

Appendix

Appendix A1 Study characteristics: Rorie, 2007

Characteristic	Description
Study citation	Rorie, L. B. (2007). An investigation of achievement in the <i>AVID (Advancement Via Individual Determination)</i> program at the high school level. <i>Dissertation Abstracts International</i> , 68(11A), 168–4657.
Participants	The researcher used retrospective archival data to construct the <i>AVID</i> and comparison groups from the graduating classes of 2005 and 2006 that had complete data from 8th through 12th grades and attended the four high schools in the study. Students in the <i>AVID</i> group attended one of the four participating high schools from one school district and had participated in the <i>AVID</i> program for four years (grades 9–12) in high school (but not in 8th grade). The study author did not describe how students chose or were chosen to participate in the program. Non- <i>AVID</i> comparison group students attended the same schools and were matched on ethnicity, gender, and age. This WWC review focuses on the matched comparison sample that equated students on their 8th-grade Colorado State Assessment Program Reading subtest scores. The final sample includes 48 students in the <i>AVID</i> group and 48 students in the comparison group, all of whom graduated from the class of either 2005 or 2006.
Setting	Participating students attended four high schools in the Pine View School District in suburban Colorado. These schools had been implementing the <i>AVID</i> program for seven or more years. The school district had experienced a 40% increase in enrollment in the past decade, 31% of the district students were minority, 16% spoke a primary language other than English, and 16% of students were eligible to receive free or reduced-priced lunch.
Intervention	<i>AVID</i> students participated in the <i>AVID</i> elective class, and a majority of their content classes were taught by <i>AVID</i> -trained teachers. The study reported 9th- and 10th-grade student reading test score outcomes, thus measuring program effects after one to two years of participation in the intervention.
Comparison	Non- <i>AVID</i> students attended the same schools as the <i>AVID</i> students for all four years of high school, graduated from that high school during the same time period, and did not participate in any <i>AVID</i> electives. However, these students may or may not have been enrolled in classes taught by <i>AVID</i> -trained teachers. Since the <i>AVID</i> program seeks to promote whole school improvement through professional development of school and district personnel, the comparison group in this study may have been exposed to elements of the <i>AVID</i> program even if they did not participate in <i>AVID</i> electives (for example, if they were enrolled in a class taught by an <i>AVID</i> -trained teacher).
Primary outcomes and measurement	Reading comprehension was measured using the Colorado Student Assessment Program (CSAP) Reading subtest. Pre-intervention scores were from 8th grade, and outcomes were from 9th and 10th grades. ¹ For a more detailed description of these outcome measures, see Appendix A2.
Staff/teacher training	Teachers were trained to implement <i>AVID</i> , but no details are available concerning this training.

1. This study also measured the effects on the PLAN (a 10th-grade measure developed by the American College Testing [ACT] organization as a pre-ACT measure) and the COACT (a Colorado state version of the American College Testing program administered to 11th-grade students). These outcomes were not included in the WWC analysis because we do not have pre-intervention measures of these outcomes to determine whether the intervention and comparison samples were initially equivalent. The study also measured math and writing outcomes and overall GPA scores, none of which are eligible for review in the WWC Adolescent Literacy topic area.

Appendix A2 Outcome measure for the comprehension domain

Outcome measure	Description
The Colorado Student Assessment Program (CSAP) Reading subtest	This criterion-referenced assessment measures adequate yearly progress toward Colorado state standards. The Reading subtest in grades 4–10 consists of 56 multiple choice and 14 constructed response questions (which require the student to answer in complete sentences). Assessments for grades 9 and 10 are designed to be given in three 60-minute sessions. Each session includes four types of items in which students (1) read and demonstrate their understanding of a variety of materials; (2) apply thinking skills to their reading, writing, speaking, listening, and viewing (such as analyzing a text's main idea and differentiating fact from opinion); (3) read to locate, select, and make use of relevant information from a variety of media, references, and technological sources; and (4) read and recognize literature as a “record of human experience” (such as identifying the theme of text, developing a thesis statement for text, and applying literary techniques to understand text). Content areas addressed in the CSAP Reading subtest include fiction, nonfiction, vocabulary, and poetry (as cited in Rorie, 2007; Colorado Department of Education, 2009 ¹).

1. Colorado Department of Education. (2009, February). *Fact sheet for reading/Lectura CSAP—Grades 3–10*. Retrieved June 23, 2010, from http://www.cde.state.co.us/cdeassess/documents/csap/csap_frameworks.html.

