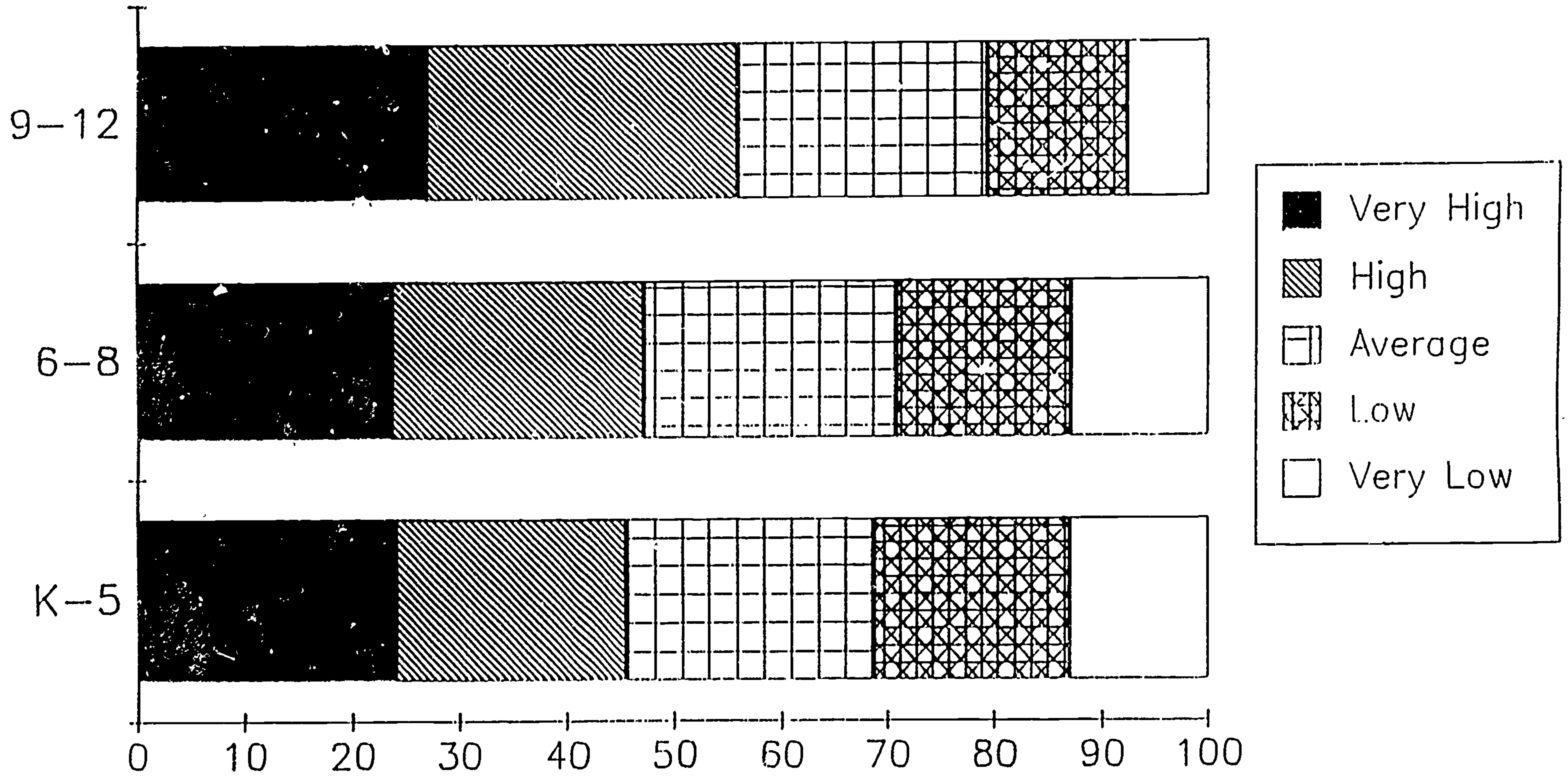


Fig. 5: How Most-Affected Students Were Helped
(by Grade Level & Ability)

