Washington State Institute for Public Policy
110 Fifth Avenue SE, Suite 214 ● PO Box 40999 ● Olympia, WA 98504 ● 360.664.9800 ● www.wsipp.wa.gov

June 2018

Updated Inventory of Evidence- and Research-Based Practices: *Washington's K–12 Learning Assistance Program*

Washington State provides funding to school districts to help underachieving students through the Learning Assistance Program (LAP).[1] The 2013 Washington State Legislature directed three efforts to identify effective practices for helping students served through LAP.

- ✓ The Washington State Institute for Public Policy (WSIPP) was directed to "prepare an inventory of evidence-based and research-based effective practices, activities, and programs for use by school districts in the learning assistance program."[2] The updated inventory is displayed starting on page 10 of this report and is also available online.[3]
- ✓ The Office of Superintendent of Public Instruction (OSPI) was directed to "convene a panel of experts, including the Washington state institute for public policy, to develop additional state menus of best practices and strategies for use in the learning assistance program to assist struggling students at all grade levels in English language arts and mathematics and reduce disruptive behaviors in the classroom."[4]
- ✓ OSPI was also directed to "convene a panel of experts, including the Washington state institute for public policy, to develop a state menu of best practices and strategies for intensive reading and literacy improvement designed to assist struggling students in reaching grade level in reading by the end of fourth grade."[5]

The table below summarizes the implementation timeline.

Legislative assignment	Due date	Follow up
WSIPP to develop a LAP inventory of evidence- and research-based practices, activities, and programs	August 1, 2014	Update every two years thereafter*
OSPI to convene a panel of LAP experts to develop a menu of best practices and strategies	July 1, 2015	Update each July 1st thereafter
OSPI to convene a panel of English language arts (ELA) experts to develop a menu of best practices and strategies to help students reach grade level in reading by the end of 4th grade	July 1, 2014	Update each July 1st thereafter

Note:
*WSIPP updated the LAP inventory in July 2015 to align with OSPI's menu timeline.

[1] The Office of Superintendent of Public Instruction's website for the Learning Assistance Program.
[2] Engrossed Substitute Senate Bill 6002, Chapter 221, Sec. 609(3), Laws of 2014.
[3] Cramer, J., Bitney, K., & Wanner, P. (2018). *Updated inventory of evidence- and research-based practices: Washington's K–12 Learning Assistance Program* (Doc. No. 18-06-2201). Olympia: Washington State Institute for Public Policy.
[4] Engrossed Substitute Senate Bill 5946, Chapter 18, Sec. 206(3), Laws of 2013.
[5] Ibid.

These legislative assignments are similar but distinct. Staff members from WSIPP and OSPI coordinate the inventory and menu processes to ensure that the results are consistent. WSIPP staff serve as non-voting members on the expert panels and provide research findings that inform the panels' deliberations. The expert panels help WSIPP identify topics for analysis and review topics for inclusion on the inventory when no evaluation research is available. The final classifications on the LAP inventory and expert panel menus, however, reflect each group's independent judgment.

The expert panel's menu of best practices is presented in a separate OSPI report, which can be found on OSPI's website.[6]

Creating the LAP Inventory

Washington State's Learning Assistance Program was created by the legislature in 1987 to assist underachieving students. In the 2017-19 biennium, the legislature appropriated $671 million for LAP.[7] The funds can be used for a variety of practices, strategies, and activities in K–12 schools, including:

- Tutoring support,
- Extended learning time,
- Professional development,
- Consultant teachers,
- Parent outreach,
- Community-based partnerships,
- Addressing disruptive behavior in the classroom, and
- Services for 8th, 11th, and 12th graders.[8]

WSIPP consulted with legislative staff, OSPI, and members of the expert panel to develop a list of the highest priority topics to investigate for this inventory. To date, we have reviewed and included 58 interventions.

[6] http://www.k12.wa.us/SSEO/Menus.aspx.
[7] Engrossed Substitute Senate Bill 6032, Chapter 299, Sec. 515, Laws of 2018.
[8] RCW 28A.165.035.

