DOCUMENT RESUME

ED 293 274 EC 202 473

AUTHOR Fuchs, Douglas; Fuchs, Lynn S.
TITLE An Evaluation of the Adaptive Learning Environments Model.
PUB DATE Apr 87
NOTE 40p.; A version of this paper was presented in a symposium at the Annual Meeting of the American Educational Research Association (Washington, DC, April 1987).
AVAILABLE FROM Douglas Fuchs, Box 328, Department of Special Education, George Peabody College, Vanderbilt University, Nashville, TN 37203.
PUB TYPE Information Analyses (070) -- Viewpoints (120) -- Speeches/Conference Papers (150)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Academic Achievement; Classroom Environment; *Disabilities; Elementary Secondary Education; *Evaluation Methods; Instructional Effectiveness; *Mainstreaming; Models; *Program Effectiveness; *Program Evaluation
IDENTIFIERS *Adaptive Learning Environments Model; Regular and Special Education Relationship

ABSTRACT

The paper examined empirical evidence (published through 1986) concerning the effectiveness of the Adaptive Learning Environments Model (ALEM), a multifaceted program designed to integrate handicapped pupils and remedial and compensatory education students in general education. Evaluation of three evaluative studies by the program's originators claiming to support ALEM effectiveness resulted in the following conclusions: (1) it is unknown whether ALEM can be implemented in classroom settings differing in aims, needs, and contextual characteristics; (2) it is probably not true that hypothesized patterns of classroom process occur when ALEM dimensions are in place; and (3) it is unclear whether implementation of the ALEM leads to important improvement in student academic performance. It is concluded that currently there is insufficient cause to view ALEM as a successful, large-scale, full-time mainstreaming program. Results have implications for implementation of the General Education Initiative which has assumed scientific validity of ALEM as an example of large scale mainstreaming. (DB)

An Evaluation of the Adaptive Learning Environments Model

Douglas Fuchs and Lynn S. Fuchs

George Peabody College

Vanderbilt University

A version of this paper was presented in a symposium at the annual meeting of the American Educational Research Association, Washington, DC, April 1987.

Requests for reprints should be sent to Douglas Fuchs, Box 328, Department of Special Education, George Peabody College, Vanderbilt University, Nashville, TN 37203.

Running head: Evaluation of the ALEM

Abstract

Perhaps the most multifaceted and visible ongoing program to integrate handicapped pupils and remedial and compensatory education students in general education is the Adaptive Learning Environments Model (ALEM). Despite frequent expressions of support for the effectiveness of the ALEM, none has been based on an independent, systematic, and comprehensive review of empirical evidence. The purpose of this paper was to present such a critique. An evaluation of empirical evidence published in journal articles and a book chapter through 1986 indicated insufficient cause to view the ALEM as a successful, large-scale, full-time mainstreaming program. Implications of this finding are discussed in terms of the desirability of a proposed merger between special and general education.

An Evaluation of the Adaptive Learning Environments Model

More and more professionals believe major problems exist with service delivery in special education. First, special education is viewed as one of too many categorical programs for students with unique learning and management needs. The proliferation of programs associated with the Education for All Handicapped Children Act, Chapter I, migrant education, and low English proficiency has been characterized by Reynolds, Wang, and Walberg (1987) as "disjointed incrementalism." This refers to what happens when "a series of narrowly framed programs is launched one by one, each program well-justified in its own time and way, but based on the assumption that it does not interact with others" (p. 392).

Results of this disjointedness can be (a) overlapping programs that spawn confusion about which one is most appropriate for a given low-achieving student (Wang, Reynolds, & Walberg, 1985), (b) excessive rules and regulations that decrease school efficiency and increase administrative costs (Wang et al., 1985), and (c) frequent removal, or "pull-out," of pupils from regular classrooms, which represents discontinuity in the instructional-learning process (Will, 1986). According to Pugach and Lilly (1984), such fragmented service delivery has contributed to a "myth of differentness," which exaggerates differences among students, teaching practices, and curricula in special and general education. Belief in these questionable differences, in turn, may contribute to an apparent breakdown in professional relationships in local, state, and federal agencies and institutions of higher education (Stainback & Stainback, 1984).

Paralleling the belief that there are too many programmatic "boxes" for difficult-to-teach pupils is an increasingly negative sentiment toward use of traditional categories of exceptionality, or "boxes" within the "box" of special education. Such compartmentalization frequently is described as exacerbating service delivery problems. Categories of exceptionality are associated closely with use of labels that may (a) fallaciously become explanatory constructs, a circuitous evolution from their intended use as descriptive devices (e.g., Goodman, 1977; Marston, 1987), (b) encourage the erroneous belief that poor school performance and adjustment are attributable solely to characteristics of the student rather than to the quality of the learning environment (e.g., Gottlieb, 1975; Will, 1986), (c) stigmatize some to whom they are applied (e.g., Jones, 1972; Reynolds & Balow, 1972), and (d) reduce teacher expectations (e.g., Reschly, 1979). Moreover, many labels may be assigned incorrectly, since numerous test instruments, checklists, and observation procedures used in the referral-to-placement process lack demonstrated validity for use as classificatory devices (e.g., Fuchs, Fuchs, Benowitz, & Barringer, in press; Ysseldyke & Shinn, 1981).

Yet another related concern is that empirical evidence on the effectiveness of pull-out programs is mixed (e.g., Leinhardt & Pallay, 1982; Semmel, Gottlieb, & Robinson, 1979). Less ambiguous, but equally troubling, are data indicating that skills and knowledge learned in resource rooms often fail to transfer to regular classrooms when handicapped pupils are returned to general education (e.g., Anderson-Inman, in press; Anderson-Inman, Walker, & Purcell, 1984; Wehman, Abramson, & Norman, 1977), increasing the likelihood that placement in special education is a terminal assignment in many handicapped children's educational careers.

Responding to these numerous and serious concerns, Wang et al. (1985) introduced an ambitious two-part initiative, which is often called the General

Education Initiative. First, it involves the "joining of demonstrably effective practices from special, compensatory, and general education to establish a general education system that is more inclusive and that better serves all students, particularly those who require greater-than-usual educational support" (Reynolds et al., 1987, p. 394). Second, the initiative encourages the federal government to collaborate with states and local school districts in supporting experimental trials of integrated forms of education for students who currently are segregated for service in special and remedial education programs. Thus, by merging special, remedial, and general educators' expertise, and incorporating many special and remedial education resources under the aegis of general education, the Wang et al. plan aims to facilitate new partnerships in education and enhance classroom teachers' capacity to accommodate diverse groups of students.