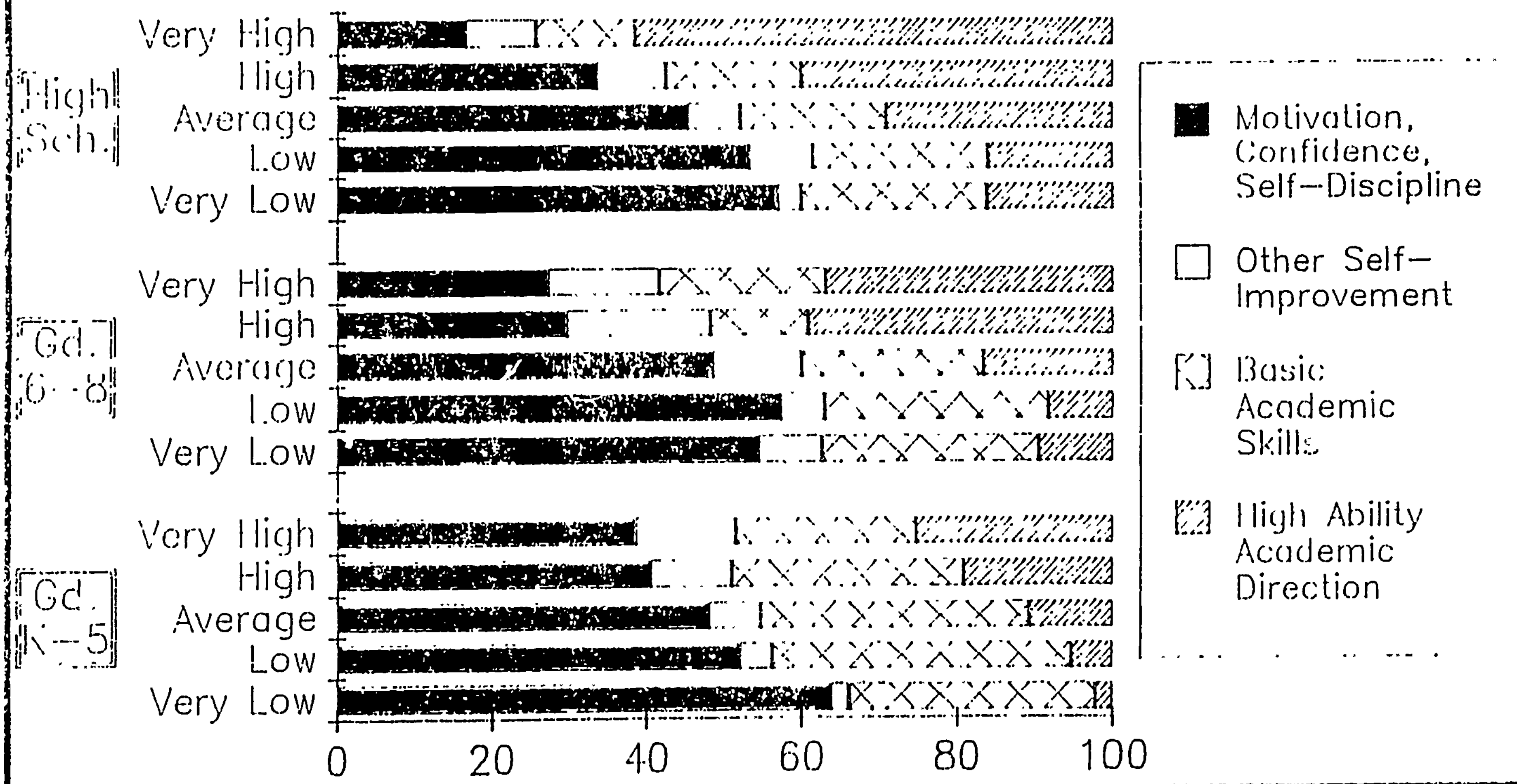


Table 3: Distribution of Computer Use, by Subject and Ability Level, Upper-Elementary through High School, Spring, 1985.*

