

What Works Clearinghouse



Children Classified as Having an Emotional Disturbance

November 2011

The Incredible Years

Program Description¹

The Incredible Years is composed of training programs for children, parents, and teachers. The child program is designed for children (ages 0–12) with challenging behaviors and focuses on building social and emotional skills. Lessons can be delivered to children referred for difficult behavior or to an entire classroom as a preventative measure. The program consists of 20- to 30-minute lessons two to three times a week; these lessons are reinforced by small-group activities, practicing skills throughout the day, and communicating with parents. Lessons cover recognizing and understanding feelings, getting along with friends, anger management, problem solving, and behavior at school. Parent training programs focus on positive discipline, promoting learning and development, and involvement in children’s life at school.

Research²

One study of *The Incredible Years* that falls within the scope of the Children Classified as Having an Emotional Disturbance review protocol meets What Works Clearinghouse (WWC) evidence standards. This one study included 51 four- to eight-year-old children with oppositional defiant disorder who attended school in Washington state. Based on this one study, the WWC considers the extent of evidence for *The Incredible Years* on children classified as having an emotional disturbance (or children at risk for classification) to be small for the external behavior and social outcomes domains.

Effectiveness

The Incredible Years was found to have potentially positive effects on external behavior and potentially positive effects on social outcomes for children classified as having an emotional disturbance.

Report Contents	
Overview	p. 1
Program Information	p. 2
Research Summary	p. 3
Effectiveness Summary	p. 4
References	p. 5
Research Details for Each Study	p. 12
Outcomes Measures for Each Domain	p. 14
Findings Included in the Rating for Each Outcome Domain	p. 15
Partial Implementation Findings for Each Outcome Domain	p. 17
Rating Criteria	p. 18
Endnotes	p. 19
Glossary of Terms	p. 20

Table 1. Summary of findings³

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of Studies	Number of Students	Extent of evidence
		Average	Range			
External behavior	Potentially positive effects	+20	+18 to +22	1	51	Small
Social outcomes	Potentially positive effects	+18	na	1	51	Small

na = not applicable

Program Information

Background

The Incredible Years was developed by Carolyn Webster-Stratton, a professor and director of the Parenting Clinic at the University of Washington, and is distributed by The Incredible Years. Address: 1411 8th Avenue West, Seattle, WA 98119. Email: incredibleyears@incredibleyears.com. Web: <http://www.incredibleyears.com/>. Telephone: (888) 506-3562 or (206) 285-7565. Fax: (888) 506-3562.

Program details

The Incredible Years is designed for children (ages 0–12) with challenging behaviors and their teachers and parents. The program has been used with children diagnosed with conduct problems (e.g., having high rates of aggression, defiance, and oppositional and impulsive behaviors) and attention deficit disorder. It also has been used with culturally diverse groups, including Hispanic/Latino, Asian-American, African-American, and new immigrant families.

The Incredible Years is composed of training programs for children, parents, and teachers. The child program focuses on building social and emotional skills during preschool through early elementary school and can be delivered to children referred for difficult behavior or to an entire classroom as a preventive measure. In a classroom setting, the teacher presents 20- to 30-minute lessons two to three times a week during circle time, which are reinforced by small-group activities; encouragement of skills practice throughout the day; and communication with parents, including home activities in which they can participate. Lessons cover recognizing and understanding feelings, getting along with friends, anger management, problem solving, and behavior at school. In a clinical setting, referred students complete activities during 18 to 20 two-hour weekly small-group meetings. Ideally, parents of children in the clinical program are involved in the parent training module. Parent training programs focus on parenting skills, including positive discipline, promoting learning and development, and ways to be involved in their children's lives at school. *The Incredible Years* also includes two programs for teachers. The first addresses classroom management as a means to improve student behavior and learning. The second is a training program for teachers who will deliver the program in their classrooms.

Cost⁴

Total per-child costs range from \$1,164 to \$3,003, depending on which components are used.

Research Summary

Seventy-seven studies reviewed by the WWC investigated the effects of *The Incredible Years* on children classified as having an emotional disturbance (or children at risk for classification). One study (Webster-Stratton, Reid, & Hammond, 2004) is a randomized controlled trial that meets WWC evidence standards. The remaining 76 studies do not meet either WWC eligibility screens or evidence standards. (See references beginning on page 5 for citations for all 77 studies.)

Table 2. Scope of reviewed research

Grade	PK, K, 1, 2
Delivery method	Small group, Individual
Program type	Curriculum, Supplement
Studies reviewed	77
Meets WWC standards	1 study
Meets WWC standards with reservations	0 studies

Summary of studies meeting WWC evidence standards without reservations

Webster-Stratton et al. (2004) randomly assigned 159 families to one of six conditions:

- Parent training alone (PT)
- Child training alone (CT)
- Parent training plus teacher training (PT + TT)
- Child training plus teacher training (CT + TT)
- Parent and child training combined with teacher training (PT + CT + TT)
- Wait-list comparison group

Twenty-five students received the full version of *The Incredible Years* (PT + CT + TT), and 26 students were in the comparison group.⁵ Participants were recruited from families requesting treatment at the University of Washington Parenting Clinic. The primary referral problem was child misconduct (e.g., noncompliance, aggression, oppositional behaviors) that had been occurring for at least six months. Families entered the study in three cohorts (50 to 55 families per cohort) in the fall of 1995, 1996, and 1997. Random assignment was conducted by lottery after all families in the cohort had completed baseline assessments. The sample included students in preschool, kindergarten, first grade, and second grade. The study reported outcomes after approximately six months (mid-November to April) of program implementation.⁶

Summary of studies meeting WWC evidence standards with reservations

No studies of *The Incredible Years* met 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 other academic performance. The one study that influences the findings in this report covers two domains: external behavior and social outcomes.⁷ The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *The Incredible Years* on children classified as having an emotional disturbance. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see Appendix E.

