Ninth-Grade Remediation Programs

A Synthesis of Evidence-Based Research

Sarah Calderon
Steven Klein
Robert Fitzgerald
Rachael Berger
MPR Associates, Inc.

June 2005

A report prepared for
MPR Associates, Inc., Berkeley, CA
under contract to the Office of Vocational and Adult Education,
U.S. Department of Education

Foreword

In 2003, in association with and under contract to the Office of Vocational and Adult Education (OVAE) within the U.S. Department of Education, MPR Associates, Inc. commissioned four papers on topics related to improving secondary education and student achievement. The topics included block scheduling, smaller learning communities, remediation or assistance programs aimed at ninth-graders, and school choice.

The authors of each paper applied a recently developed review and synthesis tool proffered by the What Works Clearinghouse, established by the Institute of Education Sciences. The What Works Clearinghouse gathers studies of the effectiveness of educational interventions, reviews the studies with the strongest designs, and reports on the strengths and weaknesses of those studies against a specific set of Evidence Standards.

The resulting set of four research syntheses document the degree to which each area of study includes research that achieves the level of rigor required to meet the standards, and whether the available research provides the clear evidentiary foundation necessary for drawing conclusions about each intervention's efficacy.

A subtask within contract ED-99-CO-0160 (Richard Smith, OVAE project director) funded the development of these papers. Opinions expressed and conclusions drawn in the research syntheses do not represent official U.S. Department of Education position or policy, nor that of MPR Associates, Inc.

Table of Contents

List of Tables

Page

Executive Summary

The first year of high school is a critical period in the educational development of adolescents. When youth enter an unfamiliar educational setting and face increased expectations for their academic performance, many experience initial declines in their grades and attendance. The transition to ninth grade can be particularly daunting for students who lack the academic preparation for high school success. Over time, without outside intervention, these youth are likely to lag behind their peers and are at greater risk of dropping out of school.

Ninth-grade remediation programs are intended to help underprepared students acquire the academic knowledge and skills they need to function in an increasingly standards-based environment. Although educators have experimented with many promising initiatives, only a handful possess the scientific rigor called for by the No Child Left Behind Act. To support efforts to identify educational programs and practices that have been proven effective, the Office of Vocational and Adult Education (OVAE) commissioned a review of studies that measured the effect of ninth-grade remediation programs on student achievement. Using the What Works Clearinghouse's (WWC) Study Design and Implementation Assessment Device (DIAD) as a filter, this report provides a catalog and assessment of the quantitative research studies published in the past 10 years. The analysis also takes into consideration the Cumulative Research Evidence Assessment Device (CREAD) used to summarize the validity characteristics of the evidence base.

This research review answers the following questions:

1. Do remediation programs increase academic achievement, as measured by test scores and grades, or have any effect on ninth-graders' attendance and dropout rates?
2. What program characteristics are associated with positive performance outcomes?
3. What are the strengths and weaknesses of the existing evidence base, and to what extent do studies that meet DIAD criteria address the educational issues that characterize the field?

Methodological Approach

The first stage of the study involved analyzing the evidence base in which researchers adopted the literature review strategy detailed in the DIAD instrument. To locate remediation studies published in books and refereed journals, we initially conducted online searches of four electronic databases

spanning the secondary literature, using Boolean logic statements to organize search requests. We then identified citations for 598 documents for which abstracts were collected.

During the second round of the search process, we reviewed abstracts to identify studies quantifying the effectiveness of remediation programs and, in particular, those targeting ninth-grade students (or those making a transition to the ninth grade). We also tracked references within qualifying studies to identify other possible studies for inclusion. This review process yielded a total of 72 articles. Research staff ordered full-text copies of each article, read, and classified studies using a key word extraction template. To warrant inclusion, a study had to have used either an experimental or quasi-experimental design, met criteria for the length of intervention, and included measurable, quantitative outcomes of student achievement. A total of 22 papers satisfied these conditions.

During the final stage, we reviewed the 22 studies for the quality of their research design, based on guidelines contained within the DIAD instrument developed by the WWC. We subsequently entered answers to design and implementation questions into a spreadsheet programmed to calculate responses to the eight composite and four global questions contained in the DIAD. A total of 10 reports met the minimum inclusion criteria for all composite questions and underwent the final stage of review.

Study Findings

Interpreting the 10 articles that passed through the DIAD review is complicated due to differences among studies in the intervention or treatment. Unlike school reform efforts that involve a relatively discrete, easily identifiable intervention, broad reform efforts such as ninth-grade remediation often include multiple components that may differ substantially in their delivery and effect. For example, although both ninth-grade academies and supplemental academic instruction are classified as remediation programs, the two interventions operate in very different ways: one approach groups freshmen into a school-within-a-school for instructional purposes, and the other involves providing additional coursework in one or more subject areas. Compounding this problem is the fact that researchers differ in their clarity of describing and testing the causal linkages across the different interventions included in a study, which can make generalization across studies difficult. Furthermore, because educators may apply multiple program strategies within a given reform approach (e.g., supplemental coursework and peer tutoring within a school restructuring model), it is often impossible to determine which component should be credited with changes in student outcomes.

Due to the small number of studies that used similar interventions in the evidence base, the limited number of student outcomes shared across studies, and differences in how studies scored on the DIAD rating system, we determined that the cumulative evidence would not support meta-analysis

of statistical findings, an essential part of the WWC's CREAD instrument. Although we rejected conducting a meta-analysis (including the calculation of effect sizes) because the evidence base as it currently exists is insufficient to estimate reliably the direction and magnitude of ninth-grade remediation programs, the 10 studies that passed the DIAD review do provide some useful, albeit general, insights into the impact of different remediation efforts.

Evidence on Academic Support Services

While the provision of supplemental academic instruction was a component in nearly all of the interventions included in our research base, four studies emphasized the use of academic support services as a reform strategy. Although the scope and implementation of services varied across sites (ranging from multi-year, comprehensive services to point-in-time changes in curriculum in a single subject), study results indicate that supplemental academic services were associated with improvements in student learning. In some programs, however, these positive effects declined over time.

Of the four programs, Upward Bound (Myers & Shirm, 1999) was the most comprehensive, serving a large population and providing a variety of academic interventions including tutoring and study skills, Saturday enrichment classes, ACT/PSAT workshops, a six-week summer enrichment program, and college preparatory workshops. Using random assignment and controlling for student ability, the researchers found that, on average, lower-performing ninth-grade students earned more credits throughout high school than those in the control group and that participating in Upward Bound reduced the probability that higher- and lower-performing students would drop out before graduating, with gains substantially larger for students at greater risk. These positive findings are tempered by other findings suggesting that, with the exception of at-risk students, most participants realized little academic benefit.

In a study focused on curriculum and teaching, Talent Development High School's ninth-grade instructional program in reading and mathematics was assessed to determine the effect of offering specialized curricula to ninth-grade students. Balfanz, Legters, and Jordan (2003) found that students receiving instruction provided in the context of a Ninth-Grade Success Academy (a school-within-a-school for freshmen) significantly outperformed students in control groups in both their overall achievement level and performance gains. Researchers also found that students at all levels of ability obtained larger achievement gains as part of the program. The evaluation design used in this study, however, was not based on random assignment. Students at three high schools in the intervention group were matched to similar high schools in Baltimore based on student demographic characteristics.

In a more narrowly defined investigation of a supplemental academic service program, Gamoran, Porter, Smithson, and White (1997) looked at the effects of "transition" mathematics courses used to increase the rigor of instruction for ninth-grade students. The study design used by Gamoran et al. was also not based on random assignment, but differences among students in the treatment groups were statistically controlled using hierarchical linear modeling. The researchers reported that the rate of increase in test scores for students participating in treatment groups was higher than for students in general track courses, but lower than for students taking advanced coursework. They also found that achievement growth was greatest in classes with more content coverage, suggesting that providing students with more rigorous curricula may be an effective way of improving student outcomes.

Finally, Woodruff, Schumaker, and Deshler (2002), in an evaluation of the effect of a Word Identification Strategy used to assist ninth-grade students reading below grade level, found that after a four- to six-week intensive academic intervention, treatment students evidenced mean gains in their reading decoding as large as six grade levels. Students in a matched comparison group evidenced only minimal changes in performance.

Evidence on School Restructuring

Four of the 10 studies forming the evidence base focused on school restructuring, although the type of services offered varied across studies, ranging from clustering ninth-grade students and teachers in Career Academies to conducting comprehensive restructuring efforts involving all grade levels. Findings on program impacts were mixed, with two studies offering positive, but weak, evidence of short-lived improved academic performance among participating students. For example, positive outcomes were associated with Project Transitions, a ninth-grade reform that emphasized clustering teachers and students into learning teams within core academic subjects. The evaluation used a cohort comparison design to compare outcomes for students who were in the ninth grade when the program was implemented to students who were in the ninth grade before the program began. When assessing program effects, the evaluators used regression methods to account for differences between students in the two cohorts. Program outcomes differed across the two sites studied, with one school noting small increases in course passing rates and the percentage of students with a GPA above grade D.

Slight reductions in both high school dropout and increased average attendance rates were also noted for high-risk students randomly assigned to participate in Career Academy programs (Kemple & Snipes, 2000); however, standardized test scores in reading and mathematics indicate that Academy students did not obtain any long-term academic benefit from participation, even when con-

trolling for level of student risk. In the aggregate, Academy students exhibited only small, statistically insignificant reductions in dropout rates, progress toward high school graduation, and increases in career-related course taking and participation in youth development activities.[1]

In contrast, an analysis of high school dropout prevention programs focusing on ninth-grade students, one involving an alternative school and another involving a school undergoing comprehensive restructuring, provided no evidence of positive program effects (Dynarski, Gleason, Rangarajan, & Wood, 1998a, 1998b). Specifically, absenteeism and dropout rates were essentially the same in the treatment and comparison schools, while standardized mathematics and reading scores, though slightly higher at the beginning of both tenth and eleventh grades for the treatment school, demonstrated little improvement between the ninth and tenth grades. Dropout prevention programs that focused on school restructuring posted similar findings; specifically, absenteeism was essentially the same in restructuring and comparison schools, while dropout rates for ninth-graders in the treatment group were essentially the same as those of the control group by the end of eleventh grade.

Evidence on Summer Remediation

Two of the 10 studies in our evidence base focused on remediation programs offered in the summer months. Conflicting findings were noted across the two programs, which differed substantially in design. Positive outcomes were noted for Summer Bridge, a program designed to assist eighth-graders in obtaining academic skills and passing test scores needed for promotion to the ninth grade (Roderick, Engel, & Nagaoka, 2003). Based on pre- and post-test comparisons of student performance, cohort analyses, and regression discontinuity design, Summer Bridge was found to be effective in the short run across demographic and achievement groups in producing test score gains, and students increased their test scores during the regular school year. Over four years (1997–2000), eighth-graders on average increased their ITBS scores by approximately six months in reading and nearly five months in mathematics. In addition, passing rates were high among eighth-graders, with approximately half meeting the promotional cutoffs in both subjects by the end of the summers.

In contrast, few positive outcomes were associated with the Summer Training and Education Program (STEP), a program spanning 15 months (Grossman & Sipe, 1992). Students were randomly assigned to treatment and control groups. The treatment group received focused instruction during the summer before the freshman (or for some students, sophomore) year continuing through the

[1] Until the number of studies increases, it will be impossible to assess whether interventions like these have measurable effects on student outcomes. A meta-analysis with enough cases would allow us to estimate if the weak, non-statistically significant effects reported in any one study are, in fact, supportive of positive student outcomes overall.

following academic year and summer. Although the treatment group evidenced gains in academic skills and personal behaviors, these improvements were not sustained over time.

Summary and Recommendations

Although we found a substantial number of studies (598) during our first review of the published literature, these articles varied in their quantitative focus, methodological rigor, and substantive definition. "Ninth-grade remediation" is an umbrella term encompassing many distinct, often unrelated reform practices. Analysis of the evidence base indicates that, while numerous studies of remediation programs have been attempted, only a small number meet the minimal criteria for sound quantitative evidence set forth in the WWC's Study DIAD. Even among those studies that met the DIAD criteria, differences in the types of interventions (e.g., "academic support" or "school restructuring"), the multiplicity of program components among studies that appear to emphasize the same intervention, and the absence of fully specified causal linkages and tests of these linkages in the published research means that too few studies exist to justify use of the CREAD.

While there appear to be several well-conducted studies showing positive outcomes associated with certain types of interventions, it is impossible using the current evidence base to draw reliable and comprehensive conclusions about the direction and magnitude of effect sizes. Given the breadth of remediation strategies, the small number of scientifically rigorous studies that currently exist, and the lack of replicability across studies of similar design, ninth-grade remediation strategies, as currently defined, do not merit further consideration as a topic for a *WWC Evidence Report*.

The Rationale for Ninth-Grade Remediation

Youth who enter high school lacking the expected academic knowledge and skills are more likely to fall behind their peers and eventually drop out. Ninth-grade remediation programs are intended to remedy students' skill deficits by providing focused instructional and support services, often targeted on youth at high risk of school failure. Although the importance of helping all ninth-graders get off to a good start in high school is well recognized, a clear consensus about how best to achieve this result has yet to emerge. This section provides a brief introduction to ninth-grade remediation services and summarizes strategies commonly used by educators to redress students' skill deficits.

The Challenge of High School Entry

The freshman year in high school can be one of the most emotionally difficult, academically challenging events in a young person's life. Occurring during the middle stage of adolescent development, when youth begin to exert their own independence and complete their physical development, high school entry often introduces a unique set of challenges that can adversely affect the learning process. As youth are plunged into an unfamiliar, often chaotic school environment, unsure of institutional norms and expectations, and granted an unprecedented degree of independence, many initially struggle to adapt to their new surroundings.

At the onset, social changes associated with high school matriculation can be disconcerting to youth, who find themselves thrust into a complex new environment often filled with a larger, more diverse student population. This change from being the one of oldest, most knowledgeable students to one of the youngest, least experienced can contribute to feelings of anxiety and isolation (Blyth, Simmons, & Carlton-Ford, 1993). Because high schools are seldom structured to ensure that students receive support services and individualized attention, students who lack the social skills and motivation can begin school at a disadvantage (Alliance for Excellent Education, 2003).

Academic demands also increase, with students facing new expectations that require more advanced learning skills. In particular, students are expected to master a host of critical thinking, written, mathematics, and communication skills that form the basis for their subsequent education (Roderick, 1996). While average class sizes in high school are generally similar to those in middle school, ninth-graders often perceive high school as larger, more competitive, and more grade-oriented than middle or elementary school (Alt & Choy, 2000; Eccles, Midgley, & Adler, 1984). Not all students respond well to these new conditions, and their transition from middle to high school

is associated with declines in academic achievement for nearly all youth (Alspaugh, 1998). For example, Roderick (1993) documented that, on average, students experience an 18 percent drop in grades following high school entry, a fall equivalent to two letter grades. Research has also documented an increased incidence of course failure and declines in attendance (Roderick & Camburn, 1999; Blyth, Simmons, & Carlton-Ford, 1993).

Students who enter high school lacking the foundation skills for success are at increased risk of scholastic failure, which often precedes dropping out of high school. In particular, Roderick and Camburn (1999) reported that students who fail coursework during their first semester of high school are likely to fall into a cycle of failure from which they are unlikely to recover. Indeed, one study of high school dropouts revealed that 60 percent of students who left school had not completed at least 25 percent of their ninth-grade credits (Roderick, 1993).

Accountability requirements introduced in the *No Child Left Behind Act* have required that all states establish performance standards in mathematics, reading, and science. Detailing the knowledge and skills that all students are expected to know and be able to do and introducing or emphasizing statewide standards both promise to help close the achievement gap among students by ensuring that all youth obtain the skills necessary for success in school. As new, more intellectually demanding coursework becomes the norm, increased performance expectations may have implications for the academic progress of all ninth-graders and, in particular, of those who enter school lacking the basic skills needed for success.

Remediation Strategies

Ninth-grade remediation programs describe a wide range of educational interventions intended to support students as they enter high school.[2] One complicating factor in assessing these programs is that the term "ninth-grade remediation" applies to a host of interventions targeting academic, social, and affective behaviors—all of which can play a mediating role in students' transition to secondary education. As a consequence, substantial variation exists in the manner in which programs are structured and the performance outcomes that are measured. For example, although both ninth-grade academies and supplemental academic coursework are classified as remediation programs, the two interventions operate in very different ways, making generalization across studies difficult. Furthermore, because multiple interventions may be used within a given reform approach (e.g., supplemental coursework or peer tutoring within a whole-school reform effort), assigning causality to

[2] A substantial literature also exists on remediation programs offered within the middle grades. These programs, intended to anticipate and resolve issues that can complicate students' transition to high school, are not addressed in this research report, which focuses on programs offered at the high school level for students who enter lacking the necessary skills for success.

observed effects can be complicated when studies employ differing intervention techniques for a given reform strategy.