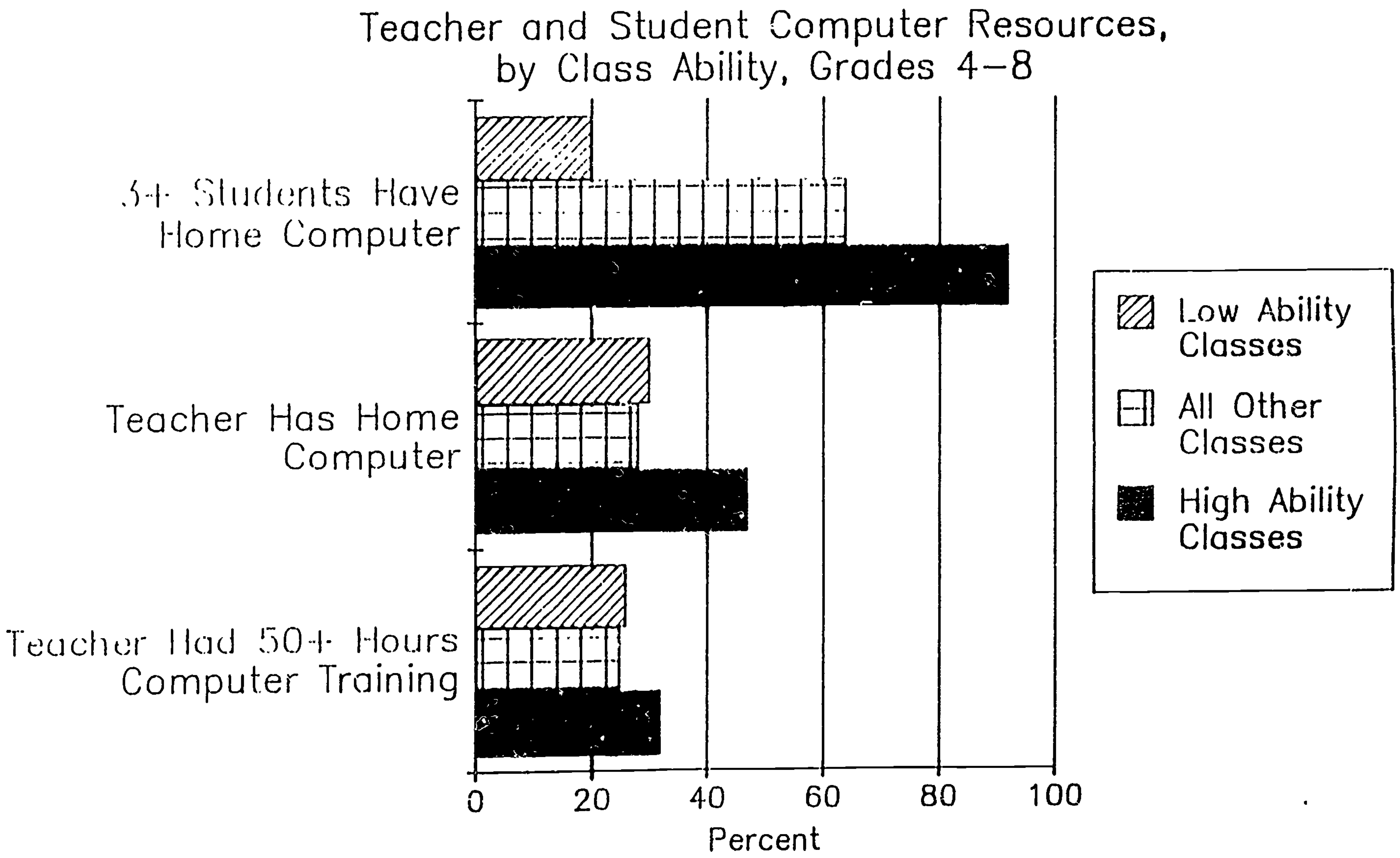
SUBJECT (from course title, activity, and software in use)	GRADE AND ABILITY LEVELS OF STUDENTS IN CLASS					
	Grades 4-8			Grades 9-12		
	Low(29%)	Mixed(59%)	High(12%)	Low(13%)	Mixed(78%)	High(9%)
Mathematics**						
Topics below algebra or unspecified math	36%	27%	9%	32%	1%	0%
Algebra, geometry	2%	1%	3%	1%	1%	2%
Trig,. advanced math	0%	0%	0%	0%	1%	3%
Subtotal: traditional math subjects	30%	28%	12%	32%	3%	5%
English						
Language arts and spelling	20%	10%	4%	16%	2%	1%
Reading	19%	5%	2%	13%	0%	2%
Subtotal: language arts	39%	15%	6%	29%	2%	3%
Writing	5%	1%	6%	4%	2%	9%
Word processing in English class	1%	3%	4%	5%	1%	5%
Subtotal: writing	6%	4%	10%	9%	3%	14%
Computers and Problem-Solving						
Math topics: problem-solving, logo, programming activities	4%	6%	10%	3%	1%	2%
Programming*** as specific topic or course	2%	14%	25%	4%	50%	32%
Logo as specific topic****	3%	6%	8%	2%	1%	1%
Computer literacy as specific topic or course	4%	12%	18%	6%	5%	4%
Subtotal: computers & problem-solving	13%	38%	61%	15%	57%	40%
Business and Word Processing						
Business, accounting, secretarial, other than word processing	0%	1%	0%	0%	7%	2%
Word processing, other than in English	2%	1%	2%	6%	13%	11%
Subtotal: business	2%	2%	2%	6%	20%	13%
Other Subjects						
Science and nutrition	1%	5%	2%	2%	6%	18%
Social studies	0%	6%	6%	3%	1%	3%
Industrial arts and agriculture	0%	1%	0%	2%	7%	0%
Others	2%	3%	3%	3%	2%	3%
Subtotal: other subjects	3%	14%	10%	9%	15%	24%
Total for All Subjects	100%	100%	100%	100%	100%	100%