The Adaptive Learning Environments Model

Perhaps the most multifaceted and visible ongoing effort to integrate handicapped pupils and remedial and compensatory education students in the regular classroom is Wang's Adaptive Learning Environments Model (ALEM). The ALEM evolved from an individualized approach to instruction developed more than a decade ago by Glaser and associates at the University of Pittsburgh's Learning Research and Development Center (LRDC; see Wang & Stiles, 1976, pp. 172-174, for a description of this model). As with the LRDC Instructional Model, the ALEM's overall goal is to "provide effective school environments that maximize the outcomes of learning for individual children -- environments where each child can effectively master basic skills in academic subjects...while becoming confident in his or her ability to learn and to cope with the social and physical classroom surroundings" (Wang, 1980, p. 126).

More specifically, according to Wang (1980), the ALEM includes (a) a prescriptive learning component comprising a series of hierarchically organized curricula for basic skills development; (b) a more open-ended exploratory learning component; (c) classroom management procedures to facilitate implementation of both the prescriptive and exploratory components and an organizational plan to maximize use of available resources; (d) a family involvement program; and (e) a multi-age and team-teaching organization to increase flexibility in use of teacher and student resources.

Whereas the LRDC Instructional Model typically was implemented in inner-city schools serving low-income minority pupils, the ALEM has been implemented in more varied settings, including rural and suburban schools with middle-class, gifted, and, most notably, handicapped students. Early versions of the ALEM incorporated specialized reading and math curricula developed in the LRDC. Wang now seeks activities and materials for the ALEM that are more compatible with extant school programs, hoping to increase chances of successful implementation (N. Zigmond, personal communication, June, 1987).

In the professional literature, one frequently finds statements, some cautious and some not, supporting the efficacy of the ALEM as a full-time, large-scale mainstreaming program (e.g., Ammer, 1984; Biklen & Zollers, 1986; Haggerty & Abramson, 1987; Maheady, Towne, Algozzine, Mercer, & Ysseldyke, 1983; Mcdowell, 1986; Nevin & Thousand, 1986; Prasse & Reschly, 1986; Stainback & Stainback, 1984; Stainback, Stainback, Courtnage, & Jaben, 1985; Zane, 1987). Such pronouncements create an impression that educators _right now_ possess the technical know-how to transform regular classrooms into settings in which most handicapped students may be productive and socially accepted. Given this upbeat and apparently popular perspective, it is no wonder that some special educators confidently and impatiently exhort us to "move off the dime" and work toward a merger between special and general education. For example, Stainback and Stainback (1984) write that a dual system is no longer needed, and that "the time has arrived for special and regular education to merge into one unified system structured to meet the unique needs of all children" (p. 102).

Purpose

Despite frequent expressions of support for the effectiveness of the ALEM as a full-time, large-scale mainstreaming program, none has been based on an independent, systematic, and comprehensive review of empirical evidence. The purpose of this paper is to present such a critique. Its importance extends beyond its potential simply to uphold or question the validity of the model. The value of the critique also derives from the fact that, as a perceived success, the ALEM tends to legitimize the claim that a special education-general education merger currently is feasible on technical grounds. Thus, the outcome of this evaluation should hold relevance for how one views both the effectiveness of the ALEM and the practicality and desirability of an immediate special education-general education merger, as some have proposed.

Method

Literature Search

To establish reasonable parameters for this review, we delimited our effort in four important ways. First, we included only evaluations of the ALEM, distinguishing such studies from those of the ALEM's predecessor, the LRDC Instructional Model. Second, we excluded investigations of isolated components of the ALEM such as the Self-Schedule System (see Wang & Stiles, 1976); rather, we sought to include only evaluative studies in which the ALEM had been implemented fully. Third, we did not include unpublished reports such as LRDC publications or ERIC documents. Finally, our search did not extend beyond 1986. Thus, we may have omitted pertinent investigations that are unpublished and/or were completed after 1986.

We began the search with a convenience sample of 14 papers written by Wang and associates on the ALEM and related topics. These 14 papers were either published as journal articles or book chapters between 1976 and 1985 inclusive. Next, among these papers, references to additional published work were collected. We also conducted a computer search of the Current Index of Journals in Education (CIJE) database for 1982 to 1986 inclusive, using the descriptors "adaptive-learning-environments-model" and "adaptive-learning-environments-program" and Wang's first and last name. Pertinent documents to emerge from this search were reviewed. Finally, a hand search was made of all issues of Exceptional Children published from 1982 through 1986. References in articles identified by this search also were obtained.

We found four studies exploring the effectiveness of the ALEM. The first, a pilot investigation involving one classroom during 1977-78, was reported in The Elementary School Journal (Wang, 1981). A second study was implemented during the 1980-81 school year, and described in five journals: American Educational Research Journal (Wang & Walberg, 1983a); Curriculum Inquiry (Wang, Nojan, Strom, & Walberg, 1984); Educational Evaluation and Policy Analysis (Wang & Walberg, 1983b), Exceptional Children (Wang & Birch, 1984a); and Teacher Education and Special Education (Wang & Gennari, 1983). The third study, also conducted in 1980-81, was reported in Exceptional Children (Wang & Birch, 1984b), and the fourth investigation, implemented in 1982-83, was described in Educational Leadership (Wang, Rubenstein, & Reynolds, 1985), Exceptional Children (Wang & Reynolds, 1985), Remedial and Special Education (Wang, Peverly, & Randolph, 1984), and TEACHING Exceptional Children (Wang, Vaughan, & Dytman, 1985). The last three implementations also were described in a book chapter by Wang, Gennari, and Waxman (1985).

Analysis of Data-Based Articles

Among the four investigations of the ALEM, the pilot study, reported in The Elementary School Journal (Wang, 1981), was eliminated from further consideration because of its very modest scope. In the remaining studies, three basic questions guided evaluation of the model (see, for example, Wang, Gennari, & Waxman, 1985):

1. Can a high degree of implementation of the ALEM be attained in classroom settings that differ in terms of aims, needs, and contextual characteristics? In other words, is there evidence for the model's implementability or feasibility in varying school settings?

2. When the dimensions of the ALEM are in place, do hypothesized patterns of classroom process occur?

3. Does implementation of the ALEM, and resulting patterns of classroom process, lead to important improvement in student academic performance?

The salience of determining the ALEM's degree of implementation flows from a reasonable belief that the ALEM can be connected to measured student performance only if evidence of program implementation is provided (see Wang & Walberg, 1983b). Wang judges degree of implementation with The Implementation Assessment Battery for Adaptive Instruction (Wang cited in Wang & Birch, 1984a). The Battery consists of 96 performance indicators that assess whether, and if so to what extent, 12 classroom dimensions are present. When 85% or more of the performance indicators in a given dimension are present, the degree of implementation of that dimension is considered "high;" 50% to 84% and less than 50% of these indicators signals an "average" and "low" degree of implementation, respectively (Wang & Gennari, 1983).

