Pivotal Response Training

Intervention Description

Pivotal response training (PRT) is an intervention designed for children with autism spectrum disorders. This practice focuses on pivotal (core) areas affected by autism, such as communication and responding to environmental stimuli. PRT sessions typically begin with a parent or teacher providing clear instructions to a child, having the child help choose a stimulus (such as a toy), and focusing the child’s attention. The parent or teacher then encourages the desired behavior (for example, asking for the toy or choosing “toy” from a list of words) by providing rewards if the child implements or attempts to implement the desired behavior. Parents and teachers often model the appropriate behavior or use the stimulus with the child. Activities that maintain existing behaviors are interspersed with activities eliciting new behaviors. The complexity of the required responses increases as training progresses. Parents, teachers, and peers collaboratively implement the practice at school, at home, and in the community. PRT can be used with autistic children aged 2–18. PRT is also known as Pivotal Response Therapy, Pivotal Response Treatment®, or Natural Language Paradigm.

Research

The What Works Clearinghouse (WWC) identified two group design studies of PRT that fall within the scope of the Children and Students with an Autism Spectrum Disorder topic area and meet WWC group design standards without reservations (no studies meet WWC group design standards with reservations). These studies included 85 children between the ages of 20 months and 4 years in one academic medical center and two university sites.

The WWC considers the extent of evidence for PRT on children and students with an autism spectrum disorder to be small for one outcome domain—communication/language competencies. There were no studies that meet WWC group design standards with outcomes in the 15 other domains, so this intervention report does not summarize the effectiveness of PRT for those domains. (See the Effectiveness Summary on p. 6 for more details of effectiveness by domain.)

The WWC also reviewed 37 eligible studies against pilot single-case design standards. Three of these studies meet WWC pilot single-case design standards without reservations, and one meets WWC pilot single-case design standards with reservations. The results from single-case design studies only affect the WWC effectiveness rating for an outcome domain if the studies with outcomes in that domain collectively meet a set of threshold criteria, reflecting replication across different studies, research teams, and cases. The evidence from the four single-case design studies does not reach the threshold to include single-case design evidence in the effectiveness ratings for any of the outcome domains in this report. (See the Research Summary on p. 4 for the rationale behind this threshold and a description of the criteria.)
Effectiveness

PRT was found to have no discernible effects on communication/language competencies for children and students with an autism spectrum disorder.

Table 1. Summary of findings

<table>
<thead>
<tr>
<th>Outcome domain</th>
<th>Rating of effectiveness</th>
<th>Improvement Index (percentile points)</th>
<th>Number of studies</th>
<th>Number of students</th>
<th>Extent of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication/language competencies</td>
<td>No discernible effects</td>
<td>1 – 9 to 10</td>
<td>2</td>
<td>85</td>
<td>Small</td>
</tr>
</tbody>
</table>
Intervention Information

Background


Intervention details

To begin a PRT session, a parent or teacher provides instructions that the child can easily understand. The child then chooses the stimulus, or the parent or teacher selects a stimulus that the child has indicated an interest in—for example, by gazing at the stimulus—or that the parent or teacher knows the child likes. The stimulus can be a toy or other object found in the child’s regular environment. Once the child responds appropriately or attempts to respond appropriately, the parent or teacher provides a reward that is related to the desired behavior. For example, a parent or teacher can present a ball to the child and prompt the child to ask “What’s that?” When the child answers the question or attempts to answer the question (for example, saying “ba, ba”), the parent or teacher verbally identifies the item and hands the item to the child for interaction. The stimuli used and the desired behaviors are regularly varied to improve motivation. As the child progresses, the focus of each session changes to accommodate more advanced goals and needs.

To consistently provide PRT throughout the day, parents, teachers, and peers collaborate on implementation, with parents often serving as the primary intervention agents. Implementation should incorporate the family context, such as household routines, and school context, enabling children to participate in a natural setting. At school, PRT does not require a separate curriculum; instead, PRT procedures are integrated into the regular curriculum.

By improving functioning and response in the core or pivotal areas affected by autism, PRT aims to develop other important social and academic skills, creating generalized improvement. PRT can be used to teach functioning in a new area or reinforce or maintain existing functioning, and has been used to target pivotal areas, such as motivation to respond to social and environment stimuli, self-initiation of social interactions, responsiveness to multiple cues, self-management, and empathy. PRT uses both a development approach and applied behavior analysis procedures.

Cost

As of December 2016, Pivotal Response Treatment® training manuals range in price from $9.95 to $20.00 and can be purchased from www.autismprthelp.com. A 2-day introductory training conducted onsite in the school or district is available from Koegel Autism Consultants for about $2,500 per day plus expenses. Training on more advanced techniques—taking 2 or more days—is available for $2,500 per day plus expenses.
Research Summary

The WWC identified three eligible group design studies and 37 eligible single-case design studies that investigated the effects of PRT on children and students with an autism spectrum disorder. An additional 161 studies were identified but do not meet WWC eligibility criteria for review in this topic area. Citations for all 201 studies are in the References section, which begins on p. 7.

The WWC reviewed three eligible studies against group design standards. Two of the eligible group design studies are randomized controlled trials that meet WWC group design standards without reservations. These studies are summarized in this report. The other eligible group design study does not meet WWC group design standards.

The WWC reviewed 37 eligible studies against pilot single-case design standards. Three of the eligible studies meet WWC pilot single-case design standards without reservations, and one of the eligible studies meets WWC pilot single-case design standards with reservations. The remaining 33 eligible single-case design studies do not meet WWC pilot single-case design standards. More details on the four studies that meet pilot single-case design standards with or without reservations can be found in Appendix E. The results from single-case design studies only affect the WWC effectiveness rating for an outcome domain if the studies with outcomes in that domain collectively meet a set of threshold criteria. (See the box below for the rationale behind this threshold and a description of the criteria.) The evidence from these four single-case design studies does not reach the threshold to include single-case design evidence in the effectiveness ratings for any of the outcome domains in this report.

Threshold to include single-case design evidence in WWC effectiveness ratings

All single-case design experiments presented in the same research article are characterized as one study. Results from single-case design studies contribute to the WWC effectiveness rating for an outcome domain only if the studies with outcomes in that domain meet a set of threshold criteria, reflecting replication across different studies, research teams, and cases.

Specifically, these criteria are: (1) at least five studies that examine the intervention must meet WWC pilot single-case design standards without reservations or meet WWC pilot single-case design standards with reservations, and (2) the single-case design studies must be conducted by at least three different research teams with no overlapping authorship at three different institutions, and (3) the combined number of cases (i.e., participants, classrooms) must total at least 20.

For more information, please refer to the Pilot Single-Case Design Standards in Appendix E of the WWC Procedures and Standards Handbook (version 3.0).

Summary of studies meeting WWC group design standards without reservations

Hardan et al. (2015) used a randomized controlled trial to compare the effectiveness of PRT to psychoeducation group training on outcomes in the communications/language competencies domain. Twenty-seven children were randomly assigned to the intervention group and 26 children were randomly assigned to the comparison group. All children had been diagnosed with autism based on the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) criteria, the Autism Diagnostic Interview–Revised (ADI-R), the Autism Diagnostic Observation Schedule (ADOS), and expert clinical judgment. The average participant age was 4 years 1 month. Parents were trained to implement PRT once a week for 12 consecutive weeks using the manual How to Teach Pivotal Behaviors to Children with Autism (Koegel et al., 1989), and were asked to implement the intervention daily with their child. Forty-seven children had outcome data measuring: the total number of utterances (and number of specific types of utterances); performance on the communications subscale from the Vineland Adaptive Behavior Scales, 2nd Edition; and performance on the expressive communication subscale of the Preschool Language Scale, 4th Edition.
Schreibman and Stahmer (2014) used a randomized controlled trial to compare the effectiveness of PRT to Picture Exchange Communication System (PECS) on outcomes in the communication/language competencies domain. Forty-one children and their families were recruited from two different university sites and randomly assigned to either PRT or PECS. The children in the study sample were between the ages of 20 and 45 months and had been diagnosed with an autism spectrum disorder using the ADI-R and the Autism Diagnostic Observation Schedule–Generic. On average, participants received 247 hours of the assigned intervention. The intervention was administered by student therapists in the children’s homes; parents also received training to implement PRT. Thirty-eight children completed the Mullen Scales of Early Learning (MSEL) Expressive Language Scale before the intervention (baseline), after the intervention (postintervention), and 3 months later (follow-up).

