

What Works Clearinghouse



Early Risers

Program Description¹

Early Risers is a multi-year prevention program for elementary school children demonstrating early aggressive and disruptive behavior. The intervention model includes two child-focused components and two parent/family components. The Child Skills component is designed to teach skills that enhance children’s emotional and behavioral self-regulation, positive peer relationships, and academic success. The Child School Support component aims to identify areas of difficulty in the classroom and creates individualized plans to address those difficulties during the course of normal school activities. The Parent Skills component is delivered in “family night” group sessions and is intended to promote parents’ abilities to support their children’s healthy development by teaching skills that address positive parent–child relations, effective discipline practices, and parent involvement in school. The Family Support component, which is delivered via home visits, identifies basic needs and health concerns and then implements plans designed to assist families in achieving and maintaining healthy lifestyles.

Research²

Two studies of *Early Risers* that fall within the scope of the Children Classified as Having an Emotional Disturbance review protocol meet What Works Clearinghouse (WWC) evidence standards without reservations. The two studies included 30 schools with 389 children from kindergarten to second grade exhibiting signs of early aggressive behavior. Based on these two studies, the WWC considers the extent of evidence for *Early Risers* on children classified with an emotional disturbance (or children at risk for classification) to be medium to large for three domains: external behavior, social outcomes, and academic performance, and small for one domain: emotional/internal behavior. Three other domains are not reported in this intervention report. (See the Effectiveness Summary for further description of all domains.)

Effectiveness

Early Risers was found to have no discernible effects on external behavior and emotional/internal behavior, and potentially positive effects on social outcomes and academic performance for children classified as having an emotional disturbance.

Report Contents

Overview	p. 1
Program Information	p. 3
Research Summary	p. 4
Effectiveness Summary	p. 5
References	p. 8
Research Details for Each Study	p. 10
Outcome Measures for Each Domain	p. 14
Findings Included in the Rating for Each Outcome Domain	p. 16
Supplemental Findings for Each Outcome Domain	p. 20
Endnotes	p. 22
Rating Criteria	p. 24
Glossary of Terms	p. 25

Table 1. Summary of findings³

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students ⁴	Extent of evidence
		Average	Range			
External behavior	No discernible effects	+4	+1 to +12	2	380	Medium to large
Emotional/ internal behavior	No discernible effects	+7	+5 to +9	1	181	Small
Social outcomes	Potentially positive effects	+7	-4 to +14	2	380	Medium to large
Academic performance	Potentially positive effects	+3	-5 to +12	2	389	Medium to large

Program Information

Background

Early Risers was developed by Gerald J. August, George M. Realmuto, and Michael L. Bloomquist at the Center for Prevention and Children's Mental Health at the University of Minnesota. *Early Risers* is distributed by the Division of Child and Adolescent Psychiatry, University of Minnesota, F256/2B West, 2450 Riverside Avenue, Minneapolis, MN 55454-1495. Email: earlyrisers@umn.edu. Web: <http://www.med.umn.edu/psychiatry/research/earlyrisers/home.html>. Telephone: (612) 273-9711.

Program details

Early Risers is designed for elementary school children ages 6 to 12 with aggressive or otherwise disruptive behavior.

Early Risers uses integrated child-, school-, and family-focused interventions aimed at altering the developmental trajectory of children with early onset aggressive behavior. The program is a fluid intervention with modifications and improvements made continuously based on research results, and generally includes social skills groups and a summer school program for students, teacher consultation and mentoring, and parent training. A "family advocate" (someone with a bachelor's degree and experience working with children/parents) coordinates the child- and family-focused components. During the regular school year, the family advocate consults with classroom teachers to assess the child's academic progress, peer relations, classroom behavior, and emotional regulation, and then collaborates with teachers on appropriate intervention options.

Cost⁵

The total annual cost to administer the program is approximately \$1,500 to \$2,500 per student.

Research Summary

Thirteen studies reviewed by the WWC investigated the effects of *Early Risers* on children classified as having an emotional disturbance (or children at risk for classification). Two studies (August, Hektner, Egan, Realmuto, & Bloomquist, 2002; August, Lee, Bloomquist, Realmuto, & Hektner, 2003) are randomized controlled trials that meet WWC evidence standards without reservations. Those two studies are summarized in this report. The remaining 11 studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on p. 8 for citations for all 13 studies.)

Table 2. Scope of reviewed research

Grade	1, 2, 3
Delivery method	Individual/Small group/ Whole class
Program type	Supplement
Studies reviewed	13
Meets WWC standards without reservations	2 studies
Meets WWC standards with reservations	0 studies

Summary of studies meeting WWC evidence standards without reservations

August et al. (2002) measured the effects of *Early Risers* on a sample of students with early-onset aggressive behavior. Students were drawn from 20 schools that were randomly assigned to either *Early Risers* or a comparison condition. During the baseline year, all students were in kindergarten. Intervention effects on external behavior, social outcomes, and academic performance outcomes were measured twice, after students in the *Early Risers* condition had received two and three years of implementation.⁶ Parent nurturance, distress, and effective discipline were also measured; these outcomes are not presented in this report because they do not fall under a domain specified in the protocol. The analysis sample after three years of implementation included 199 students: *Early Risers* group (n = 100) and comparison group (n = 99).

August et al. (2003) randomly assigned kindergarten and first-grade students who displayed aggressive behaviors to one of three conditions: full intervention/*Early Risers*–Child Skills and Family Support (referred to as CORE+FLEX in August et al., 2003), partial intervention/*Early Risers*–Child Skills (referred to as CORE in August et al., 2003), or a no-treatment comparison condition.⁷ The study collapsed the two *Early Risers* groups in the analyses and reported student outcomes on external behavior, emotional/internal behavior, social outcomes, and academic performance outcomes after two years of implementation. Parenting stress and negative parenting style were also measured; these outcomes are not presented in this report because they do not fall under a domain specified in the protocol. The analysis sample after two years of implementation included 190 students: *Early Risers* group (n = 127) and comparison group (n = 63).⁸

Summary of studies meeting WWC evidence standards with reservations

No studies of *Early Risers* meet WWC evidence standards with reservations.