Eight dimensions presumed critical to effective adaptive instruction are creating and maintaining instructional materials; record keeping; diagnostic testing; prescribing; monitoring and diagnosing; interactive teaching; instructing; and motivating. An additional four factors address classroom management and support. They are arranging space and facilities; establishing and communicating rules and procedures; managing aides; and developing student self-responsibility (see Wang & Birch, 1984; Wang, Nojan, Strom, & Walberg, 1984).

The Student Behavior Observation Schedule (SBOS; Wang, 1976) is employed to obtain information on five major categories of classroom process: (a) nature of interactions between students and teachers; (b) nature of peer interactions; (c) settings in which learning activities occur; (d) types of tasks or activities on which students work; and (e) manner in which students spend class time (see Wang & Birch, 1984a). Finally, student academic achievement typically is measured by standardized achievement tests in reading and math.

Since Wang and colleagues have explored the effectiveness of the ALEM in terms of degree of implementation, classroom process, and student achievement, these three evaluative dimensions constitute the focus of our critique and represent the structure for its presentation.

Results

Study 1 (1980-81)

Participants. A total of 138 general educators participated in this study. They taught 156 kindergarten through third grade regular classrooms, containing 35 handicapped and 17 gifted pupils. The teachers represented 10 school districts, of which 6 and 4 were participating in the National Follow Through Program (Follow Through) and Handicapped Children's Model Program (Mainstream), respectively. Wang and associates did not statistically analyze possible differences between Follow Through and Mainstream districts, although

Table 2 on p. 609 of Wang and Walberg (1983a) indicates they differed in two potentially important ways.

First, five of six Follow Through districts were described as "rural," "semirural," or "urban/rural," whereas all four Mainstream districts were characterized as suburban. Second, the median percentage of Title I eligible students in Follow Through districts was more than twice as large as the corresponding percentage in Mainstream districts (26% vs. 12%). Additionally, Mainstream sites, by definition, served handicapped children, while it is unclear whether Follow Through districts did so. Perhaps most importantly, "certain components of the ALEM (were) in operation at some of the Follow Through sites for nearly 10 years" (Wang & Walberg, 1983a, p. 609), and, in at least four of six Follow Through districts, an integrated version of the ALEM had been implemented in 1979, one year before it was implemented in Mainstream sites participating in this study (see Wang & Walberg, 1983a, p. 615). Thus, in contrast to Follow Through school districts, Mainstream sites seem to have had no prior experience with the ALEM and served handicapped pupils from wealthier, suburban families.

Implementation. Wang and Birch (1984a) write, "Of 138 teachers included in the study across the 10 districts, 96.4% were found to have either an average or high degree of implementation" (p. 395). Referring to this finding, Wang and Walberg (1983a) conclude, "These results suggest that the 12 critical dimensions of the ALEM can be implemented with a high degree of treatment fidelity in many classes that include poor and handicapped students and are located in schools with varying characteristics and constraints" (p. 611). This conclusion, however, is based on an undifferentiated aggregation of implementation data from all 10 school districts, which ignores apparent differences between Follow Through and Mainstream districts, and possibly masks systematic differences between the two sets of districts in degree of implementation.

To explore a possible relation between type of school district and degree of implementation, we employed Wang, Nojan, Strom, and Walberg's Table 3 (1984, p. 269), which displays percentages of classrooms in each of the 10 districts achieving high, average, and low levels of implementation. By transforming this table into a three (High vs. Average vs. Low implementation) x two (Follow Through vs. Mainstream) contingency table, we discovered that 53 (45.30%) of the classrooms in Follow Through districts evidenced a high degree of implementation, whereas only 2 (9.52%) of the classes at Mainstream sites did so (see Table 1 of this article). A chi-square analysis indicated degree of implementation was indeed significantly related to type of district, $\chi^2(2, \underline{N} = 138) = 9.54$, $\underline{p} < .01$, strengthening the conclusion that classrooms in Follow Through districts were more likely than those at Mainstream sites to achieve a high degree of implementation of the ALEM.

Insert Table 1 about here

Does this mean school sites that accommodate handicapped pupils from wealthier suburban families, like the Mainstream districts, will be less likely to implement the model with a high degree of fidelity, and, more generally, that the ALEM is not equally implementable across different settings? Not necessarily, since Mainstream districts, in contrast to Follow Through sites, had no prior experience with the ALEM, and, on this basis alone, could have been expected to show lower degrees of implementation. Our inability to make further sense of the implementation data reflects the absence of analysis that (a) stratifies extent of implementation on dimensions like proportions of handicapped pupils at participating school sites and (b) controls for potentially confounding factors such as the amount of previous experience districts had with the ALEM. Without such analysis, we believe it is premature and potentially misleading to state that, "adaptive instruction (e.g., the ALEM) can be established and effectively maintained in many diverse school sites that include students from poor families and students with special learning needs" (Wang & Walberg, 1983a, pp. 621-622).

Classroom processes. Before reviewing these data, there are at least three methodological issues warranting discussion. Each concerns the SBOS, which is used to obtain data on classroom processes. First, Wang and associates state that interrater agreement for the SBOS is consistently above 85%. This figure appears to represent an average across the various categories constituting the observation system. We have found only one published document that presents interrater agreement for each SBOS category (see Wang, 1976). In her Table 1, Wang (1976, pp. 371-372) indicates that 9 of 35 categories (26%) were associated with levels of agreement below 85%, and, for six variables, percentage of agreement ranged between .48 and .68. This information seems important, since Wang and colleagues report their SBOS data at the category, rather than overall or summary, level. Furthermore, the 35 categories listed by Wang (1976) represent more than twice the number of variables constituting the SBOS version used in this study. Wang and associates, it appears, have melded many of the original categories, creating qualitatively different ones. However, we failed to find percentages of interrater agreement for these new categories.

Second, in this study, SBOS data were collected only in first and second grade classes, representing 72 of 138 participating teachers. Of these 72 classrooms, 61 (85%) and 11 (15%) were in Follow Through and Mainstream sites, respectively. Thus, SBOS data heavily reflect dynamics in settings in which teachers and students were comparatively experienced with the ALEM. Such data probably should not be viewed as (a) indicative of the types of classroom changes to expect during a first year of implementation, and (b) a "conservative" test of hypothesized relations between implementation of the model and classroom processes. Finally, students in the 72 classrooms were observed on one occasion for five consecutive 1-minute intervals, raising the question whether such an abbreviated sample of behavior constitutes meaningful indication of a student's typical classroom behavior.