Summary of studies meeting WWC group design standards with reservations

No studies of PRT met WWC group design standards with reservations.
Effectiveness Summary

The WWC review of PRT for the Children and Students with an Autism Spectrum Disorder topic area includes student outcomes in 16 domains: alphabetics, communication/language competencies, community, general reading achievement, math achievement, problem behavior, reading comprehension, reading fluency, school engagement, science achievement, self-care/daily living, self-determination, social-emotional competence, social studies achievement, vocational/occupational, and writing achievement. The two group design studies of PRT that meet WWC group design standards reported findings in one of the 16 domains: communication/language competencies. The findings below present the authors’ estimates and WWC-calculated estimates of the size and statistical significance of the effects of PRT on children and students with an autism spectrum disorder directly following the intervention. Additional comparisons are presented as supplemental findings in Appendix D. The supplemental findings do not factor into the intervention’s rating of effectiveness. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 35.

The findings from single-case design research do not meet the threshold to include single-case design evidence in the effectiveness ratings in this report. For a more detailed description of the rating of effectiveness for single-case design studies and extent of evidence criteria, see the WWC Rating Criteria that starts on p. 35.

Summary of effectiveness for the communication/language competencies domain

Table 3. Rating of effectiveness and extent of evidence for the communication/language competencies domain

<table>
<thead>
<tr>
<th>Rating of effectiveness</th>
<th>Criteria met</th>
</tr>
</thead>
<tbody>
<tr>
<td>No discernible effects</td>
<td>In the two studies that reported findings, the estimated impact of the intervention on outcomes in the communication/language competencies domain was neither statistically significant nor large enough to be substantively important.</td>
</tr>
<tr>
<td>No affirmative evidence of effects.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extent of evidence</th>
<th>Criteria met</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>Two studies that included 85 students reported evidence of effectiveness in the communication/language competencies domain.</td>
</tr>
</tbody>
</table>

Two studies that meet WWC group design standards without reservations reported findings in the communication/language competencies domain.

Hardan et al. (2015) did not report the statistical significance of the impact of PRT from baseline to posttest for three measures of communication/language competencies: the total number of utterances; the communications subscale from the Vineland Adaptive Behavior Scales, 2nd Edition; and the expressive communication subscale of the Preschool Language Scale, 4th Edition. The WWC-calculated $p$-values were not statistically significant for any of the outcomes, and the WWC-calculated average effect size across all findings was not large enough to be considered substantively important. The WWC characterizes this study as having an indeterminate effect because the effect is neither statistically significant nor substantively important. The authors also reported the impact of PRT on five subscales that measure the frequency of specific types of utterances, which are presented as supplemental findings in Appendix D and do not factor into the intervention’s rating of effectiveness. After correcting for multiple comparisons, the WWC-calculated $p$-values were statistically significant for the following subscales: imitative utterances, nonverbally prompted utterances, and spontaneous utterances.

Schreibman and Stahmer (2014) found no statistically significant effect of PRT on one measure of communication/language competencies: the MSEL Expressive Language Scale. The WWC confirmed the lack of statistical significance of this finding, and the WWC-calculated effect size was not large enough to be considered substantively important. The WWC characterizes this study as having an indeterminate effect because the effect is neither statistically significant nor substantively important.

Thus, for the communication/language competencies domain, no studies found statistically significant or substantively important effects of PRT. This results in a rating of no discernible effects, with a small extent of evidence.
References

Studies that meet WWC group design standards without reservations


Additional source:


Studies that meet WWC group design standards with reservations

None.

Studies that meet WWC pilot single-case design standards without reservations

The evidence from the single-case design studies on PRT does not reach the threshold to include single-case design evidence in the effectiveness ratings for any of the eligible outcome domains.


Additional source:


Study that meets WWC pilot single-case design standards with reservations

The evidence from the single-case design studies on PRT does not reach the threshold to include single-case design evidence in the effectiveness ratings for any of the eligible outcome domains.


Study that does not meet WWC group design standards


Additional source:

Studies that do not meet WWC pilot single-case design standards


Bruinsma, Y. E. M. (2005). *Increases in the joint attention behavior of eye gaze alternation to share enjoyment as a collateral effect of pivotal response treatment for three children with autism* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3145711) The study does not meet WWC pilot single-case design standards because there are insufficient data to evaluate the attempts to demonstrate an intervention effect.


Dyer, K., & Karp, R. (2013). A staff-training program to increase spontaneous vocal requests in children with autism. *Behavior Analysis in Practice, 6*(2), 42–60. The study does not meet WWC pilot single-case design standards because there are insufficient data to evaluate the attempts to demonstrate an intervention effect.


O’Hara, K. S. (2015). A comparison of PIPRT to VMO to increase social play skills in children with autism. *UNLV Theses, Dissertations, Professional Papers, and Capstones. Paper 2402.* The study does not meet WWC pilot single-case design standards because there are insufficient data to evaluate the attempts to demonstrate an intervention effect.


*Additional source:*


Pierce, K., & Schreibman, L. (1997). Multiple peer use of pivotal response training social behaviors of classmates with autism: Results from trained and untrained peers. *Journal of Applied Behavior Analysis, 30*(1), 157–160. The study does not meet WWC pilot single-case design standards because the measures of effectiveness cannot be attributed solely to the intervention.

*Additional sources:*


Robinson, S. E. (2011). Training paraprofessionals of students with autism to implement pivotal response treatment in inclusive school settings using a brief video feedback training package. *Focus on Autism and Other Developmental Disabilities, 26*(2), 105–118. The study does not meet WWC pilot single-case design standards because there are insufficient data to evaluate the attempts to demonstrate an intervention effect.

Additional source:


Additional sources:


**Studies that are ineligible for review using the Children and Students with an Autism Spectrum Disorder Evidence Review Protocol**


Barton, E. E., Lawrence, K., & Deurloo, F. (2012). Individualizing interventions for young children with autism in preschool. *Journal of Autism and Developmental Disorders, 42*(6), 1205–1217. This study is ineligible for review because it does not use an eligible design.


Bondy, A., & Weiss, M. J. (2013). *Teaching social skills to people with autism: Best practices in individualizing interventions* (1st ed.). Bethesda, MD: Woodbine House. This study is ineligible for review because it does not use an eligible design.

Boyd, B. A., Odom, S. L., Humphreys, B. P., & Sam, A. M. (2010). Infants and toddlers with autism spectrum disorder: Early identification and early intervention. *Journal of Early Intervention, 32*(2), 75–98. This study is ineligible for review because it does not use an eligible design.


Carr, M. E., Moore, D. W., & Anderson, A. (2014). Self-management interventions on students with autism: A meta-analysis of single-subject research. *Exceptional Children, 81*(1), 28–44. This study is ineligible for review because it does not use an eligible design.


Castorina, L. L., & Negri, L. M. (2011). The inclusion of siblings in social skills training groups for boys with Asperger syndrome. *Journal of Autism and Developmental Disorders, 41*(1), 73–81. This study is ineligible for review because it is out of the scope of the protocol.

Cohen, S. (2011). Commentary on providing services to students with autism spectrum disorders. *Journal of Visual Impairment and Blindness, 105*(6), 325–329. This study is ineligible for review because it does not use an eligible design.


Crosland, K., & Dunlap, G. (2012). Effective strategies for the inclusion of children with autism in general education classrooms. *Behavior Modification, 36*(3), 251–269. This study is ineligible for review because it does not use an eligible design.


Du Bose, L. V. M. (2012). Using evidence-based practices addressing language and communication in students with autism and developmental disabilities: Do special education teachers determine approaches consistent with characteristics of students? San Bernadino: California State University. This study is ineligible for review because it does not use a sample aligned with the protocol.


Ferraioli, S. J., & Harris, S. L. (2011). Effective educational inclusion of students on the autism spectrum. Journal of Contemporary Psychotherapy, 41(1), 19–28. This study is ineligible for review because it does not use an eligible design.

Ferraioli, S. J., & Harris, S. L. (2011). Teaching joint attention to children with autism through a sibling-mediated behavioral intervention. Behavioral Interventions, 26(4), 261–281. This study is ineligible for review because it does not use an eligible design.


Flynn, L., & Healy, O. (2012). A review of treatments for deficits in social skills and self-help skills in autism spectrum disorder. Research in Autism Spectrum Disorders, 6(1), 431–441. This study is ineligible for review because it does not use an eligible design.