While all ninth-grade remediation programs intend to help students—particularly those lacking fundamental skills—make a successful transition into high school, the outcomes of program interventions can differ. Some focus on improving students' academic performance; others address affective factors, such as students' attitudes or beliefs about school. Although the latter outcomes are important, this review is confined to intervention strategies intended to bring about increases in students' academic performance and engagement in school. To be included in the evidence base, selected studies must have reported quantitative data on at least one of the following specific student outcome measures: test scores, grades, attendance, or ninth-grade dropout rates.

A review of the extant literature indicates that ninth-grade remediation programs come in a variety of forms. To improve students' academic performance, some educators have developed intervention programs that primarily focus on students' social and emotional needs, in the belief that learning occurs when students are confident and motivated to achieve. These interventions, which can include mentoring or social skill development, are intended to assist students in making a successful transition to high school. Justified by studies documenting the benefits of positive adult influences and individualized attention (Flaxman, 1992; Jekielek, Moore, Hair, & Scarupa, 2002), these approaches are often designed to change student attitudes, motivation, or beliefs about schools.

Academic support services involve providing supplemental instruction to help lower-achieving students catch up to their higher-achieving peers. Academic delivery can take many forms including using peer or adult tutors to provide more individualized instruction (Cooledge & Wurster, 1985; Wasik & Slavin, 1990; Wilks & Clarke, 1988); providing students with an extra academic course (often called "double-dosing"); and offering pull-out services during or after school. Other forms of remediation programs may also be offered during the summer months (Cooper, Charlton, Valentine, & Muhlenbruck, 2000). In addition, specialized academic transition programs offered during or near the end of middle school and overlapping with the freshman year may also be introduced to curtail the achievement loss that students experience when making the transition from middle school to high school, especially for students who did not perform well in the middle grades (Alspaugh, 1998; Irvin, 1997; Barone, Aquirre-Deandreis, & Trickett, 1991; Reyes, Gillock, & Kobus, 1994; Roderick, 1993; Felner et al., 1993; Seidman, Aber, Allen, & French, 1996).

More intensive approaches involve restructuring schools into smaller, self-contained units, such as schools-within-schools or academies. Reorganized schools may include all students, the freshman class, or only entering students who are at risk of school failure. School restructuring strategies draw on research literature suggesting that smaller learning environments can be more conducive to student learning (Cotton, 1996, 2001; Irmsher, 1997; Klonsky, 1998; Wasley & Lear, 2001; Lee &

Smith, 1995; Howley, Strange, & Bickel, 2000; Page, Layzer, Schimmenti, Bernstein, & Horst, 2002; Raywid, 1995; Wasley et al., 2000).

What Constitutes an Effective Ninth-Grade Remediation Program?

This research synthesis is intended to offer a comprehensive, systematic review of the literature on the effectiveness of ninth-grade remediation programs and practice and, in particular, on efforts to improve the academic knowledge and skills of youth at high risk of school failure. While there are also compelling reasons for providing students with personal and social supports, the reality is that many students enter high school lacking the academic skills required to meet or exceed state academic performance benchmarks and fall further behind their peers over time. Attention to self-esteem and school attitudes may be warranted, but a more immediate concern is ensuring that all students have the minimum academic knowledge and skills to graduate.

The push for higher secondary standards, manifested in increased state promotion and graduation requirements, has considerable implications for secondary educators. Although performance expectations are changing, high schools have not traditionally been structured to provide ninth-grade students with supplemental services to strengthen (or impart) foundation academic skills that they may lack (Balfanz et al., 2003).

This paper is intended to assist educators in identifying effective ninth-grade academic remediation strategies that can be used to put underachieving ninth-grade students back on track to graduate. Using the WWC's DIAD as a filter, this report arrays the range of quantitative research studies that have been published in the past 10 years and summarizes findings from the studies that meet or exceed the minimum level of acceptability for inclusion in this report.

Key questions that this research review seeks to answer include the following:

1. Do remediation programs increase student academic achievement, as measured by test scores and grades, or have any effect on ninth-grade attendance and dropout rates?
2. What program characteristics are associated with positive performance outcomes?
3. What are the strengths and weaknesses of the existing research base, and to what extent do studies that meet or exceed DIAD review criteria address the larger educational issues that characterize the field?

To provide a context for the discussion of the study findings, the following section describes the review methodology used to conduct this research synthesis.

Research Methodology

Recognizing the importance of supporting students in their first year of high school, researchers have invested substantial resources documenting outcomes associated with ninth-grade remediation programs. To navigate the voluminous literature, MPR researchers devised an analysis strategy to structure literature searches and identify studies worthy of follow-up. The following section reviews the study methodology and describes how studies were identified and reviewed in preparation for DIAD and CREAD analysis.

Surveying the Literature

To locate remediation studies published in books and refereed journals, MPR researchers conducted online searches of four electronic databases spanning the secondary literature: Educational Resources Information Center (ERIC), Dissertation Abstracts International (DAI), American Psychological Association (PsycINFO), and the University of California Systemwide Online Catalog (Melvyl). Researchers accessed the library of stored documents within each database, which consisted of full-text reports, abstracts, or bibliographic citations of research findings.

To access unpublished materials, including information posted by professional associations, state and local governments, local educational agencies, and public and private firms, MPR researchers also surveyed the Google and Yahoo Internet search engines. The researchers also conducted a review of resources posted on the U.S. Department of Education's Web site. Because each database is independently maintained and contains overlapping records, search results frequently produced multiple hits on a given document.

Database searches were organized using Boolean logic statements, cross-referencing the term *remediation* with related concepts, including *ninth grade*, *high school*, and *academic achievement*. (See Appendix A for a list of searches conducted.) Search years were restricted to include studies completed between 1993 to the present, with the exception of seminal studies occurring before 1993. This general search strategy produced citations for 598 documents for which abstracts were collected.

During the second round of the search process, the researchers reviewed abstracts to identify studies quantifying the effectiveness of remediation programs and, in particular, those targeting ninth-grade students (or those making a transition to ninth grade). Rejected papers included remediation projects serving unrelated populations (e.g., elementary or college students) and descriptive reports lacking a research base. The researchers also reviewed references within qualifying studies to identify

other possible studies for inclusion. This review process yielded a total of 72 articles. (See Appendix B for list of these papers.)

During the third stage, research staff ordered full-text copies of each of the 72 articles and read and classified studies using a key wording template. (A full-text version of this template is included in Appendix C.) Before using the instrument, research staff reviewed the template instructions for clarity and tested them for reliability. To warrant inclusion in the DIAD process, a study had to have used either an experimental or a quasi-experimental design.

Application of the Study DIAD

Studies that did not provide comparative data (i.e., quantitative studies lacking a control or self-comparison group) or that used qualitative analyses were rejected. Studies that did not include populations drawn from the high school setting were also discarded, as well as studies that did not contain measurable outcomes of student achievement (e.g., standardized tests, GPA, curriculum and/or end-of-course tests, exit exams, classroom tests) or quantifiable indicators of student engagement (e.g., attendance, dropout rates). A total of 22 papers met these minimal criteria for study inclusion.

During the fourth stage, the researchers reviewed 22 studies for the quality of their research design based on guidelines contained within the DIAD instrument developed by the WWC. The DIAD is a system for objectively assessing the extent to which the design and implementation of an individual study allows a researcher to make conclusions about the causal effects of an intervention. Before this coding process began, senior staff at MPR developed criteria for applying and interpreting design and implementation questions, based upon the theoretical literature, and established statistical practice by creating a data extraction template (Appendix D). Staff then entered answers to design and implementation questions into a spreadsheet (Appendix E) programmed to calculate responses to the eight composite and the four global questions according to DIAD algorithms. A total of 10 reports met the minimum inclusion criteria for all composite questions and were sent to the final stage of review. (See Appendix F for the 12 studies that did not make it through the DIAD.)

Research Findings

Interpretation of study results is complicated due to differences in the focus and construction of ninth-grade remediation programs. Unlike educational improvement programs such as high-stakes testing or raising graduation requirements, which describe a relatively narrow, easily identifiable intervention, broad reform efforts such as ninth-grade remediation often include a range of approaches using substantially different program components. For example, although both ninth-grade academies and academic "double-dosing" are classified as ninth-grade remediation strategies, the two interventions operate under very different program structures, making generalization difficult when only one or two studies examine a particular type of intervention. Furthermore, because multiple interventions may be used within a given reform approach (e.g., supplemental coursework and peer tutoring within a whole-school reform effort), assigning causality to observed effects can be complicated if researchers do not control for each program component.

Reported outcomes are often as variable as program interventions. For instance, among the four studies passing through the DIAD review that focused on academic support, the outcome measures and the number of studies sharing that outcome included the following: honors/AP credits (1), high school completion status (1), course taking (1), GPA (1), dropout rate (1), algebra I pass rate (1), reading achievement test score (2), other achievement test (2). Alternatively, among programs that used summer remediation—of which there were two—the outcome measures and the number of studies using each outcome measure were as follows: standardized test scores (2), graduation credits (1), dropout rate (1), graduation rate (1), absenteeism (1). In short, outcomes for most studies were unique.

The literature on ninth-grade remediation, although voluminous, has relatively few methodologically rigorous studies (10 of 598 reviewed), and they vary greatly in focus, implementation, and measured outcomes. The combination of these factors suggests that there is little utility in subjecting even the best studies to a rigorous meta-analysis and consideration using the WWC's CREAD process. Nonetheless, the 10 studies that are methodologically sound according to the DIAD criteria are worthy of notice: they can serve as the foundation for later treatment of this subject area when more studies, equal in methodological rigor, are published.

Academic Support Services

Supplemental academic instruction was provided in nearly all of the interventions included in our research base, but four studies emphasized the use of academic supports as a reform strategy. Although the scope and implementation of services varied across sites, study results indicate that supplemental academic services were associated with improved student learning.

In their evaluation of Upward Bound—a federal program designed to help disadvantaged students enter and succeed in college—Myers and Shirm (1999) collected longitudinal data on approximately 1,500 students randomly assigned to treatment groups and 1,300 students assigned to control groups from 67 Upward Bound projects operating across the country. The researchers collected baseline data and made assignments to treatment or control groups from 1992 through 1994 and conducted follow-up surveys of both groups in 1994 and 1996. The evaluators assessed student attitudes, school experiences, and other outcomes and collected high school transcripts to determine changes in students' academic performance.

Although the sample was not limited to ninth-grade students—secondary students in all grade levels participated in the intervention—investigators analyzed the effects of the Upward Bound program on key subgroups, including ninth-graders with varying levels of academic risk (based on ninth-grade GPA and academic credits) and students who entered Upward Bound in the ninth grade compared to students who entered in the tenth grade. Myers and Schirm found that lower-performing ninth-graders earned, on average, 1.3 more credits throughout high school than those in the control group and that Upward Bound reduced the probability that higher- and lower-performing students would drop out before graduating, with gains substantially larger for students at greater risk. Lower-performing participants were also more likely to attend a four-year college and to earn a greater number of non-remedial college credits than those in the control group.

When comparing students who entered Upward Bound in the ninth grade to students who entered in the tenth grade, researchers found that the ninth-graders and their parents had higher educational expectations, but they found few significant results for these students in terms of the number of non-remedial credits they earned in high school. However, when controlling for duration (time in the program), ninth-graders who participated for more than 24 months had significantly higher numbers of earned credits compared to control groups. Overall, those who participated for more than 24 months earned 2.4 more total credits (remedial and non-remedial) in high school.

These positive findings are tempered by a number of factors. On average, the impact of participation was similar for students applying for admission in the ninth and tenth grades. And given that most students participate in Upward Bound for only a short time, even though program benefits are di-

rectly related to how long they participate, only a small proportion of all entering students ultimately benefit.

To assess the effect of offering subject area remediation, Balfanz et al. (2003) examined the Talent Development High School's ninth-grade instructional program in reading and mathematics. In this program, ninth-graders receive double doses of mathematics and English instruction, are scheduled into a 4x4 block schedule, take three specially designed courses during the first semester to increase their academic preparation, and attend a Ninth-Grade Success Academy (a school-within-a-school for freshmen) with a separate principal and team of teachers. Teachers also receive necessary support and extra professional development to help them implement the program. The first-semester reading/English intervention—Strategic Reading—emphasizes skill development in reading fluency, writing, and comprehension; gives students opportunities to work cooperatively on reading and discussing novels and plays; and provides reading and writing activities chosen by students from classroom libraries. The first-semester mathematics intervention—Transition to Advanced Mathematics—covers five pre-algebra units with an emphasis on problem-based and contextual teaching and learning.

To evaluate the program, Balfanz et al. (2003) performed one primary study and two supplemental studies. The primary study looked at the performance of three non-selective neighborhood high schools in Baltimore compared to the performance of three control schools with similar student demographic characteristics and standardized test scores in mathematics and reading. Treatment schools received the full ninth-grade instructional program, while control schools implemented double-dosing (90 minutes a day of mathematics and English instruction for the full year, as chosen by the school). Like the treatment schools, students in the control schools were also part of ninth-grade academies, which had their own teaching staff.

The evaluation examined a number of outcome measures. Students in both treatment and control schools took the Comprehensive Test of Basic Skills (CTBS)-5 Terra Nova achievement test in reading and mathematics in February and again in May of the 1999–2000 school year. Other data sources and measures included opinion surveys completed by teachers in the experimental schools and students in the control and treatment schools, the schools' performance on State Functional exams, students' scores on the eighth-grade CTBS test, and algebra 1 pass rates calculated from school records.

The evaluators estimated a number of least squares regression models comparing student scores on the February and May (ninth-grade) CTBS assessments. For both the reading and mathematics models, students' eighth-grade test scores were used to control for prior achievement. Balfanz et al. (2003) reported that students in the experimental schools significantly outperformed the students in the control schools in both their overall level of and gains in achievement. These results held in

both the eighth grade to May and the February to May comparisons. No statistically significant differences in achievement gains were found between the eighth grade to February test administration comparisons.

The first supplemental study included three Philadelphia high schools implementing the same program for ninth-graders compared to three matched control schools. All students were given the Stanford-9 assessment in April of their eighth-grade year and again in May of their freshman year.[3] Balfanz et al. (2003) found that students in TDHS schools performed better on the Stanford-9 in both mathematics and reading than students in the comparison schools, even when controlling for attendance, gender, and age. The second supplemental study examined the effects of the first-term classes—Strategic Reading and Transition to Advanced Mathematics—on student outcomes. Students in eight TDHS high schools with ninth-grade instructional programs were given pre- and post-tests (the CTBS-5 for mathematics and the Gates-McGinitie for reading).[4] The average gain for students in mathematics was eight months over four months of instruction time. In reading, 51 percent of students gained five or more months, while 35 percent gained a year or more between September and January.

In a more narrowly defined investigation of a supplemental academic service program, Woodruff et al. (2002) evaluated the effect of a Word Identification Strategy used to assist at-risk and learning disabled ninth-grade students who were reading one or more grade levels below the ninth-grade level. A total of 124 students participated in the study, with 62 students in a treatment school and 62 students matched by demographic (e.g., gender, age, race), and where possible, educational characteristics (e.g., grade level, grade-equivalent reading scores).

Students in both the treatment and control groups were administered the Slossen Diagnostic Battery at the start and end of the research period. Students in the treatment group were taken out of their ninth-grade English classes in small groups to receive Word Identification Strategy instruction—a specialized approach to decoding multi-syllabic words—while control group students attended their normally scheduled classes. Upon completion of the post-test, Woodruff et al. (2002) found that after the four- to six-week intensive academic intervention, treatment students had gains in reading decoding as large as 6 grade levels. Male students in the treatment school had average gains in reading decoding ranging between 2.8 and 3.8 grade levels, while female students had gains ranging between 2.8 and 3.4 grade levels. Across genders, African-American students made the largest and Hispanics the lowest mean gains. Students in a matched comparison group evidenced only minimal changes. Although this study demonstrates that the intervention was effective, it is limited because

[3] Unlike the main study, students in the comparison schools did not have mathematics and English double-dose classes. Other program components remained intact.

[4] For mathematics, eight high schools across three cities were tested, and for reading, eight high schools across four cities were tested.

students were not randomly selected into treatment and control groups, and it tested only the short-term achievement gains of students.