To investigate whether degree of implementation is related to anticipated changes in classroom dynamics, SBOS data were examined for classrooms displaying high, average, and low levels of ALEM practice. It appears that multiple (and seemingly interdependent) one-way analyses of variance (ANOVAs) were used to compare the three types of classrooms on each and every SBOS category separately. Despite Wang and associates' statement that, "Significant differences were observed in the patterns of classroom processes among the three degree of implementation groups" (Wang, Nojan, Strom, & Walberg, 1984, p. 271), we did not find one F or p value associated with these ANOVAs. We conclude that the Wang et al. statement was based on the one inferential statistic we could find related to SBOS data; namely, a canonical correlation ($r = .36$, $p < .01$), which reportedly indicated only a nonspecific relation between "an ungrouped set of implementation measures and classroom processes" (Wang & Walberg, 1983a, p. 614).

An examination of descriptive, rather than inferential, differences between high, average, and low implementation classrooms reveals a similar story. High compliance classrooms tended to demonstrate "desirable" classroom processes more often than average and low implementation classes, and these lower implementation classrooms exhibited "undesirable" dynamics more frequently than the highest compliance settings. Yet, a majority of these differences are of questionable educational importance. For example, as reported by Wang and Walberg (1983a), the frequency with which teacher-student instructional interactions occurred in high, average, and low implementation classes was 93.3%, 91.7%, and 90.0%, respectively (see Table 5, p. 615). Thus, despite the fact that SBOS data were collected in a majority of ALEM-experienced settings, high, average, and low compliance classes seem indistinguishable in terms of their respective process characteristics, and this comparability appears to be signaled by inferential and descriptive statistics.

Student achievement. Two types of reading and math achievement data were collected: (a) number of objectives mastered in the ALEM's curriculum and (b) percentile ranks attained on standardized achievement tests. Wang, Nojan, Strom, and Walberg (1984) state the number of objectives mastered was obtained from all 138 participating teachers (see their Table 5, p. 273); Wang and Birch (1984a), on the other hand, indicate these data were collected from only 72 teachers (see p. 396), presumably those in whose classrooms the SBOS data were generated. Regarding the standardized achievement data, Wang and Walberg (1983a) make clear these scores were generated on students from only 4 of 6 Follow Through districts (see p. 611). Thus, we are uncertain whether Wang and associates obtained the number of objectives mastered on handicapped pupils; we are more confident they did not collect standardized achievement data from handicapped students.

Wang, Nojan, Strom, and Walberg (1984) display the achievement data for high, average, and low implementation classrooms in their Table 5 (p. 273), which is reproduced here as Table 2. Descriptively, a consistent trend favored higher implementation classrooms. However, no disparity seems educationally important. For example, high, average, and low implementation classrooms mastered an average 29.77, 28.97, and 20.21 math objectives, respectively; their percentile scores in math were 53.49, 52.16, and 48.00, respectively. Furthermore, as indicated in a footnote to this Table, no significant differences separated types of classroom on numbers of objectives mastered or percentile ranks attained. The failure of pupils in higher compliance classes to distinguish themselves on standardized achievement tests may be explained on grounds that the tests did "not closely match the goals of the school districts or the ALEM" (Wang & Walberg, 1983a, p. 611). It appears more difficult to explain an absence of reliable and important between-classroom divergence on objectives mastered, since this index, unlike the percentile ranks, seems to possess a high degree of instructional validity (see McClung cited in Yalow & Popham, 1983).

Insert Table 2 about here

Wang and Walberg (1983a) approached the standardized achievement data from another angle. Since improvement in achievement was expected as pupils spent more time in ALEM settings, Wang and Walberg compared reading and math percentiles of pupils in the ALEM kindergartens and first and second grades in Spring 1980 to the same groups' scores one year later. Wang and Walberg write, "As expected, comparisons...showed increases in mean percentile ranks in reading of from 7 to 9 points" (p. 616). Their Figure 1 (p. 617) indicates results were less consistent for math. It also reveals that 1766 scores were obtained in Spring 1980, whereas data were collected on only 1350 students in Spring 1981, a 1-year attrition rate of 24%. Since Wang and associates did not use a control group, a plausible alternate explanation for higher reading scores in Spring is that more capable students from comparatively stable homes maintained enrollment in the ALEM schools; less capable or less motivated students from more transient families left. The legitimacy of this "competing hypothesis" receives support from Seitz, Apfel, and Efron (1978) who discovered that academically competent girls were more likely than less competent girls to remain in a Follow Through project in New Haven.

In Study 1, two achievement-related conclusions appear warranted. First, the bulk of Wang and associates' evidence on the relation between the ALEM and pupils' academic achievement appears nonsignificant, both statistically and practically. Second, among the reviewed journal articles and book chapter describing the investigation, no specific mention is made of handicapped students' school achievement.

Study 2 (1980-81)

Participants. This study involved one school and 179 kindergarten through third grade students; 108 pupils in K-3 and 71 students in grades 1-3 were assigned randomly to ALEM and non-ALEM classes, respectively. A footnote to Table 1 (Wang & Birch, 1984b, p. 36) indicates there were 11 handicapped pupils in both ALEM and non-ALEM classrooms on whom complete data were obtained. This information prompts two questions. First, despite that handicapped pupils were assigned randomly to ALEM and non-ALEM classes, there are no data documenting the groups' comparability. Underscoring the importance of such documentation is the small number of handicapped participants. Second, we are uncertain of the attrition rate, let alone whether it may have been selective in nature, since no information is provided on the number of handicapped pupils participating in the study when it began in Fall 1980. We have a similar concern about nonhandicapped participants.

Additionally, Wang and Birch do not describe (a) how the one study school was selected, (b) how many ALEM and non-ALEM classrooms were involved, and (c) how many regular classroom teachers participated and how they were selected. Finally, only one resource room teacher participated, which seriously limits the generalizability of results (Haynes & Jenkins, 1986).

Implementation. No data are provided on the degree to which the ALEM was implemented in ALEM-designated classrooms.

Classroom processes. These data were collected during observations conducted in Fall 1980 and Spring 1981. Wang and Birch do not describe the frequency or duration of these observations. They also do · t report who performed them or whether interrater agreement was obtained.

The process data are presented in three contrasts. The first included morning ALEM and non-ALEM classes. One concern with this comparison addresses the fact that handicapped students assigned to non-ALEM classrooms spent entire mornings in a resource room, whereas their counterparts in ALEM settings remained in the mainstream. Thus, this first contrast pitted handicapped and nonhandicapped students in ALEM settings against only nonhandicapped pupils in non-ALEM classrooms.

Whereas the first contrast analyzed classroom activity and behavior of different groups in similar environments, a second comparison looked at similar groups in dissimilar school settings. Specifically, this second contrast included handicapped pupils in ALEM classes and handicapped students in the one resource room. Results from this second contrast are displayed in Wang and Birch's Table 1 (p. 36). Three behaviors were observed in the ALEM and resource room settings: teacher-directed activity and students' independent work and on-task behavior. From Fall to Spring, handicapped children in both settings (a) participated significantly less often in teacher-directed activities and (b) engaged in significantly more frequent independent work. The single difference between the two groups was that handicapped students in ALEM classrooms significantly increased on-task behavior, while their counterparts in the resource room did not.