Gianoumis, S., Seiverling, L., & Sturmey, P. (2012). The effects of behavior skills training on correct teacher implementation of natural language paradigm teaching skills and child behavior. Behavioral Interventions, 27(2), 57–74. This study is ineligible for review because it is out of the scope of the protocol.
Hansen, B. D., Wadsworth, J. P., Roberts, M. R., & Poole, T. N. (2014). Effects of naturalistic instruction on phonological awareness skills of children with intellectual and developmental disabilities. Research in Developmental Disabilities, 35(11), 2790–2801. This study is ineligible for review because it does not use a sample aligned with the protocol.


Jang, J., Dixon, D. R., Tarbox, J., Granpeesheh, D., Kornack, J., & de Nocker, Y. (2012). Randomized trial of an eLearning program for training family members of children with autism in the principles and procedures of applied behavior analysis. Research in Autism Spectrum Disorders, 6(2), 852–856. This study is ineligible for review because it is out of the scope of the protocol.


Kalyva, E. (2011). Autism: Educational and therapeutic approaches (English language ed.). Los Angeles, CA: SAGE. This study is ineligible for review because it does not use an eligible design.


Koegel, R. L., Shirotova, L., & Koegel, L. (2009). Brief report; Using individualized orienting cues to facilitate first-word acquisition in nonresponders with autism. *Journal of Autism and Developmental Disorders, 39*(11), 1587–1592. This study is ineligible for review because it is out of the scope of the protocol.


Koenig, K. (2012). *Practical social skills for autism spectrum disorders: Designing child-specific interventions* (p. 239). New York, NY: W. W. Norton & Co. This study is ineligible for review because it does not use an eligible design.

Kuhn, L. D. (2010). *Evaluation of a public school group-based applied behavioral analysis program for elementary students with autism* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3391519) This study is ineligible for review because it does not use an eligible design.

Lang, R., Machalicek, W., Rispoli, M., & Regester, A. (2009). Training parents to implement communication interventions for children with autism spectrum disorders (ASD): A systematic review. *Evidence-Based Communication Assessment and Intervention, 3*(3), 174–190. This study is ineligible for review because it does not use an eligible design.

Law, J., Plunkett, C. C., & Stringer, H. (2012). Communication interventions and their impact on behaviour in the young child: A systematic review. *Child Language Teaching and Therapy, 28*(1), 7–23. This study is ineligible for review because it does not use an eligible design.


Luiselli, J. K. (2014). *Children and youth with autism spectrum disorder (ASD): Recent advances and innovations in assessment, education, and intervention*. New York, NY: Oxford University Press. This study is ineligible for review because it does not use an eligible design.


Mahoney, G. (2013). Assimilative practice and developmental intervention. *International Journal of Early Childhood Special Education, 5*(1), 45–65. This study is ineligible for review because it does not use an eligible design.

Masiello, T. (2003). Effectiveness of pivotal response training as a behavioral intervention for young children with autism spectrum disorders. *Winterberry Research Syntheses, 1*(14). This study is ineligible for review because it does not use an eligible design.


McLay, L. (2012). Inconclusive evidence for the effectiveness of pivotal response treatment (PRT) for the infant siblings of children with ASD. *Evidence-Based Communication Assessment and Intervention, 6*(4), 195–200. This study is ineligible for review because it does not use a sample aligned with the protocol.


Mesibov, G. B., & Shea, V. (2011). Evidence-based practices and autism. *Autism, 15*(1), 114–133. This study is ineligible for review because it does not use an eligible design.


Odom, S. L., Hume, K., Boyd, B., & Stabel, A. (2012). Moving beyond the intensive behavior treatment versus eclectic dichotomy: Evidence-based and individualized programs for learners with ASD. *Behavior Modification, 36*(3), 270–297. This study is ineligible for review because it does not use an eligible design.

Ostryn, C., & Wolfe, P. S. (2011). Teaching children with autism to ask “what’s that?” using picture communication with vocal results. *Infants & Young Children, 24*(2), 174–192. This study is ineligible for review because it does not use an eligible design.


Rinald, K., & Mirenda, P. (2012). Effectiveness of a modified rapid toilet training workshop for parents of children with developmental disabilities. *Research in Developmental Disabilities, 33*(3), 933–943. This study is ineligible for review because it does not use a sample aligned with the protocol.

Rispoli, M., Neely, L., Lang, R., & Ganz, J. (2011). Training paraprofessionals to implement interventions for people autism spectrum disorders: A systematic review. *Developmental Neurorehabilitation, 14*(6), 378–388. This study is ineligible for review because it does not use an eligible design.


Schmidt, C., & Stichter, J. P. (2012). The use of peer-mediated interventions to promote the generalization of social competence for adolescents with high-functioning autism and Asperger’s Syndrome. *Exceptionality, 20*(2), 94–113. This study is ineligible for review because it does not use an eligible design.


Seiverling, L., Pantelides, M., Ruiz, H. H., & Sturmey, P. (2010). The effect of behavioral skills training with general-case training on staff chaining of child vocalizations within natural language paradigm. *Behavioral Interventions, 25*(1), 53–75. This study is ineligible for review because it does not use an eligible design.


Shannon, J. B. (2011). *Autism and pervasive developmental disorders sourcebook: Basic consumer health information about autism spectrum disorders (ASD) including autistic disorder, Asperger syndrome, Rett syndrome, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified (PDDNOS): Along with facts about causes, symptoms, assessment, interventions, treatments, and education, tips for family members and teachers on the transition to adulthood* ... (2nd ed.). Detroit, MI: Omnigraphics. This study is ineligible for review because it does not use an eligible design.

Sisavath, J. (2014). *A review of parent training interventions for children with autism spectrum disorder and proposed guidelines for choosing best practices* (Unpublished doctoral dissertation). University of Texas, Austin. This study is ineligible for review because it does not use an eligible design.


Stahmer, A. C. (1999). Using pivotal response training to facilitate appropriate play in children with autistic spectrum disorders. *Child Language Teaching and Therapy, 15*(1), 29–40. This study is ineligible for review because it does not use an eligible design.


Stahmer, A. C., & Gist, K. (2001). The effects of an accelerated parent education program on technique mastery and child outcome. *Journal of Positive Behavior Interventions, 3*(2), 75. This study is ineligible for review because it does not use an eligible design.

Stahmer, A. C., Schreibman, L., & Cunningham, A. B. (2011). Toward a technology of treatment individualization for young children with autism spectrum disorders. *Brain Research, 1380*, 229–239. This study is ineligible for review because it does not use an eligible design.

(pp. 397–412). New York, NY: Springer Science + Business Media. This study is ineligible for review because it does not use an eligible design.


Suhrheinrich, J. B. (2010). *A sustainable model for training teachers to use pivotal response training* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3404295) This study is ineligible for review because it does not use a sample aligned with the protocol.


Tran, Q. H. (2008). *Using a self-assessment procedure to improve parent implementation of intervention for children with autism* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3274397) This study is ineligible for review because it does not use an eligible design.

Trowbridge, M. D. (2012). *Professional development needs of intervention specialists in the area of evidence-based practices for students with autism* (Unpublished doctoral dissertation). Northern Kentucky University, Highland Heights. This study is ineligible for review because it is out of the scope of the protocol.


Ventola, P. E., Oosting, D. R., Keifer, C. M., & Friedman, H. E. (2015). Toward optimal outcome following pivotal response treatment: A case series. *Yale Journal of Biology and Medicine, 88*(1), 37–44. This study is ineligible for review because it does not use an eligible design.


Volkmar, F. R., Reichow, B., & Doehring, P. (2011). Evidence-based practices in autism: Where we are now and where we need to go. In B. Reichow, P. Doehring, & D. Cicchetti (Eds.), *Evidence-based practices and treatments for children with autism* (pp. 365–391). New York, NY: Springer Science + Business Media. This study is ineligible for review because it does not use an eligible design.


Wang, S., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skills interventions for children with autism spectrum disorders: A meta-analysis in single-case research using HLM. *Research in Autism Spectrum Disorders, 5*(1), 562–569. This study is ineligible for review because it does not use an eligible design.

Wang, S., Parrila, R., & Cui, Y. (2013). Meta-analysis of social skills interventions of single-case research for individuals with autism spectrum disorders: Results from three-level HLM. *Journal of Autism and Developmental Disorders, 43*(7), 1701–1716. This study is ineligible for review because it does not use an eligible design.