Summary of effectiveness for the external behavior domain

Webster-Stratton et al. (2004) found, and the WWC confirmed, three statistically significant differences between the treatment and comparison group on Child Conduct Problems (CCP). These measures are (1) CCP at Home: Mother Report, (2) CCP at Home: Father Report, and (3) CCP at School. Thus, for the external behavior domain, this one study with a strong design showed statistically significant, potentially positive effects.⁸ This results in a rating of potentially positive effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the external behavior 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>The Incredible Years</i> had one study showing a statistically significant positive effect, no studies showing a statistically significant or substantively important negative effect, and no studies showing indeterminate effects.
Extent of evidence	Criteria met
Small	The review of <i>The Incredible Years</i> had a) one study, AND b) an unknown number of schools, AND c) 51 students.

Summary of effectiveness for the social outcomes domain

Webster-Stratton et al. (2004) found, and the WWC confirmed, one statistically significant difference between the treatment and comparison groups on Child Social Competence (CSC) with Peers. Thus, for the social outcomes domain, this one study with a strong design showed a statistically significant, potentially positive effect. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 4. 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>The Incredible Years</i> had one study showing a statistically significant positive effect, no studies showing a statistically significant or substantively important negative effect, and no studies showing indeterminate effects.
Extent of evidence	Criteria met
Small	The review of <i>The Incredible Years</i> had a) one study, AND b) an unknown number of schools, AND c) 51 students.

References

Study that meets WWC evidence standards without reservations

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology, 33*(1), 105–124.

Additional source:

Reid, M. J., Webster-Stratton, C., & Hammond, M. (2003). Follow-up of children who received The Incredible Years intervention for oppositional-defiant disorder: Maintenance and prediction of 2-year outcome. *Behavior Therapy, 34*(4), 471–491.

Studies that do not meet WWC evidence standards

Reid, M. J., Webster-Stratton, C., & Hammond, M. (2007). Enhancing a classroom social competence and problem-solving curriculum by offering parent training to families of moderate- to high-risk elementary school children. *Journal of Clinical Child and Adolescent Psychology, 36*(4), 605–620. The study does not meet WWC evidence standards because it is a randomized controlled trial in which the combination of overall and differential attrition rates exceeds WWC standards for this area, and the subsequent analytic intervention and comparison groups are not shown to be equivalent.

Webster-Stratton, C., Reid, M. J., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of The Incredible Years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry, 49*(5), 471–488. The study does not meet WWC evidence standards because it is a randomized controlled trial in which the combination of overall and differential attrition rates exceeds WWC standards for this area, and the subsequent analytic intervention and comparison groups are not shown to be equivalent.

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

Barrera, M. J., Biglan, A., Taylor, T. K., Gunn, B. K., Smolkowski, K., Black, C., . . . Fowler, R. C. (2002). Early elementary school intervention to reduce conduct problems: A randomized trial with Hispanic and non-Hispanic children. *Prevention Science, 3*(2), 83–94. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample includes less than 50% students at-risk for emotional disturbance or classified as emotionally disturbed.

Beauchaine, T. P., Webster-Stratton, C., & Reid, M. J. (2005). Mediators, moderators, and predictors of 1-year outcomes among children treated for early-onset conduct problems: A latent growth curve analysis. *Journal of Consulting & Clinical Psychology, 73*(3), 371–388. 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.

Borden, L. A., Schultz, T. R., Herman, K. C., & Brooks, C. M. (2010). The Incredible Years parent training program: Promoting resilience through evidence-based prevention groups. *Group Dynamics, 14*(3), 230–241. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Brotman, L. M., Gouley, K. K., Chesir-Teran, D., Dennis, T., Klein, R., & Shrout, P. (2005). Prevention for preschoolers at high risk for conduct problems: Immediate outcomes on parenting practices and child social competence. *Journal of Clinical Child and Adolescent Psychology, 34*(4), 724–734. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

Brotman, L. M., Klein, R. G., Kamboukos, D., Brown, E. J., Coard, S., & Sosinsky, L. S. (2003). Preventive intervention for urban, low-income preschoolers at familial risk for conduct problems: A randomized pilot study. *Journal of Child Psychology & Psychiatry, 32*(2), 246–257. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