Wang and Birch's last contrast was based on ALEM and non-ALEM handicapped groups' behavior and activity in mainstream classrooms. Unlike the preceding two contrasts, observations were conducted during afternoons when a "district-based program" (p. 36) was implemented, involving all classrooms in the study. Their Table 1 indicates that, from Fall to Spring, handicapped students in ALEM-designated classrooms were significantly more likely to be involved in independent work and on-task behavior during implementation of this afternoon program. No corresponding significant change was found for handicapped pupils in non-ALEM classrooms, leading Wang and Birch to conclude that "positive changes in behavior from October to April during the a.m. sessions were transferred to the p.m. sessions only for the ALEM students. In other words, the effects of the resource room program were not generalized to the district's regular program, while transfer effects were observed for the ALEM students" (p. 37).

Maybe so, but at least two facts weaken this interpretation. First, as already mentioned, Wang and Birch provide no evidence that the ALEM-designated classrooms implemented the model with fidelity; it is an assumption, not a fact, that this study was associated with a bonafide treatment. Second, there is no definition or description of the district's regular afternoon program, and no effort to document that it was implemented similarly across ALEM and non-ALEM settings. Thus, a reasonable alternate explanation of reported group differences is that teachers in ALEM and non-ALEM classrooms implemented the district program differently.

Finally, statistical analyses of the classroom process data lack incisiveness. Rather than integrate important study dimensions such as treatment (ALEM vs. non-ALEM classes), time (Fall vs. Spring), class sessions (a.m. vs. p.m.), and type of pupils (handicapped vs. nonhandicapped) into a single analysis, which would have facilitated direct comparisons of these factors and explorations of possible interactions among them, Wang and Birch conducted separate analyses for (a) the three experimental contrasts, (b) the three behavior categories on which the observations were conducted, and (c) the two handicapped groups. This resulted in use of numerous *t* tests, increasing the likelihood of Type I error. Moreoever, it seems Wang and Birch used *uncorrelated*, rather than more appropriate *correlated t* tests for each analysis.

Student achievement. Wang and Birch present Fall and Spring reading and math achievement scores for six groups: handicapped, nonhandicapped, and (heretofore unmentioned) gifted pupils in ALEM and non-ALEM classrooms. These data, displayed in our Table 3, represent *raw*, not *standard*, scores on the Stanford Achievement Test, which raises the question whether all K-3 pupils were administered the same level of the test. Although Wang and Birch do not report statistical analyses, there are several apparent trends. First, gifted students seem to have demonstrated greater improvements in reading and math than handicapped and nonhandicapped pupils. Second, and more important, ALEM handicapped pupils appear to have out-gained non-ALEM handicapped students in reading, whereas the two groups made comparable gains in math. Third, ALEM nonhandicapped and gifted students showed the same Fall-to-Spring improvements in reading and math as exhibited by their respective counterparts in non-ALEM settings. Thus, this study provides some, albeit inconsistent, evidence that the ALEM enhances handicapped students' achievement relative to controls. No such effect was observed among nonhandicapped and gifted study participants.

Insert Table 3 about here

Study 3 (1982-83)

Participants. Three articles describing this study state 26 classrooms participated (see Wang, Peverly, & Randolph, 1984; Wang, Rubenstein, & Reynolds, 1985; Wang, Vaughan, & Dytman, 1985); however, Wang, Gennari, and Waxman (1985) and Wang and Reynolds (1985) indicate the number was 28. In any case, the classrooms were associated with five schools in three community school districts of the New York City Public Schools. According to Wang, Vaughan, and Dytman (1985), these classrooms represented grades one through four; Wang, Peverly, and Randolph, however, indicate only grades two through four were involved. Teachers volunteered or were selected by their principals to participate. Class size ranged from 21 to 31 students; no mean or standard deviation for class size, or total number of pupil participants, is presented.

Wang, Peverly, and Randolph state 69 students were handicapped (see p. 23). In another apparent contradiction, Wang, Vaughan, and Dytman write that, "Approximately five students in each classroom were identified as educable mentally retarded, learning disabled, or socially-emotionally disturbed" (p. 117), which means there were at least 130 (26 classes x 5 students), not 69, handicapped subjects. If, in fact, 130 handicapped pupils were enrolled in the study classes, but only 69 participated in this investigation, it is unclear (a) how the subset of handicapped students was selected, (b) how they differed, if at all, from those not selected, and (c) how many study classrooms were represented. Finally, there are no data on handicapped or nonhandicapped children's race, ethnicity, SES, primary language spoken at home, or other demographic characteristics.

Implementation. Wang, Peverly, and Randolph's (1984) Table 1 (p. 25) presents evidence for increasing levels of implementation of the ALEM from Fall to Spring. For example, across all dimensions and classrooms, mean percentages of implementation increased from 76.42% (not 63.08% as erroneously reported in the Table) in October to 92.72% in February to 96.72% in May. However, as in Study 1, these data are presented in aggregate, and may not be viewed as necessarily indicative of the degree of implementation achieved at each study site. It remains unclear whether the ALEM is equally implementable across sites differing on SES, ethnocultural, or other potentially important pupil characteristics.

Classroom processes. These data are displayed in Tables 2 and 4 of Wang, Peverly, and Randolph (1984). Their Table 2 shows that many process variables changed in hypothesized directions, and that the size of some changes was impressive. For example, teacher-pupil instructional interactions increased from .70 to .86 to .93 in Fall, Winter, and Spring, respectively. However, just as many SBOS variables were associated with scant change, such as pupils' distractability, which only decreased from .16 to .14 to .15 across the three rounds of data collection. In fact, as suggested by column five of their Table 2, the median absolute value of the correlations between SBOS variables and the three periods of data collection (1 = Fall, 2 = Winter, 3 = Spring) is only .13. The sixth column of the same Table shows correlations between classroom processes and degree of implementation. Again, the median absolute value of these correlations is only .15. The low magnitude of these medians suggests weak overall relations between classroom processes and (a) time of year and (b) degree of implementation of the ALEM.

Wang and associates also explored classroom processes separately for handicapped and nonhandicapped pupils. This comparison was based on a contention that, "if instructional programs are adaptive to student differences, then all students, in spite of varied learning needs, should exhibit the behaviors hypothesized as classroom process outcomes of the ALEM" (Wang, Peverly, & Randolph, 1984. p. 28). In short, handicapped and nonhandicapped students were not expected to differ on the SBOS variables. Indeed, a glance at their Table 4 (p. 29) indicates few significant disparities between the two groups in Spring, corroborating the belief that successful implementation affects the two groups similarly. However, more deliberate study of this Table reveals that the groups also were not different in Fall. Specifically, in October, after just 1 month of ALEM implementation (see Wang, Vaughan, & Dytman, 1985, p. 117 and p. 119), handicapped and nonhandicapped pupils were evidencing similar behavior in 19 of 20 SBOS categories. Thus, if the two groups were essentially indistinguishable barely 1 month into the study, it is difficult to perceive the ALEM as contributing to their observed similarities in Spring.