Gamoran et al. (1997) looked at the effects of "transition" mathematics courses used to increase the rigor of mathematics instruction for ninth-grade students. Three transition courses—offered in California via the University of Chicago School Mathematics Project (UCSMP) and in New York via Stretch Regents—were analyzed using a three-level hierarchical linear model based on data from 882 students (498 students were tested three times, and 384 students were tested twice) in 48 mathematics classes in seven high schools.[5] Students were not randomly assigned to each type of instruction, but statistical controls were applied to control for differences among students (e.g., race, ethnicity, prior math grades, and a rough measure of socioeconomic status). Researchers measured individual achievement growth over time for each student, differences between students within classes, and estimated differences between classes to assess differences among classes in average achievement growth.

In comparison to students in New York's Regents classes, "students in general math and pre-algebra learned significantly less over the course of the year (Gamoran et al., 1997, p. 333)."[6] Achievement growth for students in the treatment mathematics groups, algebra, MathA/B/UCSMP, and Stretch Regents fell in the middle—between Regents and general mathematics. The more content that was covered in a class, the greater the student achievement, and when controlling for content coverage, the smaller the difference in achievement between Regents classes and other classes, implying that the difficulty of the curriculum explains much of the variation found in test scores. The results for Stretch Regents and UCSMP outcomes also supported the notion that students would benefit if these programs increased their content to correspond more closely to that of Regents Math.

School Restructuring

Four of the 10 studies that provide our evidence base focus on restructuring schools, although the actual manner in which these programs operate varied across studies. Quint, Miller, Pastor, and Cytron (1999) evaluated Project Transition, a ninth-grade reform effort that emphasized clustering teachers and students into learning teams within core academic subjects. In contrast, Dynarski et al. (1998a, 1998b) focused on whole-school reform and alternative school organization, and Kemple and Snipes (2000) focused on career academy initiatives. Both of the latter authors evaluated com-

[5] The original concept of the study was to involve eight high schools, with the goal of including two high schools from two districts in two states—San Francisco and San Diego in California and Buffalo and Rochester in New York. However, only one high school in Buffalo met the study criteria.

[6] Regents classes in New York State are college-preparatory classes that include an end-of-course exam (called a "Regents Exam") as a requirement to pass the class. Students must pass a required number of Regents classes (and associated exams) in order to graduate with a Regents Diploma.

prehensive school restructuring initiatives that included ninth-grade students as part of an effort to accomplish whole school transformation.

Findings on school restructuring were generally mixed, with two of the four studies offering positive, albeit weak, short-lived evidence of improved academic performance among participating students. Of the four projects, Project Transition was the only one that specifically targeted ninth-grade students. Project Transition's three reform elements included the creation of student-teacher teams in four core academic subjects consisting of approximately 120 students who shared similar classes; daily teacher team meetings for collaboration on professional development; and creation of coaches and other supports to aid in teacher professional development. Two demonstration high school sites implemented the program, both of which were urban school districts, with similar attendance, grades, and student characteristics. The program was evaluated using a cohort comparison design, with outcomes for students who were ninth-graders when Project Transition was implemented (treatment group) assessed against outcomes for a group of ninth-graders enrolled before the program was implemented (control group). School records and student surveys were collected to assess student outcomes.

Outcomes varied by school. In one site, the program improved the quality of relationships program participants had with peers compared to an earlier group of ninth-grade students who served as the control, but had no effect on other measurable student outcomes. In the second school, program participants reported improved relationships with teachers, increased feelings of autonomy, and increased self-reported engagement, all of which were significantly higher than for members of the control group. There were, however, few positive academic benefits. A small increase in course passing rates and the percentage of students with a GPA above grade D was found in one school, with effects greatest for those participants who had had relatively low attendance rates in middle school. No statistically significant differences were found between program participants and control group members on attendance rates or average GPA in either school.

To assess the effect of Career Academies on student performance, Kemple and Snipes (2000) analyzed data on 1,764 students at the end of the eighth or ninth grade who applied for enrollment in one of nine Career Academies. Career Academies are school-within-school structures that combine academic and career education courses in order to create smaller, more personalized learning communities and contextual teaching and learning.[7] The nine sites had fully implemented the Career Academy model. About half of the applicants were randomly assigned to participate in the program, with the remainder assigned to the control group. Students were categorized as being at *high risk*,

[7] Career Academies are often a whole-school restructuring reform strategy. This particular evaluation of career academies was included in this review because the academies were implemented in the ninth grade.

medium risk, or *low risk* of dropping out of school. The evaluation followed participants and control group students through high school until just before their expected date of graduation.

When averaged across all students, the outcome of Academy participation was inconclusive: Academy students exhibited only small, statistically insignificant reductions in dropout rates, increased progress toward high school graduation, and increased career-related course taking and participation in youth development activities. However, the evaluation found more pronounced effects for Academy students at high risk of school failure: high-risk Academy students had lower dropout rates and higher average attendance rates, and they were more likely to earn sufficient credits to meet district graduation requirements compared to high-risk students in the control group. However, Academy participants, regardless of risk category, did not score higher on standardized tests in reading and mathematics compared to students in the control group.

Much of what is understood about dropout prevention programs comes from evaluations of a series of programs funded under the U.S. Department of Education's School Dropout Demonstration Assistance Program in the mid-1990s (Dynarski et al., 1998a). Many of these dropout programs target ninth-grade students at risk of leaving high school. These evaluations used randomly assigned treatment and control groups to assess changes in dropout rates, as well as in students' academic performance and behavior. Dynarski and his colleagues examined 16 dropout prevention programs that were supported by grants from the Department between 1991 and 1995. For purposes of this review, we focused on the evaluations of two programs: one school-within-a-school for ninth-grade students in a Chicago high school serving predominantly Hispanic students and one alternative high school in Las Vegas that worked with predominantly white ninth- or tenth-graders performing below grade level.[8] The evaluation collected baseline and follow-up data from student records and questionnaires. Regardless of intervention type, no statistically significant difference was found in dropout rates or on a measure of locus of control for treatment and control group members; however, treatment group participants in the Chicago high school had significantly lower levels on a measure of self-esteem than control group students.

In a related set of studies, Dynarski et al. (1998b) also studied schools in five districts undergoing schoolwide restructuring activities funded under the same federal School Dropout Demonstration Assistance Program. For purposes of this literature review, we selected the site whose primary focus was on improving the academic experience of ninth-graders: the intervention school consisted of a ninth-grade "enclave" that provided block scheduling, small classes, and interdisciplinary instruction. In addition, the intervention included additional services provided by three community-based

[8] Other studies included in Dynarski's evaluation looked at whole-school dropout prevention programs that included ninth-graders in the sample; these studies were excluded because disaggregated data on ninth-graders were not provided.

organizations. This high school was matched to another school in the district that served students similar to those attending the intervention school.

The evaluation collected outcome measures for students, teachers, and parents. Outcomes for students—the focus of this review—included attendance and dropout rates, test scores, and attitudes and perceptions about school and self (Dynarski et al., 1998b, p. 10). Few student outcomes differed between the treatment and control group students. Absenteeism and dropout rates were essentially the same in restructuring and comparison schools, while standardized mathematics and reading scores, although slightly higher at the beginning of both tenth and eleventh grades for the treatment school, showed little improvement between the ninth and tenth grades (Dynarski et al., p. 32).

Summer Remediation Programs

Two of the 10 studies in our evidence base evaluate summer remediation programs. The Chicago Public Schools Summer Bridge Program was developed as part of an effort to end social promotion. Every third-, sixth-, and eighth-grade student with scores below the passing cutoff on the Iowa Test of Basic Skills (ITBS) is subject to being retained in grade. District policy requires every student with scores below the cutoff point to attend Summer Bridge (although not all do). Our review of this program's evaluation focuses on the results of participation in Summer Bridge for eighth-graders, specifically efforts to improve students' performance on an end-of-summer re-test that would allow them to enter the ninth grade in the fall. The second program, Summer Training and Education Program (STEP), integrates reading and mathematics remediation with life skills training and work experience.

Summer Bridge provides eighth-graders with four hours of summer classes per day for seven weeks. Classes are smaller than those students attend during the regular school year, and teachers follow a strictly defined curriculum. The evaluation used two sets of data: student test scores and transcripts that covered participants' entire school careers (to measure previous achievement) and a teacher survey. Although this study did not use random assignment, because attendees were selected based on school district policy, the researchers were able to use statistical techniques (hierarchical linear modeling) to discern differences in achievement for students who attended Summer Bridge.

Short-run improvements in test scores occurred across demographic and achievement groups. The rate of learning during the summer was higher than the rate during the course of the regular academic year. Over four years (1997–2000), eighth-graders increased their ITBS scores by approximately six months in reading and nearly five months in mathematics, on average. Higher test scores were associated with end-of-summer promotion: approximately half of the eighth-grade students

met the promotional cutoffs in both mathematics and reading and were eligible to begin high school in the fall.

The Summer Training and Education Program (STEP) spanned a 15-month period beginning the summer before the freshman (or possibly sophomore) year and continuing into the following summer (Grossman & Sipe, 1992). STEP targeted at-risk youth ages 14–15 who were JTPA-eligible and residing within five cities. STEP had two principal foci: academic remediation and "life skills and opportunities" (intended to address issues of substance abuse and sexuality). We focused on the academic outcomes for program participants for this research review. Approximately 3,000 youth were randomly assigned to treatment and control groups. Students in the treatment group received half-time work combined with half-time classes during the summers and were paid for their attendance in school and their employment. Academic remediation included an individually paced curriculum emphasizing critical thinking skills and hands-on learning (Grossman & Sipe, p. 12). Control group members received a summer job for one summer.

Evaluators tracked changes in educational outcomes over the two summers and intervening school year. Outcome measures included performance on pre- and post-tests of the reading and mathematics subtests of the Intermediate Level Metropolitan Achievement Test (MAT) Survey Battery, student youth surveys (mostly addressing fertility-related behavior and employment) and follow-up interviews, and student transcripts (Grossman & Sipe, 1992, pp. 19–23). Program results were mixed. Treatment group students had higher dropout rates than control group students in the first two terms after participation, although the difference between the groups was not statistically significant after the first year (pp. 37–39). Similarly, students in one of the two treatment groups had a lower graduation rate than students in the control group. Initial gains in academic skills and personal behaviors for program participants compared to those of the control group had dissipated by the time a long-term follow-up was conducted (two to three years after the program).

Table 1. Summary of Studies in the Evidence Base

Author & Publication Date	Sample	Type of Design	Treatment	Measures	Results
Balfanz, R., Legters, N., & Jordan, W. (2003, May). *Catching up: Impact of the Talent Development Ninth Grade Instructional interventions on reading and mathematics in high poverty high schools.* In Press.	Three TDHS treatment high schools, three matched comparison control high schools. TDHS, math = 140 students, reading = 257 students Control, math = 233 students, reading = 200 students.	Pre- and post-test with matched control comparison group	*Academic Support Services* Talent Development High Schools (TDHS) Ninth-Grade Instructional Program in Reading and Mathematics includes the following: Double dose of math and English instruction in 4x4 block scheduling, three special courses,[9] teachers receive intensive and sustained professional development, context of a Ninth-Grade Success Academy located in separate area of school with own Academy principal.	*Academic Achievement*–Standardized test scores *Student and Teacher Perceptions* –Student and teacher surveys	Students in the treatment group schools significantly outperformed students in the control schools in terms of both overall level of and gains in achievement. In addition, students at these schools passed algebra I at a higher rate and performed just as well on the state's functional math, reading, and writing exams as students in control schools.
Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998a, June). *Impacts of dropout prevention programs: Final report.* Princeton, NJ: Mathematica Policy Research, Inc.	Entire study consisted of eight middle school programs and five high school programs.[10] Analysis focuses on two high schools, ninth- and tenth-graders, two cohorts—one beginning in 1992–1993 and followed through 1994–1995, one beginning in 1993–1994 and followed through 1994–1995. Chicago school had 106 treatment and 65 control students; Las Vegas school had 287 treatment and 197 control students.	Random assignment	*Restructuring Schools* One high school was a school-within-a-school model for ninth-grade students including team teaching, small class sizes, and support services. Another high school was an alternative school for ninth- and tenth-graders behind grade level or with low grades or test scores focused on small group instruction, project based learning, and support services.	*Student Achievement*–Test scores *Student Engagement*–High school records, attendance, GPA *Student Perception* –Questionnaires	Programs had little effect on dropping out.

[9] First semester includes Strategic Reading, Transition to Advanced Mathematics, and Freshman Seminar. Second semester students take algebra I, English I, and U.S. history along with either science or an elective.

[10] Concentrated only on high school programs, and within those high school programs, only those that concentrated on ninth-grade remediation strategies.

Author & Publication Date	Sample	Type of Design	Treatment	Measures	Results
Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998b, June). *Impacts of school restructuring initiatives: Final report.* Princeton, NJ: Mathematica Policy Research, Inc.	313 ninth-graders in treatment high school and 331 ninth-graders in comparison high school.[11]	Matched comparison of schools with random samples of students within schools	*Restructuring Schools* Ninth-grade enclave with block scheduling, smaller classes, and interdisciplinary instruction	*Student Achievement*–Test scores *Student Engagement*–Dropout rates and attendance *Student Perceptions*–Questionnaires	Students did not improve on most student outcomes—absenteeism was essentially the same in restructuring and comparison schools. First cohort dropout rates at the end of tenth grade were somewhat higher for the restructuring school than the comparison school but were essentially the same by the end of the eleventh grade. Test scores improved little from the ninth to the tenth grade.
Gamoran, A., Porter, A., Smithson, J., & White, P.A. (1997, Winter). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis, 19*(4), 325-338.	882 ninth- and tenth-graders in 48 classes in seven high schools. Selected two school districts in New York and California and two schools within each district.[12]	Multiple time series design with comparison groups	*Academic Support Services* Three mathematics transition courses–Math A, Stretch Regents, and UCSMP Math[13]	*Academic Achievement*-Mathematics achievement test[14]	Rigorous content coverage distinguishes college-preparatory math classes from general-track math classes, and students learn more in college-preparatory classes. Transition courses are located in between in both coverage and achievement. Learning gains were greatest for Regents classes, followed by algebra, Stretch Regents, Math A/B/UCSMP, and general math/prealgebra.
Grossman, J.B., & Sipe, C.L. (1992, Winter). *Summer Training and Education Program (STEP).* Philadelphia: Public/Private Ventures.	1,514 students in Cohort II – 637 students in treatment group, 626 students in control group, and 1,454 students in Cohort III – 684 students in treatment group, 663 students in control group.	Random assignment	*Summer Remediation* Program had four key components: 1. Remediation–90 hours/summer in reading and math skills 2. Life Skills and Opportunities–18 hours/summer 3. Work Experience–80 hours/summer 4. School Year Support–between summer of first year and summer of second year in program	*Academic Achievement*–Intermediate Level Metropolitan Achievement Test Survey Battery and school transcripts Education attainment, employment, and fertility-related behavior Interviews	STEP had little or no positive impact on youth's educational experience—the treatment and control groups were equally likely to drop out of school, performed similarly on standardized tests, completed the same number of grades, and earned similar number of graduation credits.

[11] Longitudinal data were collected from more than 7,000 students in 21 schools in five school districts; for purposes of this review, the focus was on the one high school initiating ninth-grade reforms.

[12] Chose only one school from the Buffalo School District because only one high school met criteria.

[13] Math A was intended to replace general math courses in California and serves as a bridge to the college-preparatory courses. Stretch Regents, Regents Math I-II-III is a three-year sequence of college-preparatory mathematics in which algebra, geometry, and higher mathematics are integrated. UCSMP curriculum is a six-year sequence designed to begin in grade seven.

[14] Created a test from the National Assessment of Educational Progress (NAEP) 1990 public-release items. The test consisted of 26 problems; 75 percent of items were multiple-choice and 25 percent were short answers.