- Brotman, L. M., O'Neal, C. R., Huang, K.-Y., Gouley, K. K., Rosenfelt, A., & Shrout, P. E. (2009). An experimental test of parenting practices as a mediator of early childhood physical aggression. *Journal of Child Psychology & Psychiatry*, 50(3), 235–245. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Carlson, J., Tired, H., Bender, S., & Benson, L. (2011). The influence of group training in the Incredible Years teacher classroom management program on preschool teachers' classroom management strategies. *Journal of Applied School Psychology*, 27(2), 134–154. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Chang, M., Park, B., & Kim, S. (2009). Parenting classes, parenting behavior, and child cognitive development in Early Head Start: A longitudinal model. *School Community Journal*, 19(1), 155–174. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Daley, D., Jones, K., Hutchings, J., & Thompson, M. (2009). Attention deficit hyperactivity disorder in pre-school children: Current findings, recommended interventions and future directions. *Child: Care, Health & Development*, 35(6), 754–766. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Dionne, R., Davis, B., Sheeber, L., & Madrigal, L. (2009). Initial evaluation of a cultural approach to implementation of evidence-based parenting interventions in American Indian communities. *Journal of Community Psychology*, 37(7), 911–921. 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.
- Emond, S. R. (2008). School readiness and delayed entry: The effect of parent training on perceived school readiness. *Dissertation Abstracts International*, 69(12B), 125-7853. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Foster, E. M., Olchowski, A. E., & Webster-Stratton, C. (2007). Is stacking intervention components cost-effective? An analysis of The Incredible Years program. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(11), 1414–1424. 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.
- Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *Journal of Clinical Child & Adolescent Psychology*, 39(4), 568–580. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Gross, D., Fogg, L., Webster-Stratton, C., Garvet, C., Julian, W., & Grady, J. (2003). Parent training with families of toddlers in day care in low-income urban communities. *Journal of Consulting and Clinical Psychology*, 71(2), 261–278. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Hartman, R. R., Stage, S., & Webster-Stratton, C. (2003). A growth curve analysis of parent training outcomes: Examining the influence of child factors (inattention, impulsivity, and hyperactivity problems), parental and family risk factors. *The Child Psychology and Psychiatry Journal*, 44(3), 388–398. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Jones, K., Daley, D., Hutchings, J., Bywater, T., & Eames, C. (2007). Efficacy of The Incredible Years basic parent training programme as an early intervention for children with conduct problems and ADHD. *Child: Care, Health & Development*, 33(6), 749–756. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Kennedy, M. D. (2004). Implementation of The Incredible Years model program in three Colorado communities: A case study. *Dissertation Abstracts International*, 65(08A), 250-3157. The study is ineligible for review because it does not use a comparison group design or a single-case design.

- Kim, E., Cain, K. C., & Webster-Stratton, C. (2008). The preliminary effect of a parenting program for Korean American mothers: A randomized controlled experimental study. *International Journal of Nursing Studies*, 45(9), 1261–1273. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Lau, A. S., Fung, J. J., & Yung, V. (2010). Group parent training with immigrant Chinese families: Enhancing engagement and augmenting skills training. *Journal of Clinical Psychology*, 66(8), 880–894. 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.
- Lavigne, J. V., Lebaillly, S. A., Gouze, K. R., Cicchetti, C., Pochyly, J., Arend, R., . . . Binns, H. J. (2008). Treating oppositional defiant disorder in primary care: A comparison of three models. *Journal of Pediatric Psychology*, 33(5), 449–461. 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.
- Letarte, M.-J., Normandeau, S., & Allard, J. (2010). Effectiveness of a parent training program “Incredible Years” in a child protection service. *Child Abuse & Neglect*, 34(4), 253–261. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample includes less than 50% students at-risk for emotional disturbance or classified as emotionally disturbed.
- Levac, A. M., McCay, E., Merka, P., & Reddon-D’Arcy, M. L. (2008). Exploring parent participation in a parent training program for children’s aggression: Understanding and illuminating mechanisms of change. *Journal of Child & Adolescent Psychiatric Nursing*, 21(2), 78–88. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Linares, L. O., Montalto, D., Li, M., & Oza, V. S. (2006). A promising parenting intervention in foster care. *Journal of Consulting and Clinical Psychology*, 74(1), 32–41. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Marcynyszyn, L. A., Maher, E. J., & Corwin, T. W. (2011). Getting with the (evidence-based) program: An evaluation of the Incredible Years parenting training program in child welfare. *Children and Youth Services Review*, 33(5), 747–757. The study is ineligible for review because it does not include an outcome within a domain specified in the protocol.
- McIntyre, L. L. (2008). Adapting Webster-Stratton’s Incredible Years parent training for children with developmental delay: Findings from a treatment group only study. *Journal of Intellectual Disability Research*, 52(12), 1176–1192. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- McIntyre, L. L. (2008). Parent training for young children with developmental disabilities: Randomized controlled trial. *American Journal on Mental Retardation*, 113(5), 356–368. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- McMenamy, J., Sheldrick, R. C., & Perrin, E. C. (2011). Early intervention in pediatrics offices for emerging disruptive behavior in toddlers. *Journal of Pediatric Healthcare*, 25(2), 77–86. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Melhuish, E. C. (2007). Parenting training improves problem behaviour in children at risk of conduct disorder. *Evidence-Based Mental Health*, 10(4), 125. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Miller, L. S., & Rojas-Flores, L. (1999). *Preventing conduct problems in urban, Latino preschoolers through parent training: A pilot study*. New York: New York University Child Study Center. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Neill, T. K. (2006). *Helping others help children: Clinical supervision of child psychotherapy*. Washington, DC: American Psychological Association. 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.