Handicapped and nonhandicapped pupils did exhibit significant divergence over time on a couple of SBOS variables. Whereas in Fall there were no reliable differences between the groups' constructive and disruptive behaviors, handicapped pupils were significantly less constructive and more disruptive in Spring. However, this undesirable change may be artifactual, since an additional 39 nonhandicapped and 9 handicapped pupils participated in Spring than Fall. In addition to this possible "apples vs. oranges" situation, the number of students involved in Fall data collection (<u>N</u> = 196) represented 42% of the students on whom data in Wang and associates' Table 2 were collected, and only 36% of a <u>minimum</u> estimate of original pupils involved in the study (26 classes x 21 students = 546 students). Thus, the representativeness of the classroom process data also should be questioned.

Student achievement. Wang and her colleagues explored whether both handicapped and nonhandicapped pupils in the ALEM classrooms made expected or greater-than-expected gains in reading and math. "Expected gain" was defined for both groups as a 1-year increase in grade equivalence on standardized achievement tests. For handicapped students, achievement also was compared to a gain of 6 months, which, according to Wang, Peverly, and Randolph (1984), represented "the average achievement gain for students in the three school districts with similar special education classifications" (p. 25). Table 4, which summarizes data discussed by Wang, Peverly, and Randolph on page 25, shows that nonhandicapped students' reading and math achievement was significantly greater than a 1-year increase. Handicapped pupils, while achieving significantly smaller gains in math than nonhandicapped students, did manage a 1-year gain in both academic areas. Moreover, this improvement was significantly greater than the purported average annual gain of "comparably" handicapped children.

Insert Table 4 about here

Reviewing these results, Wang, Vaughan, and Dytman (1985) conclude that the ALEM "was found to have a positive impact on the learning outcomes...of both the general education and the special education students" (p. 119). Is this conclusion justified? We think not, and base our skepticism on a single important fact: Wang and associates failed to use control teachers and students, which necessarily precludes an opportunity to investigate experimentally whether the ALEM was a causal factor in students' reading and math gains.

Secondarily, much is made of the fact that ALEM handicapped pupils' achievement in reading and math outstripped the 6 month expected gain. As mentioned, this expected gain is described as an average annual increase for learning disabled, educable mentally retarded, and socially and emotionally disturbed pupils across the three participating school districts in reading and math. Although this index seems based on no less than 18 separate averaged gain scores (2 content areas x 3 handicapping conditions x 3 school districts), undoubtedly requiring numerous calculations, we are told nothing of its derivation. Moreover, gain scores are represented in analyses as grade equivalents. Grade equivalents are ordinal rather than equal interval data, and should not be used in inferential analyses. Thus, we believe there are salient design and statistical problems with this investigation, which represent serious constraints on the conclusions that Wang and associates legitimately may draw.

Discussion

Let us return to the three fundamental questions guiding Wang and associates' implementation studies and this critique. First, "Can a high degree of implementation of the ALEM be attained in classroom settings that differ in terms of aims, needs, and contextual characteristics?" Answer: We don't know. The reason for uncertainty is that, in Studies 1 and 3, Wang and associates present implementation data in aggregate form, rather than analyzing such information separately for school districts known to differ on factors like geography, income, racial-ethnic identity, and prior experience with the ALEM. (No implementation data were presented in Study 2.)

Second question: "When dimensions of the ALEM are in place, do hypothesized patterns of classroom process occur?" On balance, we conclude, "probably not." In Study 1, high, average, and low implementation classes appeared indistinguishable when compared statistically and descriptively on classroom process, or SBOS, variables. In Study 3, the median absolute value of the correlations between ALEM implementation and SBOS variables was a weak .15. Only in St[illegible] were certain SBOS factors associated exclusively with ALEM students. From Fall to Spring, handicapped pupils in ALEM classes (a) lengthened on-task behavior, while their counterparts in a resource room did not, and (b) worked with increasing independence and concentration during an afternoon district program, whereas no corresponding change was observed among handicapped students in non-ALEM classes. As mentioned above, however, it is difficult to ascribe these behavior changes to the ALEM, since Wang and associates did not document whether, and if so to what extent, the ALEM had been implemented.

The last question is, "Does implementation of the ALEM, and presumed resulting patterns of classroom process, lead to important improvement in student academic performance?" At best, the jury is still out. In Study 1, pupils in high, average, and low implementation classrooms performed equally in terms of (a) number of reading and math objectives mastered in the ALEM curriculum and (b) percentiles attained on standardized achievement tests of reading and math skills. In Study 2, ALEM handicapped pupils seemed to outgain non-ALEM handicapped students in reading, but not in math; no differences appeared to separate ALEM from non-ALEM groups among nonhandicapped and gifted study participants. Adding to the tenuousness of these findings was that the achievement data were not subjected to statistical analysis. In Study 3, handicapped ALEM pupils demonstrated 1-year gains in reading and math, which were significantly greater than an expected gain of 6 months. Nonhandicapped ALEM students in this study also exhibited improvements in reading and math that reliably outstripped an expected 1-year gain. However, since there was no control group, it is difficult to impute the apparent academic improvement to the ALEM.

Contributing to these largely equivocal findings is that the investigations often do not provide readers with sufficient information to understand the nature of the evaluative effort. Consider Study 2. There is no information on such salient matters as: (a) the comparability of ALEM and non-ALEM handicapped subjects; (b) the number of ALEM and non-ALEM handicapped and nonhandicapped pupils at the start of the study and degree of attrition sustained; (c) the frequency and duration of classroom observations, who performed them, and whether interrater agreement was obtained; (d) the number of teachers participating and the manner in which they were chosen; and (e) which level(s) of the Stanford Achievement Test was administered to the K-3 study participants.

Nevertheless, Wang and associates appear secure in the belief that the ALEM has been successful in most major respects. In summarizing effects of the model across several studies, Wang, Gennari, and Waxman (1985) claim that results provide "substantial support" for three main conclusions. "First, it is possible to establish and maintain average to high degrees of implementation of the ALEM on a large-scale basis in a variety of school settings.[1] Second, as critical features of the ALEM are established, so are classroom processes that are hypothesized to facilitate effective adaptive instruction in classroom settings. Finally, implementation of the ALEM and the presence of desired classroom processes of adaptive instruction seem to facilitate student achievement" (p. 228). Such purported findings "demonstrate the possibility that students with poor prognoses for academic achievement can succeed in their school learning through the provision of the type of adaptive instruction imbedded in the design of the ALEM. Thus, despite the limitations of attempting to generalize...from studies of a single program, there seems to be substantial evidence (our italics) to support making educational provisions for individual differences in regular classroom settings" (pp. 228-229).