Author & Publication Date	Sample	Type of Design	Treatment	Measures	Results
Kemple, J.J., & Snipes, J.C. (2000, March). *Career Academies: Impacts on students' engagement and performance in high school.* New York: Manpower Demonstration Research Corporation.	1,764 students from nine sites: 959 are in Academy group (treatment) and 805 in non-Academy group (control).	Random assignment	*Restructuring Schools* Career academies include school-within-a-school design, integrated, career-based curriculum, employer partnerships	*Student Achievement*– School transcripts and achievement test scores *Student Perceptions*– Twelfth-grade survey	Academies significantly reduced the percentage of students in the high-risk subgroup who dropped out of high school. Thirty-two percent of the non-Academy group dropped out of high school before the end of twelfth grade, while 21 percent of the high-risk Academy group dropped out before the end of their twelfth-grade year. Academies also improved attendance rates, particularly by reducing chronic absenteeism. Academies also increased the total number of course credits earned by high-risk students. Career Academies had no effect on standardized measures of student achievement in math and reading.
Myers, D., & Shirm, A. (1999, April). *The impacts of Upward Bound: Final report for Phase I of the National Evaluation.* Washington, DC: Mathematica Policy Research, Inc.	1,500 students in treatment group and 1,300 students in control group, with two-thirds low-income and potential first-generation students.	Random assignment	*Academic Support Services* Instruction, tutoring, and counseling during the academic year and an intensive instructional program that meets daily for about six weeks during the summer	*Student Engagement*– College attendance, postsecondary credits earned, high school completion, high school course taking, school-related behavior, participation in high school activities	For students who had lower-academic performance as high school freshmen, Upward Bound (UB) had substantial impacts on students' course taking, with increases in all core subjects except science, and on earning credits in AP/honors classes, and it reduced the dropout rate by six points for this group. There was also a substantial increase in attendance at four-year colleges for this group. Also, for students who entered in ninth grade and participated for less than two years, UB had no impact on participants' educational experiences. For all students, there was a small, but significant, difference in the percentage of UB students who had graduated, with fewer UB students graduating than control group students. The program also had only small impacts on the number of credits students had earned and had no impact on extracurricular or community activities.
Quint, J.C., Miller, C., Pastor, J.J., & Cytron, R. (1999, April). *Project Transition: Testing an intervention to help high school freshmen succeed.* New York: MDRC.	Two high schools, 2,101 students, ninth-graders for the year treatment was implemented (the program group), and ninth-graders who were students the year before implementation (the comparison group).	Cohort comparison design	*Restructuring Schools* Teacher-student clusters of teams in core academic subjects, daily teacher team meetings, and teacher coach	*Student Achievement*– School transcripts *Student Perceptions*– Student surveys	At one high school, Project Transition improved the quality of students' relationships with their peers but produced no impacts on any other student outcomes. At the other high school, Project Transition improved students' relationships with their teachers, increased students' feelings of autonomy in school, and increased their reported engagement in school. It also resulted in a small increase in the rate at which students passed their courses, increasing their average number of credits earned. Neither school experienced increases in attendance rates or average GPA.

Author & Publication Date	Sample	Type of Design	Treatment	Measures	Results
Roderick, M., Engel, M., & Nagaoka, J. (2003, February). *Ending social promotion: Results from Summer Bridge.* Chicago: Consortium on Chicago School Research.	297 schools ran eighth-grade Summer Bridge, on average, each year; From 1997 through 2000, 6,804 eighth-graders did not meet cutoff on Iowa Test of Basic Skills (ITBS) and were required to attend Summer Bridge.	Pre- and post-comparisons, cohort analyses, and regression discontinuity designs	*Summer Remediation* Summer Bridge[15] is a mandatory summer school program that provides extra instructional time in reading and mathematics. The program provides four hours per day of instruction for seven weeks, a total of 140 hours of instruction, and has small classes (on average, 16 students).	*Student Achievement* – ITBS, passing and promotion rates *Teacher Perceptions* – Teacher surveys *Student Perceptions* – Student surveys	Between 1997 and 1999, the average eighth-grade Summer Bridge student increased her ITBS reading scores, using adjusted measure over six months. In the year 2000, adjusted Summer Bridge test score gains declined in reading for eighth-graders. In general, however, adjusted gains suggest that eighth-graders who attended Summer Bridge experienced substantial test score increases over the summer. Eighth-graders, on average, increased their ITBS test scores by approximately six months in reading and nearly five months in mathematics across all four years (1997–2000). Eighth-grade students in Summer Bridge increased their reading achievement at over three times the rate of the academic year. Between 1997 and 2000, approximately half of eighth-graders required to attend Summer Bridge met the promotional cutoffs in both subjects by the end of the summer.
Woodruff, S., Schumaker, J.B., & Deshler, D.D. (2002). *The effects of an intensive reading intervention on the decoding skills of high school students with reading deficits.* Institute for Academic Access	124 ninth-grade students, 62 students in each group–control and treatment	Pre- and Post-Matched comparison	*Academic Support Services* Students taught Word Identification Strategy, a strategy for decoding words, for 4–6 weeks.	*Student Achievement*– Word Identification Subtest of the Slosson Diagnostic Screening Test for Reading, Form A	Strategy was effective in producing mean gains in reading decoding as large as six grade levels. In addition, students with and without learning disabilities benefited from the instruction.

[15] This study looked at third-, sixth-, and eighth-grade Summer Bridge programs, but due to the restraints of this review, we look only at the results and design of the eighth-grade Summer Bridge program.

Overview and Summary

The U.S. Department of Education established the WWC to serve as an independent, centralized source of scientific evidence on what works in education. To fulfill that mission, the Clearinghouse has developed a number of resources, including the DIAD and CREAD instruments, to assist educators, policymakers, and the public in obtaining objective, scientifically rigorous research on education reform strategies. This section summarizes results from the review of the literature on ninth-grade remediation programs and closes with an assessment of the merits of "ninth-grade remediation" programs as a subject for a *WWC Evidence Report*.

Confidence in Causal Inferences

The CREAD is a coding instrument that assists researchers in expressing their confidence in any conclusions that they make about the causality of an intervention, based on the entire body of accumulated evidence they have amassed. Designed to be completed by the evidence report team after reviewing all studies that successfully pass the DIAD assessment, the CREAD provides a standardized approach for reporting the validity characteristics of the reported evidence base.

Applying the CREAD is contingent on the depth, breadth, and consistency of studies that have met the minimum criteria for study inclusion, as defined in the Study DIAD. Use of a key wording extraction template—which enabled reviewers to assess the extent to which program outcomes were clearly defined; how participants were selected; and whether factors outside the study affected observed outcomes—provided researchers with some assurance that the *breadth* and *consistency* of selected studies fell within desired confidence levels. Specifically, efforts to restrict selections to analyses that focused on students currently in or entering ninth grade, examined outcome measures relating to a discrete set of academic and student engagement factors, and involved equally matched treatment and comparison populations (although methodological approaches varied) helped to ensure that a majority of studies were both valid and representative of the evidence base. (One study—Summer Bridge—did not have a comparison group, per se.)

Unfortunately, small numbers of overlapping studies in the evidence base, differences in how studies scored on the DIAD, and variation among studies in the mechanism of operation indicate that the *depth* of cumulative evidence will not support meta-analysis of statistical findings. Although a total of 10 ninth-grade remediation studies were identified, these studies described three very different types of program interventions (i.e., academic support, school restructuring, and summer reme-

diation programs). Studies also differed within intervention types: for example, among the four studies examining school restructuring, one focused on a ninth-grade academy limited to freshmen, one on a Career Academy model extending across four years of high school, one on a school-within-a-school targeting dropout prevention, and the final study on a school undergoing comprehensive schoolwide reorganization.

Project components also varied across interventions, with some providing multiple services with no attempt to control for the marginal contribution of individual factors. While individual studies were generally well formulated to allow researchers to assess differences between treatment and control groups within studies, cross-study comparisons were hindered by a lack of replicability (of common treatment and common outcome measures).

Table 2. Study DIAD Summary Report

Study DIAD Dimension	Yes		Maybe Yes		Maybe No	
	#	%	#	%	#	%
Composite Questions						
1. Construct Validity–Intervention	8	80			2	20
2. Construct Validity–Outcomes						
Academic Support Services	4	40				
Restructuring Schools	4	40				
Summer Remediation	2	20				
3. Internal Validity–Selection	4	40	6	60		
4. Internal Validity–Contamination	3	30	7	70		
5. External Validity–Sampling			8	80	2	20
6. External Validity–Testing Within Subgroups	4	40	3	30	3	30
7. Statistical Conclusion Validity– Effect Size Estimation	8	80	2	20		
8. Statistical Conclusion Validity– Completeness of Report	7	70	3	30		
Global Questions						
1. Construct Validity	7	70			2	20
2. Internal Validity	2	20	8	80		
3. External Validity	6	60			4	40
4. Statistical Conclusion Validity	6	60	3	30	1	10

While the lack of *depth* across studies precludes calculation of effect sizes, a review of studies that met or exceeded Study DIAD parameters provides some useful, albeit general, insights into the impact of different remediation strategies.

Evidence Supporting Ninth-Grade Remediation Strategies

In the preliminary review of the literature, MPR researchers identified an array of strategies that educators commonly use to remedy the educational deficiencies of high school freshmen. While many of these strategies enjoy widespread acceptance in the field, few have been subjected to scientifically rigorous examination to objectively assess their merit. As this review confirms, only a few published studies are methodologically strong: synthesis researchers moved only 10 of the 598 studies initially found into the DIAD review process, and of these, not all were judged sound on all study parameters (see Table 2).

Although much scientifically rigorous research has yet to be performed, findings from studies that met or exceeded the DIAD inclusionary criteria indicate that some ninth-grade remediation strategies are associated with positive outcomes for students and, in particular, for students at high risk of dropping out or academic failure. Positive outcomes were noted in six studies in the evidence base. These studies were generally associated with academic support programs that provided supplemental academic services, specifically in mathematics and reading, using accelerated or specially designed curriculum.

Although these positive findings are encouraging, it is not clear whether observed gains are either cost effective or persist over time. As noted by Myers and Shirm (1999) in their evaluation of Upward Bound—one of the few studies for which robust samples and longitudinal data were available—long-term outcomes for ninth-grade students were no different than for students who entered Upward Bound in the tenth grade. Similarly, Grossman and Sipe (1992) noted that initial program gains attributed to the STEP program were no longer evident at long-term follow-up.

What is taught may also be as important as how the information is communicated. Both Woodruff et al. (2002) and Gamoran et al. (1997) reported that treatment groups exposed to supplemental academic services, using specially designed curricula, outperformed control groups receiving traditional services. Moreover, in his evaluation of the Talent Development High School model, Balfantz et al. (2003) demonstrated that double-dosing itself is more effective when implemented with specially designed curricula delivered by trained instructors.

The research also indicates that school restructuring initiatives, creating alternative schools, or grouping students in themed Career Academies offers little, if any, lasting benefits for students. For example, though there is some evidence of improved attitudes and expectations among students participating in Career Academies, no lasting academic benefits were observed.

Assessment of the Subject for Review

While there appear to be relatively small positive outcomes associated with certain types of interventions, it is impossible, using the current evidence base, to draw reliable conclusions about the direction and magnitude of effect sizes. Given the breadth of remediation strategies, the small number of scientifically rigorous studies that currently exist, and the lack of replicability across studies of similar design, it is not apparent that ninth-grade remediation strategies, as currently defined, merit further consideration as a topic for a *WWC Evidence Report*.

References

Achieve. (2001). *Standards: How high is enough?* Achieve Policy Brief #3. Washington, DC: Author.

Alliance for Excellent Education. (2003). Retrieved September 15, 2005, from http://www.all4ed.org/about/.

Alspaugh, J. W. (1998, September/October). Achievement loss associated with the transition to middle school and high school. *The Journal of Educational Research, 92*(1), 20–25.

Alspaugh, J. W. (2000, Fall). The effect of transition grade to high school, gender, and grade level upon dropout rates. *American Secondary Education, 29*(1), 2–9.

Alt, M. N., & Choy, S. P. (2000). *In the middle: Characteristics of public schools with a focus on middle schools* (NCES 2000–312). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Ascher, C. (1987, August). *The ninth grade—a precarious time for the potential dropout*. ERIC Digest. Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education. (ERIC ED284922)

Balfanz, R., Legters, N., & Jordan, W. (2003, May). *Catching up: Impact of the talent development ninth grade instructional interventions on reading and mathematics in high poverty high schools*. In Press.

Barone, C., Aguirre-Deandreis, A. I., & Trickett, E. J. (1991). Mean-ends problem-solving skills, life stress, and social support as mediators of adjustment in the normative transition to high school. *American Journal of Community Psychology, 19*(2), 207–225.

Blyth, D. A., Simmons, R. G., & Carlton-Ford, S. (1993). The adjustment of early adolescents to school transitions. *Journal of Early Adolescence, 3*(12), 105–120.

Braddock, J. H., & McPartland, J. M. (1993). Education of early adolescents. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 19, pp. 135–170). Washington, DC: American Educational Research Association.

Bryk, A. S. (1994). More good news that school organization matters. In *Issues in restructuring schools. Issues report no. 7*. Madison, WI: Center on Organization and Restructuring Schools.

Bryk, A. S., & Driscoll, M. E. (1988). *The high school as communities: Theoretical foundations, contextual influences, and consequences for students and teachers*. Madison, WI: National Center on Effective Secondary Schools, University of Wisconsin.

Bryk, A. S., & Thum, Y. M. (1989). The effects of high school organization on dropping out: An exploratory investigation. *American Educational Research Journal, 26*(3), 353–383.

Carnegie Council on Adolescent Development, Task Force on Education of Young Adolescents. (1989). *Turning points: Preparing American youth for the 21st century*. Washington, DC: Author.

Cave, G., & Quint, J. (1990). *Career beginnings impact evaluation: Findings from a program for disadvantaged high school students*. New York: Manpower Demonstration Research Corporation. (ERIC 325598)

Committee for Economic Development. (2000). *Measuring what matters: Using assessment and accountability to improve student learning*. New York: Author.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin Company.

Cooledge, N. J., & Wurster, S. R. (1985). Intergenerational tutoring and student achievement. *Reading Teacher, 39*, 343–346.

Cooper, H., Charlton, K., Valentine, J. C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development, 65*(1), 1–117.

Cotton, K. (1996). *School size, school climate, and student performance*. Close-up Number 20. Portland, OR: Northwest Regional Educational Laboratory. Retrieved September 30, 2005, from http://www.nwrel.org/scpd/sirs/10/c020.html.

Cotton, K. (2001, December). *New small learning communities: Findings from recent literature*. Portland, OR: Northwest Regional Educational Laboratory.

Devin-Sheehan, L., Feldman, R., & Allen, V. (1976). Research on children tutoring children: A critical review. *Review of Educational Research, 436*, 355–385.

Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998a, June). *Impacts of dropout prevention programs: Final report*. Princeton, NJ: Mathematica Policy Research, Inc.

Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998b, June). *Impacts of school restructuring initiatives: Final report*. Princeton, NJ: Mathematica Policy Research, Inc.

Eccles (Parsons), J., Midgley, C., & Adler, T. F. (1984). Grade-related changes in the school environment: Effects on achievement motivation. In J. G. Nicholls (Ed.), *Advances in motivation and achievement* (Vol. 3, pp. 283–331). Greenwich, CT: JAI Press.

Evans, W., & Bechtel, D. (1997). *Extended school day/year programs: A research synthesis*. *Spotlight on student success*. Philadelphia: Laboratory for Student Success. (ERIC ED461695)

Felner, R. D., Brand, S., Adan, A. M., Mulhall, P. F., Flowers, N., Sartain, B., & DuBois, D. L. (1993). Restructuring the ecology of the school as an approach to prevention during school transitions: Longitudinal follow-ups and extensions of the School Transitional Environment Project (STEP). *Prevention in Human Services, 10*(2), 103–136.

Flaxman, E. (1992). *Evaluating mentoring programs*. New York: Institute for Urban and Minority Education, Teachers College, Columbia University.

Flaxman, E., & Ascher, C. (1992). *Mentoring in action*. New York: Institute for Urban and Minority Education, Teachers College, Columbia University.

Gamoran, A., Porter, A. C., Smithson, J., & White, P. A. (1997, Winter). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis, 19*(4), 325–338.

Green, P., & Scott, L. (1995). *"At-risk" eighth-graders four years later* (NCES 95–736). U.S. Department of Education. Washington, DC: National Center for Education Statistics Statistics in Brief. (ERIC ED386496)

Grossman, J. B., & Sipe, C. L. (1992, Winter). *Summer Training and Education Program (STEP)*. Philadelphia, PA: Public/Private Ventures.

Hertzog, C. J., & Morgan, P. L. (1999, January–February). Making the transition from middle level to high school. *High School Magazine, 6*(4), 26–30.

Hertzog, C. J., & Morgan, P. L. (1999). *Transition: A process not an event*. Reston, VA: National Association of Secondary School Principals.

Howley, C., Strange, M., & Bickel, R. (2000). *Research about school size and school performance in impoverished communities*. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools. Retrieved September 30, 2003, from http://ael.org/eric/digests/edorc0010.htm.

Irmsher, K. (1997). *School size*. Eugene, OR: ERIC Clearinghouse on Educational Management. (ERIC ED414615)

Irvin, J. (Ed.). (1997). *What current research says to the middle level practitioner*. Columbus, OH: National Middle School Association.

Jekielek, S. M., Moore, K. A., Hair, E. C., & Scarupa, H. J. (2002). *Mentoring: A promising strategy for youth development*. Washington, DC: Child Trends.

Johnston, J. H. (1992). Climate and culture as mediators of school values and collaborative behavior. In J. L. Irvin (Ed.), *Transforming middle level education: Perspectives and possibilities*. Boston: Allyn and Bacon.