Appendix A3 Summary of study findings included in the rating for the comprehension domain¹

Outcome measure	Study sample	Sample size (students)	Author's findings from the study					
			Mean outcome (standard deviation) ²		WWC calculations			
			AVID group	Comparison group	Mean difference ³ (AVID – comparison)	Effect size ⁴	Statistical significance ⁵ (at $\alpha = 0.05$)	Improvement index ⁶
Rorie, 2007⁷								
CSAP–Reading subtest	High school sample	96	nr (nr)	nr (nr)	nr	nr	ns	nr
Domain average for comprehension (Rorie, 2007)						nr	ns	nr

ns = not statistically significant

nr = not reported

CSAP = Colorado Student Assessment Program

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the comprehension domain.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are: a smaller standard deviation on a given measure would indicate that participants had more similar outcomes. For Rorie (2007), standard deviation information was not available.
3. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Mean scores on the Colorado Student Assessment Program (CSAP) and mean differences were not reported in Rorie (2007). Instead, the study reported results from a doubly repeated measures analysis of variance for CSAP Reading subtest, which included two dependent variables (9th-grade and 10th-grade CSAP scores) and three independent variables (participation in AVID intervention, grade level, and the AVID*grade level interaction). The author reported no significant effect for AVID (partial eta-squared effect size = <.001) or the AVID*grade level interaction (partial eta-squared effect size = .01). Based on the partial eta-squared effect size and non-significant p-values reported in the study, the WWC deems these results to be neither statistically significant nor substantively important. For a discussion of the comparability of partial eta-squared effect sizes and standardized mean differences, see Barnette, J. J. (2006). *Effect size and measures of association*. 2006 Summer Evaluation Institute sponsored by the American Evaluation Association and the Centers for Disease Control and Prevention, June 14, 2006.
4. For an explanation of the effect-size calculation, see WWC Procedures and Standards Handbook, Appendix B. For Rorie (2007), the WWC was unable to calculate an effect size due to lack of sufficient information reported.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results for the intervention group. The improvement index is not available, as Rorie (2007) did not provide sufficient information to calculate an effect size and improvement index using standard WWC methods.
7. The level of statistical significance was reported by the study authors or, when necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For the formulas the WWC used to calculate the statistical significance, see WWC Procedures and Standards Handbook, Appendix C for clustering and WWC Procedures and Standards Handbook, Appendix D for multiple comparisons. In the case of Rorie (2007), no corrections for clustering or multiple comparisons were needed.

Appendix A4 AVID (Advancement Via Individual Determination) rating for the comprehension domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.¹ For the comprehension outcome domain, the WWC rated AVID (Advancement Via Individual Determination) as having no discernible effects for adolescent learners.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: No studies showing a statistically significant or substantively important effect, either *positive* or *negative*.

Met. No studies showed a statistically significant or substantively important effect, either positive or negative. The one study showed indeterminate effects.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a *strong* design.

Not met. No studies showed a statistically significant positive effect.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

Met. The one study did not show a statistically significant or substantively important negative effect.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

Not met. No study showed a statistically significant or substantively important positive effect.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

Not met. No studies showed a statistically significant or substantively important negative effect. The one study showed indeterminate effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

Not met. No studies showed a statistically significant or substantively important positive or negative effect.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

Not met. No studies showed a statistically significant or substantively important positive or negative effect.

(continued)

Appendix A4 AVID (Advancement Via Individual Determination) rating for the comprehension domain (continued)

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: One study showing a statistically significant or substantively important *negative* effect and no studies showing a statistically significant or substantively important *positive* effect.

Not met. No studies showed a statistically significant or substantively important effect, either negative or positive.

OR

- Criterion 2: Two or more studies showing statistically significant or substantively important *negative* effects, at least one study showing a statistically significant or substantively important *positive* effect, and more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

Not met. No studies showed a statistically significant or substantively important effect, either negative or positive.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a *strong* design.

Not met. No studies showed a statistically significant negative effect.

AND

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

Met. The one study did not show a statistically significant or substantively important positive effect.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. For a complete description, see the WWC Procedures and Standards Handbook, Appendix E.

Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence ¹
		Schools	Students	
Alphabetics	na	na	na	na
Reading fluency	na	na	na	na
Comprehension	1	4	96	Small
General literacy achievement	na	na	na	na

na = not applicable/not studied

1. A rating of “medium to large” requires at least two studies and two schools across studies in one domain and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is “small.” For more details on the extent of evidence categorization, see the WWC Procedures and Standards Handbook, Appendix G.