Effectiveness Summary

The WWC review of interventions for Children Classified as Having an Emotional Disturbance addresses student outcomes in seven domains: external behavior, emotional/internal behavior, social outcomes, reading achievement/literacy, math achievement, school attendance, and academic performance. The two studies that contribute to the effectiveness rating in this report cover four domains: external behavior, emotional/internal behavior, social outcomes, and academic performance. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Early Risers* on children classified as having an emotional disturbance. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 24.

Summary of effectiveness for the external behavior domain

Two studies reported findings in the external behavior domain.

August et al. (2002) found, and the WWC confirmed, no statistically significant differences between the intervention and comparison groups on aggression, hyperactivity, or impulsivity composite measures. The average effect size across the outcomes was not substantively important according to WWC criteria (that is, at least 0.25 standard deviations).

August et al. (2003) found a statistically significant difference between the intervention and comparison groups on the Externalizing Problems Teacher Report measure. Based on WWC calculations, the effect was neither statistically significant nor large enough to be substantively important according to WWC criteria. August et al. (2003) also found, and the WWC confirmed, no statistically significant differences between the intervention and comparison groups on the Externalizing Problems Parent Report and School Adjustment Teacher Report measures. The average effect size from these three outcomes was not large enough to be considered substantively important by the WWC.

Thus, for the external behavior domain, no studies showed statistically significant effects or mean effect sizes large enough to be considered substantively important. This results in an intervention rating of no discernible effects, with a medium to large extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the external behavior domain

Rating of effectiveness	Criteria met
No discernible effects <i>No affirmative evidence of effects.</i>	The review of <i>Early Risers</i> in the external behavior domain had no studies showing a statistically significant or substantively important effect, either positive or negative.
Extent of evidence	Criteria met
Medium to large	The review of <i>Early Risers</i> in the external behavior domain was based on two studies that included 30 schools and 380 students.

Summary of effectiveness for the emotional/internal behavior domain

One study reported findings in the emotional/internal behavior domain.

August et al. (2003) found, and the WWC confirmed, no statistically significant difference between the intervention and comparison groups on the Internalizing Problems Parent Report and Internalizing Problems Teacher Report measures. The average effect size across the outcomes was not substantively important according to WWC criteria.

Thus, for the emotional/internal behavior domain, no studies showed statistically significant or substantively important effects. This results in an intervention rating of no discernible effects, with a small extent of evidence.

Table 4. Rating of effectiveness and extent of evidence for the emotional/internal behavior domain

Rating of effectiveness	Criteria met
No discernible effects <i>No affirmative evidence of effects.</i>	The review of <i>Early Risers</i> in the emotional/internal behavior domain had no studies showing a statistically significant or substantively important effect, either positive or negative.
Extent of evidence	Criteria met
Small	The review of <i>Early Risers</i> in the emotional/internal behavior domain was based on one study that included 10 schools and 181 students.

Summary of effectiveness for the social outcomes domain

Two studies reported findings in the social outcomes domain.

August et al. (2002) found, and the WWC confirmed, a positive and statistically significant difference between the intervention and comparison groups on the Social Skills Composite measure, and no statistically significant difference between the intervention and comparison groups on the Adaptability Composite measure.

August et al. (2003) found, and the WWC confirmed, no statistically significant difference between the intervention and comparison groups on the Social Competence Parent Report or Social Competence Teacher Report measures. The mean effect from these two outcomes was not large enough to be considered substantively important.

Thus, for the social outcomes domain, one study with a strong design showed a statistically significant positive effect. This results in an intervention rating of potentially positive effects, with a medium to large extent of evidence.

Table 5. Rating of effectiveness and extent of evidence for the social outcomes domain

Rating of effectiveness	Criteria met
Potentially positive effects <i>Evidence of a positive effect with no overriding contrary evidence.</i>	The review of <i>Early Risers</i> in the social outcomes domain had one study showing a statistically significant positive effect, no studies showing a statistically significant or substantively important negative effect, and one study showing an indeterminate effect.
Extent of evidence	Criteria met
Medium to large	The review of <i>Early Risers</i> in the social outcomes domain was based on two studies that included 30 schools and 380 students.

Summary of effectiveness for the academic performance domain

Two studies reported findings in the academic performance domain.

August et al. (2002) found, and the WWC confirmed, a positive and statistically significant difference between the intervention and comparison groups on the Academic Achievement Composite and no statistically significant difference between the intervention and comparison groups on the Concentration Problems Composite.

August et al. (2003) found, and the WWC confirmed, no statistically significant difference between the intervention and comparison group on the Academic Achievement Composite. The effect size for this outcome was not substantively important according to WWC criteria.

Thus, for the academic performance domain, one study with a strong design showed a statistically significant positive effect. This results in an intervention rating of potentially positive effects, with a medium to large extent of evidence.

Table 6. Rating of effectiveness and extent of evidence for the academic performance domain

Rating of effectiveness	Criteria met
Potentially positive effects <i>Evidence of a positive effect with no overriding contrary evidence.</i>	The review of <i>Early Risers</i> in the academic performance domain had one study showing a statistically significant positive effect, no studies showing a statistically significant or substantively important negative effect, and one study showing an indeterminate effect.
Extent of evidence	Criteria met
Medium to large	The review of <i>Early Risers</i> in the academic performance domain was based on two studies that included 30 schools and 389 students.

References

Studies that meet WWC evidence standards without reservations

August, G. J., Hektner, J. M., Egan, E. A., Realmuto, G. M., & Bloomquist, M. L. (2002). The Early Risers longitudinal prevention trial: Examination of 3-year outcomes in aggressive children with intent-to-treat and as-intended analyses. *Psychology of Addictive Behaviors, 16*(4, Suppl), S27–S39.

Additional sources:

August, G. J., Egan, E. A., Realmuto, G. M., Hektner, J. M., & Haaga, D. A. F. (2003). Four years of the Early Risers early-age-targeted preventive intervention: Effects on aggressive children's peer relations. *Behavior Therapy, 34*(4), 453–470.