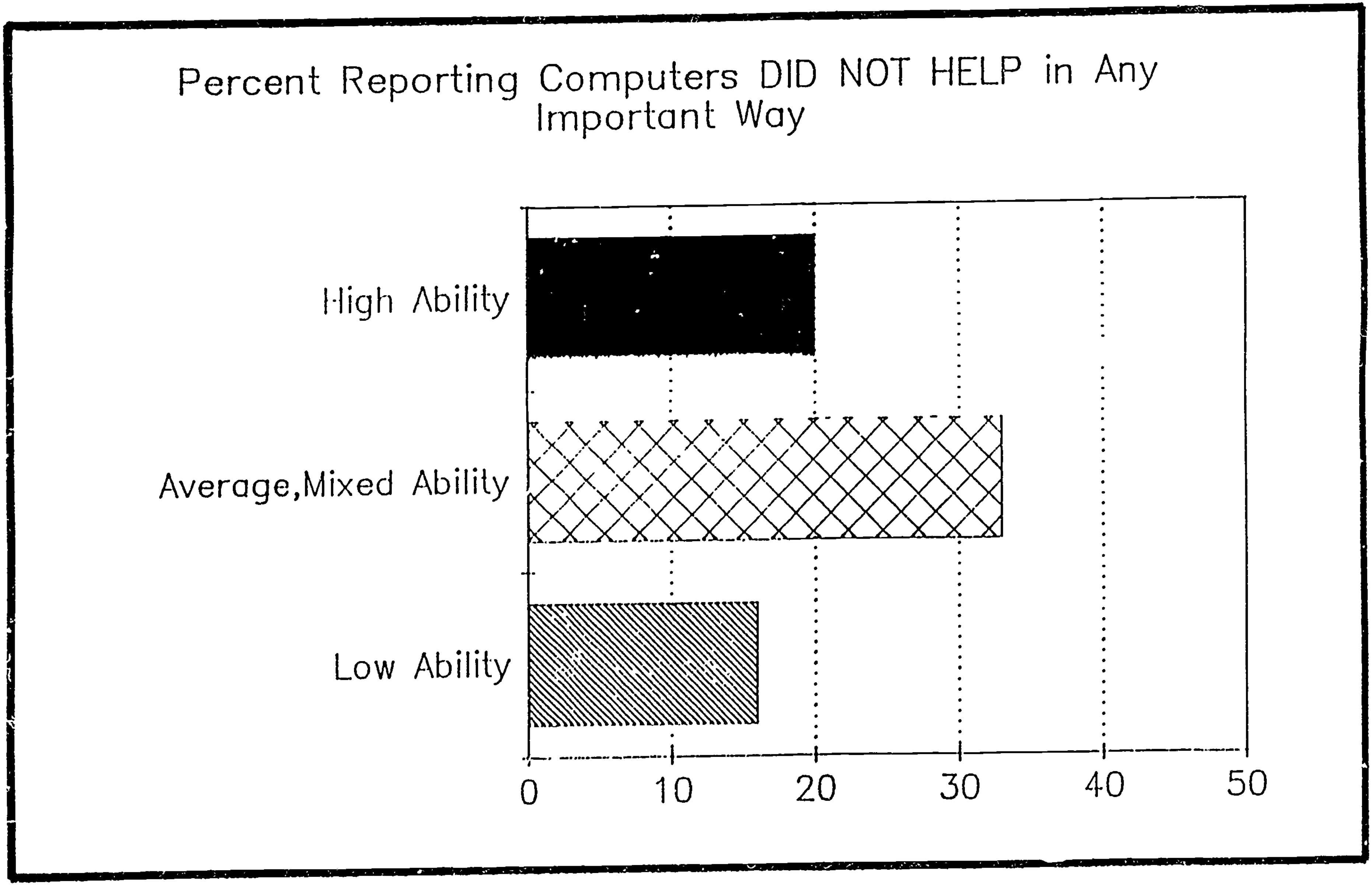
*Sum of individual entries may not equal subtotal entry because of rounding. Similarly, subtotals may not sum to grand total.