Kemple, J. (2001, December). *Career academies: Impacts on students' initial transitions to post-secondary education and employment*. New York: Manpower Demonstration Research Corporation.

Kemple, J., & Snipes, J. (2000, February). *Career academies: Impacts on students' engagement and performance in high school*. New York: Manpower Demonstration Research Corporation.

Kerr, K. A. (2002, Summer). An examination of approaches to promote ninth-grade success in Maryland public high schools. *Educational Research Service Spectrum*. Retrieved September 17, 2003, from http://www.ers.org/spectrum/sum02a.htm.

Klonsky, M. (1998). *Small schools: The numbers tell a story. A review of the research and current experiences*. Chicago: University of Illinois at Chicago, Small Schools Workshop. Retrieved September 30, 2005, from http://www.smallschoolsworkshop.org/klonsky.html.

Lee, V. E., Bryk, A. S., & Smith, J. B. (1993). The organization of effective secondary schools. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 19, pp. 135–170). Washington, DC: American Educational Research Association.

Lee, V. E., & Smith, J. B. (1995, October). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education, 68*(4), 241–270.

Lee, V. E., & Smith, J. B. (1997). High school size: Which works best, and for whom? *Educational Evaluation and Policy Analysis, 19*(3), 205–227.

Legters, N. E. (2000). Small learning communities meet school-to-work: Whole-school restructuring for urban comprehensive high schools. In M. G. Sanders (Ed.), *Schooling students placed at risk: Research, policy, and practice in the education of poor and minority adolescents*. Mahwah, NJ: Erlbaum Associates, Inc.

Legters, N. E., & Kerr, K. A. (2001, January 13). *Easing the transition to high school: An investigation of reform practices to promote ninth grade success*. Paper presented at the Dropouts in America: How Severe Is the Problem?

What Do We Know About Intervention and Prevention? A Forum Convened by the Harvard Civil Rights Project, Cambridge, MA.

Lindstrom, J., & Raddatz, A. (2003). Characteristics of middle adolescents. In *Volunteer training models: Washington State University 4-H youth development*. Pullman, WA: Washington State University Cooperative Extension.

Lucas, S. R. (1999). *Tracking inequality: Stratification and mobility in American high schools*. New York: Teachers College Press.

Mac Iver, D. J. (1991, October). *Helping students who fall behind: Remedial activities in the middle grades*. Report 9. Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Mayer, L. (1995, May). Bringing 'em up right: Making school transitions a success. *Schools in the Middle* 4(4), 41–42.

McPartland, J. M., & Nettles, S. M. (1991). Using community adults as advocates or mentors for at-risk middle school students: A two-year evaluation of Project Raise. *American Journal of Education, 99*(4), 568–586.

Mizelle, N. B. (1999, August). *Helping middle school students make the transition into high school*. ERIC Digest. Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education. (ERIC ED432411)

Myers, D., & Shirm, A. (1999, April). *The impacts of Upward Bound: Final report for phase I of the National Evaluation*. Washington, DC: Mathematica Policy Research, Inc.

National Association of Secondary School Principals. (1985). *How fares the ninth grade? A day in the life of a ninth grader*. Reston, VA: Author.

National Association of Secondary School Principals. (1996). *Breaking ranks: Changing an American institution*. Reston, VA: Author.

National Association of State Boards of Education. (2000). *Failure is not an option: The next stage of education reform. The report of the NASBE Study Group on Alternatives to Retention and Social Promotion*. Retrieved September 17, 2003, from http://www.nasbe.org/Educational_Issues/Students_High.html.

Page, L., Layzer, C., Schimmenti, J., Bernstein, L., & Horst, L. (2002, February). *National evaluation of smaller learning communities: Literature review*. Washington, DC: Abt Associates, Inc.

Pecori, J. (2000, Autumn). *State innovations: State responses to the need for student remediation*. National Association of State Boards of Education. Retrieved September 18, 2003, from http://www.nasbe.org/Educational_Issues/Students_High.html.

Powell, A. G., Farrar, E., Cullen, D. K., the National Association of Secondary School Principals (U.S.), & the National Association of Independent Schools. (1985). *The shopping mall high school: Winners and losers in the educational marketplace*. Commission on Educational Issues. Boston: Houghton Mifflin.

Quint, J. C., Miller, C., Pastor, J. J., & Cytron, R. (1999, April). *Project transition: Testing an intervention to help high school freshmen succeed*. New York: MDRC.

Raywid, M. A. (1995). *Taking stock: The movement to create mini-schools, schools-within-schools, and separate small schools*. New York: ERIC Clearinghouse on Urban Education, Teachers College, Columbia University.

Reyes, O., Gillock, K., & Kobus, K. (1994). A longitudinal study of school adjustment in urban, minority adolescents: Effects of a high school transition program. *American Journal of Community Psychology, 22*(3), 341–369.

Rice, J. K. (1997). Explaining the negative impact of the transition from middle to high school on student performance in math and science. *Educational Administration Quarterly, 37*(3), 372–400.

Roderick, M. (1993). *The path to dropping out: Evidence for intervention*. Westport, CT: Auburn House.

Roderick, M. (1995). School transitions and school dropout. In K. K. Wong (Ed.), *Advances in educational policy* (pp. 135–185). Greenwich, CT: Jai Press.

Roderick, M. (1996, September 26). Roderick on school reform. *The University of Chicago Chronicle*.

Roderick, M., & Camburn, E. (1996). Academic difficulty during the high school transition. In P. Sebring, A. S. Bryk, M. Roderick, & E. Camburn (Eds.), *Charting reform in Chicago: The students speak* (pp. 47–65). Chicago: The University of Chicago, Consortium on Chicago School Research.

Roderick, M., & Camburn, E. (1999, Summer). Risk and recovery from course failure in the early years of high school. *American Educational Research Journal, 36*(2), 303–343.

Roderick, M., Engel, M., & Nagaoka, J. (2003, February). *Ending social promotion: Results from Summer Bridge*. Chicago: Consortium on Chicago School Research.

Seidman, E., Aber, J. L., Allen, L., & French, S. E. (1996). The impact of the transition to high school on the self-system and perceived social context of poor urban youth. *American Journal of Community Psychology, 24*(4), 489–515.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.

Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence: The impact of pubertal change and school context*. New York: Aldine de Gruyter.

Smith, J. B. (1997, January/February). Effects of eighth-grade transition programs on high school retention and experiences. *Journal of Educational Research, 90*(3), 144–152.

Somers, C. L., & Piliawsky, M. (2002, April). *Drop-out prevention among urban, African-American, ninth grade adolescents*. Paper presented at the American Educational Research Association conference, Wayne State University, Detroit, MI.

Valentine, J. C., & Cooper, H. (2003). *What Works Clearinghouse study design and implementation assessment device (version 1.0)*. Washington, DC: U.S. Department of Education.

Wasik, B. A., & Slavin, R. E. (1990). *Preventing early reading failure with one-to-one tutoring: A best evidence synthesis*. Report No. 6. Baltimore: Johns Hopkins University, Center for Research on Effective Schooling for Disadvantaged Students.

Wasley, P., & Lear, R. (2001). Small schools, real gains. *Educational Leadership, 58*(6), 22–27.

Wasley, P. A., Fine, M., King, S. P., Powell, L. C., Holland, N. E., Gladden, R. M., & Mosak, E. (2000). *Small schools: Great strides. A study of new small schools in Chicago*. New York: The Bank Street College of Education. Retrieved September 30, 2003, from http://www.bankstreet.edu/news/SmallSchools.pdf.

Wilks, R. T. J., & Clarke, V. A. (1988). Training vs. nontraining of mothers as home reading tutors. *Perceptual and Motor Skills, 67*, 135–142.

Woodruff, S., Schumaker, J. B., & Deshler, D. D. (2002). The effects of an intensive reading intervention on the decoding skills of high school students with reading deficits. Lawrence, KS: Institute for Academic Access.

Appendix A: Boolean Logic Search

Search Engine: ERIC, http://askeric.org/Eric/

Search 1: remediation and high school, 1993–2003
Findings: 456 documents

Search 2: remediation and high school and ninth grade, 1993–2003
Findings: 12 documents

Search 3: remediation and high school and ninth grade and achievement, 1993–2003
Findings: 7 documents

Search 4: remediation and high school and academic achievement, 1993–2003
Findings: 88 documents

Search 5: remediation and ninth grade, 1993–2003
Findings: 14 documents

Search 6: remediation and ninth grade and achievement, 1993–2003
Findings: 8 documents

Search 7: ninth grade and academy, 1993–2003
Findings: 13 documents

Search 8: ninth grade and academy, 1993–2003
Findings: 3 documents

Search 9: freshman and academy and high school, 1993–2003
Findings: 8 documents

Search 10: double dose, 1993–2003
Findings: 1 document

Search 11: ninth grade and transitions, 1993–2003
Findings: 27 documents

Search 12: ninth grade and transitions and achievement, 1993–2003
Findings: 11 documents

Search 13: ninth grade and transitions and remediation, 1993-2003
Findings: 2 documents

Search 14: transitions and remediation, 1993–2003
Findings: 132 documents

Search 15: transitions and remediation and high school, 1993–2003
Findings: 40 documents

Search 16: summer bridge and ninth grade, 1993–2003
Findings: 0 documents

Search 17: ninth grade and magnet program, 1993–2003
Findings: 0 documents

Search 18: ninth grade and mentoring programs, 1993–2003
Findings: 3 documents

Search 19: ninth grade and mentoring and remediation, 1993–2003
Findings: 0 documents

Search 20: ninth grade and small schools, 1993–2003
Findings: 0 documents

Search 21: ninth grade and school-within-a-school, 1993–2003
Findings: 23 documents

Search 22: ninth grade and school-within-a-school and achievement, 1993–2003
Findings: 14 documents

Search 23: transition and middle and high school and achievement, 1993–2003
Findings: 47 documents

Search 24: random experiment and ninth grade and remediation, 1993–2003
Findings: 0 documents

Search 25: quantitative and ninth grade and remediation, 1993–2003
Findings: 0 documents

Search 26: after school and remediation and high school, 1993–2003
Findings: 7 documents

Search 27: summer school and remediation and high school, 1993–2003
Findings: 6 documents

Search 28: year-round school and remediation and high school, 1993–2003
Findings: 5 documents

Search 29: upward bound and achievement and evaluation
Findings: 35 documents

Search 30: quantum opportunities
Findings: 7 documents

Search 31: school transitional environment
Findings: 0 documents

Search 32: ninth grade and success academy
Findings: 2 documents

Search 33: ninth grade and dropout prevention
Findings: 62 documents

Search 34: ninth grade and dropout prevention and achievement
Findings: 13 documents

Search Engine: Google, www.google.com and Yahoo, www.yahoo.com, PsychInfo, and Dissertation Abstracts International

In addition to searching with terms listed above, also searched the following:

Search 1: "ninth grade" remediation programs evaluation

Search 2: ninth-grade remediation study, OR evaluation, OR summer OR school, OR after OR school, OR double OR dose, OR mentoring, OR tutoring, OR transitions, OR remediation OR program, OR year-round "academic achievement" -postsecondary -higher –college

Search 3: remediation high school study evaluation study, OR effective, OR evaluation, OR summer OR school, OR after OR school, OR double OR dose, OR mentoring, OR tutoring, OR transitions, OR remediation OR program, OR year-round "academic achievement, ninth grade" -postsecondary -higher –college

Many of these searches produced the same results that were found in ERIC. References within studies that were acquired were also searched for additional key words and studies. We then looked for these studies or keywords in the search engines listed above.

Appendix B: Bibliography for 72 Pre-DIAD Studies

During the third stage, full-text copies of each of the 72 articles were ordered and research staff read and classified studies using a keywording extraction template. Prior to using the instrument, template instructions were reviewed for clarity and tested for reliability by research staff trained in its use. To warrant inclusion in the DIAD process, a study had to have employed either an experimental design (i.e., random assignment, pretest/posttest, Solomon four-group, or posttest only control group design); or a quasi-experimental design: (i.e. non-randomly assigned treatment and comparison groups, time-series experiments, equivalent time-samples design, equivalent materials design, nonequivalent control group design, counterbalanced designs, separate-sample pretest/posttest, separate-sample pretest/posttest control group design, multiple time-series design, recurrent institutional cycle design, or regression-discontinuity analysis).

Advances in Education Research. (1999, Fall). *Advances in education research*. Issue 4. (ERIC ED439660)

Baenen, N., & Lloyd, W. (2000). *Is summer school effective for remediation in algebra I? Research watch. E&R report.* Raleigh, NC: Wake County Public School System, Department of Evaluation and Research. (ERIC ED446114)

Baenen, N., & Yaman, K. (2000–2001). *A structure of Accelerated Learning Program (ALP) efforts*. Raleigh, NC: Wake County Public School System, Department of Evaluation and Research. (ERIC ED457228)

Balfanz, R., Legters, N., & Jordan, W. (2003, May). *Catching up: Impact of the talent development ninth grade instructional interventions on reading and mathematics in high poverty high schools*. In Press.

Bauer, R., & Michael, R. (1993, April 12–16). *They're still in school: Results of an intervention program for at-risk high school students.* Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta.

Bradford Jr., J. C. (1995, February 11–15). *Year-round schools: A twenty-year follow-up study of a nationally recognized single track four-quarter plan at the high school level.* Paper presented at the Annual Meeting of the National Association for Year-round Education, San Diego. (ERIC ED381855)

Bradford Jr., J. C. (1996, April 8–12). *Year-round schools: A twenty-year follow-up study of a nationally recognized single track four-quarter plan at the high school level.* Paper presented at the Annual Meeting of the American Educational Research Association, New York. (ERIC ED396405)

Castellano, M., Stringfield, S., Stone, J., & James III, R. (2002). *Helping disadvantaged youth succeed in school: Second-year findings from a longitudinal study of CTE-based whole-school reforms*. Columbus, OH: National Dissemination Center for Career and Technical Education. (ERIC ED465063)

Center for Applied Research and Educational Improvement. *Saint Paul extended day and week programs*. Retrieved September 30, 2003, from http://education.umn.edu/CAREI/Reports/calendars/stpaul_eval/volumei/sectionIV.html.

Charles, B. (2000). *Hartford public high school: Analysis of 1998–1999 first quarter ninth grade student records*. Chaplin, CT: Curriculum Research and Evaluation. (ERIC ED438330)

Ciaccio, L. A., & Sanders, J. W. (1996). *Project Discovery II. FIPSE project*. Washington, DC: Fund for the Improvement of Postsecondary Education.

Clark, A. (1993). *Project Reach final evaluation report, 1992–1993*.

Corbett, H. D., & Wilson, B. L. (2001). *Sustaining reform: Students' appraisals of the second year in talent development high schools in Philadelphia, 2000–2001*. Philadelphia: Philadelphia Education Fund. (ERIC ED460221)

DaGiau, B. J. (1997). *A program of counseling and guidance to facilitate the transition from middle school to high school*. (ERIC ED413562)

Dangler, S. A. (1994). *Intervention and remediation strategies for the Ohio ninth grade proficiency tests: Implementation and perceived success in northwest Ohio secondary schools*. (ERIC ED382673)

Duke, D. L., Bourdeaux, J., Epps, B., & Wilcox, T. (1998). *Ninth grade transition programs in Virginia*. Charlottesville: Thomas Jefferson Center for Educational Design, University of Virginia.

Dynarski, M., & Gleason, P. (1998). *How can we help? What we have learned from evaluations of federal dropout-prevention programs*. Princeton, NJ: Mathematica Policy Research, Inc.

Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998a, June). *Impacts of dropout prevention programs: Final report*. Princeton, NJ: Mathematica Policy Research, Inc.

Dynarski, M., Gleason, P., Rangarajan, A., & Wood, R. (1998b, June). *Impacts of school restructuring initiatives: Final report*. Princeton, NJ: Mathematica Policy Research, Inc.

Elliott, M. N., Hanser, L. M., & Gilroy, C. L. (2000). *Evidence of positive student outcomes in JROTC career academies*. Santa Monica, CA: RAND National Defense Research Institute.

Fashola, O., & Slavin, R. (1997, March). *Effective dropout prevention and college attendance programs for Latino students*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago. (ERIC ED410084)

Fashola, O. S. (1998). *Review of extended day and after school programs and their effectiveness*. Report No. 24. Baltimore: Johns Hopkins University, Center for Research on the Education of Children Placed at Risk (CRESPAR). Retrieved September 30, 2003, from http://www.gse.uci.edu/asp/aspeval/resources/Report_24_Fashola.pdf.

Gamoran, A., Porter, A. C., Smithson, J., & White, P. A. (1997, Winter). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis, 19*(4), 325–338.