- 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.
- Nilsen, W. (2007). Fostering futures: A preventive intervention program for school-age children in foster care. *Clinical Child Psychology and Psychiatry, 12*(1), 45–63. 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.
- Olchowski, A. E., Foster, E. M., & Webster-Stratton, C. (2007). Implementing behavioral intervention components in a cost-effective manner: Analysis of The Incredible Years program. *Journal of Early & Intensive Behavior Intervention, 3–4*(4–1), 284–304. 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.
- Pearl, E. (2009). Parent management training for reducing oppositional and aggressive behavior in preschoolers. *Aggression and Violent Behavior, 14*(5), 295–305. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Phaneuf, L., & McIntyre, L. L. (2007). Effects of individualized video feedback combined with group parent training on inappropriate maternal behavior. *Journal of Applied Behavior Analysis, 40*(4), 737–741. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Piquero, A. R., Farrington, D. P., Welsh, B. C., Tremblay, R., & Jennings, W. G. (2009). Effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology, 5*(2), 83–120. 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.
- Powell, D., & Dunlap, G. (2009). *Evidence-based social-emotional curricula and intervention packages for children 0–5 years and their families (Roadmap to Effective Intervention Practices)*. Tampa: University of South Florida, Technical Assistance Center on Social Emotional Intervention for Young Children. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly, 23*, 10–26. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Reid, M. J., & Webster-Stratton, C. (2001). The Incredible Years parent, teacher, and child intervention: Targeting multiple areas of risk for a young child with pervasive conduct problems using a flexible, manualized treatment program. *Cognitive and Behavioral Practice, 8*(4), 377–386. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Reid, M. J., Webster-Stratton, C., & Baydar, N. (2004). Halting the development of conduct problems in Head Start children: The effects of parent training. *Journal of Clinical Child and Adolescent Psychology, Division 53, 33*(2), 279–291. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Reid, M. J., Webster-Stratton, C., & Beauchaine, T. P. (2001). Parent training in Head Start: A comparison of program response among African American, Asian American, Caucasian, and Hispanic mothers. *Prevention Science, 2*(4), 209–227. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Rinaldi, J. (2001). Long-term outcomes of parent training and predictors of adolescent adjustment. *Dissertation Abstracts International, 62*(5), 2498. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

- Ritch, C. M. (2005). Evaluation of a parent training intervention: Parent and teacher assessment of perceived changes in children's at-risk behaviors. *Dissertation Abstracts International*, 66(06B), 91-3395. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Rogers, R. D. (2007). Evaluating the effectiveness of The Incredible Years parenting education and support program. *Dissertation Abstracts International*, 68(08B), 72-5591. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Schmidt, F., & Taylor, T. K. (2002). Putting empirically supported treatments into practice: Lessons learned in a children's mental health center. *Professional Psychology: Research and Practice*, 33(5), 483-489. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Shernoff, E. S. (2004). Transporting an evidence-based classroom management program for preschoolers with disruptive behavior problems to a school: An analysis of implementation, outcomes, and contextual variables. *Dissertation Abstracts International*, 65(11A), 223-4106. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Shernoff, E. S., & Kratochwill, T. R. (2007). Transporting an evidence-based classroom management program for preschoolers with disruptive behavior problems to a school: An analysis of implementation, outcomes, and contextual variables. *School Psychology Quarterly*, 22(3), 449-472. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Shriver, M. D., & Allen, K. D. (2008). *Working with parents of noncompliant children: A guide to evidence-based parent training for practitioners and students*. Washington, DC: American Psychological Association. 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.
- Sougstad, J. R. (2010). Transforming everyday practices using scientific evidence: Meta-analysis of a parent training program. *Dissertation Abstracts International*, 72(08). 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.
- Spaccarelli, S., Cotler, S., & Penman, D. (1992). Problem-solving skills training as a supplement to behavioral parent training. *Cognitive Therapy and Research*, 16, 1-18. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Steiman, M. (2005). Parent training with children with conduct problems: The role of the marital relationship and parental adjustment. *Dissertation Abstracts International*, 65(7-B), 3727. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Stewart, L. S., & Carlson, J. S. (2010). Investigating parental acceptability of The Incredible Years self-administered parent training program for children presenting externalizing behavior problems. *Journal of Applied School Psychology*, 26(2), 162-175. The study is ineligible for review because it does not include a student outcome.
- Taylor, T. K., Webster-Stratton, C., Feil, E. G., Broadbent, B., Widdop, C. S., & Sevenson, H. H. (2008). Computer-based intervention with coaching: An example using The Incredible Years program. *Cognitive Behaviour Therapy*, 37(4), 233-246. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Walcott, C. M., Carlson, J. S., & Beamon, H. L. (2009). Effectiveness of a self-administered training program for parents of children with ADHD. *School Psychology Forum*, 3(1), 43-61. 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.
- Webster-Stratton, C. (1990). Enhancing the effectiveness of self-administered videotape parent training for families with conduct-problem children. *Journal of Abnormal Child Psychology*, 18(5), 479-492. 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.
- Webster-Stratton, C. (1992). Individually administered videotape parent training: "Who benefits?" *Cognitive Therapy and Research*, 16(1), 31-35. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.