Our approach to creating the inventory is the same approach we use for legislatively directed inventories in other policy areas.[9] We first use a rigorous, three-step research approach to assess the evidence, economics, and risk for each program. Then, using information derived from the three-step approach, we classify all programs according to standard definitions.
WSIPP's three-step approach is as follows:

1) Identify what works (and what does not). For each program under consideration, we systematically review all rigorous research evidence and estimate the program's effect on a desired outcome or set of outcomes like high school graduation rates or student test scores.[10] The evidence may indicate that a program worked (i.e., had a desirable effect on outcomes), caused harm (i.e., had an undesirable effect on outcomes), or had no detectable effect one way or the other.

2) Assess the return on investment. Given the estimated effect of a program from Step 1, we estimate—in dollars and cents—how much the program would benefit people in Washington were it implemented and how much it would cost the taxpayers to achieve this result. We use WSIPP's benefit-cost model to develop standardized, comparable results for all programs that illustrate the expected returns on investment. We present these results as net present values on a per-participant basis. We also consider how monetary benefits are distributed across program participants, taxpayers, and other people in society.

3) Determine the risk of investment. We allow for uncertainty in our estimates by calculating the probability that a program will at least "break even" if critical factors—like the actual cost to implement the program and the precise effect on the program—are lower or higher than our estimates.

We follow a set of standardized procedures (see Exhibit 1) for each of these steps. These standardized procedures support the rigor of our analyses and allow programs to be compared on an "apples-to-apples" basis. For full detail on WSIPP's methods, see WSIPP's Technical Documentation.[11]

[9] EBPI, & WSIPP. (2017). Updated inventory of evidence-based, research-based, and promising practices: For prevention and intervention services for children and juveniles in the child welfare, juvenile justice, and mental health systems. (Doc. No. E2SHB2536-8). Olympia: Washington State Institute for Public Policy.
[10] For the inventory, we looked for studies measuring outcomes related to the goals of LAP (to assist underachieving students and reduce disruptive behaviors in the classroom—RCW 28A.165.005). For example, we included studies that measured changes in test scores, graduation rates, grade point average, attendance, and suspensions/expulsions. We did not include studies that measured outcomes that may or may not be related to the change in students' educational outcomes (such as teacher or student satisfaction) if the studies did not also measure the outcomes of interest.
[11] Washington State Institute for Public Policy (December 2017). *Benefit-cost technical documentation.* Olympia, WA: Author.

Exhibit 1

WSIPP's Three-Step Approach

Step 1: Identify what works (and what does not)

We conduct a meta-analysis—a quantitative review of the research literature—to determine if the weight of the research evidence indicates whether desired outcomes are achieved, on average.

WSIPP follows several key protocols to ensure a rigorous analysis for each program examined. We:

- Search for all studies on a topic—We systematically review the national and international research literature and consider all available studies on a program, regardless of their findings. That is, we do not "cherry pick" studies to include in our analysis.
- Screen studies for quality—We only include rigorous studies in our analysis. We require that a study reasonably attempt to demonstrate causality using appropriate statistical techniques. For example, studies must include both treatment and comparison groups with an intent-to-treat analysis. Studies that do not meet our minimum standards are excluded from analysis.
- Determine the average effect size—We use a formal set of statistical procedures to calculate an average effect size for each outcome, which indicates the expected magnitude of change caused by the program (e.g., tutoring by adults) for each outcome of interest (e.g., standardized test scores).

Step 2: Assess the return on investment

WSIPP has developed, and continues to refine, an economic model to provide internally consistent monetary valuations of the benefits and costs of each program on a per-participant basis.

Benefits to individuals and society may stem from multiple sources. For example, a program that reduces the need for publicly funded health care services decreases taxpayer costs. If that program also improves participants' educational outcomes, it will increase their expected labor market earnings. Finally, if a program reduces crime, it will also reduce expected costs to crime victims.

We also estimate the cost required to implement an intervention. If the program is operating in Washington State, our preferred method is to obtain the service delivery and administrative costs from state or local agencies. When this approach is not possible, we estimate costs using the research literature, using estimates provided by program developers, or using a variety of sources to construct our own cost estimate.