Williams, C. M., Fan, W., & Goodman, G. (2011). Preliminary analysis of the “survey of educators' knowledge and value of research-based practices for students with autism.” *Assessment for Effective Intervention, 36*(2), 113–130. This study is ineligible for review because it is out of the scope of the protocol.

Wilson, N. L. (2010). *Life skills acquisition curricula for children with autistic disorder* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3382673) This study is ineligible for review because it does not use an eligible design.


Wong, C., & Kasari, C. (2012). Play and joint attention of children with autism in the preschool special education classroom. *Journal of Autism and Developmental Disorders, 42*(10), 2152–2161. This study is ineligible for review because it does not use a sample aligned with the protocol.

Yetter, G. (2014). Improving educational outcomes for students on the autism spectrum. *PsycCRITIQUES, 59*(36). This study is ineligible for review because it does not use an eligible design.

Appendix A.1: Research details for Hardan et al. (2015)


### Table A1. Summary of findings

<table>
<thead>
<tr>
<th>Outcome domain</th>
<th>Sample size</th>
<th>Average improvement index (percentile points)</th>
<th>Statistically significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication/language competencies</td>
<td>47 children</td>
<td>10</td>
<td>No</td>
</tr>
</tbody>
</table>

### Setting
Parents were trained to implement *PRT* at a large academic medical center (the location is not reported). Parents implemented *PRT* in their home, and occasionally with their child at the medical center.

### Study sample
Twenty-seven children were randomly assigned to the intervention group and 26 children were randomly assigned to the comparison group; 25 children in the intervention group and 22 in the comparison group had observed outcomes. About 75% of the participating children were boys, and the average age was 4 years 1 month. All children had been diagnosed with autism based on the *Diagnostic and Statistical Manual of Mental Disorders-IV* (DSM-IV) criteria, the Autism Diagnostic Interview–Revised (ADI-R), the Autism Diagnostic Observation Schedule (ADOS), and expert clinical judgment.

### Intervention group
Parents in the intervention group received training once a week for 12 consecutive weeks. Trainers used the manual *How to Teach Pivotal Behaviors to Children with Autism*. Eight of the sessions were 90 minutes and involved small groups of four to six parents, led by one or two psychologists. The other four sessions lasted 60 minutes, and each session included a parent and child meeting with a psychologist. Parents were asked to implement the intervention daily with their child.

### Comparison group
Parents in the comparison group received psychoeducation group (PEG) training. Sessions were conducted once a week for 12 consecutive weeks. Ten of the sessions were 90 minutes and involved groups of parents taught by clinical psychology graduate students and supervised by a psychologist. Two sessions lasted 60 minutes, and each session included a parent and his or her child meeting with a psychologist. The curriculum was based on an existing autism parent psychoeducation program at the medical center and covered the following topics: diagnosis and symptoms of autism spectrum disorders; neurobiology of autism; basic behavior management strategies; common autism interventions; evaluating treatment effectiveness; service systems; and strategies for improving social skills and child stress reduction.

### Outcomes and measurement
The study included three outcomes that fall into the communication/language competencies domain: the total number of utterances; the communications subscale from the Vineland Adaptive Behavior Scales, 2nd Edition; and the expressive communication subscale of the Preschool Language Scale, 4th Edition. For a more detailed description of these outcome measures, see Appendix B.
The study also included five supplemental subscale outcomes that fall in the communication/language competencies domain that measure the frequency of specific types of utterances: (a) unintelligible utterances, (b) imitative utterances, (c) verbally prompted utterances, (d) nonverbally prompted utterances, and (e) spontaneous utterances. The sum of these outcomes is the total number of utterances, and so these outcomes are reported as supplemental findings in Appendix D and do not factor into the intervention's rating of effectiveness. For a more detailed description of these outcome measures, see Appendix B.

The study included four parent-reported outcomes that do not meet review requirements: the Social Responsiveness Scale and three measures based on the MacArthur-Bates Communicative Development Inventories. Because each child is assessed by their parent, there is not a consistent assessment of all children, and these outcomes do not meet review requirements. The study included two other measures that do not meet reliability requirements: the Vineland Adaptive Behavior Scales (second edition) expressive raw score and receptive raw score. Two other outcomes, the Clinical Global Impression Scale–Severity and the Clinical Global Impression Scale–Improvement, are not in an eligible domain.

Support for implementation

Parents did not receive any additional support, aside from the training.

Appendix A.2: Research details for Schreibman and Stahmer (2014)


<table>
<thead>
<tr>
<th>Outcome domain</th>
<th>Sample size</th>
<th>Average improvement index (percentile points)</th>
<th>Statistically significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication/language competencies</td>
<td>38 children</td>
<td>–9</td>
<td>No</td>
</tr>
</tbody>
</table>

Table A2. Summary of findings Meets WWC group design standards without reservations

Setting Undergraduate student therapists provided the assigned intervention to children in their homes. Parents were also trained on using the intervention and participated in education sessions with their children in small playrooms at two universities and in their homes.

Study sample Forty-one children were randomly assigned to either PRT or PECS, and 39 participated in the study. Two families, one in each condition, discontinued participation during the initial phase of the intervention; one family moved, and one family preferred to receive the nonassigned intervention. The children were diagnosed with autistic disorder using the Autism Diagnostic Interview–Revised and the Autism Diagnostic Observation Schedule–Generic. To participate, children also had to speak fewer than nine intelligible words. There were 34 male children and five female children; all were aged 20–45 months. The final analytic sample included 38 children, with either 20 or 19 children in the PRT group (the study does not report analytic sample sizes by condition).
Intervention group

Children in the intervention condition received PRT. Parents and therapists who provided PRT were trained in accordance with the PRT training manual. Children in the intervention group received between 181 and 263 hours of PRT in their home.

Comparison group

Children in the comparison condition received the PECS intervention, an intervention that teaches children to exchange picture icons to communicate. Parents and therapists who provided PECS were trained in accordance with the PECS training manual. Children in the comparison group received between 181 and 263 hours of PECS in their home.

Outcomes and measurement

There is one child outcome that meets review requirements: the Expressive Language Scale from the Mullen Scales of Early Learning. This outcome falls under the communication/language competencies domain. The measure was administered by trained staff not involved with the intervention procedures. For a more detailed description of this outcome measure, see Appendix B.

The study measured this outcome postintervention and at follow-up 3 months later. The immediate posttest is used to determine the evidence rating for this review. The follow-up data are reported in Appendix D as supplemental findings that do not factor into the rating of effectiveness.

Findings from three other eligible outcomes are not included in this review because there was insufficient information on sample sizes to assess attrition and demonstrate equivalence for the analytic sample: The MacArthur Communicative Developmental Inventory (CDI), the Expressive One-Word Picture Vocabulary Test–Revised (EOWPVT), and the Vineland Adaptive Behavior Scale, 2nd Edition (VABS). Augmentative communication is not included in the review because the measure was only administered to children in the PECS group. Parent satisfaction, which does not fall in a protocol outcome domain, is also not eligible for review.

Support for implementation

During the study period, undergraduate student therapists and parents received training in accordance with the PRT manual. During the first 15 weeks, parents participated in 2-hour education sessions in the laboratory with their children twice a week. During the subsequent 8 weeks, they received 2-hour education sessions once a week and 2-hour sessions in the home twice a week. The implementation process was identical in the comparison group for PECS.
### Appendix B: Group design outcome measures for the communication/language competencies domain