- Webster-Stratton, C. (1994). Advancing videotape parent training: A comparison study. *Journal of Consulting and Clinical Psychology, 62*(3), 583–593. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample is not within the specified age or grade range.
- Webster-Stratton, C. (1998). Preventing conduct problems in Head Start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology, 66*(5), 715–730. The study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol.
- Webster-Stratton, C. (2001). The Incredible Years: Parents, teachers, and children training series. *Residential Treatment for Children & Youth, 18*(3), 31–45. 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.
- Webster-Stratton, C. (2005). The Incredible Years: A training series for the prevention and treatment of conduct problems in young children. In E. D. Hibbs & P. S. Jensen (Eds.), *Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice* (2nd ed.; pp. 507–555). Washington, DC: American Psychological Association. 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.
- Webster-Stratton, C., & Hammond, M. (1990). Predictors of treatment outcome in parent training for families with conduct problem children. *Behavior Therapy, 21*, 319–337. The study is ineligible for review because it does not use a comparison group design or a single-case design.
- Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology, 65*, 93. 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.
- Webster-Stratton, C., & Herman, K. C. (2008). The impact of parent behavior-management training on child depressive symptoms. *Journal of Counseling Psychology, 55*(4), 473–484. 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.
- Webster-Stratton, C., & Herman, K. C. (2010). Disseminating Incredible Years series early-intervention programs: Integrating and sustaining services between school and home. *Psychology in the Schools, 47*(1), 36. 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.
- Webster-Stratton, C., & Reid, M. J. (2003) The Incredible Years parents, teachers, and children training series: A multifaceted treatment approach for young children with conduct problems. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 224–240). 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.
- Webster-Stratton, C., & Reid, M. J. (2003). Treating conduct problems and strengthening social and emotional competence in young children: The Dina Dinosaur treatment program. *Journal of Emotional and Behavioral Disorders, 11*(3), 130–143. 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.
- Webster-Stratton, C., & Reid, M. J. (2010). A school-family partnership: Addressing multiple risk factors to improve school readiness and prevent conduct problems in young children. In S. L. Christenson & A. L. Reschly (Eds.), *Handbook on school-family partnerships* (pp. 204–227). New York: Routledge/Taylor and Francis. 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.
- Webster-Stratton, C., & Reid, M. J. (2010). The Incredible Years parents, teachers, and children training series: A multifaceted treatment approach for young children with conduct problems. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2nd ed.; pp. 194–210). 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.

Webster-Stratton, C., Hollinsworth, T., & Kolpacoff, M. (1989). The long-term effectiveness and clinical significance of three cost-effective training programs for families with conduct-problem children. *Journal of Consulting and Clinical Psychology, 57*(4), 550–553. 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.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in Head Start. *Journal of Clinical Child Psychology, 30*(3), 283–302. The study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol.

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Social skills and problem-solving training for children with early-onset conduct problems: Who benefits? *Journal of Child Psychology and Psychiatry, 42*(7), 943–952. 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.

Webster-Stratton, C., Rinaldi, J., & Reid, J. M. (2011). Long-term outcomes of Incredible Years parenting program: Predictors of adolescent adjustment. *Child and Adolescent Mental Health, 16*(1), 38–46. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Webster-Stratton, C. H., & Reid, M. J. (2010). The Incredible Years program for children from infancy to pre-adolescence: Prevention and treatment of behavior problems. In R. Murrihy, A. Kidman, & T. Ollendick (Eds.), *Clinician's handbook for the assessment and treatment of conduct problems in youth* (pp. 117–138). New York: Springer 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.

Webster-Stratton, C. H., Reid, M. J., & Beauchaine, T. (2011). Combining parent and child training for young children with ADHD. *Journal of Clinical Child and Adolescent Psychology, 40*(2), 191–203. The study is ineligible for review because it does not use a sample aligned with the protocol—the sample includes less than 50% students at-risk for emotional disturbance or classified as emotionally disturbed.

Appendix A: Research details for Webster-Stratton et al., 2004

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology*, 33(1), 105–124.

Table A1. Summary of findings

Meets WWC evidence standards

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
External behavior	51 students	+20	Yes
Social outcomes	51 students	+18	Yes

Setting The teacher, parent, and child trainings were primarily instituted in the University of Washington Parenting Clinic. The children assigned to treatment conditions came to the clinic’s “Dinosaur School.” Participants then practiced their skills at home and in the classroom.

Study sample A sample of 159 families was randomly assigned to one of six conditions: parent training alone (PT; n = 31); child training alone (CT; n = 30); parent training plus teacher training (PT + TT; n = 24); child training plus teacher training (CT + TT; n = 23); parent and child training combined with teacher training (PT + CT + TT; n = 25); or a wait-list comparison group (n = 26). The final sample sizes for each of the six conditions differ by outcome measure. Participants were recruited from families requesting treatment at the University of Washington Parenting Clinic. Families were self-referred or referred by professionals in the community (20% by teachers and 38% by physicians). The primary referral problem was child misconduct (e.g., noncompliance, aggression, oppositional behaviors) that had been occurring for at least six months. Families entered the study in three cohorts (50 to 55 families per cohort) in the fall of 1995, 1996, and 1997. Random assignment was conducted by lottery after all families in the cohort had completed baseline assessments. The student sample was predominantly European American (79%), 90% were boys, and the mean age was 71 months. The sample consisted of students in preschool, kindergarten, first grade, and second grade.