**See also first entry under 'Computers.'

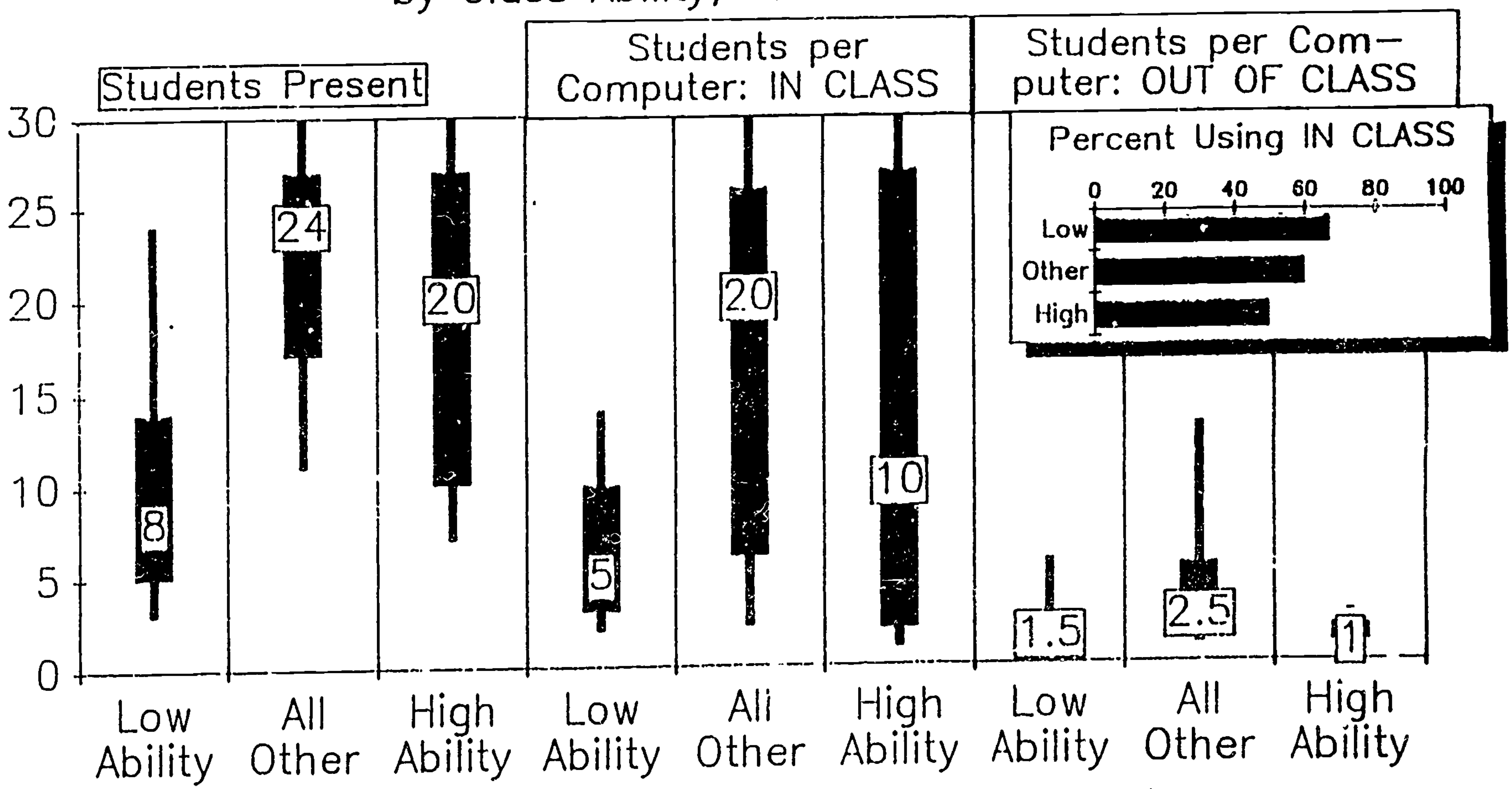
***Excludes specific mentions of Logo.