Grossman, J. B., Price, M. L., Fellerath, V., Jucovy, L. Z., Kotloff, L. J., Raley, R., & Walker, K. E. (2002, June). *Multiple choices after school: Findings from the extended-service schools initiative*. Public Private Ventures and ManPower Demonstration Research Corporation. Retrieved September 15, 2003, from http://www.mdrc.org/publications/48/full.pdf.

Grossman, J. B., & Sipe, C. L. (1992, Winter). *Summer Training and Education Program (STEP)*. Philadelphia, PA: Public/Private Ventures.

Hahn, A. (1992, June). *Evaluation of the quantum opportunities program: Interim impacts covering the ninth-tenth grades (1989–1991)*. Waltham, MA: Brandeis University, Center for Human Resources.

Ham, S., Doolittle, F. C., Ventura, A. M., & Jackson, R. (2000). *Project Grad Newark: Building the foundation for improved student performance.* Manpower Demonstration Research Corporation. Retrieved September 30, 2003, from http://www.mcrc.org/publications/42-full.pdf.

Honig, B. (1987). *Year-round education: Year round opportunities. A study of year-round education in California.* California State Department of Education. Retrieved September 30, 2003, from http://www.mofo.com/decentschools/studies/Year_Round_Education.pdf.

Horn, L., & Chen, X. (1998). *Toward resiliency: At-risk students who make it to college*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Huang, D., Gribbons, B., Kim, K. S., Lee, C., & Baker, L. B. (2000). *A decade of results: The impact of the LA's Best after school enrichment program on subsequent student achievement and performance*. Los Angeles: UCLA Center for the Study of Evaluation (CSE).

Hudley, C. (1996, April 8–12). *Educational alternatives for at-risk adolescent learners: Two case examples*. Paper presented at the Annual Meeting of the American Educational Research Association, New York.

Johnson, A. (1999). *Sponsor-A-Scholar: Long-term impacts of a youth mentoring program on student performance*. Mathematica Policy Research, Inc. Retrieved September 30, 2003, from http://www.cmwf.org/programs/youth/johnson_sp-s-sch_355.asp.

Jones, H. (1995, April 28–30). *Ninth grade repeaters–why they did better in summer school.* Paper presented at the Multidisciplinary Conference on the Scholarship and Creativity of African Americans, Towson, MD.

Kemple, J. J., & Snipes, J. C. (2000, March). *Career academies: Impacts on students' engagement and performance in high school*. New York: Manpower Demonstration Research Corporation.

Kerr, K. (2002, Summer). An examination of approaches to promote ninth-grade success in Maryland public high schools. *Education Research Service Spectrum.* Retrieved September 30, 2003, from http://www.ers.org/spectrum/sum02a.htm.

Marnell, E. R. (2000). *Administration of multi-faceted reading skills programs: A dual text program for ninth graders; applying reading strategies in social studies content area and teaching reading skills in reading class*. Decatur, GA: DeKalb County School System. (ERIC ED445314)

Matos, C. M. (1998). *Decreasing early school withdrawal and retention of Hispanic to other Limited English Proficient (LEP) students through counseling and arts/career exploration*. Dissertation, Nova Southeastern University. (ERIC ED428308)

McMillan, B. J. (1999). *Achievement in year-round schools*. Raleigh, NC: North Carolina Department of Public Instruction. Retrieved September 30, 2003, from http://www.geocities.com/weswalker99/NCStudy/McMillen1.htm.

McPartland, J. M., Legters, N., Jordan, W., & McDill, E. L. (1996). *The talent development high school: Early evidence of impact on school climate, attendance, and student promotion*. Report No. 2. Baltimore: Johns Hopkins University and Howard University, Center for Research on the Education of Students Placed at Risk (CRESPAR).

Metis Associates, Inc. (2002, March). *Evaluation report. New York City board of education summer school 2001*. New York: Author.

Morris, L. (1998). *The effects of integrated curriculum on 9th grade at-risk students*. Dissertation, Saint Xavier University and IRI/Skylight. (ERIC ED422419)

Munoz, M. A. (2002). *Facing the challenges of at-risk students in urban school districts: The impact of an attendance and dropout prevention program in non-traditional school*. Louisville, KY: Jefferson Public Schools.

Myers, D., & Shirm, A. (1999, April). *The impacts of Upward Bound: Final report for phase I of the National Evaluation*. Washington, DC: Mathematica Policy Research, Inc.

Neild, R., & Weis, C. (1999). *The Philadelphia Education Longitudinal Study (PELS): Report on the transition to high school in the school district of Philadelphia*.

North Carolina State Department of Public Instruction, Raleigh Division of Accountability/Testing. (1998). *Report of student performance on North Carolina competency standard 1997–98*. (ERIC ED425183)

North East Association for Institutional Research. (1999, November 13–16). *Sailing into the new millennium: Charting the course for institutional research.* Paper presented at the Annual Conference Proceedings, Newport, RI. (ERIC ED439148)

Portz, D. J. (2000). *For the first time ever: The extraordinary efforts in Massachusetts schools to get help to the students who need it most.* Mass Insight Education. Retrieved September 30, 2003, from http://www.massinsight.org/pubs.htm.

Portz, D. J., & Guadet, R. (2000). *A new commitment: Effective remediation strategies for high school students*. Mass Insight Education. Retrieved September 30, 2003, from http://www.massinsight.org/pdf/A%20New%20Commitment.pdf.

Powers, S., & McConner, S. (1997). *Project SOAR 1996–1997. Evaluation report*. Tucson, AZ: Creative Research Associates, Inc.

Quint, J. C., Miller, C., Pastor, J. J., & Cytron, R. (1999, April). *Project transition: Testing an intervention to help high school freshmen succeed*. New York: MDRC.

Roderick, M., Engel, M., & Nagaoka, J. (2003, February). *Ending social promotion: Results from Summer Bridge*. Chicago: Consortium on Chicago School Research.

Rossman, S., & Morley, E. (1995). *The national evaluation of cities and schools: Executive summary*. Washington, DC: The Urban Institute.

Schalago-Schirm, C. (1995, April). *Does the computer-assisted remedial mathematics program at Kearny High School lead to improved scores on the N.J. Early Warning Test?* Paper in partial fulfillment of the requirements for the Masters of Arts Degree, Kean College of New Jersey.

Southern Regional Education Board (SREB). (2002). *Opening doors to the future: Preparing low-achieving middle grades students to succeed in high school. 2002 outstanding practices*. Washington, DC: Office of Educational Research and Improvement. (ERIC ED469956)

Syropoulos, M. (1997a). *Evaluation of the 1996–97 ninth grade restructuring program. Area A*. Detroit, MI: Research and Evaluation Specialists, Inc., Detroit Public Schools, MI. Office of Research, Evaluation and Testing. (ERIC ED417246)

Syropoulos, M. (1997b). *Evaluation of the 1996–97 ninth grade restructuring program. Area B*. Detroit, MI: Research and Evaluation Specialists, Inc., Detroit Public Schools, MI. Office of Research, Evaluation and Testing. (ERIC ED417246)

Syropoulos, M. (1997c). *Evaluation of the 1996–97 ninth grade restructuring program. Area C*. Detroit, MI: Research and Evaluation Specialists, Inc., Detroit Public Schools, MI. Office of Research, Evaluation and Testing. (ERIC ED417246)

Syropoulos, M. (1997d). *Evaluation of the 1996–97 ninth grade restructuring program. Area D*. Detroit, MI: Research and Evaluation Specialists, Inc., Detroit Public Schools, MI. Office of Research, Evaluation and Testing. (ERIC ED417246)

Syropoulos, M. (1997e). *Evaluation of the 1996–97 ninth grade restructuring program. Area E*. Detroit, MI: Research and Evaluation Specialists, Inc., Detroit Public Schools, MI. Office of Research, Evaluation and Testing. (ERIC ED417246)

Syropoulos, M. (1997f). *Evaluation of the 1996–97 ninth grade restructuring program. Area F*. Detroit, MI: Research and Evaluation Specialists, Inc., Detroit Public Schools, MI. Office of Research, Evaluation and Testing. (ERIC ED417246)

Syropoulos, M. (1998). *Academic and support ninth-grade restructuring programs as reported by the ninth-grade administrators*. Detroit, MI: Research and Evaluation Specialists, Inc., Detroit Public Schools, MI. Office of Research, Evaluation and Testing. (ERIC ED419866)

Texas Center for Educational Research. (2002, March). *At-risk students and the transition to high school: Texas' efforts to support ninth grade successes*. Austin, TX: Texas Education Agency. Retrieved September 30, 2003, from http://www.tcer.org/tcer/publications/ninthgrade_final_report.pdf.

Tierney, J. P., & Grossman, J. B. (2000). *Making a difference: An impact study of Big Brothers and Big Sisters*. Public/Private Ventures. Retrieved September 30, 2003, from http://www.ppv.org/pdffiles/mad.pdf.

Till, F. (2002). *Ninth grade intensive reading program, 2001–2002*. The School Board of Broward County, FL.

Titus, D. (1994). *It works! Vocational exploration as drop out prevention*. Viewpoints. (ERIC ED391074)

Tuss, P. (2003). *Evaluation of the summer intervention programs: Short term effects*. San Juan Islands, WA: San Juan Unified School District, Accountability and Organizational Evaluation Department.

Useem, E., Neild, R. C., & Morrions, W. (2001). *Philadelphia's talent development high schools: Second year results, 2000–01*. Philadelphia and Baltimore: Philadelphia Education Fund and Johns Hopkins University. (ERIC ED467132)

Valdez, V. (2003a). *An evaluation of the counseling services at four Latino plurality high schools in the Chicago public schools*. Chicago: Chicagoland Latino Educational Research Institute (CLERI). Retrieved September 30, 2003, from http://il.aspira.org/pdf/Research-1.pdf.

Valdez, V. (2003b). *How effective are schools tutoring services?* Chicago: Chicagoland Educational Research Institute (CLERI). Retrieved September 30, 2003, from http://www.aspirail.org/pdf/Issue_Brief.pdf.

Visher, M., Bhandari, R., & Blakely, K. (2002). *What good high schools do: Results from a survey of the principals and teachers from 54 nationally recognized high schools*. Berkeley, CA: MPR Associates, Inc.

White, W. D. (1993, February 6–11). *Educational benefits in year-round high schools*. Paper presented at the annual meeting of the National Association for Year-Round Education, Las Vegas, NV. (ERIC ED359660)

Woodruff, S., Schumaker, J. B., & Deshler, D. D. (2002). *The effects of an intensive reading intervention on the decoding skills of high school students with reading deficits*. Lawrence, KS: Institute for Academic Access.

Yanez, D. A., & Wenrick, M. (2000). *Improving algebra I end-of-course exam scores: Evidence from the field*. Austin, TX: The Charles A. Dana Center, University of Texas at Austin. Retrieved September 30, 2003, from http://www.tenet.edu/teks/math/resources/algeocdraft.pdf.

Appendix C: Ninth-Grade Remediation Keywording Guidelines

A.1 Source of report *Where was the report/article obtained?*	**A.1.1 Electronic database (please specify):** Article/report is identified from an electronic search (e.g., ERIC, Psych Info, Medline, DAI, SSCI, other). **A.1.2 Other:** Write in source of article.
A.2 Status *What is the publication status of the report/article?*	**A.2.1 Published:** The article/report is published in a journal or book that has an ISSN or ISBN number. **A.2.2 In press:** The article/report will be published in a journal or book that has an ISSN or ISBN number. **A.2.3 Other:** The article/report was posted or circulated electronically (e.g., website) or in print. **A.2.4 Unpublished:** The article/report is not published in a journal or book that has an ISSN or ISBN number.
A.3 Linked reports *Is the report/article one of at least two reports/articles from a singular study (that have already been obtained)?*	**A.3.1 Not linked:** The article/report is the only report obtained from a singular study. **A.3.2 Linked (please provide details):** The article/report is one of at least two reports obtained from a singular study (e.g., an article and an executive report both report the findings of a study conducted under the same grant number). **A.3.3. Don't know**
A.4 Type of treatment *What type of remediation treatment is researched in the report/article?*	**A.4.1 Ninth-grade remediation:** The intervention is focused on a ninth-grade remediation strategy that includes restructuring of curriculum or grouping of students (e.g., double-dosing, freshman academies, career academies, etc.). **A.4.2 Mentoring/Social skill development:** The intervention is focused on the implementation of non-academic support. **A.4.3 Tutoring/Extra help:** The intervention is focused on the implementation of academic support. **A.4.4: Middle grades transition programs:** The intervention is focused on preparing middle graders to enter high school. **A.4.5: Extended day/year:** The intervention is focused on increasing the time students spend in academic coursework (e.g., extended day, before school and after school, or extended year programs). **A.4.6: Summer school:** The intervention is focused on summer remediation strategies (not year-round school). **A.4.7. Other (please specify):**
A.5 Intensity of remediation *How many years was the remediation treatment that was researched in the report/article?*	**A.5.1 One-time intervention (less than six months)** **A.5.2 Six months to one year** **A.5.3 More than one year** **A.5.4 Unclear**

A.6 Type of design *What is the overall research design for the report/article?*	**A.6.1 Experimental study** (study includes randomly assigned subjects to a control group and an experimental group selected from students within the school or neighboring schools) **A.6.2 Quasi-experimental study** (study includes treatment and comparison point – can be longitudinal within groups or comparison between groups – no random assignment) **A.6.3 Descriptive study** (study reports quantitative outcomes but has no control or self-comparison group) **A.6.4 Qualitative study** (study is qualitative in design).
A.7 Type of sample *What school level are the participants being researched in the report/article?*	**A.7.1 Sample is drawn from high school setting** (participants are in grades 9–12) **A.7.2 Sample is drawn from ninth-grade setting** (participants are in grade 9) **A.7.3 Sample is drawn from K–12 school setting** (data is not disaggregated for high-school students) **A.7.4 Sample is drawn from elementary, junior high, or college setting** (participants are not in grades 9–12)
A.8 Type of outcomes *What are the outcomes of the remediation treatment being researched in the report/article?*	**A.8.1 Outcomes focused on measures of student achievement** (e.g., standardized tests, GPA, curriculum tests, high-stakes tests, classroom tests). **A.8.2 Outcomes focused on student engagement** (e.g., attendance, dropout rates). **A.8.3 Outcomes focused on student perceptions of changes in academic or non-academic performance** (e.g., questionnaire). **A.8.4 Outcomes focused on teacher or administrators' perceptions of changes in academic or non-academic performance** (e.g., survey).
A.9 Exclusionary criteria *Should the report/article be included in this study?* **Reason #1** – Treatment is not at least six months of school year. **Reason #2** – Study methodology is not experimental or quasi-experimental in nature. **Reason #3** – Sample does not include eighth- or ninth-grade students in experimental cohort. **Reason #4** – Study is not focused on student academic achievement OR student engagement outcomes.	**A.9.1 Yes** **A.9.2 No** **If no, enter exclusionary code(s)**: _____ _____ _____ _____

Appendix D: Ninth-Grade Remediation Data Extraction Guidelines

Intervention's Relevance to the Review

This section assesses information about the nature of the intervention(s) and ends with a judgment of DIAD Composite Question #1 about the intervention's relevance to the review.

1.1 Characteristics that <u>must</u> be part of the intervention. a. Ninth grade b. At-risk students c. Type of intervention (see 1.1a–e) i. **Ninth-grade remediation:** The intervention is focused on a ninth-grade remediation strategy that includes restructuring of curriculum or grouping of students (e.g., double-dosing, freshman academies, career academies, etc.). ii. **Mentoring/Social skill development:** The intervention is focused on the implementation of non-academic support. iii. **Tutoring/Extra help:** The intervention is focused on the implementation of academic support. iv. **Middle grades transition programs:** The intervention is focused on preparing middle graders to enter high school. v. **Extended day/year:** The intervention is focused on increasing the time students spend in academic coursework (e.g., extended day, before school and after school, or extended year). vi. **Summer school:** d. Student engagement and/or student achievement outcomes measured.	**Response Categories:** **Fully:** Study contains characteristics a–d. **Not at all:** Study does not contain all four characteristics a–d.
1.2 Details about intervention that must be available in report a. Details of intervention easily replicable and are described in the report. b. Details of intervention are easily accessible through mass media/literature.	**Yes:** The intervention was sufficiently described (by either 1.2a or 1.2b) at a level that would allow relatively easy and thorough replication by other implementers. **No:** The authors of the study omit important descriptive information concerning the essential elements of the intervention such that its replication would be impossible.