August, G. J., Realmuto, G. M., Hektner, J. M., & Bloomquist, M. L. (2001). An integrated components preventive intervention for aggressive elementary school children: The Early Risers program. *Journal of Consulting & Clinical Psychology, 69*(4), 614–626.

Bernat, D. H., August, G. J., Hektner, J. M., & Bloomquist, M. L. (2007). The Early Risers preventive intervention: Testing for six-year outcomes and mediational processes. *Journal of Abnormal Child Psychology, 35*(4), 605–617.

Endicott, L. G. (2003). Reducing risk for antisocial behavior via protective factor development: The Early Risers prevention trial. *Dissertation Abstracts International, 64*(04B), 89-1924.

August, G. J., Lee, S. S., Bloomquist, M. L., Realmuto, G. M., & Hektner, J. M. (2003). Dissemination of an evidence-based prevention innovation for aggressive children living in culturally diverse, urban neighborhoods: The Early Risers effectiveness study. *Prevention Science, 4*(4), 271–286.

Additional source:

August, G. J., Lee, S. S., Bloomquist, M. L., Realmuto, G. M., & Hektner, J. M. (2004). Maintenance effects of an evidence-based prevention innovation for aggressive children living in culturally diverse, urban neighborhoods: The Early Risers effectiveness study. *Journal of Emotional & Behavioral Disorders, 12*(4), 194–205.

Studies that are ineligible for review using the Children Classified as Having an Emotional Disturbance Evidence Review Protocol

August, G. J., Bloomquist, M. L., Lee, S. S., Realmuto, G. M., & Hektner, J. M. (2006). Can evidence-based prevention programs be sustained in community practice settings? The Early Risers' advanced-stage effectiveness trial. *Prevention Science, 7*(2), 151–165. The study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review.

August, G. J., Bloomquist, M. L., Realmuto, G. M., & Hektner, J. M. (2007). The Early Risers "Skills for Success" program: A targeted intervention for preventing conduct problems and substance abuse in aggressive elementary school children. In P. Tolan, J. Szapocznik, & S. Sambrano (Eds.), *Preventing youth substance abuse: Science-based programs for children and adolescents* (pp. 137–158). Washington, DC: American Psychological Association. The study is ineligible for review because it does not use a comparison group design or a single-case design.

August, G. J., Egan, E. A., Realmuto, G. M., & Hektner, J. M. (2003). Parceling component effects of a multifaceted prevention program for disruptive elementary school children. *Journal of Abnormal Child Psychology, 31*(5), 515–527. The study is ineligible for review because it does not use a comparison group design or a single-case design.

August, G. J., Realmuto, G. M., Mathy, R. M., & Lee, S. (2003). The Early Risers FLEX program: A family-centered preventive intervention for children at-risk for violence and antisocial behavior. *Behavioral Analyst Today, 4*, 26–33. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

- August, G. J., Winters, K. C., Realmuto, G. M., Tarter, R., Perry, C., & Hektner, J. M. (2004). Moving evidence-based drug Abuse1 prevention programs from basic science to practice: "Bridging the efficacy-effectiveness interface." *Substance Use & Misuse*, 39(10–12), 2017–2053. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Carter, E. W., Lane, K. L., Crnobori, M., Bruhn, A. L., & Oakes, W. P. (2011). Self-determination interventions for students with and at risk for emotional and behavioral disorders: Mapping the knowledge base. *Behavioral Disorders*, 36(2), 100–116. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Gewirtz, A., & August, G. J. (2008). Incorporating multifaceted mental health prevention services in community sectors of care. *Clinical Child and Family Review*, 11, 1–11. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Hektner, J. M., August, G. J., & Realmuto, G. M. (2003). Effects of pairing aggressive and nonaggressive children in strategic peer affiliation. *Journal of Abnormal Child Psychology*, 31(4), 399. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Lee, C. S., Anderson, J. R., Horowitz, J. L., & August, G. J. (2009). Family income and parenting: The role of parental depression and social support. *Family Relations*, 58, 417–430. The study is ineligible for review because it does not include a student outcome.
- Nelson, W. M., III, & Schultz, J. R. (2009). Managing anger and aggression in students with externalizing behavior problems: Focus on exemplary programs. In M. J. Mayer, J. E. Lochman, & R. Van Acker (Eds.), *Cognitive-behavioral interventions for emotional and behavioral disorders: School-based practice* (pp. 143–170). New York: Guilford Press. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.
- Realmuto, G. M., August, G. J., & Egan, E. A. (2004). Testing the goodness-of-fit of a multifaceted preventive intervention for children at risk for conduct disorder. *Canadian Journal of Psychiatry*, 49(11), 743–752. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Appendix A.1: Research details for August et al., 2002

August G. J., Hektner, J. M., Egan, E. A., Realmuto, G. M., & Bloomquist, M. L. (2002). The Early Risers longitudinal prevention trial: Examination of 3-year outcomes in aggressive children with intent-to-treat and as-intended analyses. *Psychology of Addictive Behaviors, 16*(4, Suppl), S27–S39.

Table A1. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
External behavior	20 schools/199 students	+3	No
Social outcomes	20 schools/199 students	+4	Yes
Academic performance	20 schools/199 students	+4	Yes

Setting The study was conducted in two semirural sites in Minnesota, characterized by families of low and middle socioeconomic status.

Study sample The sample consisted of students from 20 schools that were randomly assigned to either the *Early Risers* condition (n = 10 schools) or the comparison condition (n = 10 schools). Within these schools, 95% of kindergarten students were screened using teacher ratings of aggressive-disruptive behavior on the 25-item Aggression Scale of the Child Behavior Checklist–Teacher Rating Form (Achenbach, 1991).⁹ Children who obtained a t-score greater than 58 on the Aggression Scale (using gender-specific norms) or who were at or above the 85th percentile relative to all kindergarten students in their school without dropping below a t-score of 55 were eligible. Students were excluded from the study if their IQ was less than 80 or if they had a pervasive developmental disorder that required special education placement. Using these criteria, 341 children were screened in as potential participants; this initial sample consisted of 173 students in intervention schools and 168 students in comparison schools.¹⁰ During the baseline year, all students were in kindergarten. Intervention effects were measured after students in the *Early Risers* group had received two and three years of implementation.¹¹ The analysis sample after three years of implementation included 199 students: *Early Risers* group (n = 100) and comparison group (n = 99). Gender information was not available for the analysis sample. Of the sample of children who received initial parental consent prior to assignment (n = 245), 69% were boys and 31% were girls. Race and ethnicity information for the study sample was not presented.

