What Works Clearinghouse

ies INSTITUTE OF EDUCATION SCIENCES

July 2012

WWC Review of the Report "Randomized, Controlled Trial of the LEAP Model of Early Intervention for Young Children With Autism Spectrum Disorders"[1]

The findings from this review do not reflect the full body of research evidence on *Learning Experiences and Alternative Program for Preschoolers and Their Parents (LEAP).*

What is this study about?

The study examined whether a full two-year *Learning Experiences and Alternative Program for Preschoolers and Their Parents (LEAP)* training and support model for teachers was superior in improving outcomes of preschool children with Autism Spectrum Disorder (ASD), as compared to a limited *LEAP* model in which no professional development support was provided.

Fifty preschool classrooms from 16 districts and seven states across the United States were randomly assigned to either intervention or comparison conditions. Teachers in classrooms assigned to the intervention group received written *LEAP* materials, skill area presentations and demonstrations by *LEAP* trainers, onsite training and support, and maintenance checks. Comparison classroom teachers received written *LEAP* materials and videos, but no in-person training or ongoing support. In total, there were 177 children with ASD in the full *LEAP* training classrooms and 117 children with ASD in the comparison classrooms.

The study assessed the effectiveness of the full *LEAP* training and support model by comparing average classroom-level measures of behavior, communication, intellectual, and social outcomes of the two conditions after two years of implementation.[2]

WWC Rating

The research described in this report meets WWC evidence standards without reservations

Strengths: This study is a well-implemented randomized controlled trial.

Features of *LEAP*

Learning Experiences and Alternative Program for Preschoolers and Their Parents (LEAP) involves immersing children with ASD into preschool classrooms with typically-developing children. The typically-developing peers are trained on how to communicate and interact with children with ASD. Classroom adaptations and support are provided to children with ASD to facilitate this immersion. Teachers receive written materials and onsite training and support, and then collect and review data on children's progress toward generalized behavioral changes and modify the intervention as needed. Families of children with ASD also receive training in behavioral teaching strategies.

What did the study find?

The study found a positive impact after two years of the full *LEAP* training and support model on the average classroom-level measures of behavior, communication, intellectual outcomes, and social outcomes of preschool children. The average effect sizes for all four domains are statistically significant.

Appendix A: Study details

Strain, P. S., & Bovey II, E. H. (2011). Randomized, controlled trial of the LEAP model of early intervention for young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 31*(3), 133–154.

Setting

The study was conducted in 50 preschool classrooms located in 16 school districts and seven states across the United States over a period of four years. Two cohorts of classrooms participated in the study for two years each; the first cohort of classrooms participated in the first two years of the study, and the second cohort of classrooms participated in the third and fourth years of the study.

Study sample

The study recruited preschool classrooms that had previously expressed interest in the *LEAP* model. Fifty-six classrooms were selected and agreed to participate across the two cohorts. A clustered randomized comparison design was used to match classrooms on program dimensions, such as number of program days per week and adult–child ratio. Classrooms were then randomized into intervention (n = 28) and comparison (n = 28) groups, so that each district contained at least one intervention classroom and one comparison classroom.

Following random assignment, one intervention classroom and five comparison classrooms were deemed ineligible for the study because they had no children with Autism Spectrum Disorder (ASD). The analysis sample consisted of 27 intervention classrooms and 23 comparison classrooms. Across the two cohorts, a total of 177 children with ASD were included in the intervention condition, and 117 children with ASD were included in the comparison condition. Approximately 52% of the classrooms were located in metropolitan areas, 36% were in suburban areas, and 12% were located in rural areas. The average student age at baseline was 50.1 months in the intervention group and 50.7 months in the comparison group.

Intervention group

Teachers in intervention condition classrooms received written *LEAP* materials and two years of training and mentoring, including skill-area presentations and demonstrations by *LEAP* trainers, and onsite training and support. Typically-developing peers were trained on how to communicate and interact with children with ASD. Classroom adaptations and support were provided to children with ASD to facilitate their immersion. An average of 87% of *LEAP* program components were implemented in intervention group classrooms by the end of the second year of implementation.

Comparison group

Preschool teachers in classrooms assigned to the comparison condition were provided with *LEAP* intervention manuals, videos, and training presentation materials for family skills training, social skills training, and design and operation of an inclusive classroom. No mentoring or onsite training was provided to these teachers. An average of 38% of *LEAP* program components were implemented in comparison group classrooms by the end of the second year of implementation.