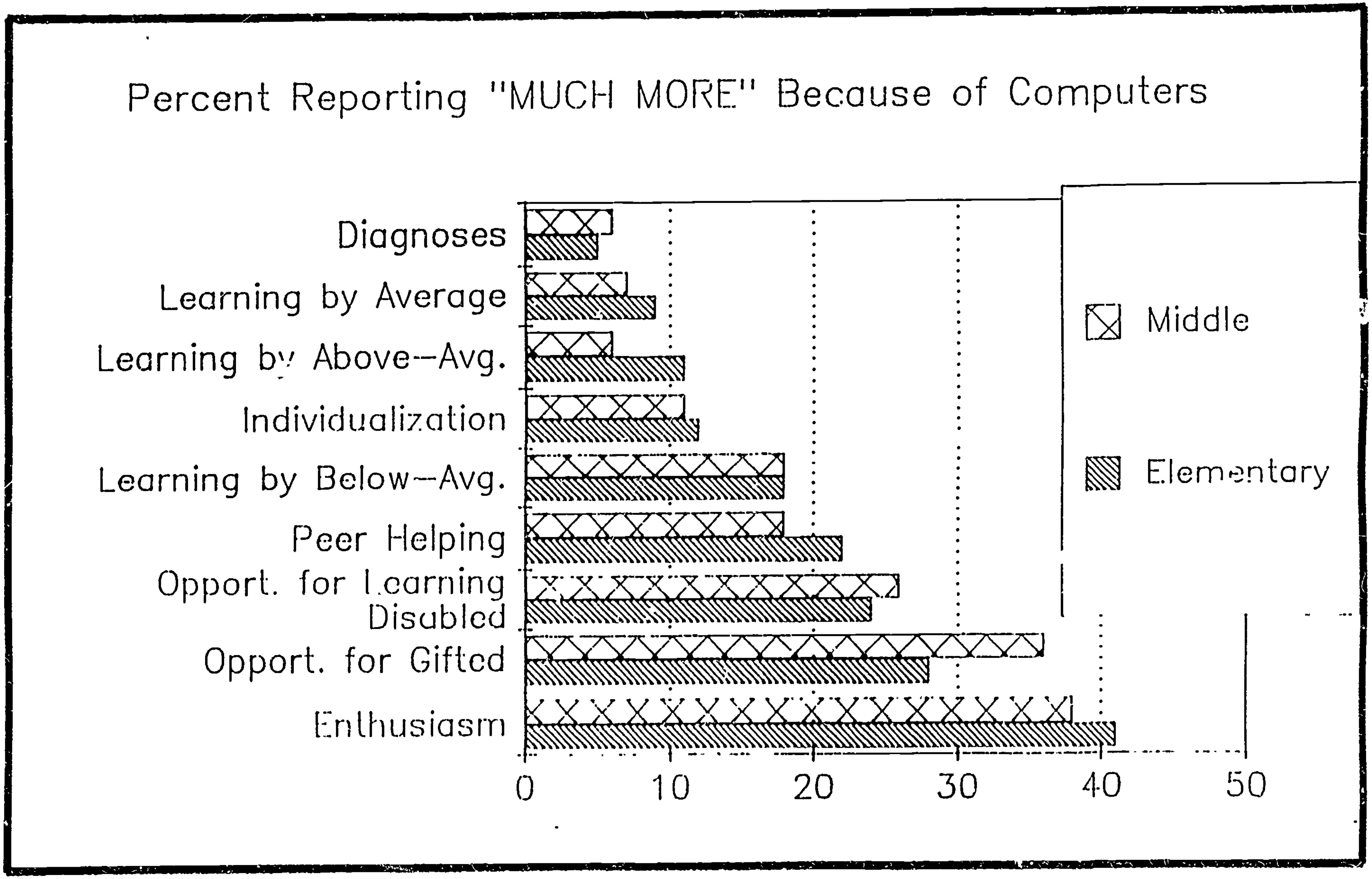
****Includes some general problem-solving not classified elsewhere.


Teacher and Student Computer Resources, by Class Ability, Grades 4–8
5+ Students Have Home Computer
Teacher Has Home Computer
Teacher Had 50+ Hours Computer Training
0
20
40
60
80
100
Percent
Low Ability Classes
All Other Classes
High Ability Classes


Percent Reporting Computers DID NOT HELP in Any Important Way
High Ability
Average,Mixed Ability
Low Ability
0
10
20
30
40
50

Students in Class and Student-to-Computer Ratios by Class Ability, Grades 4-8




Percent Reporting "MUCH MORE" Because of Computers
Diagnoses
Learning by Average
Learning by Above-Avg.
Individualization
Learning by Below-Avg.
Peer Helping
Opport. for Learning Disabled
Opport. for Gifted
Enthusiasm
0
10
20
30
40
50
Middle
Elementary