Intervention group

The current report focuses on impacts after two (Appendix D) and three (Appendix C) years of implementation.¹²

The Child Skills component included a Monitoring and Mentoring School Consultation Program during the school year and an annual six-week, full-day summer school program. The Monitoring and Mentoring School Consultation Program consisted of teacher consultation and student mentoring. The summer school program began in the summer following kindergarten and included academic learning centers; training in social skills, art, drama, and sports; large-group recreation; lunch; recess; and the use of peer mentors. A structured behavior modification program was implemented across all daily activities.

The Family Program consisted of separate but concurrent parent and child sessions held on evenings or weekends from October through May. During the first three years, parent sessions addressed topics such as use of praise and discipline, involvement in schoolwork and learning at home, self-control and problem solving, communication skills, stress management, and social support. The child sessions focused on emotion regulation, conflict resolution, social skills, and understanding school rules. Session content was delivered using video modeling, fantasy play, and role-plays. Home visits, modeled after home-based wraparound mental health service programs, also were used to meet family goals.

Comparison group

Children in the comparison condition did not participate in any aspect of the *Early Risers* program.

Outcomes and measurement

This study included measures of social skills, adaptability, academic achievement, aggression, hyperactivity, and impulsivity after two and three years of implementation. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Staff members were required to participate in a formal program of education and training prior to the implementation of each intervention component. Intervention manuals were obtained from the original program developers, who also served as project consultants. Staff members, who received ongoing supervision during the implementation phase, were required to demonstrate mastery of content and delivery methods.

Appendix A.2: Research details for August et al., 2003

August, G. J., Lee, S. S., Bloomquist, M. L., Realmuto, G. M., & Hektner, J. M. (2003). Dissemination of an evidence-based prevention innovation for aggressive children living in culturally diverse, urban neighborhoods: The Early Risers effectiveness study. *Prevention Science, 4*(4), 271–286.

Table A2. Summary of findings

Meets WWC evidence standards without reservations

Outcome domain	Sample size ⁴	Study findings	
		Average improvement index (percentile points)	Statistically significant
External behavior	181 students	+6	No
Emotional/internal behavior	181 students	+7	No
Social outcomes	181 students	+9	No
Academic performance	190 students	+1	No

Setting The study took place in two neighborhood family centers and 10 affiliated elementary schools in a large midwestern city.

Study sample Kindergarten and first-grade students from 10 elementary schools were included in the study. Children were recruited for participation in two cohorts; parents were informed that some students would be assigned via a lottery procedure to participate in the two-year intervention and others would be involved in assessments only. Children who received parental consent were then screened on the 25-item Aggression Scale of the Child Behavior Checklist–Teacher Rating Form (Achenbach, 1991). Students who received a t-score greater than or equal to 55 were eligible for the study, unless they had a pervasive developmental disorder or serious emotional-behavioral disorder that required special education placement. A total of 327 students were eligible for the study and were randomized into three groups: full intervention (n = 107), partial intervention (n = 111), and comparison group (n = 109). The two intervention groups were collapsed by the researchers.¹³ The final sample included 190 students: *Early Risers* group (n = 127) and comparison group (n = 63). The *Early Risers* group consisted mostly of African-American (86%) and male (59%) students. The comparison group also consisted mostly of African-American (80%) and male (55%) students.

Intervention group

Children were originally assigned to two intervention groups (full and partial strength). Both groups received Child Skills components for two years, and the full-strength group also received the Family Support component. The Child Skills component included a summer program, an after-school program, and a Monitoring and Mentoring School Consultation Program. For two consecutive summers, the summer program activities took place over a six-week period and focused on social skills, creative arts, physical fitness, and recreation. The after-school program took place one day a week over a two-year period (from October to May) and included small-group social skills instruction, homework assistance, and recreational activities. The first year of the after-school program focused on social, emotional, and problem-solving skills, whereas the second year focused on empathy and anger management. Fifty percent of children attended at least 48% of the summer program and 43% of the after-school sessions. Formal fidelity assessment was conducted on a scale of 1 (not at all) to 4 (most of the time); means scores for the after-school and summer program ranged from 3.5 to 4.

Beginning midway through Year 1 and continuing through Year 2, students received support at their regular school through the Monitoring and Mentoring School Consultation Program; this component involved monitoring student attendance, behavior, homework completion, and academic performance through consultation with each child's teacher. When a domain was flagged as being problematic, a school advocate would meet with the teacher to develop a plan for one-to-one mentoring at the school. The amount of mentoring time received by individual students varied across schools and classrooms.

The Family Support component included home-based therapy delivered by family advocates who were required to make a minimum of three bimonthly contacts in the first year and six contacts in the second year. The program was adjusted to the needs of each family, and there was much variability in the amount of contact time families accumulated. Some families did not meet the minimum contact time requirements, whereas other families received many more contacts. The average amount of contact time per family was 9.6 hours. The Family Support program was utilized primarily by highly stressed families to help find housing, health care, employment, and child care.

Comparison group

Children in the comparison condition did not participate in any aspect of the *Early Risers* program.

Outcomes and measurement

This study included measures of academic achievement, externalizing problems, school adjustment, social competence, and internalizing problems. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Two employees at each neighborhood center served as family/school advocates and coordinated the summer program and after-school components. Staff received an intensive training program prior to the start of each component and received weekly structured supervision by center supervisors. Adherence to content and delivery specifications was monitored periodically via unannounced observations made by fidelity technicians who observed sessions. School advocates were available to consult with students' classroom teachers upon request. Two of the four original family advocates left the program after Year 1. One of these positions experienced two additional personnel changes.