<table>
<thead>
<tr>
<th>Communication/language competencies</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Imitative utterances</strong></td>
<td>The number of spoken expressions that repeat parent expressions (as cited in Hardan et al., 2015). This outcome is included as a supplemental finding in Appendix D.</td>
</tr>
<tr>
<td><strong>Mullen Scales of Early Learning (MSEL), Expressive Language Scale</strong></td>
<td>The MSEL (Mullen, 1995) measures the cognitive ability of children from birth to 68 months. The MSEL covers a variety of domains of cognitive ability, including the Expressive Language Scale used in this study. The Expressive Language Scale has a mean of 50 and standard deviation of 10 (as cited in Schreibman &amp; Stahmer, 2014).</td>
</tr>
<tr>
<td><strong>Nonverbally prompted utterances</strong></td>
<td>The number of spoken expressions that occurred after a clear nonverbal prompt from the parent, such as the parent pausing and waiting expectantly for a response (as cited in Hardan et al., 2015). This outcome is included as a supplemental finding in Appendix D.</td>
</tr>
<tr>
<td><strong>Preschool Language Scale, 4th Edition (expressive communication subscale)</strong></td>
<td>A standardized, norm-referenced measure that measures expressive communication (as cited in Hardan et al., 2015).</td>
</tr>
<tr>
<td><strong>Spontaneous utterances</strong></td>
<td>The number of spoken expressions that occurred without any effort from the parent to elicit a response (as cited in Hardan et al., 2015). This outcome is included as a supplemental finding in Appendix D.</td>
</tr>
<tr>
<td><strong>Total utterances</strong></td>
<td>The number of total spoken expressions during parent–child interactions after 12 weeks. Raters reviewed 10-minute videos and recorded the child’s functional verbal utterances (as cited in Hardan et al., 2015).</td>
</tr>
<tr>
<td><strong>Unintelligible utterances</strong></td>
<td>The number of spoken expressions that cannot be understood (as cited in Hardan et al., 2015). This outcome is included as a supplemental finding in Appendix D.</td>
</tr>
<tr>
<td><strong>Verbally prompted utterances</strong></td>
<td>The number of spoken expressions that occurred after a clear verbal prompt from the parent (as cited in Hardan et al., 2015). This outcome is included as a supplemental finding in Appendix D.</td>
</tr>
<tr>
<td><strong>Vineland Adaptive Behavior Scales, 2nd edition (communications subscale)</strong></td>
<td>A standardized, norm-referenced measure that measures expressive, receptive, and written communication (as cited in Hardan et al., 2015).</td>
</tr>
</tbody>
</table>
Appendix C: Group design findings included in the rating for the communication/language competencies domain

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study sample</th>
<th>Sample size</th>
<th>Mean (standard deviation)</th>
<th>WWC calculations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intervention group</td>
<td>Comparison group</td>
</tr>
<tr>
<td>Hardan et al. (2015)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>63.40 (11.60)</td>
<td>63.00 (13.40)</td>
</tr>
<tr>
<td>Preschool Language Scale, 4th Edition (expressive communication subscale)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total utterances</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>60.00 (28.90)</td>
<td>51.40 (33.80)</td>
</tr>
<tr>
<td>Vineland Adaptive Behavior Scales, 2nd edition (communications subscale)</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>80.60 (18.90)</td>
<td>72.80 (16.50)</td>
</tr>
<tr>
<td>Domain average for communication/language competencies (Hardan et al., 2015)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schreibman &amp; Stahmer (2014)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>20–45 months</td>
<td>38 children</td>
<td>24.30 (7.00)</td>
<td>26.70 (12.70)</td>
</tr>
<tr>
<td>Mullen Scales of Early Learning (MSEL), Expressive Language Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain average for communication/language competencies (Schreibman &amp; Stahmer, 2014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domain average for communication/language competencies across all studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual’s percentile rank that can be expected if the individual is given the intervention. The statistical significance of the study’s domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. nr = not reported. na = not applicable.

<sup>a</sup> For Hardan et al. (2015), the study-reported p-values were based on an analysis that included an additional time period before the intervention was complete and are not included in this report. The WWC-calculated p-values were adjusted for multiple comparisons and were not statistically significant for any outcome. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information. This study is characterized as having an indeterminate effect because the mean effect reported is neither statistically significant nor substantively important. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

<sup>b</sup> For Schreibman and Stahmer (2014), no corrections for clustering or multiple comparisons were needed. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information. This study is characterized as having an indeterminate effect because the estimated effect for the only measure within the domain is neither statistically significant nor substantively important. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.
## Appendix D.1: Group design supplemental subscale findings for the communication/language competencies domain

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study sample</th>
<th>Sample size</th>
<th>Mean (standard deviation)</th>
<th>WWC calculations</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intervention group</td>
<td>Comparison group</td>
<td>Mean difference</td>
</tr>
<tr>
<td>Hardan et al. (2015)a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.70</td>
</tr>
<tr>
<td>Imitative utterances</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>15.80 (14.20)</td>
<td>7.10 (7.80)</td>
<td>8.70</td>
</tr>
<tr>
<td>Nonverbally prompted utterances</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>3.00 (3.80)</td>
<td>0.10 (0.30)</td>
<td>2.90</td>
</tr>
<tr>
<td>Spontaneous utterances</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>1.30 (1.40)</td>
<td>0.40 (0.60)</td>
<td>0.90</td>
</tr>
<tr>
<td>Unintelligible utterances</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>21.16 (14.90)</td>
<td>24.60 (23.30)</td>
<td>−3.44</td>
</tr>
<tr>
<td>Verbally prompted utterances</td>
<td>4 years, 1 month</td>
<td>47 children</td>
<td>17.70 (15.20)</td>
<td>16.00 (17.00)</td>
<td>1.70</td>
</tr>
</tbody>
</table>

### Table Notes:
The supplemental findings presented in this table are additional findings from a study in this report that meets WWC design standards without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. nr = not reported.

a For Hardan et al. (2015), the study-reported p-values were based on an analysis that included an additional time period before the intervention was complete and are not included in this report. After correction for multiple comparisons, the WWC-calculated p-values were statistically significant for the following outcomes: imitative utterances, nonverbally prompted utterances, and spontaneous utterances. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.

## Appendix D.2: Group design supplemental follow-up test findings in the communication/language competencies domain

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study sample</th>
<th>Sample size</th>
<th>Mean (standard deviation)</th>
<th>WWC calculations</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intervention group</td>
<td>Comparison group</td>
<td>Mean difference</td>
</tr>
<tr>
<td>Schreibman &amp; Stahmer (2014)a</td>
<td>20–45 months</td>
<td>38 children</td>
<td>25.50 (11.20)</td>
<td>28.70 (16.50)</td>
<td>−3.20</td>
</tr>
<tr>
<td>Mullen Scales of Early Learning (MSEL), Expressive Language Scale</td>
<td>20–45 months</td>
<td>38 children</td>
<td>25.50 (11.20)</td>
<td>28.70 (16.50)</td>
<td>−3.20</td>
</tr>
</tbody>
</table>

### Table Notes:
The supplemental findings presented in this table are additional findings from a study in this report that met WWC design standards without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. nr = not reported.

a For Schreibman and Stahmer (2014), no corrections for clustering or multiple comparisons were needed. The WWC calculated the intervention group mean using a difference-in-differences approach by adding the impact of the intervention (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information.
Appendix E: Single-case design studies that meet WWC pilot standards

Table E.1. Research details for single-case design studies that meet WWC pilot standards