We have suggested how and why Wang and associates' own data do not support these conclusions. Assuming we have not missed reports of successful implementations of the ALEM, we believe currently there is insufficient cause to view it as a successful, large-scale, full-time mainstreaming program. The importance of this belief pivots on the fact that many advocates of a General Education Initiative view the ALEM as a symbol of scientific readiness; a "tried and true" technology of large-scale mainstreaming. While we share with these advocates (a) the belief that mainstreaming is a very important goal and (b) the frustration engendered by the infrequency with which it is undertaken, we reject their assumption that we currently know how to establish full-time mainstreaming on a large scale. Before endorsing a merger between special and general education, we hope parents, teachers, researchers, and policymakers insist on additional empirical studies of full-time, large-scale mainstreaming and pursuasive evidence that such programs indeed work as their creators claim they do. If these programs are implemented widely without sufficient validation, we fear many handicapped children and teachers may suffer. Sound policy on service delivery, as in other areas, depends on the inspiration of advocates and the perspiration of researchers. To date, there has been too little of the latter.

References

Ammer, J.J. (1984). The mechanics of mainstreaming: Considering the regular educators' perspective. Remedial and Special Education, 5, 15-20.

Anderson-Inman, L. (in press). Consistency of performance across classrooms: Instructional materials versus setting as influencing variables. The Journal of Special Education.

Anderson-Inman, L., Walker, H., Purcell, J. (1984). Promoting the transfer of skills across settings: Transenvironmental programming for handicapped students in the mainstream. In W.L. Heward, T.E. Heron, & J. Trap-Porter (Eds.), Focus on behavior analysis in education (pp. 17-38). Columbus, OH: Merrill.

Biklen, D., & Zollers, N. (1986). The focus of advocacy in the LD field. Journal of Learning Disabilities, 19, 579-586.

Fuchs, D., Fuchs, L.S., Benowitz, S.A., & Barringer, K. (in press). Norm-referenced tests: Are they valid for use with handicapped students? Exceptional Children.

Goodman, J.F. (1977). The diagnostic fallacy: A critique of Janer Mercer's concept of mental retardation. Journal of School Psychology, 15, 197-206.

Gottlieb, J. (1975). Public, peer, and professional attitudes toward mentally retarded persons. In M.J. Begab & S.A. Richardson (Eds.), The mentally retarded and society: A social science perspective (pp. 99-125). Baltimore: University Park Press.

Haggerty, G.J., & Abramson, M. (1987). Impediments to implementing national policy change for mildly handicapped students. Exceptional Children, 53, 315-323.

Haynes, M.C., & Jenkins, J.R. (1986). Reading instruction in special education resource rooms. American Educational Research Journal, 23, 161-190.

Jones, R.L. (1972). Labels and stigma in special education. Exceptional Children, 38, 553-564.

Leinhardt, G., & Pallay, A. (1982). Restrictive educational settings: Exile or haven? Review of Educational Research, 52, 557-578.

Maheady, L., Towne, R., Algozzine, B., Mercer, J., & Ysseldyke, J. (1983). Minority overrepresentation: A case for alternative practices prior to referral. Learning Disability Quarterly, 6, 448-456.

Marston, D. (1987). Does categorical teacher certification benefit the mildly handicapped child? Exceptional Children, 53, 423-431.

Mcdowell, F.E. (1986). Adaptive learning model fosters both equity and excellence. School Administrator, 43, 20-23.

Nevin, A., & Thousand, J. (1986). What research says about limiting or avoiding referrals to special education. Teacher Education And Special Education, 9, 149-161.

Prasse, D.P., & Reschly, D.J. (1986). Larry P.: A case of segregation, testing, or program efficacy? Exceptional Children, 52, 333-346.

Pugach, M., & Lilly, S.M. (1984). Reconceptualizing support services for classroom teachers: Implications for teacher education. Journal of Teacher Education, 35, 48-55.

Reschly, D.J. (1979). Nonbiased assessment. In G.D. Phye & D.J. Reschly (Eds.), School psychology: Perspectives and issues (pp. 215-253). New York: Academic Press.

Reynolds, M.C., & Balow, B. (1972). Categories and variables in special education. Exceptional Children, 38, 357-366.

Reynolds, M.C., Wang, M.C., & Walberg, H.J. (1987). The necessary restructuring of special and regular education. Exceptional Children, 53, 391-398.

Seitz, V., Apfel, N., & Efron, C. (1978). Long-term effects of early intervention: The New Haven Project. In B. Brown (Ed.), Found: Long-term gains from early intervention (pp. 79-110). Boulder, CO: Westview.

Semmel, M., Gottlieb, J., & Robinson, N. (1979). Mainstreaming: Perspectives on educating handicapped children in the public schools. In D. Berliner (Ed.), Review of research in education (pp. 223-279). Washington, DC: American Educational Research Association.

Stainback, W., & Stainback, S. (1984). A rationale for the merger of special and regular education. Exceptional Children, 51, 102-111.

Stainback, W., Stainback, S., Courtnage, L., & Jaben, T. (1985). Facilitating mainstreaming by modifying the mainstream. Exceptional Children, 52, 144-152.

Wang, M.C. (1976). The use of observational data for formative evaluation of an instructional model. Instructional Science, 5, 365-369.

Wang, M.C. (1980). Adaptive instruction: Building on diversity. Theory into Practice, 19, 122-128.

Wang, M.C. (1981). Mainstreaming exceptional children: Some instructional design and implementation considerations. The Elementary School Journal, 81, 195-221.

Wang, M.C., & Birch, J.W. (1984a). Effective special education in regular classes. Exceptional Children, 50, 391-398.

Wang, M.C., & Birch, J.W. (1984b). Comparison of a full-time mainstreaming program and a resource room approach. Exceptional Children, 51, 33-40.

Wang, M.C., & Gennari, P. (1983). Analysis of the design, implementation, and effects of a data-based staff development program. Teacher Education And Special Education, 6, 211-224.

Wang, M.C., Gennari, P., & Waxman, H.C. (1985). The adaptive learning environments model: Design, implementation, and effects. In M.C. Wang & H.J. Walberg (Eds.), Adapting instruction to individual differences. Berkeley, CA: McCutchan.

Wang, M.C., Nojan, M., Strom, C.D., & Walberg, H.J. (1984). The utility of degree of implementation measures in program implementation and evaluation research. Curriculum Inquiry, 14, 249-286.

Wang, M.C., Peverly, S., & Randolph, R. (1984). An investigation of the implementation and effects of a full-time mainstreaming program. Remedial and Special Education, 5, 21-32.