Appendix B: Outcome measures for each domain

External behavior	
<i>Aggression Composite</i>	This composite score is the mean of the aggressive behavior scores from four instruments: the Teacher Observation of Class Adaptation–Revised (TOCA-R), the Parent Observation of Class Adaptation (POCA), the Behavioral Assessment System for Children–Teacher Rating Scale (BASC-TRS), and the Behavioral Assessment System for Children–Parent Rating System (BASC-PRS). The TOCA-R is a teacher-report measure using Likert scales to describe the frequency of classroom behaviors in the previous four weeks; aggressive and disruptive behavior is measured via 14 items. The POCA uses an identical structure to the TOCA-R but is based on parent reports of aggressive and disruptive behavior, using 14 items (as cited in August et al., 2002). The teacher version of the BASC-TRS Aggression scale consists of 14 items, and the parent version consists of 13 items (as cited in August et al., 2002).
<i>Externalizing Problems Parent Report</i>	This composite score was computed as the mean of the aggression, conduct, depression, and hyperactivity scales from the BASC-PRS (as cited in August et al., 2003).
<i>Externalizing Problems Teacher Report</i>	This composite score was computed as the mean of the adaptability, aggression, and conduct scales from the BASC-TRS (as cited in August et al., 2003).
<i>Hyperactivity Composite</i>	This composite score was computed as the mean of the BASC-TRS and BASC-PRS hyperactivity scales. The teacher version consists of 13 items, and the parent version consists of 10 items (as cited in August et al., 2002).
<i>Impulsivity Composite</i>	The Impulsivity Composite was computed as the mean of three impulsivity items from the TOCA-R and the three impulsivity items from the POCA measure (as cited in August et al., 2002).
<i>School Adjustment Teacher Report</i> ¹⁴	This composite score was computed as the mean of the attention problems, learning problems, study skills, cognitive competence, and school adjustment scales from the BASC-TRS (as cited in August et al., 2003).
<i>Self-Regulation Problems</i>	This composite score is based on the aggressive-disruptive scales, hyperactivity scales, and impulsivity scales from both the TOCA-R and the POCA, as well as the aggression and hyperactivity scales from the externalizing domain of the BASC-TRS and the BASC-PRS (as cited in August et al., 2001).
Emotional/internal behavior	
<i>Internalizing Problems Parent Report</i>	This composite was calculated as the mean of the anxiety, somatization, and withdrawal scales from the BASC-PRS (as cited in August et al., 2003).
<i>Internalizing Problems Teacher Report</i>	This composite was calculated as the mean of the anxiety, somatization, and withdrawal scales from the BASC-TRS (as cited in August et al., 2003).
Social outcomes	
<i>Adaptability Composite</i>	In both August et al. (2001) and August et al. (2002), this composite score was calculated as the mean of the adaptability scales from the BASC-TRS and BASC-PRS. It assessed the child’s ability to share, adapt to changes in activities and routines, and take setbacks in stride (as cited in August et al., 2002).
<i>Social Competence Composite</i>	This composite score was based on the six-item social acceptance scale from the Teacher’s Scale of Child’s Actual Competence (Harter, 1985) ¹⁵ and three scales (adaptability, social skills, and leadership) from both the BASC-TRS and BASC-PRS (as cited in August et al., 2001).
<i>Social Competence Parent Report</i>	This composite score was computed as the mean of the adaptability, leadership, and social skills scales from the BASC-PRS (as cited in August et al., 2003).
<i>Social Competence Teacher Report</i>	This composite score was computed as the mean of the leadership, social skills, peer acceptance, and peer-preferred scales from the BASC-TRS (as cited in August et al., 2003).
<i>Social Skills Composite</i>	In both August et al. (2001) and August et al. (2002), this composite score was calculated as the mean of the social skills and leadership scales in the BASC-TRS and BASC-PRS. The social skills scale assesses behaviors such as helping and encouraging others, beginning conversations appropriately, and admitting mistakes. The leadership scale assesses the ability to offer good suggestions and make decisions easily (as cited in August et al., 2002).

Academic performance

Academic Achievement Composite

In August et al. (2002), this variable is the mean of four scales: the broad reading and applied problems composite scores from the Woodcock-Johnson Tests of Achievement–Revised (WJ-R), the learning problems scale of the BASC-TRS, and the cognitive competence scale of the Teacher’s Scale of Child’s Actual Competence and Social Acceptance (Harter, 1985).¹⁵ The WJ-R is a standardized measure of reading and arithmetic skills, and the two other scales are completed by the teacher (as cited in August et al., 2002).

In August et al. (2003), this variable is the mean of the basic reading, broad reading, and applied learning problems scales from the WJ-R. Scale scores are standardized with a mean of 100 and a standard deviation of 15 (as cited in August et al., 2003).

Academic Competence Composite

This composite is the mean of the broad reading and applied problems scores from the WJ-R (as cited in August et al., 2001).

Concentration Problems Composite

This composite is the mean of the Concentration Problems scales from the TOCA-R and the POCA. The scale includes items that assess the child’s ability to complete assignments, pay attention, and stay on task. The TOCA-R items ask teachers to rate the frequency of behaviors within the previous four weeks, using a Likert scale. The POCA uses parent ratings to assess child behavior in the home (as cited in August et al., 2002).