<table>
<thead>
<tr>
<th>Study</th>
<th>Study sample, setting, comparison sessions, and intervention sessions</th>
</tr>
</thead>
</table>
| **Feldman & Matos (2012), Meets WWC Pilot Single-Case Design Standards Without Reservations** | This study included three children independently diagnosed with autism. The children spent at least 75% of their time in general education classrooms and received support from a paraprofessional. The intervention took place in a public elementary school in general education classrooms.  
During baseline sessions, students had the usual interactions with and support from their paraprofessional. During intervention sessions, paraprofessionals used the following PRT procedures: following the child’s lead to identify activities, providing clear instructions focused on the activity, using an appropriate prompting hierarchy when the desired behavior does not occur, providing contingent rewards related to the activity, and moving farther away from the child after correct behavior. Paraprofessionals were trained using a PRT manual and received feedback during implementation.  
For all students, the outcome was reciprocal interactions with their peers. This measure falls under the social-emotional competence domain. For a more detailed description of this outcome measure, see Appendix E, Table E.2. Maintenance data consisting of one data point for each student were collected 3–7 weeks after the end of the intervention. The maintenance data indicate that all students continued at about the same level as during the intervention phase. |
| **Kim (2015), Meets WWC Pilot Single-Case Design Standards With Reservations** | The study included four Korean-American children independently diagnosed with an autism spectrum disorder under the Diagnostic and Statistical Manual of Mental Disorders-IV Text Revision (DSM-IV-TR) or DSM-V criteria. All participating children had difficulties with social interaction and play, and regularly attended community-based programs (e.g., church, recreational program). The intervention took place in a classroom at a Korean culture and language school (held in a church) for three of the four children and at home for the fourth child.  
During baseline sessions, which lasted for 10 minutes and occurred twice a week over 2–7 weeks, the researcher observed the children with an autism spectrum disorder and their peer mediators; the researcher did not provide any prompts or instructions to the children. For the intervention, peers of autistic children were trained to implement four PRT strategies: (1) incorporating mutually preferred items/activities, (2) giving choices and asking questions, (3) delivering desired activities/items (natural reinforcement), and (4) sharing and take turns (shared control). Peer mediators used the social skills they learned during the training sessions to facilitate social interactions with an autistic child. Each session lasted for 10 minutes, and sessions occurred two to three times a week for between 3–8 weeks.  
For all students, the outcomes were responses to peer communication initiations and initiating social communications with peers. These measures fall in the communication/language competencies domain. For a more detailed description of these outcome measures, see Appendix E, Table E.2. Maintenance data for both outcomes, consisting of one data point for each student were collected 1 week after the end of the intervention. For both outcomes, the maintenance data indicated that all children continued at about the same level as during the intervention phase.  
The single-case design experiments for both outcomes had fewer than five data points in at least one phase; because all phases had at least three data points, these experiments meet WWC pilot single-case design standards with reservations. |
| **Schreibman et al. (2009), Meets WWC Pilot Single-Case Design Standards Without Reservations** | This study included six children diagnosed with an autism spectrum disorder by an outside professional, and the diagnosis was confirmed by the Autism Diagnostic Interview—Revised. All children were classified as unlikely to be responsive to PRT because of observed behaviors. All children had four of the following five behaviors: low levels of toy interaction or contact, low levels of social approach, low levels of verbal stereotypy, high levels of social avoidance, and high nonverbal stereotypy. The authors formed two groups based on the behaviors: (1) three children who had high toy contact behavior and the other four behaviors, and (2) three children who had low social avoidance and the other four behaviors. The study occurred in a structured laboratory environment in a large treatment room. Children participated in the intervention in their homes, and parents participated in education sessions with their children in small playrooms at the university and in their homes.  
During baseline sessions, the child had access to toys and was periodically given an opportunity to respond to therapist questions (with no reinforcement). The study also considered whether students responded to Discrete Trial Training, a structured behavioral treatment, but this component is not included in the review. During the intervention sessions, PRT was provided for 18 hours by a therapist trained in accordance with the PRT training manual. To increase expressive verbal communication skills, the therapist began by reinforcing any vocalizations, followed by reinforcing contingent vocalizations, and reinforcing specific sound imitation.  
For all students, the outcomes were prompted vocalizations and spontaneous vocalizations. These measures fall in the communication/language competencies domain. For a more detailed description of these outcome measures, see Appendix E, Table E.2. |
This study included six children who were diagnosed with autism by psychologists independent of the study. All children were classified into one of two groups based on whether they had behavior profiles that were expected to be responsive to PRT (i.e., responders and nonresponders). Relative to nonresponders, responders had higher interest in toys, greater approach behaviors with adults, fewer avoidance behaviors with adults, lower nonverbal stimulatory behavior, and higher verbal stimulatory behavior. The nonresponder had several characteristics (e.g., social avoidance) that the authors believed would make the children non-responsive to PRT.

During baseline sessions, which ranged from 4–16 weeks, toys were placed within reach of the children, and they were granted free access to them. Three sessions were conducted daily 4–5 days per week. The intervention was conducted in accordance with the PRT training manual.17 Children received PRT 4–5 times a week during 90-minute sessions broken into three 30-minute sessions that alternated between two different play rooms. A learning interaction began when the therapist presented possible activities to the child. The child chose and requested items and activities, and each time the child indicated a preference, the therapist requested a response before the child had access to the item or activity. The complexity of the required responses increased as the child developed. The responders received treatment for 6 months, and the nonresponders’ treatment was discontinued after 5 weeks because they failed to demonstrate any improvement.

For all students, the outcomes were appropriate communication, functional play behaviors, maintaining interactions with an adult, social initiations with a therapist, symbolic play behaviors, and varied play behaviors. The appropriate communication measure falls in the communication/language competencies domain, and the remaining measures fall in the social-emotional competence domain. For a more detailed description of these outcome measures, see Appendix E, Table E.2. Maintenance data for all outcomes were collected for two responders and one non-responder 6 or 9 months after the end of the intervention. For all the outcomes, the maintenance data indicate that all children continued at about the same level as, or slightly higher than, the intervention phase.

**Table E.2. Single-case design outcome measures for each domain**

<table>
<thead>
<tr>
<th>Communication/language competencies</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Appropriate communication</strong></td>
<td>The percentage of 30-second intervals that contained appropriate communication, including immediate echolalia, verbally cued speech, nonverbally cued speech, and spontaneous utterances. Calculated by totaling across all four of these communication categories (as cited in Sherer &amp; Schreibman, 2005).</td>
</tr>
<tr>
<td><strong>Prompted vocalization</strong></td>
<td>The percentage of instances where the child responded verbally to the therapist’s verbal or nonverbal prompt, including immediate echolalia. Five randomly-chosen minutes from each 20-minute session were coded and scored in 10-second intervals for occurrence/nonoccurrence of vocalizations (as cited in Schreibman et al., 2009).</td>
</tr>
<tr>
<td><strong>Responses to peer initiations</strong></td>
<td>The frequency that the autistic child responded to peer initiations during each 10 minute session, where positive nonverbal or verbal responses were scored and negative responses were not scored (as cited in Kim, 2015).</td>
</tr>
<tr>
<td><strong>Social initiations</strong></td>
<td>The frequency that the autistic child verbally or physically began a conversation or a new play theme with their peer without the peers’ prompts (as cited in Kim, 2015).</td>
</tr>
<tr>
<td><strong>Spontaneous vocalizations</strong></td>
<td>The percentage of instances where the child used appropriate vocalizations with communicative intent with at least 7 seconds between a verbal or nonverbal cue. Five randomly-chosen minutes of each 20-minute session were coded and scored in 10-second intervals for occurrence/nonoccurrence of vocalizations (as cited in Schreibman et al., 2009).</td>
</tr>
<tr>
<td><strong>Social-emotional competence</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Functional play behaviors</strong></td>
<td>The percentage of 30-second intervals in which the child engaged in functional play behaviors, defined as using an object as it was intended (e.g., rolling a toy car). Play behaviors could be spontaneous or follow a therapist’s prompt (as cited in Sherer &amp; Schreibman, 2005).</td>
</tr>
<tr>
<td><strong>Maintaining interactions–adult</strong></td>
<td>The percentage of 30-second intervals in which the child displayed continued engagement with the therapist in a verbal or nonverbal interaction. For example, complying with a therapist’s request or engaging in spontaneous eye contact was considered as maintaining an interaction, while turning away from the therapist was not scored as maintaining interaction (as cited in Sherer &amp; Schreibman, 2005).</td>
</tr>
<tr>
<td><strong>Reciprocal social interaction</strong></td>
<td>The percentage of 30-second intervals with verbal or nonverbal interactions between the child and his/her peers. Interactions were defined as instances when the peer responded verbally or nonverbally to the child (as cited in Feldman &amp; Matos, 2012).</td>
</tr>
</tbody>
</table>

Footnotes:
Social-emotional competence

Social initiations: The percentage of 30-second intervals in which the child engaged in social initiations with the therapist, defined as verbal or nonverbal requests to change the play activity or an appropriate spontaneous verbalization (e.g., “The bus is broken”). Social initiations had to involve a request for therapist involvement, not just a request for an object (as cited in Sherer & Schreibman, 2005).

Symbolic play behaviors: The percentage of 30-second intervals in which the child engaged in symbolic play behaviors, defined as using a toy as though it were another object (e.g., using a block as a bar of soap to wash a baby doll). Play behaviors could be spontaneous or follow a therapist’s prompt (as cited in Sherer & Schreibman, 2005).

Varied play behaviors: The percentage of 30-second intervals in which the child engaged in varied play behaviors, defined as playing with one toy and then playing with a different toy or playing differently with the same toy (e.g., switched from rolling a ball to putting it down a chute). Play behaviors could be spontaneous or follow a therapist’s prompt (as cited in Sherer & Schreibman, 2005).

* The authors collected inter-assessor agreement (IAA) data in each phase and on at least 20% of all sessions, but it is not clear whether IAA data were collected during 20% of the data points in each condition.