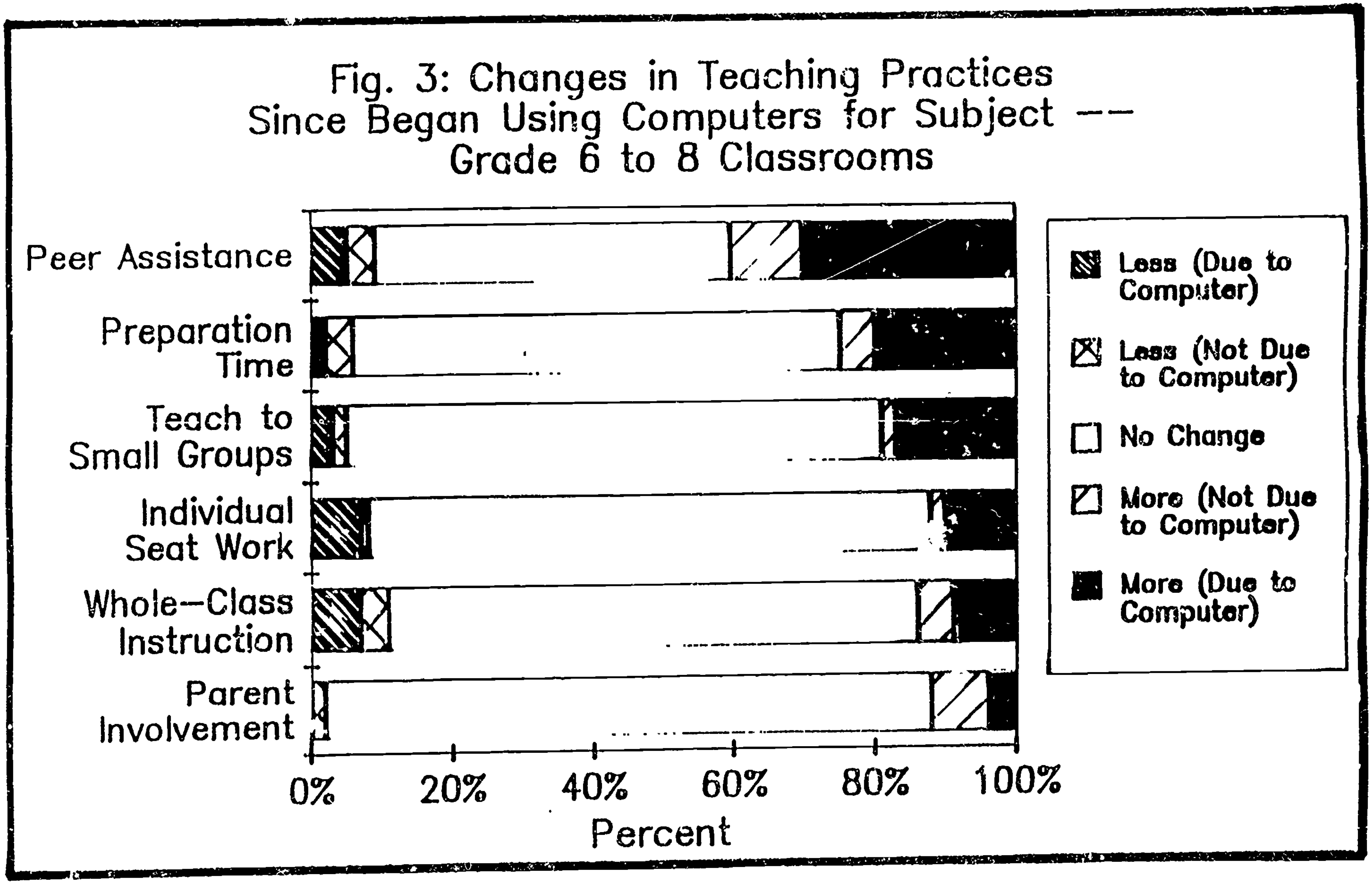


Fig. 3: Changes in Teaching Practices
Since Began Using Computers for Subject --
Grade 6 to 8 Classrooms
Peer Assistance
Preparation Time
Teach to Small Groups
Individual Seat Work
Whole-Class Instruction
Parent Involvement
0%
20%
40%
60%
80%
100%
Percent
Less (Due to Computer)
Less (Not Due to Computer)
No Change
More (Not Due to Computer)
More (Due to Computer)