Outcomes and measurement

The study measured average classroom-level behavior, communication, intellectual, and social outcomes. Specific instruments used were the Childhood Autism Rating Scale (CARS), the Mullen Scales of Early Learning (Composite score and Expressive and Receptive Language scales), Preschool Language Scale, Fourth Edition (PLS-4), and the Social Skills Rating System (SSRS) Social Skills and Problem Behavior scales. For a more detailed description of these outcome measures, see Appendix B.

The authors also collected data using the visual reception and fine motor scales of the Mullen Scales of Early Learning; the results of these individual scales are not reported here because they were deemed ineligible for review by an ASD content expert. Data from these scales are included, however, in the Mullen Scales of Early Learning composite score, as this composite was deemed eligible under the intellectual outcomes domain by the ASD content expert.

Support for implementation

Teachers in intervention condition classrooms received 23 full days of *LEAP* training and coaching support over two years, including presentations, discussions, and demonstrations by *LEAP* trainers; onsite training and support including practice sessions, observations, and feedback; and maintenance checks. They also received written *LEAP* materials and manuals.

Reason for review

This study was identified for review by the WWC because it was supported by a grant to the University of Colorado, Denver (Principal Investigator: Phillip Strain) from the National Center for Special Education Research (NCSER) at the Institute of Education Sciences (IES).

Appendix B: Outcome measures for each domain

Behavior outcomes	
Average classroom-level Childhood Autism Rating Scale (CARS)	The CARS consists of 15 items used by observers to rate children on behaviors associated with autism. The scale has a short-term test-retest reliability and internal consistency of $r = 0.86$ and interrater reliability of > 80%.
Average classroom-level Social Skills Rating System (SSRS)–Problem Behavior scale	The SSRS Problem Behavior scale is a norm-referenced assessment completed by teachers to describe the frequency of a series of a child's problem behaviors. The scale includes subscales for externalizing behaviors, internalizing behaviors, and hyperactivity. The full SSRS score (the Problem Behavior scale and the Social Skills scale) has a six-week test-retest reliability of 0.90, an internal consistency of 0.96, and interrater reliability of > 80%.
Communication outcomes	
Average classroom-level Mullen Scales of Early Learning–Expressive Language scale	The Mullen Scales of Early Learning is an individually administered norm-referenced assessment that can be used with children up to 68 months old. Reliability information was supplied by Pearson (the company that produces the assessment) at the request of the WWC; the internal consistency of the Expressive Language scale ranges from 0.80 to 0.83 for children 33 months old through 50 months old.
Average classroom-level Mullen Scales of Early Learning–Receptive Language scale	The Mullen Scales of Early Learning is an individually administered norm-referenced assessment that can be used with children up to 68 months old. Reliability information was supplied by Pearson (the company that produces the assessment) at the request of the WWC; the internal consistency of the Receptive Language scale ranges from 0.73 to 0.83 for children 33 months old through 50 months old.
Average classroom-level Preschool Language Scale, Fourth Edition (PLS-4)	The PLS-4 is an individually administered norm-referenced assessment that measures both expressive and receptive communication and can discriminate among ASD, communication disorder, and hearing impairment. The scale has a test-retest reliability and internal consistency exceeding 0.90 and an interrater reliability of > 80%.
Intellectual outcomes	
Average classroom-level Mullen Scales of Early Learning—Composite score	The Mullen Scales of Early Learning is an individually administered norm-referenced assessment that can be used with children up to 68 months old. The Composite score comprises the expressive language, receptive language, visual reception, and fine motor scales; it has a short-term test-retest reliability of 0.95, internal consistency of 0.91, and interrater reliability of > 80%.
Social outcomes	
Average classroom-level SSRS—Social Skills scale	The SSRS Social Skills scale is a norm-referenced assessment completed by teachers to describe the frequency with which a child demonstrates cooperation, assertion, self-control, empathy, and responsibility. The full SSRS score (the Problem Behavior scale and the Social Skills scale) has a six-week test-retest reliability of 0.90, an internal consistency 0.96, and interrater reliability of > 80%.