**Table E.3. Single-case design findings for the communication/language competencies domain**

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study characteristics</th>
<th>WWC summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample size (case)</td>
<td>Age(s)</td>
</tr>
<tr>
<td><strong>Kim (2015)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responses to peer initiations</td>
<td>4 (all students)</td>
<td>5–10</td>
</tr>
<tr>
<td>Social initiations</td>
<td>4 (all students)</td>
<td>5–10</td>
</tr>
<tr>
<td><strong>Schreibman et al. (2009)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cued vocalizations (% of intervals)</td>
<td>3 (HTC subgroup)</td>
<td>2</td>
</tr>
<tr>
<td>Cued vocalizations (% of intervals)</td>
<td>3 (LSA subgroup)</td>
<td>2–4</td>
</tr>
<tr>
<td>Spontaneous vocalizations (% of intervals)</td>
<td>3 (HTC subgroup)</td>
<td>2</td>
</tr>
<tr>
<td>Spontaneous vocalizations (% of intervals)</td>
<td>3 (LSA subgroup)</td>
<td>2–4</td>
</tr>
<tr>
<td><strong>Sherer &amp; Schreibman (2005)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appropriate communication (% of intervals)</td>
<td>3 (responder subgroup)</td>
<td>3</td>
</tr>
<tr>
<td>Appropriate communication (% of intervals)</td>
<td>3 (nonresponder subgroup)</td>
<td>3–5</td>
</tr>
</tbody>
</table>

*The WWC does not calculate effect sizes for single-case design (SCD) research. Characterizations of Strong and Moderate evidence, based on WWC visual analysis, indicate that the experiment demonstrated an effect of the intervention. Characterizations of No evidence indicate that the experiment did not provide at least three demonstrations of an intervention effect in the same direction. + = a positive (favorable) effect in the desired direction. The evidence from the SCD studies on PRT does not reach the threshold to include SCD evidence in the effectiveness ratings for the communication/language competencies domain. HTC = high toy contact. LSA = low social avoidance.*

* In Schreibman et al. (2009), it is unclear whether the graphical presentation of data in the original study uses a consistent display of time across cases. An author query was conducted to confirm whether the numbered treatment hours occurred at different points in time for each student, as displayed in the graphical presentation of data in the study. The WWC did not receive a reply to this query; however, the WWC characterization of evidence would not be affected by this issue.

* In Sherer and Schreibman et al. (2005), it is unclear whether the graphical presentation of data in the original study uses a consistent display of time across cases in the responder subgroup. An author query was conducted to confirm whether the numbered treatment days occurred at different points in time for each student, as displayed in the graphical presentation of data in the study. The WWC did not receive a reply to this query; however, the WWC characterization of evidence would not be affected by this issue.
### Table E.4: Single-case design findings for the social-emotional competence domain

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Study characteristics</th>
<th>WWC summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample size (case)</td>
<td>Age(s)</td>
</tr>
<tr>
<td><strong>Feldman &amp; Matos (2012)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reciprocal social interaction</td>
<td>3 (all)</td>
<td>5–9</td>
</tr>
<tr>
<td><strong>Sherer &amp; Schreibman (2005)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional play behaviors (% of intervals)</td>
<td>3 (responder subgroup)</td>
<td>3</td>
</tr>
<tr>
<td>Functional play behaviors (% of intervals)</td>
<td>3 (nonresponder subgroup)</td>
<td>3–5</td>
</tr>
<tr>
<td>Maintaining interactions–adult (% of intervals)</td>
<td>3 (responder subgroup)</td>
<td>3</td>
</tr>
<tr>
<td>Maintaining interactions–adult (% of intervals)</td>
<td>3 (nonresponder subgroup)</td>
<td>3–5</td>
</tr>
<tr>
<td>Social initiations (% of intervals)</td>
<td>3 (responder subgroup)</td>
<td>3</td>
</tr>
<tr>
<td>Social initiations (% of intervals)</td>
<td>3 (nonresponder subgroup)</td>
<td>3–5</td>
</tr>
<tr>
<td>Symbolic play behaviors (% of intervals)</td>
<td>3 (responder subgroup)</td>
<td>3</td>
</tr>
<tr>
<td>Symbolic play behaviors (% of intervals)</td>
<td>3 (nonresponder subgroup)</td>
<td>3–5</td>
</tr>
<tr>
<td>Varied play behaviors (% of intervals)</td>
<td>3 (responder subgroup)</td>
<td>3</td>
</tr>
<tr>
<td>Varied play behaviors (% of intervals)</td>
<td>3 (nonresponder subgroup)</td>
<td>3–5</td>
</tr>
</tbody>
</table>

**Table Notes:** The WWC does not currently calculate effect sizes for single-case design (SCD) research. Characterizations of Strong and Moderate evidence, based on WWC visual analysis, indicate that the experiment demonstrated an effect of the intervention. Characterizations of No evidence indicate that the experiment did not provide at least three demonstrations of an intervention effect in the same direction. + = a positive (favorable) effect in the desired direction. The evidence from the SCD studies on PRT does not reach the threshold to include SCD evidence in the effectiveness ratings for the social-emotional competence domain.

*In Sherer and Schreibman et al. (2005), it is unclear whether the graphical presentation of data in the original study uses a consistent display of time across cases in the responder subgroup. An author query was conducted to confirm whether the numbered treatment days occurred at different points in time for each student, as displayed in the graphical presentation of data in the study. The WWC did not receive a reply to this query; however, the WWC characterization of evidence would not be affected by this issue.*
Endnotes

1 The descriptive information for this intervention was obtained from publicly available sources: Koegel Autism Consultants (www.autismprthelp.com); Koegel, Koegel, & Brookman, 2003; and Koegel, Openden, Fredeen, & Koegel, 2006. The WWC requests developers review the intervention description sections for accuracy from their perspective. The intervention description was provided to the developer in May 2014; however, the WWC received no response. Further verification of the accuracy of the descriptive information for this intervention is beyond the scope of this review.

2 The literature search reflects documents publicly available by August 2015. The studies in this report were reviewed using the Standards from the WWC Procedures and Standards Handbook (version 3.0), and the Children and Students with an Autism Spectrum Disorder review protocol (version 3.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

3 Please see the Children and Students With an Autism Spectrum Disorder review protocol (version 3.0) for a list of all the outcome domains.

4 For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 35. These improvement index numbers show the average and range of individual-level improvement indices for all findings across the studies.

5 In single-case design research, a case, such as a student or classroom, is the unit of intervention administration and data analysis. A single-case design experiment is the examination of a single outcome measure repeatedly within and across different phases defined by the presence or absence of the intervention. There may be multiple experiments for a case if more than one outcome is examined, for example. All experiments within a research article comprise one single-case design study.

6 For the communication/language competencies domain, there are four studies (fewer than the five required) and 19 cases (fewer than the 20 required). For the social-emotional competence domain, there are two studies (fewer than the five required) and nine cases (fewer than 20 required). There are no studies that meet WWC pilot single-case design standards in the 14 other domains.

7 The results from single-case design studies are not used to report an intervention effectiveness rating for an outcome domain unless the studies collectively meet the threshold criteria described on pp. 4 and 36. The four single-case design studies of PRT that meet WWC pilot single-case design standards with or without reservations reported findings in two of the 16 domains: communication/language competencies and social-emotional competence. The evidence from the single-case design studies on PRT does not reach the threshold to include single-case design evidence in the effectiveness ratings for either domain.

8 The study also reported analyses for the following subgroups: students with lower visual reception scores; students with higher visual reception scores; females; males; students between the ages of 2 years and 4 years 6 months; and students between the ages 4 years 7 months and 6 years 11 months. However, the WWC cannot assess the effectiveness of the intervention with these subgroups because the authors did not respond to a request for information necessary to establish that these contrasts meet WWC group design standards with or without reservations.

9 When there is more than one single-case design experiment in a publication that does not meet WWC pilot single-case design standards, the citation list reports the disposition code that applies to the majority of single-case designs in that publication. Some single-case design experiments within a given publication might not meet WWC pilot single-case design standards for reasons other than the one listed in the citation list.


14 The results from single-case design studies are not used to report an intervention effectiveness rating for an outcome domain unless the studies collectively meet the threshold criteria described on p. 36. The evidence from the single-case design studies on PRT does not reach the threshold to include single-case design evidence in the effectiveness ratings for any of the eligible outcome domains.