Intervention group The children assigned to CT, CT + TT, and CT + PT + TT conditions came to the clinic’s Dinosaur School for 2 hours each week for 18 to 19 weeks (lasting approximately six months) and met with two therapists. The Dinosaur School program specifically addressed interpersonal difficulties that are problematic for young children with oppositional defiant disorder (ODD). Weekly letters were sent to teachers and parents explaining the key concepts and the rationale for the targeted skill (e.g., sharing, teamwork, friendly talk, listening, compliance to requests, feeling talk, and problem solving). Teachers and parents were asked to reinforce the targeted social skills whenever they noticed the child using them in the home or school, and children were given weekly homework assignments to complete with their parents. The parents assigned to PT, PT + TT, and PT + CT + TT conditions met at the clinic each week for a 2-hour session. Over the course of 22 to 24 weeks, they watched 17 videotape programs on parenting and interpersonal skills designed to reduce parents’ coercive interactions and strengthen positive interactions and relationships with their

children. Teachers in the PT + TT, CT + TT, and PT + CT + TT conditions came to the clinic for 4 full days (32 hours) of group training sequenced throughout the school year, to correspond roughly with the beginning, first quarter, second quarter, and end of the PT and CT treatments.

Comparison group

The families assigned to the comparison condition received no treatment from the Parenting Clinic and had no contact with therapists during the 8- to 9-month waiting period. These families were offered the parent training program after the outcomes from the first year of the study had been measured.

Outcomes and measurement

This study included measures of Child Conduct Problems (CCP) at Home, Child Conduct Problems (CCP) at School, and Child Social Competence (CSC) with Peers. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

The teacher curriculum targeted teachers' use of effective classroom management strategies for handling misbehavior, promoting positive relationships with difficult students, and strengthening social skills in all school settings. Workshop topics included promoting social skills through praise and encouragement, proactive teaching, using incentives to motivate children, strategies to decrease disruptive behavior, and collaborative approaches for working with parents. Teachers also learned to prevent peer rejection by helping the aggressive child learn appropriate problem-solving strategies and by helping his or her peers respond appropriately to aggression. Teachers were trained to have age-appropriate expectations and to be sensitive to individual developmental differences and biological deficits in children, and to understand the relevance of these differences for enhanced teaching efforts that are positive, accepting, and consistent. To ensure the integrity of the treatment, therapists co-led their first parent or child group with a supervisor, completed a weekly checklist of standards, and were monitored weekly. All child and parent sessions were videotaped for feedback and analyses, and the supervisor randomly selected videotapes for fidelity checks. Analyses indicated that all required videotape vignettes were shown and that all required homework was assigned.

Appendix B: Outcome measures for each domain

External behavior	
<i>Child Conduct Problems (CCP) at Home: Mother Report</i>	This composite score includes the Eyberg Child Behavior Inventory (ECBI) Total Intensity score and four independent observations of child behavior in the home using the Dyadic Parent-Child Interaction Coding System–Revised (DPICS-R) and the Coder Impressions Inventory (CII). The ECBI is a 36-item behavioral inventory of child conduct problem behavior for children ages 2 to 16 and was completed by the mother; the ECBI Total Intensity score was used as an indicator of the frequency with which problem behaviors occur. The DPICS-R, originally developed by Robinson and Eyberg (1981) and revised by Webster-Stratton (1989), is an observational measure for recording behaviors of children in their home; this composite score included the DPICS-R total child deviance (sum of whine, cry, physical negative, smart talk, yell, and destructive) plus noncompliance (child does not begin to comply to parent command within 5 seconds) variables and a one-item rating of child affect. This composite score also included two single-item variables from the CII (percentage of time child acted inappropriately and total overall poor conduct) (as cited in Webster-Stratton et al., 2004).
<i>Child Conduct Problems (CCP) at Home: Father Report</i>	This composite score was identical to the CCP at Home: Mother Report with one exception—the ECBI Total Intensity score was completed by the father for the CCP at Home: Father Report (as cited in Webster-Stratton et al., 2004).
<i>Child Conduct Problems (CCP) at School</i>	This composite score includes two teacher report variables: Teacher Assessment of School Behavior (TASB; aggressive behavior scale) and the Teacher Rating scales of Perceived Competence Scale for Young Children (PCSC; behavior conduct score). The composite score also includes two summary scores from independent observations in the classroom using Multi-Option Observation System for Experimental Studies (MOOSSES) (frequency of child negative behaviors with teachers and peers) and observation ratings of poor authority acceptance from the Social Health Profile (SHP). The composite score also includes the Dyadic Peer Interaction Scale (DPIS), a measure of total inappropriate behavior with peers (e.g., dyad was loud, physically active, impulsive, reckless), from independent observations of the child interacting with a peer in a laboratory (as cited in Webster-Stratton et al., 2004).
Social outcomes	
<i>Child Social Competence (CSC) with Peers</i>	This composite score includes two teacher-report variables: the social acceptance scores from the Teacher Assessment of School Behavior (TASB) and from the Perceived Competence Scale for Young Children (PCSC). This composite score also includes one classroom observation variable, the social contact score on the Social Health Profile (SHP), and one laboratory observation variable, the positive communication score on the Dyadic Peer Interaction Scale (DPIS) (as cited in Webster-Stratton et al., 2004).