1.3 Was there evidence that the group receiving the intervention might also have experienced a changed expectancy, novelty, and/or disruption effect not also experienced by the comparison group?	**No:** The report says that these effects did not occur. **Yes**: The report says that at least one of these effects occurred. **NR (not reported):** The report says nothing about whether these effects occurred.
1.4 Was there evidence that the intervention was implemented in a manner similar to the way it was defined?	**No:** The report says the program was changed when it was implemented. **Yes:** The report says the intervention was implemented as described. **NR (not reported):** The report gave no information about implementation beyond the program description.
Composite Question 1: Was the intervention properly defined?	**Yes:** The intervention was adequately described and it fully reflected commonly-held or theoretically derived ideas about what the intervention should be. **Maybe yes:** At a minimum the intervention was adequately described, and it at least largely reflected commonly-held or theoretically derived ideas about what the intervention should be. **Maybe no:** There were important details missing from the description of the intervention and its implementation and/or the intervention was described only as a member of a broader class of treatments. **No:** The intervention did not reflect commonly-held or theoretically derived ideas about what it should be and/or there were known problems with its implementation.

Outcome Measure's Relevance to the Review

This section assesses information about the nature of the outcome(s) of the study and ends with a judgment of DIAD Composite Question #2 about the outcome measure's relevance to the review.

2.1a Did the study identify the actual name(s) of the instruments and tools used to measure the outcome(s) in this study? *Answer the following questions for each outcome.*	**Yes** **No** **NR (not reported):** Items were not reported or could not be easily retrieved anywhere.
2.1b Do items on the outcome measure appear to represent the content of interest (i.e., have face validity)?	**Yes** **No**
2.2 Was there evidence that the outcome measure was measured with acceptable reliability? a. What was the reliability of the outcome measure? b. What type of reliability does the estimate represent)? i. Internal consistency ii. Test-retest iii. Inter-rater	**Yes:** Study addresses both 2.2a and 2.2b. **No:** Study does not address either 2.2a or 2.2b.

2.3 To what extent was the outcome measure aligned with the intervention?	**Proper alignment:** The material presented on the outcome measure is drawn from the same universe of content represented in the intervention. **Over alignment:** The materials used in the intervention are identical to material on the outcome measure. **Under alignment:** The outcome measure contains content irrelevant to the intervention.
Composite Question 2: Was the outcome measure properly defined and aligned to the intervention?	**Yes:** The study provided adequate evidence that the outcome measure was properly defined and was aligned to the intervention. **Maybe no:** There was evidence that the measure had face validity and was properly aligned to the intervention. However, evidence suggested the measure might not be reliable. **No:** It is unclear what the outcome was and how it was measured.

Clarity of Causal Inference: Fair Comparison

This section assesses information about the procedures used to select participants for the study, as well as issues regarding the assignment of participants to groups, attrition, and pre- and post-equating of participants across groups. This section ends with a judgment of DIAD Composite Question #3 about the internal validity of the selection and equating procedures.

3.1a Was any information provided on the sampling frame for this study? *How did the researchers compile the list of sampling units (i.e., students, schools) from which they drew the actual sample for this study?*	**Yes** **No** **Unclear**
3.1b What was the unit of allocation into each intervention and control/ comparison group?	**Individuals** **Intact classes** **Schools** **Unclear/Not stated**
3.1c Was there attrition in the intervention group? the comparison group? a. Sample size of the intervention group at the beginning and end of the study b. Sample size of the comparison group at the beginning and end of the study	**Intervention:** **Yes No** **Comparison:** **Yes No**
3.1d Was there differential attrition between intervention and comparison groups?	**No:** The report specifically states that the groups did not experience attrition for different reasons. **Yes:** the report specifically states that the groups experienced attrition for different reasons. **NR (not reported):** the report said nothing about the groups experiencing attrition for different reasons.
3.2 Was there severe overall attrition? *To be determined on individual basis based on report background.*	**Yes** **No**

3.3 Were adequate equating procedures used to recreate the selection model? a. What characteristics of students or schools were used to equate groups? *At a minimum, there should be attention paid to the key explanatory variables in the model and the dependent variable. This list will vary by study.* a. When was equating of groups done? i. Before ii. After iii. Both before and after iv. Could not determine b. Was there evidence that the equating procedure led to differential statistical regression? i. No, differential statistical regression is highly unlikely, given the design. ii. Yes, differential statistical regression is likely, given the design. iii. NA (not applicable), no equating was done. iv. NR (not reported), the likelihood of differential statistical regression cannot be assessed from the report.	**Yes:** All characteristics under 3.3a must be accounted for, 3.3b must be ii or iii if differential attrition occurred, and 3.3c must be i. **No** *Basically, how did they create the sample?* Note: If there is differential attrition, equating must have occurred after the attrition took place. Did they use weighting? Correct for non-response or dropping out? Balance groups to look like each other? *When they tried to make the groups comparable, could there have been bias?*
3.4a Was there attrition in the intervention group? the comparison group? a. Sample size of the intervention group at the beginning and end of the study b. Sample size of the comparison group at the beginning and end of the study	**Intervention:** **Yes No** **Comparison:** **Yes No**
3.4b Was there differential attrition between intervention and comparison groups?	**No:** The report specifically states that the groups did not experience attrition for different reasons. **Yes:** the report specifically states that the groups experienced attrition for different reasons. **NR (not reported):** the report said nothing about the groups experiencing attrition for different reasons.
3.5 Was there severe overall attrition? *To be determined on individual basis based on report background.*	**Yes** **No**
3.6a Was there attrition in the intervention group? the comparison group? a. Sample size of the intervention group at the beginning and end of the study b. Sample size of the comparison group at the beginning and end of the study	**Intervention:** **Yes No** **Comparison:** **Yes No**
3.6b Was there differential attrition between intervention and comparison groups?	**No:** The report specifically states that the groups did not experience attrition for different reasons. **Yes:** the report specifically states that the groups experienced attrition for different reasons. **NR (not reported):** the report said nothing about the groups experiencing attrition for different reasons.
3.7 Was there severe overall attrition? *To be determined on individual basis based on report background.*	**Yes** **No**

3.8 Could all observations have received the intervention had the cutoff point been set differently? (Regression discontinuity design)	**Yes** **No** **Unclear** *No random assignment to treatment and control groups, and no evidence that treatment is withheld from anyone deemed eligible. Must have administered a pre-test and specified the cut point on the pre-test for assignment to treatment or control group.*
Composite Question 3: Were the participants in the group receiving the intervention comparable to the participants in the comparison group?	**Yes:** Participants were randomly assigned to conditions, and there was not differential attrition within groups or severe overall attrition across groups or within subjects when subjects serve as their own controls. **Maybe yes:** Randomized assignment was used but there was serious differential attrition within groups or serious overall attrition across groups or within subjects, OR although random assignment was not used, reasonable attempts were made to make the groups comparable (i.e., matched sampling, use of a covariate) AND there was not serious attrition within or across groups or within subjects. **Maybe no:** Randomized assignment was not used and although steps were taken to make the groups comparable, the steps may not have been adequate. **No**: It is unlikely that the participants in the groups were comparable.

Clarity of Causal Inference: Lack of Contamination

This section assesses the steps used to minimize the probability that events alternative to the intervention in the research context could be responsible for the effects measured in the study. This section ends with a judgment of DIAD Composite Question #4 about the clarity of causal inference related to lack of contamination.

4.1 Was there evidence of a local history event? *A local history event is an event occurring between the beginning of an intervention and the posttest within the context of the intervention, outcome, time, setting, and persons studied that could have produced the observed results in the absence of the intervention.*	**Yes:** The report said a local history event occurred. **No:** The report said that no local history event occurred or the report said nothing to suggest the occurrence of a local history event.
4.2a Were the intervention and comparison groups drawn from the same local pool? *Groups are assumed to be from the same local pool unless specified otherwise in the research report.*	**Yes** **No**
4.2b Were intervention conditions known to study participants, data collectors, and/or other authorities (e.g., parents, teachers)? *This information is assumed to be available unless specified otherwise in the research report.*	**Yes** **No** **NA (not applicable)**

4.3 Did the description of the study give any other indication of the plausibility of other intervention contaminants? *Exclude from this answer the four types of events already coded (i.e., expectancy, novelty, disruption, and local history events).*	**Yes** **No** **NR (not reported)**
Composite question 4. Was the study free of events that happened at the same time as the intervention that may have confused its effect?	**Yes:** Concurrent processes and events that might be alternative explanations to the treatment effect have been ruled out, either explicitly or implicitly. **Maybe yes:** There were no identified processes or events that could be alternative explanations for the treatment effect, but some alternative explanations cannot be explicitly ruled out either because there was some evidence that alternative explanations might exist, or because no attention was given to ruling out an alternative explanation and it is reasonable to expect that one or more alternative explanations might exist. **No:** Identifiable processes or events that are described to be occurring simultaneously with the intervention may have caused the observed effect.

Generality of Findings: Inclusive Sampling

This section assesses how representative the actual participants, settings, outcomes, and data collection activities were to the theoretical population, school settings, typical measures, and appropriate measurement processes. This section ends with a judgment of DIAD Composite Question #5 about the generality of findings related to inclusive sampling.

5.1 Did the sample contain participants with the necessary characteristics to be considered part of the target population? Were the participants: a. High school setting (must include ninth grade) ninth-grade setting b. Students transitioning from eighth grade c. Students transitioning out of ninth grade	**Yes:** Must include one of a–d. **No:** Does not include any characteristics a–d.
5.2 To what extent did the sample capture variation among participants on important characteristics of the target population? What labels characterize the sample? a. At-risk b. Representative minority groups c. SES d. Gender groups	**Fully:** Includes groups a–d. **Largely:** Includes a, b, and c. **Somewhat:** Includes a and b. **Not at all:** Includes one or no label characteristics.

5.3 To what extent did the study include variation on important characteristics of the target setting? What labels characterize the target setting? a. School size (student population) b. Classroom size c. Regional characteristics of school (urban, urban fringe, rural) d. State or region in which school(s)/ school district(s) resides e. Other: Define_____	**Fully:** Includes at least two relevant characteristics of the target setting. **Largely:** Includes at least one relevant characteristic of the target setting. **Somewhat:** Includes "other" characteristic of target setting. **Not at all:** Does not discuss any relevant target setting characteristics.
5.4 To what extent were important classes of outcome measures included in this study? What are the important classes of outcomes? a. Student achievement i. Standardized tests ii. GPA iii. Curriculum tests iv. High-stakes tests v. Classroom tests b. Student engagement i. Attendance ii. Drop-out rate c. Student perceptions d. Teacher or administrator perceptions	**Fully:** Includes 1–4. **Largely:** Must include 1 or 2, and may include 3 and 4. **Not at all:** Does not include 1 or 2.
5.4.a Which method(s) was/were used to collect the data for this study)?	**Teacher-made test(s)** **Researcher-made test(s)** **Curriculum test(s)** **Standardized test(s)** **High-stakes test(s)** (state level) **Document review:** This includes policy documents and school records. **Other** (please specify_____) **Unclear** **Not stated**
5.5 Did the study measure the outcome at a time appropriate for capturing the intervention's effect? *When was the outcome measured relative to the end of the intervention? In days (or weeks, months) when was the outcome measured relative to the end of the intervention?* *Should be a reasonable length of time to measure the post-treatment effects. Ideally, looking at persistent effects of treatment. Length of time dependent on the stated purpose of the program and outcome measure. For example, summer school should allow students to pass tests/courses to promote to next level and may only need to be tested once students enter tenth grade or if students enter tenth grade.*	**Yes** ***Define:*** **No**
5.6. Was the study conducted during the time frame appropriate for the Evidence Report? *Must be published between 1993 and 2003.*	**Yes** **No**
Composite Question 5: Were targeted participants, settings, outcomes, and occasions included in the study?	**Yes** **Maybe Yes** **Maybe No** **No**

Generality of Findings: Effects Tested within Subgroups

This section assesses how broadly the intervention was tested across important sub-groups of students and across substantive variations within the intervention as a whole. This section ends with a judgment of DIAD Composite Question #6 about the generality of findings related to the effects tested within subgroups.

6.1a What is the socio-economic composition of the participants in this study?	**Low income** **Middle income** **High income** **Unclear/Not stated**
6.1b What is the predominant gender of the participants of this study? *What are the gender percentages of the sample for this study?* *Male =* *Female=*	**Female** **Male** **Unclear/Not stated**
6.1c What is the stated percentage(s) for each of the ethnic classifications of the participants in this study? *If available, use percentage calculations by the authors of the study. If feasible and not reported by the authors, calculate these percentages yourself, but specify that you calculated these percentages and not the study authors.*	**White/European** **Black or African American** **Native Hawaiian or other Pacific Islander** **Asian** **Hispanic/Latino** **Native American** **Information provided but not adequately detailed to provide percentages (please specify)** **Unclear/Not stated**
6.1d To what extent was the intervention tested for effectiveness within important subgroups of participants?	**Fully:** Effect sizes could be derived for all subgroups included in the study. **Largely:** Effect sizes could be derived for some subgroups included in the study. **Not at all:** Effect could not be derived for subgroups studied, or did not test for subgroup effects.
6.2 To what extent was the intervention tested for effectiveness within important subgroups of settings?	**Fully:** Effect sizes could be derived for all settings included. **Largely:** Effect sizes could be derived for some settings. **Not at all:** Effect could not be derived for subgroups studied.
6.3 Was the intervention tested for its effectiveness across important classes of outcomes?	**Yes:** Effect sizes could be derived for all classes of outcomes studied. **No:** Effect could not be derived for any classes of outcomes studied.
6.4 Was the time of measurement (relative to the end of the intervention) tested as an influence on the intervention's effect? *When was the outcome measured relative to the end of the intervention?*	**Yes** **No** *Want to see if they tested the strength or existence of the treatment at say, program end, after 3 months, after 6 months, and so on.*
6.5 Was the intervention tested for its effectiveness across important variations in intervention implementation?	**Yes** **No** *Did they administer the treatment differently, and if so, did they look at how the estimated effects varied by treatment modality?*

Composite Question 6: Was the intervention tested for its effectiveness within important subgroups of target participants, settings, outcomes, occasions, and intervention variations?	**Yes:** The intervention was tested for its effectiveness on targeted variations of participants, settings, outcomes, occasions, and intervention variations. **Maybe yes:** The intervention was tested for its effectiveness within most important subgroups of participants and settings. **Maybe no:** Although the intervention was tested for its effectiveness within some important subgroups of the participants and settings, many were left out. **No:** At best the intervention was only tested for its effectiveness within limited important subgroups of the participants, settings, outcomes, occasions, and intervention variations

Precision of Outcome: Effect Size Estimation

This section assesses the thoroughness with which the statistical properties of the data were reported, including important assumptions underlying the analytic techniques and the reliability of the outcome measures. This section ends with a judgment of DIAD Composite Question #7 about the precision of the outcome related to effect size estimation.

7.1a Was the assumption of independence met, or could dependence (including dependence arising from clustering) be accounted for in estimates of effect sizes and their standard errors? For this effect size: a. Did each participant appear in only one of the comparison groups? i. Yes ii. No iii. NR (not reported) b. Was the unit of assignment the same as the unit of statistical analysis? i. Yes ii. No iii. NR (not reported) c. If you answered no to either of the questions above, could dependence (including dependence arising from clustering) be accounted for in estimates of effect size and their standard errors? i. Yes ii. No iii. NA (not applicable)	**Yes:** Includes answering "Yes" to a and b, or answering "Yes" to c. **No:** Answering "No" or "NR" to a and b, or answering "No" or "NR" to c.
7.1b Which statistical techniques were used to analyze the data?	**Difference statistics** (please specify) **Associational statistics** (please specify) **Other** (please specify) **Unclear/Not stated**

7.2 Did the statistical properties of the data (e.g., distributional and variance assumptions, if any, presence of outliers) allow for valid estimates of the effect sizes? *a. Were scores roughly normally distributed within groups?* i. Yes ii. No *b. Were variances roughly equivalent across groups?* i. Yes ii. No	**Yes** **No**
7.3 Were the sample sizes adequate to provide sufficiently precise estimates of effect sizes? *What were the sample sizes for the intervention and comparison groups for the analysis of the outcome measure? Assuming a "simple" study (e.g., one which does not have repeated measures), the variance of the effect size can be estimated by the sample size, and assuming reasonable sample sizes, a confidence interval computed assuming a standard normal distribution. "Sufficiently precise" could then be evaluated as within ± 1.96* effect size. The precision you "need" is arbitrary but assessable.* *For a repeated measures design, must know the correlation among measures as well as the within group sample size.* *We will not create the effect size and standard error for effect size or 95 percent confidence interval, but will check to see if this is mentioned and then referred back to in the study.*	**Yes** **No**
7.4. Were the outcome measures sufficiently reliable to allow adequately precise estimates of the effect sizes? a. Was there evidence that the internal consistency was sufficient? i. Yes ii. No iii. NR (not reported) *A statement indicating that internal consistency was "acceptable" is also sufficient, even if the specific value was not reported. A citation to an external source is also sufficient.*	**Yes** **No** *Should mention reliability of measures in discussion of outcomes.*
Composite question 7: Were the effect sizes accurately estimated?	**Yes:** The effect sizes appear to be accurately estimated. **Maybe yes:** There was some evidence of statistical issues that may have caused the effect sizes to be inaccurately estimated, but the likely impact on inferences was minimal. **Maybe no:** There was evidence that statistical issues may have caused the effect sizes to be inaccurately estimated. **No:** The assumption of statistical independence was not met, and dependence was not accounted for in the effect sizes.