Wang, M.C., & Reynolds, M.C. (1985). Avoiding the "Catch 22" in special education reform. Exceptional Children, 51, 497-502.

Wang, M.C., Reynolds, M.C., & Walberg, H.J. (1985, December). Rethinking special education. Paper presented at the Wingspread Conference on the Education of Students with Special Needs: Research Findings and Implications for Policy and Practice. Racine, WI.

Wang, M.C., Rubenstein, J.L., & Reynolds, M.C. (1985). Clearing the road to success for students with special needs. Educational Leadership, 43, 62-67.

Wang, M.C., & Stiles, B. (1976). An investigation of children's concept of self-responsibility for their school learning. American Educational Research Journal, 13, 159-179.

Wang, M.C., Vaughan, E.D., & Dytman, J.A. (1985). Staff development: A key ingredient of effective mainstreaming. TEACHING Exceptional Children, 17, 112-121.

Wang, M.C., & Walberg, H.J. (1983). Adaptive instruction and classroom time. American Educational Research Journal, 20, 601-626.

Wang, M.C., & Walberg, H.J. (1983). Evaluating educational programs: An integrative, causal-modeling approach. Educational Evaluation and Policy Analysis, 5, 347-366.

Wehman, P., Abramson, M., & Norman, C. (1977). Transfer of training in behavior modification programs: An evaluative reiview. The Journal of Special Education, 11, 215-231.

Will, M. (1986). Educating children with learning problems: A shared responsibility. Exceptional Children, 52, 411-415.

Yalow, E., & Popham, J. (1983). Content validity at the crossroads. Educational Researcher, 12, 10-14, 21.

Ysseldyke, J.E., & Shinn, M.R. (1981). Psychoeducational evaluation. In J.M. Kauffman & D.P. Hallahan (Eds.), Handbook of special education (pp. 418-440). Englewood Cliffs, NJ: Prentice Hall.

Zane, T. (1987). Research in special education. In C.R. Reynolds & L. Mann (Eds.), Encyclopedia of special education (pp. 1344-1350). New York: John Wiley & Sons.

Footnote

[1]Note Wang and associates' apparent satisfaction with an "average to high" degree of implementation of the ALEM, while it has been a "high" degree of treatment fidelity that they consistently have sought (see, for example, Wang & Birch, 1984b, p. 393; Wang, Gennari, & Waxman, 1985, p. 197; Wang, Peverly, & Randolph, 1984, p. 22).

Table 1

Frequencies and Percentages of Classrooms at Each Site with Scores at the High, Average, and Low Degree of Implementation Levels, Spring, 1981[a]

Sites	Degree of Implementation Levels[b] High	Average	Low
Follow Through (N = 117)			
Site A (N = 22)	9 (41)	13 (59)	0 (0)
Site B (N = 22)	8 (36)	14 (64	0 (0)
Site C (N = 17)	6 (35) 53 classes (45.30%)	9 (53) 60 classes (51.28%)	2 (12) 4 classes (3.42%)
Site D (N = 19)	10 (53)	9 (47)	0 (0)
Site E (N = 11)	10 (91)	1 (9)	0 (0)
Site F (N = 26)	10 (39)	14 (54)	2 (8)
Mainstream (N = 21)			
Site G (N = 4)	0 (0)	4 (100)	0 (0)
Site H (N = 3)	2 (67)	1 (33)	0 (0)
Site I (N = 5)	0 (0) 2 classes (9.52%)	4 (80) 18 classes (85.72%)	1 (20) 1 class (4.76%)
Site J (N = 9)	0 (0)	9 (100)	0 (0)
Cross-Site (N = 138)	55 (40)	78 (56)	5 (4)

[a]Modification of Table 3 (p. 269) of Wang, Nojan, Strom, & Walberg (1984). Percentages are in parentheses.

[b]"High" implementation refers to classrooms with scores at or above the 85% criterion level in 11 or more of the critical dimensions; "Average" implementation indicates classrooms with scores at or above the 85% criterion level in 6 to 10 of the critical dimensions; and "Low" implementation signals classrooms with scores at or above the 85% criterion level in 5 or fewer critical dimensions.

Table 2

Summary of Student Achievement Outcomes for Classrooms at the High, Average, and Low Overall Degree of Implementation Levels Spring, 1981[a]

	Degree of Implementation Levels						
	High (N = 55)		Average (N = 78)		Low (N = 5)		
Variables	Mean	S.D.	Mean	S.D.	Mean	S.D.	F-test[b]
Student Learning Progress (Mean Objectives Mastered)							
Math Objectives	29.77	(17.4)	28.97	(17.6)	20.21	(15.8)	2.87
Reading Objectives	33.49	(20.6)	28.52	(19.4)	27.01	(20.1)	1.13
Standardized Achievement Scores (Percentile Rank)							
Math	53.49	(23.5)	52.16	(23.9)	48.00	(19.9)	2.10
Reading	53.59	(22.5)	52.57	(21.9)	51.02	(19.4)	1.92

[a]Reproduction of Table 5 (p. 273) of Wang, Nojan, Strom & Walberg (1984).

[b]None of the F ratios is significant at or beyond the .05 level.

Table 3

Handicapped and Nonhandicapped Pupils' Reading (R) and Math (M) Achievement in Alem and Non-Alem Settings for Fall and Spring[a]

	Handicapped (N = 22) ALEM (N = 11) R	Handicapped ALEM M	Handicapped Non-ALEM (N = 11) R	Handicapped Non-ALEM M	Nonhandicapped (N = ?) ALEM (N = ?) R	Nonhandicapped ALEM M	Nonhandicapped Non-ALEM (N = ?) R	Nonhandicapped Non-ALEM M	Gifted (N = ?) ALEM (N = ?) R	Gifted ALEM M	Gifted Non-ALEM (N = ?) R	Gifted Non-ALEM M
Fall	25	29	32	34	48	43	49	42	55	48	59	49
Spring	46	46	39	52	62	59	64	60	93	87	90	90

[a]These raw score data are taken from p. 37 of Wang and Birch (1984b).

Table 4

Obtained Gain vs. Expected Gains in Reading and Math between Spring 1982 and Spring 1983 for Handicapped and Nonhandicapped Pupils[a]

	Handicapped		Nonhandicapped	
	Reading	Math	Reading	Math
Obtained Gain	1.04*	1.08*	1.19+	1.87+
Expected Gain 1	1.00	1.00	1.00	1.00
Expected Gain 2[b]	.60	.60	--	--

[a]Data are summarized from p. 25 of Wang, Peverly, and Randolph (1984).

[b]Average expected gain for "comparable" handicapped pupils across the three school districts in reading and math.

*Obtained Gain is significantly greater ($p<.01$) than Expected Gain 2.

+Obtained Gain is significantly greater ($p<.01$) than Expected Gain 1.