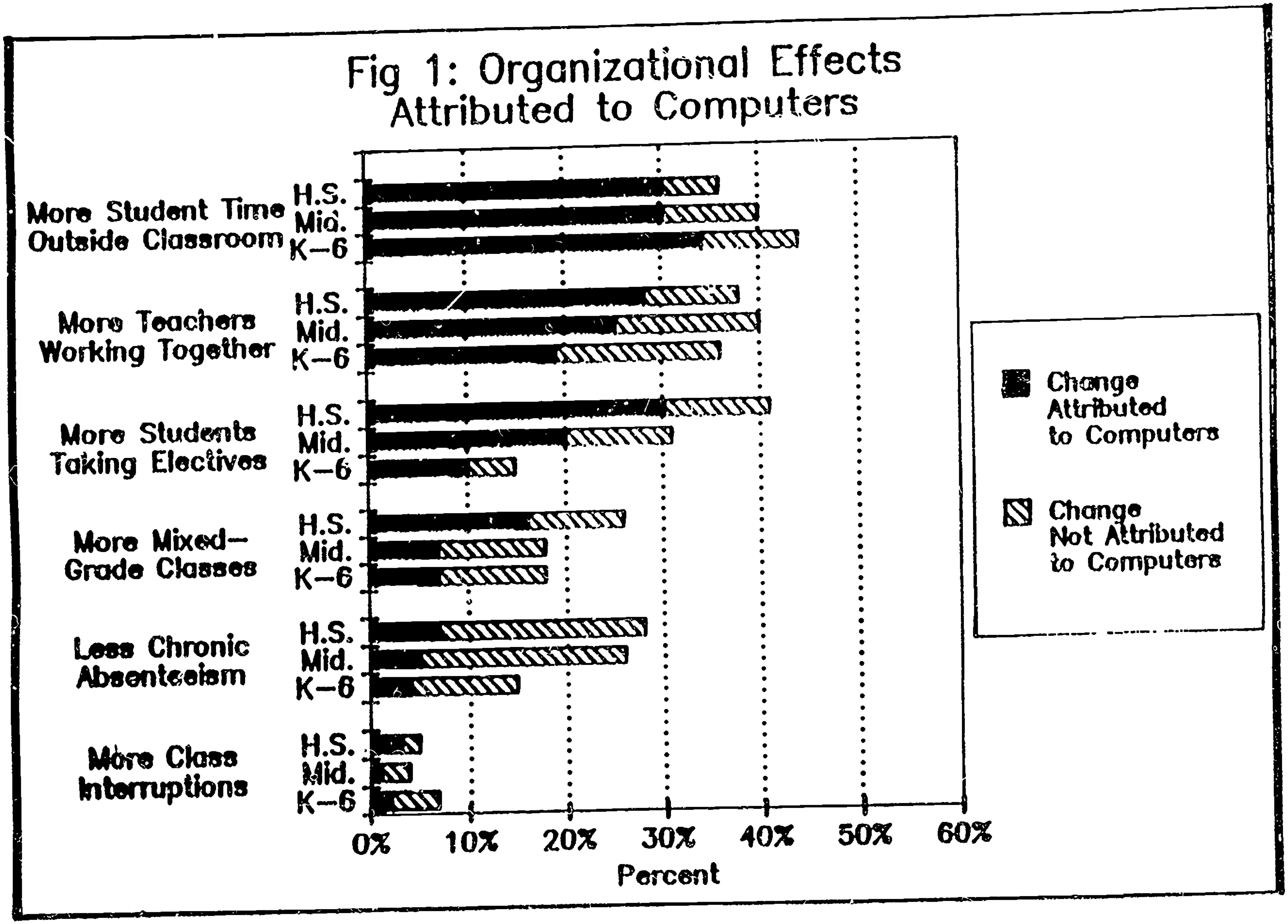


Fig 1: Organizational Effects Attributed to Computers
More Student Time Outside Classroom
More Teachers Working Together
More Students Taking Electives
More Mixed-Grade Classes
Less Chronic Absenteeism
More Class Interruptions
H.S.
Mid.
K-6
0%
10%
20%
30%
40%
50%
60%
Percent
Change Attributed to Computers
Change Not Attributed to Computers

Solutions to Which Three Problems Would Most Improve Computer Use at Your School

Problem	Percent
Not Enough Computers	62%
No Time to Develop Computer-based Activity	50%
Teachers Lack Interest in Learning about Computers	41%
Software Too Costly in Quantity Needed	35%
Not Enough Money for Teacher Training	30%
Hard to Locate or Evaluate Software Already Written	27%
Poor Quality of Software	18%
Software Not Yet Written for Topics for which it is Needed	18%
Other	9%

Capsule Summary