Appendix C.1: Findings included in the rating for the external behavior domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
August et al., 2002^a								
<i>Aggression Composite</i>	Grade 3	20 schools/ 199 students	nr	nr	nr	0.11	+4	0.29
<i>Hyperactivity Composite</i>	Grade 3	20 schools/ 199 students	nr	nr	nr	0.03	+1	0.77
<i>Impulsivity Composite</i>	Grade 3	20 schools/ 199 students	nr	nr	nr	0.07	+3	0.49
Domain average for external behavior (August et al., 2002)						0.07	+3	Not statistically significant
August et al., 2003^b								
<i>Externalizing Problems Parent Report</i>	Grades 1 and 2	132 students	0.45 (0.91)	0.59 (1.09)	0.14	0.06	+2	0.47
<i>Externalizing Problems Teacher Report</i>	Grades 1 and 2	181 students	0.67 (0.83)	0.70 (0.82)	0.03	0.07	+3	0.04
<i>School Adjustment Teacher Report</i>	Grades 1 and 2	181 students	-0.41 (0.83)	-0.69 (0.76)	0.28	0.31	+12	0.40
Domain average for external behavior (August et al., 2003)						0.15	+6	Not statistically significant
Domain average for external behavior across all studies						0.11	+4	na

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study’s domain average was determined by the WWC; a study is characterized as not statistically significant when univariate statistical tests are reported for each outcome measure and each of the effects within the domain are not statistically significant. nr = not reported. na = not applicable.

^a August et al. (2002) reported study findings after three years of implementation; study findings on this same sample after two years of implementation (from August et al., 2001) are not included in these ratings but are reported in Appendix D.1. August et al. (2002) used a two-level mixed random regression model, with time points nested within individual participants. The model included gender and severity of initial aggression as time-invariant moderators. The model that evaluated effects on the Aggression Composite outcome also accounted for the nesting of students within schools, because initial results showed significant variation between schools in intercepts and slopes on the Aggression Composite. There was no significant variation between schools in intercepts and slopes on all other variables. No corrections for clustering or multiple comparisons were needed. The effect sizes and p-values presented here were provided to the WWC by the author.

^b To make results comparable with other outcomes reported in this report, signs were reversed on the mean difference, effect size, and improvement index for *Externalizing Problems Teacher Report* and *Externalizing Problems Parent Report* (August et al., 2003). This was done to demonstrate that the intervention group was favored when negative differences were reported. August et al. (2003) analyzed outcomes with a two-level random regression model, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. Outcomes reported here represent the effect of the intervention group x time x severity of initial aggression. All variables are scaled as z-scores, standardized to the mean and standard deviation of the normative sample. The p-values presented here were reported in the original study. A correction for multiple comparisons was needed and resulted in significance levels that differ from those in the original study. As a result of the multiple comparisons adjustment, the p-value of 0.04 for the *Externalizing Problems Teacher Report* was higher than the critical p-value for statistical significance; therefore, the WWC does not find the result to be statistically significant.

Appendix C.2: Findings included in the rating for the emotional/internal behavior domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
August et al., 2003^a								
<i>Internalizing Problems Parent Report</i>	Grades 1 and 2	132 students	0.22 (0.63)	0.37 (0.88)	0.15	0.12	+5	0.21
<i>Internalizing Problems Teacher Report</i>	Grades 1 and 2	181 students	0.25 (0.74)	0.51 (1.06)	0.26	0.23	+9	0.49
Domain average for emotional/internal behavior (August et al., 2003)						0.18	+7	Not statistically significant
Domain average for emotional/internal behavior across all studies						0.18	+7	na

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study’s domain average was determined by the WWC; a study’s effect is characterized as not statistically significant when univariate statistical tests are reported for each outcome measure and each of the effects within the domain are not statistically significant. na = not applicable.

^a To make results comparable with other outcomes reported in this report, signs were reversed on the mean difference, effect size, and improvement index for both outcomes in August et al. (2003). This was done to demonstrate that the intervention group was favored when negative differences were reported. August et al. (2003) analyzed outcomes with a two-level random regression model, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. Outcomes reported here represent the effect of intervention group x time x severity of initial aggression. All variables are scaled as z-scores, standardized to the mean and standard deviation of the normative sample. No corrections for clustering or multiple comparisons were needed. The p-values presented here were reported in the original study.

Appendix C.3: Findings included in the rating for the social outcomes domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
August et al., 2002^a								
<i>Adaptability Composite</i>	Grade 3	20 schools/ 199 students	nr	nr	nr	-0.11	-4	0.14
<i>Social Skills Composite</i>	Grade 3	20 schools/ 199 students	nr	nr	nr	0.33	+13	0.02
Domain average for social outcomes (August et al., 2002)						0.11	+4	Statistically significant
August et al., 2003^b								
<i>Social Competence Parent Report</i>	Grades 1 and 2	132 students	-0.13 (0.73)	-0.27 (0.86)	0.14	0.09	+4	0.15
<i>Social Competence Teacher Report</i>	Grades 1 and 2	181 students	-0.11 (0.69)	-0.46 (0.69)	0.35	0.35	+14	0.36
Domain average for social outcomes (August et al., 2003)						0.22	+9	Not statistically significant
Domain average for social outcomes across all studies						0.17	+7	na

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC; a study is characterized as having a statistically significant positive effect when univariate statistical tests are reported for each outcome measure, the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. A study's effect is characterized as not statistically significant when univariate statistical tests are reported for each outcome measure and each of the effects within the domain are not statistically significant. nr = not reported. na = not applicable.

^a August et al. (2002) reported study findings after three years of implementation; study findings for social outcomes on this same sample after two years of implementation (from August et al., 2001) are not included in these ratings but are reported in Appendix D.2. August et al. (2002) used a two-level mixed random regression model, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. There was no significant variation between schools in intercepts and slopes on any of the variables in the social outcomes domain, so a three-level model (to account for nesting within schools) was not used. A correction for multiple comparisons was needed but did not affect significance levels. The effect size and p-value for the *Adaptability Composite* were provided to the WWC by the author. The effect size and p-value for the *Social Skills Composite* were reported in the original study.

^b August et al. (2003) used a two-level mixed random regression model, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. Outcomes from August et al. (2003) represent the effect of intervention group x time x severity of initial aggression. All variables are scaled as z-scores, standardized to the mean and standard deviation of the normative sample. A correction for multiple comparisons was needed but did not affect significance levels. The p-values presented here were reported in the original study.