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	
Webster-Stratton et al., 2004^a								
<i>Child Conduct Problems (CCP) at Home: Mother Report</i>	Ages 4–8 (PT + TT + CT)	51	36.99 (10.46)	47.28 (9.79)	10.29	0.57	22	< 0.05
<i>Child Conduct Problems (CCP) at School</i>	Ages 4–8 (PT + TT + CT)	51	29.58 (14.23)	39.55 (18.69)	9.97	0.46	18	< 0.05
Domain average for external behavior across one study						0.52	20	Statistically significant

Table Notes: This appendix reports findings considered for the effectiveness rating and the average improvement indices for the external behavior domain. Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. For the CCP at Home: Mother Report and CCP at School, signs were reversed on the mean difference, effect size, and improvement index to demonstrate that the treatment group was favored when negative differences were reported. 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 that 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; the study is characterized as having a statistically significant positive effect because 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. PT = Parent Training; TT = Teacher Training; CT = Child Training.

^a Findings from contrasts between the comparison group and other treatment conditions (PT, CT, PT + TT, and CT + TT) from Webster-Stratton et al. (2004) are not included in these ratings but are reported in Appendix D.1. The means reported here were adjusted for pretest scores by the study authors. The p-value ranges presented here were reported in the original study (Reid et al., 2003). A correction for multiple comparisons was needed but did not affect significance levels. The original study reported findings on three external behavior outcomes; the father-reported outcomes are not presented in this appendix due to high attrition. The WWC multiple comparison correction accounts for the fact that three measures were used in the article to measure external behavior.

Appendix C.2: 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	
Webster-Stratton et al., 2004^a								
<i>Child Social Competence (CSC) with Peers</i>	Ages 4–8 (PT + TT + CT)	51	53.15 (12.77)	44.13 (15.41)	9.02	0.46	18	< 0.05
Domain average for social outcomes across one study						0.46	18	Statistically significant

Table Notes: This appendix reports findings considered for the effectiveness rating and the average improvement indices for the social outcomes domain. 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 that 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. PT = Parent Training; TT = Teacher Training; CT = Child Training.

^a Findings from contrasts between the comparison group and other treatment conditions (PT, CT, PT + TT, and CT + TT) from Webster-Stratton et al. (2004) are not included in these ratings but are reported in Appendix D.1. The means reported here were adjusted for pretest scores by the study authors. The p-value range presented here was reported in the original study (Reid et al., 2003). No corrections for clustering or multiple comparisons were needed.

Appendix D.1: Partial implementation 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	
Webster-Stratton et al., 2004^a								
<i>Child Conduct Problems (CCP) at Home: Mother Report</i>	Ages 4–8 (PT)	57	35.96 (11.47)	47.28 (9.79)	11.32	0.67	25	< 0.05
<i>Child Conduct Problems (CCP) at Home: Mother Report</i>	Ages 4–8 (PT + TT)	50	39.99 (12.22)	47.28 (9.79)	7.29	0.41	16	< 0.05
<i>Child Conduct Problems (CCP) at Home: Mother Report</i>	Ages 4–8 (CT)	55	40.19 (12.19)	47.28 (9.79)	7.09	0.41	16	< 0.05
<i>Child Conduct Problems (CCP) at Home: Mother Report</i>	Ages 4–8 (CT + TT)	49	37.25 (9.83)	47.28 (9.79)	10.03	0.55	21	< 0.05
<i>Child Conduct Problems (CCP) at School</i>	Ages 4–8 (PT)	56	32.50 (14.42)	39.55 (18.69)	7.05	0.35	14	< 0.05
<i>Child Conduct Problems (CCP) at School</i>	Ages 4–8 (PT + TT)	50	31.06 (16.32)	39.55 (18.69)	8.49	0.41	16	< 0.05
<i>Child Conduct Problems (CCP) at School</i>	Ages 4–8 (CT)	56	32.23 (14.73)	39.55 (18.69)	7.32	0.41	16	< 0.05
<i>Child Conduct Problems (CCP) at School</i>	Ages 4–8 (CT + TT)	49	32.49 (15.44)	39.55 (18.69)	7.06	0.41	16	< 0.05

Table Notes: This appendix presents comparisons between children who received partial versions of *The Incredible Years* (PT, CT, PT + TT, and CT + TT) and children in the comparison group on measures that fall in the external behavior domain. These are ancillary comparisons for the purposes of this review. Comparisons between children who received the full version of *The Incredible Years* and children in the comparison group were used for rating purposes and are presented in Appendix C.1. Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. For the CCP at Home: Mother Report and CCP at School, signs were reversed on the mean difference, effect size, and improvement index to demonstrate that the treatment group was favored when negative differences were reported. 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 that 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. PT = Parent Training; TT = Teacher Training; CT = Child Training

^a The means reported here were adjusted for pretest scores by the study authors. The p-value ranges presented here were reported in the original study (Reid et al., 2003). A correction for multiple comparisons was needed and resulted in significance levels that differ from those in the original study. Due to the multiple comparisons adjustment, the p-value for the contrast between the CT and comparison group on CCP at Home: Mother Report was higher than the critical p-value for statistical significance; therefore, the WWC does not find the result to be statistically significant. The p-value for CCP at School was also higher than the critical p-value for statistical significance for all contrasts; therefore, the WWC does not find these results to be statistically significant. The original study reported findings on three external behavior outcomes; the father-reported outcomes are not presented in this appendix due to high attrition. The WWC multiple comparison correction accounts for the fact that three measures were used in the article to measure external behavior.