Precision of Outcome: Statistical Reporting

This section assesses the adequacy with which the study reports effect size estimates or the data needed to calculate effect sizes. This section ends with a judgment of DIAD Composite Question #8 about precision of the outcome related to statistical reporting.

8.1 Were sample sizes reported (or estimable) from the statistical information presented?	**Yes** (please specify) **No** (please specify)
8.2 Could directions of effects be identified for important measured outcomes? a. What was the direction of the effect?	**Yes** (please specify) **No** (please specify)
8.3a Could effect sizes be estimated for important measured outcomes? **If "Yes," answer 8.3b.**	**Yes** **No**
8.3b Could estimates of effect sizes be computed using a standard formula or its algebraic equivalent? a. If an effect size could be derived, how could it be done (e.g., standard formula, transformation of t-test, correlation)? i. Standard formula = subtract control group mean from treatment group mean and divide by standard deviation for control OR pooled standard deviation for control and treatment (square root of sum of standard deviations) ii. Algebraic equivalent of standard formula iii. Non-standard formula (very complicated design) iv. No effect size could be derived *Choose only one, they are ordered by preference	**Yes:** Study provided information on formula used. **No**
Composite question 8: Were the statistical tests adequately reported?	**Yes:** All important outcomes either have effect sizes reported by the authors or provide data to allow precise calculation of effect sizes. **Maybe yes:** Sufficient statistical information was reported to allow, at a minimum, imprecise effect sizes to be calculated for most measured outcomes. **Maybe no:** For most outcomes, effect sizes were not reported, nor is there adequate statistical information to allow effect sizes to be calculated. **No:** Sample sizes were not reported, OR neither the magnitude nor the direction of the effects could be discerned for most outcome measures.

Appendix E: Study DIAD Spreadsheet Coding

			1	2	3	4	5	6	7	8	9	10
		Type of Intervention	**Summer Remediation**	**Academic Support Services**	**Academic Support Services**	**Summer Remediation**	**School restructuring**	**Academic Support Services**	**School Restructuring**	**School Restructuring**	**School Restructuring**	**School Restructuring**
	CQ#	CQ	Roderick (2000)	Gamoran (1997)	Myers (1999)	Grossman and Sipe (1992)	Quint et al. (1999)	Woodruff, S. et al. (2002)	Dynarski et al. (1998) dropout	Dynarski et al. (1998) restructuring	Kemple (2000)	Balfanz et al. (2003)
	1.1	To what extent does the intervention reflect commonly-held or theoretically derived characteristics about what it should contain?	Fully	Fully	Fully	Fully	Fully	Fully	Fully	Fully	Fully	Fully
	1.2	Was the intervention described at a level of detail which would allow its replication by other implementers?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	1.3	Was there evidence that the group receiving the intervention might also have experienced a changed expectancy, novelty, and/or disruption effects not also experienced by the control group?	No	No	No	No	No	No	No	No	No	No
	1.4	Was there evidence that the intervention was implemented in a manner similar to the way it was defined?	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
(derived)	**Composite 1**	**Was the intervention properly defined?**	Yes	Yes	Yes	Maybe No	Yes	Yes	Yes	Maybe No	Yes	Yes
	2.1	Do items on the outcome measure appear to represent the content of interest (i.e., have face validity)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	2.2	Was there evidence of the construct validity of the outcome measure available either in the report or through other easily accessible documents?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
	2.3	To what extent was the outcome measure aligned with the intervention?	Proper Alignment	Proper Alignment	Proper Alignment	Proper Alignment	Proper Alignment	Proper Alignment	Proper Alignment	Proper Alignment	Proper Alignment	Proper Alignment
(derived)	**Composite 2**	**Was the outcome measure properly defined and aligned to the intervention?**	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Maybe No
(derived)	**Global 1**	**Construct Validity: Were the intervention and outcome measures properly defined?**	**Yes**	**Yes**	**Yes**	**Maybe No**	**Yes**	**Yes**	**Yes**	**Maybe No**	**Yes**	**Maybe No**
For randomized experiments	3.1	Was there differential attrition between intervention and comparison groups?			No	No			No		No	
	3.2	Was there severe attrition overall?			No	No			No		No	
For quasi-experimental designs	3.3	Were adequte equating procedures used to re-create the selection model?	Yes	Yes			Yes	Yes		Yes		Yes
	3.4	Was there differential attrition between intervention and comparison groups?	No	No			No	No		No		No
	3.5	Was there severe attrition overall?	No	No			No	No		No		No
For regression discontinuity designs	3.6	Was there differential attrition between intervention and comparison groups?										
	3.7	Was there severe attrition overall?										
	3.8	Could all observations have received the intervention had the cutoff point been set differently?										
(derived)	**Composite 3**	Were participants in intervention group comparable to participants in comparison group *(randomized experiments)* ?			Yes	Yes			Yes		Yes	
(derived)	**Composite 3**	Were participants in intervention group comparable to participants in comparison group *(quasi-experimental designs)* ?	Maybe Yes	Maybe Yes			Maybe Yes	Maybe Yes		Maybe Yes		Maybe Yes
(derived)	**Composite 3**	Were participants in intervention group comparable to participants in comparison group *(regression discontinuity designs)* ?										
	4.1	Was there evidence of a local history event?	No	No	No	No	No	No	No	No	No	No
	4.2a	Were the intervention and comparison groups drawn from the same local pool?	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes

			1	2	3	4	5	6	7	8	9	10
		Type of Intervention	**Summer Remediation**	**Academic Support Services**	**Academic Support Services**	**Summer Remediation**	**School restructuring**	**Academic Support Services**	**School Restructuring**	**School Restructuring**	**School Restructuring**	**School Restructuring**
	CQ#	CQ	Roderick (2000)	Gamoran (1997)	Myers (1999)	Grossman and Sipe (1992)	Quint et al. (1999)	Woodruff, S. et al. (2002)	Dynarski et al. (1998) dropout	Dynarski et al. (1998) restructuring	Kemple (2000)	Balfanz et al. (2003)
	4.2b	If yes, were intervention conditions known to study participants, providers, data collectors, and/or other authorities (e.g., parents, teachers, case managers)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	4.3	Did the description of the study give any other indication of the plausibility of other intervention contaminants?	No	No	No	No	No	No	No	No	No	No
	Composite 4	Was the study free of events that happened concurrently with the intervention that confused its effect?	Maybe Yes	Yes	Maybe Yes	Yes	Maybe Yes	Maybe Yes	Maybe Yes	Maybe Yes	Yes	Maybe Yes
(derived)	**Global 2**	**Internal Validity: Was the intervention the cause of the change in the outcome *(randomized experiments)?***			Maybe Yes	Yes			Maybe Yes		Yes	
(derived)	**Global 2**	**Internal Validity: Was the intervention the cause of the change in the outcome *(quasi-experimental designs)?***	Maybe Yes	Maybe Yes			Maybe Yes	Maybe Yes		Maybe Yes		Maybe Yes
(derived)	**Global 2**	**Internal Validity: Was the intervention the cause of the change in the outcome *(regression discontinuity designs)?***										
	5.1	Did the sample contain participants with the necessary characteristics to be considered part of the target population?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	5.2	To what extent did the sample capture variation among participants on important characteristics of the target population?	Largely	Largely	Largely	Fully	Largely	Largely	Fully	Largely	Largely	Largely
	5.3	To what extent did the study include variation on important characteristics of the target setting?	Fully	Largely	Fully	Somewhat	Fully	Largely	Somewhat	Largely	Largely	Largely
	5.4	To what extent were important classes of outcome measures included in the study?	Fully	Largely	Largely	Fully	Fully	Largely	Largely	Largely	Largely	Largely
	5.5	Did the study measure the outcome at a time appropriate for capturing the intervention's effect?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	5.6	Was the study conducted during the appropriate time frame?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(derived)	**Composite 5**	**Were targeted participants, settings, outcomes, and occasions included in the study?**	Maybe Yes	Maybe Yes	Maybe Yes	Maybe No	Maybe Yes	Maybe Yes	Maybe No	Maybe Yes	Maybe Yes	Maybe Yes
	6.1	To what extent was the intervention tested for effectiveness within important subgroups of participants?	Fully	Fully	Largely	Fully	Fully	Largely	Not at All	Largely	Fully	Largely
	6.2	To what extent was the intervention tested for effectiveness within important subgroups of settings?	Fully	Fully	Largely	Fully	Fully	Fully	Largely	Largely	Fully	Somewhat
	6.3	Was the intervention tested for its effectiveness across important classes of outcomes?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	6.4	Was the time of measurement (relative to the end of the intervention) tested as an influence on the intervention's effect?	Yes	Yes	Yes	Yes	No	No	Yes	No	Yes	Yes
	6.5	Was the intervention tested for its effectiveness across important variations in intervention implementation?	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	No

			1	2	3	4	5	6	7	8	9	10
		Type of Intervention	**Summer Remediation**	**Academic Support Services**	**Academic Support Services**	**Summer Remediation**	**School restructuring**	**Academic Support Services**	**School Restructuring**	**School Restructuring**	**School Restructuring**	**School Restructuring**
	CQ#	CQ	Roderick (2000)	Gamoran (1997)	Myers (1999)	Grossman and Sipe (1992)	Quint et al. (1999)	Woodruff, S. et al. (2002)	Dynarski et al. (1998) dropout	Dynarski et al. (1998) restructuring	Kemple (2000)	Balfanz et al. (2003)
(derived)	**Composite 6**	**Was the intervention tested for its effectiveness within important subgroups of participants, settings, outcomes, occasions, and intervention variations**	Yes	Yes	Maybe Yes	Yes	Maybe No	Maybe Yes	Maybe No	Maybe Yes	Yes	Maybe No
(derived)	**Global 3**	**External Validity: Was the intervention tested on relevant participants and environments?**	**Maybe Yes**	**Maybe Yes**	**Maybe Yes**	**Maybe No**	**Maybe No**	**Maybe Yes**	**Maybe No**	**Maybe Yes**	**Maybe Yes**	**Maybe No**
	7.1	Was the assumption of independence met, or could dependence (including dependence arising from clustering) be accounted for in estimates of effect sizes and their standard errors?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	7.2	Did the statistical properties of the data (e.g., distributional and variance assumptions, if any, presence of outliers) allow for valid estimates of the effect sizes?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	7.3	Were the sample sizes adequate to provide sufficiently precise estimates of effect sizes?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	7.4	Were the outcome measures sufficiently reliable to allow adequately precise estimates of the effect sizes?	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(derived)	**Composite 7**	**Were the effect sizes accurately estimated?**	Yes	Maybe Yes	Maybe Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	8.1	[illegible] estimable) from statistical information presented?	Fully	Fully	Largely	Largely	Fully	Fully	Fully	Largely	Fully	Fully
	8.2	Could directions of effects be identified for important measured outcomes?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	8.3a	Could effect sizes be estimated for important measured outcomes?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	8.3b	If yes, could estimates of effect sizes be computed using a standard formula or its algebraic equivalent?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(derived)	**Composite 8**	**Were the statistical tests adequately reported?**	Yes	Yes	Maybe Yes	Maybe Yes	Yes	Yes	Yes	Maybe Yes	Yes	Yes
(derived)	**Global 4**	**Statistical Conclusion Validity: Could accurate effect sizes be derived from the study report?**	**Yes**	**Maybe Yes**	**Maybe No**	**Maybe Yes**	**Yes**	**Yes**	**Yes**	**Maybe Yes**	**Yes**	**Yes**
SUMMARY	**Global 1**		**Yes**	**Yes**	**Yes**	**Maybe No**	**Yes**	**Yes**	**Yes**	**Maybe No**	**Yes**	**Maybe No**
	Global 2				**Maybe Yes**	**Yes**			**Maybe Yes**		**Yes**	
	Global 2		**Maybe Yes**	**Maybe Yes**			**Maybe Yes**	**Maybe Yes**		**Maybe Yes**		**Maybe Yes**
	Global 2											
	Global 3		**Maybe Yes**	**Maybe Yes**	**Maybe Yes**	**Maybe No**	**Maybe No**	**Maybe Yes**	**Maybe No**	**Maybe Yes**	**Maybe Yes**	**Maybe No**
	Global 4		**Yes**	**Maybe Yes**	**Maybe No**	**Maybe Yes**	**Yes**	**Yes**	**Yes**	**Maybe Yes**	**Yes**	**Yes**

Appendix F: Studies That Did Not Make It Through the DIAD

Author & Publication Date	Type of Design	Treatment	Measures	Composite Questions that Did Not Meet the Minimum for Inclusion
Hahn, A. (1992, June). *Evaluation of the Quantum Opportunities Program: Interim Impacts Covering the Ninth-Tenth Grades (1989–1991)*. Waltham, MA. Brandeis University, Center for Human Resources.	Random assignment treatment and control group	Academic support services Non-academic support services	Academic achievement Student engagement	Composite Question #7 Composite Question #8
Schalago-Schirm, C. (1995, April). *Does the Computer-Assisted Remedial Mathematics Program at Kearny High School Lead to Improved Scores on the N.J. Early Warning Test?* Paper in partial fulfillment of the requirements for the Masters of Arts Degree, Kean College of New Jersey.	Pretest/posttest comparison	Academic support services	Academic achievement	Composite Question #3 Composite Question #7
Ciaccio, L.A. and Sanders, J.W. (1996). *Project Discovery II. FIPSE Project*. Fund for the Improvement of Postsecondary Education, Washington, DC.	Pretest/posttest with matched comparison groups	Academic support services School restructuring	Academic achievement Student engagement Student perceptions	Composite Question #3 Composite Question #4 Composite Question #7 Composite Question #8
Munoz, M.A. (2002). *Facing the Challenges of At-Risk Students in Urban School Districts: The Impact of an Attendance and Dropout Prevention Program in Non-Traditional School.* Jefferson Public Schools.	Random assignment	Non-academic support services	Academic achievement Student engagement	Composite Question #1 Composite Question #2
Clark, A. (1993). *Project Reach Final Evaluation Report, 1992–1993*.	Gap reduction design, pretest/posttest with comparison students the groups on which tests were normed	Academic support services	Academic achievement	Composite Question #3 Composite Question #7

Author & Publication Date	Type of Design	Treatment	Measures	Composite Questions that Did Not Meet the Minimum for Inclusion
Elliott, M.N., Hanser, L.M., and Gilroy, C.L. (2000). *Evidence of Positive Student Outcomes in JROTC Career Academies.* Santa Monica, CA: RAND, National Defense Research Institute.	Cohort comparison design	School restructuring	Academic achievement Student engagement	Composite Question #3
Metis Associates, Inc. (2002, March). *Evaluation Report. New York City Board of Education Summer School 2001*. Author: New York, NY.	Single group pretest/posttest design	Summer remediation	Academic achievement	Composite Question #3 Composite Question #7 Composite Question #8
Bauer, R., and Michael, R. (1993). *They're Still in School: Results of an Intervention Program for At-Risk High School Students*. Paper presented at the Annual Meeting of the American Educational Research Association; Atlanta, GA, April 12–16, 1993.	Single group pretest/posttest design	Academic support services	Academic achievement Student engagement	Composite Question #7 Composite Question #8
Till, F. (2002). *Ninth Grade Intensive Reading Program, 2001–2002*. The School Board of Broward County, FL.	Unclear	Academic support services	Academic achievement	Composite Question #3
Powers, S., and McConner, S. (1997). *Project SOAR 1996–1997. Evaluation Report*. Creative Research Associates, Inc.: Tucson, AZ.	Pretest/posttest design	Academic support services Non-academic support services	Academic achievement Student engagement Student perceptions	Composite Question #2 Composite Question #3 Composite Question #4 Composite Question #7 Composite Question #8
Tuss, P. (2003). *Evaluation of the Summer Intervention Programs: Short Term Effects*. San Juan Unified School District, Accountability and Organizational Evaluation Department. San Juan Islands, WA.	Pretest/posttest design	Summer remediation	Academic achievement	Composite Question #1 Composite Question #2 Composite Question #7 Composite Question #8