Appendix C.4: Findings included in the rating for the academic performance domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
August et al., 2002^a								
<i>Academic Achievement Composite</i>	Grade 3	20 schools/199 students	nr	nr	nr	0.30	+12	0.02
<i>Concentration Problems Composite</i>	Grade 3	20 schools/199 students	nr	nr	nr	-0.12	-5	0.19
Domain average for academic performance (August et al., 2002)						0.09	+4	Statistically significant
August et al., 2003^b								
<i>Academic Achievement Composite</i>	Grades 1 and 2	190 students	-0.11 (0.79)	-0.11 (0.78)	0.00	0.07	+1	0.88
Domain average for academic performance (August et al., 2003)						0.07	+1	Not statistically significant
Domain average for academic performance across all studies						0.08	+3	na

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC; a study is characterized as having a statistically significant positive effect when univariate statistical tests are reported for each outcome measure, the effect for at least one measure within the domain is positive and statistically significant, and no effects are negative and statistically significant. A study's effect is characterized as not statistically significant when univariate statistical tests are reported for each outcome measure and each of the effects within the domain are not statistically significant. nr = not reported. na = not applicable.

^a August et al. (2002) reported study findings after three years of implementation; study findings for academic performance outcomes on this same sample after two years of implementation (from August et al., 2001) are not included in these ratings but are reported in Appendix D.3. August et al. (2002) used a two-level mixed random regression model, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. In August et al. (2002), there was no significant variation between schools in intercepts and slopes on any of the academic performance outcome variables, so a three-level model (to account for nesting within schools) was not used. A correction for multiple comparisons was needed but did not affect significance levels. The effect size and p-value for the *Concentration Problems Composite* were provided to the WWC by the author. The effect size and p-value for the *Academic Achievement Composite* were reported in the original study.

^b August et al. (2003) used a two-level mixed random regression model, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. Outcomes from August et al. (2003) represent the effect of intervention group x time x severity of initial aggression. All variables are scaled as z-scores, standardized to the mean and standard deviation of the normative sample. No corrections for clustering or multiple comparisons were needed. The p-values presented here were reported in the original study.

Appendix D.1: Summary of second-year findings for the external behavior domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
August et al., 2001^a								
<i>Aggression Composite</i>	Grade 1	20 schools/ 199 students	0.78 (0.88)	0.83 (1.05)	0.05	nr	na	0.39
<i>Hyperactivity Composite</i>	Grade 1	20 schools/ 199 students	0.83 (1.02)	0.84 (1.10)	0.01	nr	na	0.41
<i>Impulsivity Composite</i>	Grade 1	20 schools/ 199 students	0.87 (0.92)	0.95 (1.09)	0.08	nr	na	0.80
<i>Self-Regulation Problems</i>	Grade 1	20 schools/ 199 students	0.82 (0.81)	0.90 (0.95)	0.08	nr	na	0.49

Table Notes: The supplemental findings presented in this table are additional findings from the studies in this report that do not factor into the determination of the intervention rating. Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. nr = not reported. na = not applicable.

^a This appendix presents findings from August et al. (2001) after two years of implementation. To make results comparable with other outcomes in this report, signs were reversed on the mean difference, effect size, and improvement index for all findings in the external behavior domain for August et al. (2001). This was done to demonstrate that the intervention group was favored when negative differences were reported. August et al. (2001) used two-level mixed regression models, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. There was no significant variation between schools in intercepts and slopes on any of the variables in the external behavior domain, so a three-level model (to account for nesting within schools) was not used. Outcomes reported here represent the effect of intervention group x time. Effect sizes were not reported for nonsignificant findings. All variables are scaled as z-scores, relative to a normative sample in which the mean is 0 and the standard deviation is 1. Means are full-information maximum likelihood estimates provided by the author based on all available data from the full sample. No corrections for clustering or multiple comparisons were needed. The p-values presented here were reported in the original study.

Appendix D.2: Summary of second-year findings for the social outcomes domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
August et al., 2001^a								
<i>Social Competence Composite</i>	Grade 1	20 schools/ 199 students	-0.47 (0.68)	-0.60 (0.67)	0.13	nr	na	0.36
<i>Social Skills Composite</i>	Grade 1	20 schools/ 199 students	-0.34 (0.73)	-0.44 (0.74)	0.10	nr	na	0.41
<i>Adaptability Composite</i>	Grade 1	20 schools/ 199 students	-0.71 (0.71)	-0.63 (0.83)	-0.08	nr	na	0.28

Table Notes: The supplemental findings presented in this table are additional findings from the studies in this report that do not factor into the determination of the intervention rating. Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention. nr = not reported. na = not applicable.

^a This appendix presents findings from August et al. (2001) after two years of implementation. August et al. (2001) used two-level mixed regression models, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. There was no significant variation between schools in intercepts and slopes on any of the variables in the social outcomes domain, so a three-level model (to account for nesting within schools) was not used. Outcomes reported here represent the effect of intervention group x time. Effect sizes were not reported for nonsignificant findings. All variables are scaled as z-scores, relative to a normative sample in which the mean is 0 and the standard deviation is 1. Means are full-information maximum likelihood estimates provided by the author based on all available data from the full sample. No corrections for clustering or multiple comparisons were needed. The p-values presented here were reported in the original study.

Appendix D.3: Summary of second-year findings for the academic performance domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
August et al., 2001^a								
<i>Academic Competence Composite</i>	Grade 1	20 schools/ 199 students	-0.96 (1.06)	-1.13 (1.02)	0.17	0.26	+10	0.02

Table Notes: The supplemental findings presented in this table are additional findings from the studies in this report that do not factor into the determination of the intervention rating. Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student’s outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student’s percentile rank that can be expected if the student is given the intervention.

^a This appendix presents findings from August et al. (2001) after two years of implementation. August et al. (2001) used two-level mixed regression models, with time points nested within individual participants; gender and severity of initial aggression were included as time-invariant covariates. There was no significant variation between schools in intercepts and slopes on academic performance outcomes, so a three-level model (to account for nesting within schools) was not used. Outcomes reported here represent the effect of intervention group x time. All variables are scaled as z-scores, relative to a normative sample in which the mean is 0 and the standard deviation is 1. Means are full-information maximum likelihood estimates provided by the author based on all available data from the full sample. No corrections for clustering or multiple comparisons were needed. The p-values presented here were reported in the original study.