Table Notes: The authors also collected data using the visual reception and fine motor scales of the Mullen Scales of Early Learning; the results of these individual scales are not reported here because they were deemed ineligible for review by an ASD content expert. Data from these scales are included, however, in the Mullen Scales of Early Learning composite score, as this composite was deemed eligible under the intellectual outcomes domain by the ASD content expert.

Appendix C: Study findings for each domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Behavior outcomes								
Average classroom-level Childhood Autism Rating Scale (CARS)	Preschool	50 classrooms/ 294 students	31.30 (3.90)	34.60 (4.20)	3.30	0.80	+29	< 0.05
Average classroom-level Social Skills Rating System–Problem Behavior scale	Preschool	50 classrooms/ 294 students	46.40 (4.20)	49.10 (4.10)	2.70	0.64	+24	< 0.05
Domain average for behavior outcomes						**0.72**	**+26**	**Statistically significant**
Communication outcomes								
Average classroom-level Mullen Scales of Early Learning–Expressive Language scale	Preschool	50 classrooms/ 294 students	40.10 (6.40)	35.90 (4.40)	4.20	0.74	+27	< 0.05
Average classroom-level Mullen Scales of Early Learning–Receptive Language scale	Preschool	50 classrooms/ 294 students	51.90 (7.90)	40.70 (7.70)	11.20	1.41	+42	< 0.01
Average classroom-level Preschool Language Scale, Fourth Edition (PLS-4)	Preschool	50 classrooms/ 294 students	52.90 (8.10)	43.80 (7.70)	9.10	1.13	+37	< 0.01
Domain average for communication outcomes						**1.09**	**+36**	**Statistically significant**
Intellectual outcomes								
Average classroom-level Mullen Scales of Early Learning–Composite score	Preschool	50 classrooms/ 294 students	72.10 (7.50)	61.40 (9.00)	10.70	1.28	+40	< 0.01
Domain average for intellectual outcomes						**1.28**	**+40**	**Statistically significant**
Social outcomes								
Average classroom-level Social Skills Rating System–Social Skills scale	Preschool	50 classrooms/ 294 students	49.30 (12.60)	32.70 (11.90)	16.60	1.33	+41	< 0.01
Domain average for social outcomes						**1.33**	**+41**	**Statistically significant**

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. For the average classroom-level Childhood Autism Rating Scale (CARS) and average classroom-level Social Skills Rating System–Problem Behavior scale, signs were reversed on the mean differences to demonstrate that the intervention group was favored when negative differences were reported. The effect size reported here is a standardized measure of the effect of the intervention on classroom-level outcomes, representing the change (measured in standard deviations) in an average classroom's outcome that can be expected if the classroom is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average classroom's percentile rank that can be expected if the classroom is given the intervention. Because the effect sizes and improvement indexes are based on classroom-level data, they may not be directly comparable to estimates derived from student-level data. The WWC-computed average improvement index is a simple average rounded to two decimal places. The statistical significance of each domain average was determined by the WWC; the study is characterized as having a statistically significant positive effect in the behavior, communication, intellectual, and social outcomes domains because univariate statistical tests are reported for each outcome measure, the effects for all measures within each domain are positive and statistically significant, and no effects are negative and statistically significant.

Study Notes: Multivariate repeated measures analyses of variance were performed on clusters of classrooms, rather than individual child data. The means, standard deviations, and effect sizes presented here are based on classroom-level data across two cohorts. The WWC calculated the intervention group mean by adding the difference-in-differences adjusted estimate of the average impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group post-tests means. Please see the *WWC Handbook* for more information. Corrections for multiple comparisons were needed for the behavior and communication outcome domains but did not affect significance levels. The *p*-value ranges presented here were reported in the original study.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author[s]) to assess whether the study design meets WWC evidence standards. The review reports the WWC's assessment of whether the study meets WWC evidence standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the Single Study review protocol, version 2.0. The WWC rating applies only to the results that were eligible under this topic area and met WWC standards without reservations or met WWC standards with reservations, and not necessarily to all results presented in the study.

[2] Two additional outcomes were examined in this study, but are not included in this report because an ASD content expert deemed that they were ineligible for review: the visual reception and fine motor scales of the Mullen Scales of Early Learning.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2012, July). *WWC review of the report: Randomized, controlled trial of the LEAP model of early intervention for young children with autism spectrum disorders*. Retrieved from http://whatworks.ed.gov.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

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

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

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

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