Appendix D.2: Partial implementation 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	
Webster-Stratton et al., 2004^a								
<i>Child Social Competence (CSC) with Peers</i>	Ages 4–8 (PT)	56	46.12 (15.38)	44.13 (15.41)	1.99	nr	na	> 0.05
<i>Child Social Competence (CSC) with Peers</i>	Ages 4–8 (PT + TT)	49	49.05 (14.81)	44.13 (15.41)	4.92	nr	na	> 0.05
<i>Child Social Competence (CSC) with Peers</i>	Ages 4–8 (CT)	56	50.02 (14.85)	44.13 (15.41)	5.89	0.35	14	< 0.05
<i>Child Social Competence (CSC) with Peers</i>	Ages 4–8 (CT + TT)	49	49.55 (14.20)	44.13 (15.41)	5.42	0.29	11	< 0.10

Table Notes: This appendix presents comparisons between children who received partial versions of *The Incredible Years* (PT, CT, PT + TT, and CT + TT) and children in the comparison group on measures that fall in the social outcomes domain. These are ancillary comparisons for the purposes of this review. Comparisons between children who received the full version of *The Incredible Years* and children in the comparison group were used for rating purposes and are presented in Appendix C.2. 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 that 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. PT = Parent Training; TT = Teacher Training; CT = Child Training; nr = not reported; na = not applicable

^a The means reported here were adjusted for pretest scores by the study authors. The p-value range presented here was reported in the original study (Reid et al., 2003). No corrections for clustering or multiple comparisons were needed.

Appendix E1: Criteria used to determine the rating of a study

Study rating	Criteria
Meets evidence standards	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets 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.

Appendix E2: 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 meets 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 meets 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.

Appendix E3: 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.

Endnotes

¹ The descriptive information for this program was obtained from a publicly available source: the program's website (<http://www.incredibleyears.com/>, downloaded January 2011). The WWC requests developers to review the program description sections for accuracy from their perspective. The program description was provided to the developer in January 2011. 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 2.0, as described in protocol Version 2.0. The evidence presented in this report is based on available research. 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 Appendix E. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies. The one study that meets WWC evidence standards did not examine the effectiveness of *The Incredible Years* on children classified as having an emotional disturbance in the emotional/internal behavior, reading achievement/literacy, math achievement, school attendance, or other academic performance domains.

⁴ Cost information from Foster, E. M., Olchowski, A. E., & Webster-Stratton, C. (2007). Implementing behavioral intervention components in a cost-effective manner: Analysis of The Incredible Years program. *Journal of Early and Intensive Behavioral Intervention*, 4(1), 284–304.

⁵ Webster-Stratton et al. (2004) reported separate comparisons between the wait-list comparison group and each of the treatment conditions. The comparison between children who received the full version of *The Incredible Years* (PT + TT + CT) and children in the comparison group is presented in Appendices C.1 and C.2 and forms the basis of the intervention rating; this comparison was chosen as the basis of the intervention rating because it evaluates the combined effects of *The Incredible Years'* teacher, parent, and child trainings. All other comparisons are presented in Appendices D.1 and D.2 and do not contribute to the intervention rating.

⁶ Students in the treatment conditions were assessed again 1 and 2 years later. Webster-Stratton et al. (2004) and Reid, Webster-Stratton, and Hammond (2003) were not able to report comparisons between children in the treatment and comparison groups during the 1- and 2-year follow-ups because the comparison group students received *The Incredible Years* program during the period between the original posttest and the follow-up measurements. Thus, outcomes from the 1- and 2-year follow-ups are not included in this review.

⁷ Webster-Stratton et al. (2004) also reported outcomes on negative and positive parenting, classroom management, classroom atmosphere, and parent/teacher satisfaction with the program. These outcomes are not presented in this report because they do not fall under a domain specified in the protocol.

⁸ The composite score used to measure fathers' reports on Child Conduct Problems (CCP) at home was identical to the CCP at Home: Mother Report with one exception; the Eyberg Child Behavior Inventory (ECBI) Total Intensity score was completed by the father for the CCP at Home: Father Report. Mothers' reports on CCP are presented in Appendix C.1 and form the basis of the intervention rating. Comparisons based on the fathers' composite score are not included in the intervention rating. Due to high levels of attrition on the fathers' reports, comparisons between the comparison group and the PT + TT + CT, PT, and PT + TT groups meet evidence standards with reservations, and comparisons between the comparison group and the CT and CT + TT groups do not meet evidence standards.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse (2011, November). *Children Classified as Having an Emotional Disturbance intervention report: The Incredible Years*. Retrieved from <http://whatworks.ed.gov>.

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 treatment 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 standardized measure of the magnitude of an effect that is comparable across studies and outcomes.
Eligibility	A study is eligible for review if it falls within the scope of the review protocol and uses a causal design (RCT or QED).
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 Table E3.
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 treatment 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 treatment 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 Table E2.
Standard deviation	The standard deviation across all students in a group shows how dispersed the outcomes are. A measure with a small standard deviation would indicate that participants had more similar outcomes than a measure with a large standard deviation.
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 \(version 2.0\)](#) for additional details.