Recommended Citation

**WWC Rating Criteria**

### Group design

**Criteria used to determine the rating of a group design study**

<table>
<thead>
<tr>
<th>Study rating</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meets WWC group design standards without reservations</td>
<td>A group design study that provides strong evidence for an intervention’s effectiveness, such as a well-implemented RCT.</td>
</tr>
<tr>
<td>Meets WWC group design standards with reservations</td>
<td>A group design 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.</td>
</tr>
</tbody>
</table>

**Criteria used to determine the rating of effectiveness for an intervention based on group design studies**

<table>
<thead>
<tr>
<th>Rating of effectiveness</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive effects</strong></td>
<td>Two or more studies show statistically significant positive effects, at least one of which met WWC group design standards without reservations, AND No studies show statistically significant or substantively important negative effects.</td>
</tr>
<tr>
<td><strong>Potentially positive effects</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Mixed effects</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Potentially negative effects</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Negative effects</strong></td>
<td>Two or more studies show statistically significant negative effects, at least one of which met WWC group design standards without reservations, AND No studies show statistically significant or substantively important positive effects.</td>
</tr>
<tr>
<td><strong>No discernible effects</strong></td>
<td>None of the studies shows a statistically significant or substantively important effect, either positive or negative.</td>
</tr>
</tbody>
</table>

**Criteria used to determine the extent of group design evidence for an intervention**

<table>
<thead>
<tr>
<th>Extent of evidence</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medium to large</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Small</strong></td>
<td>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.</td>
</tr>
</tbody>
</table>
### WWC Rating Criteria

#### Single-case design

Criteria used to determine the rating of a study that includes single-case design experiments

<table>
<thead>
<tr>
<th>Study rating</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meets WWC pilot single-case design standards without reservations</td>
<td>A single-case design study that provides strong evidence for assessing an intervention’s effectiveness.</td>
</tr>
<tr>
<td>Meets WWC pilot single-case design standards with reservations</td>
<td>A study that provides weaker evidence for assessing an intervention’s effectiveness, such as a reversal-withdrawal design with three or four data points per phase.</td>
</tr>
</tbody>
</table>

Criteria used to determine evidence of a causal relation in a single-case design experiment

<table>
<thead>
<tr>
<th>Evidence level</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong evidence of a causal relationship</td>
<td>A single-case design study with at least three demonstrations of the intervention effect and no non-effects.</td>
</tr>
<tr>
<td>Moderate evidence of a causal relationship</td>
<td>A single-case design study with at least three demonstrations of the intervention effect and at least one non-effect.</td>
</tr>
<tr>
<td>No evidence of a causal relationship</td>
<td>A single-case design study with fewer than three demonstrations of the intervention effect.</td>
</tr>
</tbody>
</table>

Criteria used to determine whether the body of single-case design evidence for an intervention is substantive enough to summarize as evidence of intervention effectiveness for a given domain

<table>
<thead>
<tr>
<th>Threshold to include single-case design evidence</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold met</td>
<td>At least five studies examining the intervention meet WWC pilot single-case design standards without reservations or meet WWC pilot single-case design standards with reservations, AND The single-case design studies are conducted by at least three different research teams with no overlapping author- ship at three different institutions, AND The combined number of cases (i.e., participants, classrooms, etc.) totals at least 20.</td>
</tr>
</tbody>
</table>

Criteria used to determine the rating of effectiveness for an intervention based on single-case design research

<table>
<thead>
<tr>
<th>Rating of effectiveness</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive effects</td>
<td>Across all single-case design experiments, at least 80% show positive effects, AND No single-case design experiment shows negative effects, AND At least one single-case design experiment meets WWC pilot single-case design standards without reservations.</td>
</tr>
<tr>
<td>Potentially positive effects</td>
<td>Across all the single-case design experiments, 51% to 79% show positive effects, AND No single-case design experiment shows negative effects.</td>
</tr>
<tr>
<td>Mixed effects</td>
<td>At least one single-case design experiment shows positive effects AND at least one single-case design experiment shows negative effects, OR At least one single-case design experiment shows positive or negative effects AND 50% or more show indeterminate effects.</td>
</tr>
<tr>
<td>Potentially negative effects</td>
<td>Across all the single-case design experiments, 51% to 79% show negative effects, AND No single-case design experiment shows positive effects.</td>
</tr>
</tbody>
</table>
## WWC Rating Criteria

<table>
<thead>
<tr>
<th>Rating of effectiveness</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negative effects</strong></td>
<td>Across all the single-case design experiments, at least 80% show negative effects, AND</td>
</tr>
<tr>
<td></td>
<td>No single-case design experiment shows positive effects, AND</td>
</tr>
<tr>
<td></td>
<td>At least one single-case design experiment meets WWC pilot single-case design standards without reservations.</td>
</tr>
<tr>
<td><strong>No discernible effects</strong></td>
<td>None of the single-case design experiments shows effects, either positive or negative.</td>
</tr>
</tbody>
</table>

**Notes:** A single-case design experiment has all of the design elements required to meet WWC standards with or without reservations (such as three attempts to demonstrate an effect) and is presented as one experiment in a study. The WWC characterizes all single-case design experiments in the same research article as one study, and thus one study can have multiple single-case design experiments. For example, a study could include three separate ABAB design experiments for one student (across three different eligible outcomes) or could include three separate ABAB design experiments for three separate eligible students. If a study presents data for more than one outcome, the WWC classifies the single-case design for each outcome as a separate experiment. The WWC visual analysis characterizations of *Strong* and *Moderate* evidence indicate that the design demonstrated an effect of the intervention. A visual analysis rating of *No evidence* indicates that the experiment did not provide at least three demonstrations of an intervention effect in the same direction.
**Alternating treatment design**
A single-case design experiment that repeatedly introduces and withdraws the intervention(s); each phase only lasts one or two sessions.

**Attrition**
For group design research, 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. For single-case design research, attrition can occur when an individual fails to complete all required phases of a study or the case is a group and individuals attrite from the group.

**Baseline**
In a single-case design experiment, baseline is the condition when participants are not receiving the intervention.

**Case**
A case is the unit of intervention administration and data analysis in a single-case design experiment. A case may be a single participant or a cluster of participants like a classroom.

**Clustering adjustment**
In group design research, 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 (group design) or the method by which a dependent variable was repeatedly and systematically measured before, during, and after the active manipulation of an independent variable (single-case design).

**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 group design studies and outcomes.

**Eligibility**
A determination of whether a study falls within the scope of a review protocol and uses a causal 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 from group design studies supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 35.

**Fidelity**
Fidelity indicates the extent to which the intervention, as implemented, replicates the intervention's design.

**Improvement index**
Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention, using findings from group design research. As the average individual starts at the 50th percentile, the measure ranges from −50 to +50.

**Intervention**
An educational program, product, practice, or policy aimed at improving student outcomes.
Glossary of Terms

**Intervention report**  
A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.

**Maintenance probes**  
In single-case design research, maintenance probes measure outcomes after the intervention has ended.

**Multiple baseline design**  
A single-case design that staggers the introduction of the intervention to different cases or to the same case over different settings.

**Multiple comparison adjustment**  
When a group design study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

**Multiple probe design**  
A variation on the multiple baseline single-case design that features intermittent pre-intervention data collection.

**Phase**  
In single-case design research, phases are the consecutive sessions when a case receives or does not receive the intervention.

**Quasi-experimental design (QED)**  
A quasi-experimental design (QED) is a research design in which study participants 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 eligible study participants are randomly assigned to intervention and comparison groups.

**Rating of effectiveness**  
For group design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. For single-case design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the consistency of demonstrated effects. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 35.

**Reversal-withdrawal design**  
A single-case design that introduces the intervention twice and withdraws the intervention once (also known as an ABAB design). The design may be extended by adding additional baseline and/or intervention phases.

**Single-case design (SCD) experiment**  
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% \((p < .05)\).
Glossary of Terms

Substantively important  A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Systematic review  A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and, 5) summarizing the review.

Threshold to include single-case design evidence  For single-case design studies to contribute to the evidence rating, there must be a sufficient combination of participants, authors, and studies that meet evidence standards. The criteria for the threshold to include single-case design evidence are given in the WWC Rating Criteria on p. 35.

Visual analysis  A visual analysis reviews the pattern of outcome data in a single-case design experiment to determine whether a positive effect, negative effect, or no effect is demonstrated between the intervention and the outcome.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.
An **intervention report** summarizes the findings of high-quality research on a given program, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against evidence standards, and summarizes the findings of those that meet standards.

This intervention report was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.