Endnotes

¹ The descriptive information for this program was obtained from a publicly available source: the program's website (<http://www.med.umn.edu/psychiatry/research/earlyrisers/home.html>, downloaded July 2010). The WWC requests developers to review the program description sections for accuracy from their perspective. The program description was provided to the developer in November 2010, and we incorporated feedback from the developer. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review. The literature search reflects documents publicly available by August 2011.

² The studies in this report were reviewed using WWC Evidence Standards, version 2.1, as described in the Children Classified as Having an Emotional Disturbance review protocol version 2.0. The evidence presented in this report is based on available research and focuses on results that fall in domains that are described in the Children Classified as Having an Emotional Disturbance review protocol. Findings and conclusions may change as new research becomes available.

³ For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 24. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies. The two studies that meet WWC evidence standards did not examine the effectiveness of *Early Risers* on children classified with an emotional disturbance in the reading achievement/literacy, math achievement, or school attendance domains.

⁴ Outcome data were not always provided for all 190 students in August et al. (2003), so sample sizes vary slightly for each variable and domain. The student sample sizes listed throughout this report for August et al. (2003) are based on the outcome with the largest sample size within each domain.

⁵ Cost information was obtained from the Substance Abuse and Mental Health Services Administration National Registry of Evidence-based Programs and Practices: <http://www.nrepp.samhsa.gov/ViewIntervention.aspx?id=137>.

⁶ The contrasts between children who received three years of *Early Risers* and children in the comparison group (August et al., 2002) are presented in Appendices C.1 to C.4 and form the basis of the intervention rating; the comparison after three years was chosen as the basis of the intervention rating because authors reported that differences had not yet emerged on several targeted outcomes after two years of intervention. Comparisons on this same sample of students after two years of implementation (August et al., 2001) are presented in Appendices D.1 to D.3 and do not contribute to the intervention rating. In later reports of the same students, impacts were assessed after four years (August et al., 2003; Endicott, 2003) and six years (Bernat, August, Hektner, & Bloomquist, 2007) of implementation. Findings from August et al. (2003), Endicott (2003), and Bernat et al. (2007) do not meet WWC evidence standards and thus are not presented in this report.

⁷ There were no significant group differences between the two *Early Risers* conditions; August et al. (2003) interpret this to mean that the Family Support component did not yield an effect over and above the Child Skills component. August et al. (2003) also state that participation in Family Support (or FLEX) was low, and families that did use the component were "high stress" and focused on basic needs, such as child care and housing, in lieu of program features, such as self-sufficiency and empowerment. Thus, August (who is also the primary developer of the intervention) and colleagues elected to collapse the two intervention conditions in the analyses, yielding a single intervention group. On the basis that the developer stated that Family Support was underutilized and judged it appropriate to collapse the two conditions, the WWC elected to report the combined *Early Risers* intervention group analyses presented in the article.

⁸ August et al. (2004) followed this sample of students and measured outcomes again in Year 3, after one year of no *Early Risers* participation; no significant differences between students in the combined intervention group and students in the comparison group were found. The rate of attrition was 14% during this follow-up year.

⁹ Achenbach, T. M. (1991). *Manual for the teacher's report form and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.

¹⁰ A total of 96 eligible children were dropped from the study after screening and random assignment of schools because parents could not be contacted or refused the invitation to participate. The research team informed parents of their child's treatment condition when seeking permission to participate in the study. After excluding these 96 children without parental consent, 245 children remained in the initial study sample.

¹¹ In later reports of the same students, impacts were assessed after four years (August et al., 2003; Endicott, 2003) and six years (Bernat et al., 2007) of implementation. Findings from August et al. (2003) and Bernat et al. (2007) do not meet WWC evidence standards and are not presented in this report due to high levels of attrition and large baseline differences on student aggression. Impact findings on peer rejection and aggression outcomes from Endicott (2003) do not meet WWC evidence standards and are not presented in this report as a result of high levels of attrition and large baseline differences. Findings on academic and social competence from Endicott (2003) are not presented in this report because these outcomes were treated as mediators (or protective factors) in the statistical models, as opposed to targeted outcomes of *Early Risers*.

¹² The contrasts between children who received three years of *Early Risers* and children in the comparison group (August et al., 2002) are presented in Appendices C.1 to C.4 and form the basis of the intervention rating; this comparison after three years was chosen as the basis of the intervention rating because authors reported that differences had not yet emerged on several targeted outcomes after two years of intervention. Comparisons on this same sample of students after two years of implementation (August et al., 2001) are presented in Appendices D.1 to D.3 and do not contribute to the intervention rating.

¹³ There were no significant group differences between the two *Early Risers* conditions; thus, August and colleagues elected to collapse the two intervention conditions in the analyses, yielding a single intervention group. On the basis that the developer stated that Family Support was underutilized and judged it appropriate to collapse the two conditions, the WWC elected to report the combined *Early Risers* intervention group analyses presented in the article.

¹⁴ The School Adjustment measure used in August et al. (2003) captures findings in more than one domain but is best aligned with the external behavior domain.

¹⁵ Harter, S. (1985). *Self-perception profile for children*. Denver, CO: University of Denver.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, June). *Children Classified as Having an Emotional Disturbance intervention report: Early Risers*. Retrieved from <http://whatworks.ed.gov>.

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows 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, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC evidence standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition	Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.
Clustering adjustment	If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.
Confounding factor	A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.
Design	The design of a study is the method by which intervention and comparison groups were assigned.
Domain	A domain is a group of closely related outcomes.
Effect size	The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.
Eligibility	A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.
Equivalence	A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.
Extent of evidence	An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 24.
Improvement index	Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from -50 to +50.
Multiple comparison adjustment	When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.
Quasi-experimental design (QED)	A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.
Randomized controlled trial (RCT)	A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.
Rating of effectiveness	The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 24.
Single-case design	A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.
Standard deviation	The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.
Statistical significance	Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).
Substantively important	A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the [WWC Procedures and Standards Handbook \(version 2.1\)](#) for additional details.