Step 3: Determine the risk of investment

Any tabulation of benefits and costs involves a degree of uncertainty about the inputs used in the analysis, as well as the bottom-line estimates. An assessment of risk is expected in any investment analysis, whether in the private or public sector.

To assess the riskiness of our conclusions, we look at thousands of different scenarios through a Monte Carlo simulation. In each scenario we vary a number of key factors in our calculations (e.g., expected effect sizes, program costs), using estimates of error around each factor. The purpose of this analysis is to determine the probability that a particular program or policy will produce benefits that are equal to or greater than costs if the real-world conditions are different than our baseline assumptions.

Classifying Practices as Evidence-Based, Research-Based, or Promising

The legislative assignment directs WSIPP to identify evidence- and research-based practices for LAP. Washington's K–12 laws do not define these terms. The adult behavioral health statutes, however, do provide definitions, and WSIPP recently published an adult behavioral health inventory using these definitions.[12] WSIPP has also used these statutory definitions to guide classifications for inventories in the areas of children's services, adult corrections, and cannabis prevention. For the LAP inventory, we use the same definitions to maintain consistency across policy areas (see Exhibit 2).

Some programs are classified as "promising practices" when the OSPI-convened expert panel and/or the research evidence suggest the practice might improve student outcomes, but the topics did not meet the criteria for classification as evidence- or research-based.

Exhibit 2

Legislative Definitions of Evidence-Based, Research-Based, and Promising Practices

Evidence-based

A program or practice that has been tested in heterogeneous or intended populations with multiple randomized, or statistically controlled evaluations, or both; or one large multiple site randomized, or statistically controlled evaluation, or both, where the weight of the evidence from a systemic review demonstrates sustained improvements in at least one outcome. "Evidence-based" also means a program or practice that can be implemented with a set of procedures to allow successful replication in Washington and, when possible, is determined to be cost-beneficial.

Research-based

A program or practice that has been tested with a single randomized, or statistically controlled evaluation, or both, demonstrating sustained desirable outcomes; or where the weight of the evidence from a systemic review supports sustained outcomes [. . .] but does not meet the full criteria for evidence-based.

Promising practice

A practice that, based on statistical analyses or a well-established theory of change, shows potential for meeting the evidence-based or research-based criteria [. . .].

RCW 71.24.025

[12] RCW 71.24.025. WSIPP's adult behavioral health inventory can be found on our website; EBPI & WSIPP (2017).

For each program where research is available, we use the results of our meta-analysis (Step 1) and benefit-cost analysis (Steps 2 & 3) to inform classifications. To assemble the inventory, we operationalize each criterion in the statutory definitions. These are the same criteria WSIPP has used in assembling inventories in children's mental health, child welfare, juvenile justice, and adult behavioral health. The criteria are as follows:

- Heterogeneity. To be designated as evidence-based, the state statute requires that a program has been tested on a "heterogeneous" population. We operationalize heterogeneity in two ways. First, the proportion of program participants belonging to racial and ethnic minority groups must be greater than or equal to the proportion of minority children aged 0 to 17 in Washington. From the 2010 Census, among Washington children aged 0 through 17, 68% were white and 32% were from minority backgrounds.[13] Thus, if at least an average[14] of 32% of program participants in the outcome evaluations are from minority backgrounds, then the program is considered to have been tested in a heterogeneous population.

  Second, the heterogeneity criterion can also be achieved if at least one of a program's outcome evaluations was conducted with K–12 students in Washington and a subgroup analysis demonstrates the program is effective for racial and ethnic minority populations.

  Programs whose evaluations do not meet either of these two criteria do not meet the heterogeneity definition.
- Weight of evidence. We use the results of our meta-analysis from Step 1 to evaluate this criterion. To meet the evidence-based definition, results from a random effects meta-analysis of multiple evaluations or one large multiple-site evaluation must indicate the practice achieves the desired outcome(s) (p-value < 0.20).[15] To meet the research-based definition, one single-site evaluation must indicate the practice achieves desired outcomes (p-value < 0.20).
- Benefit-cost. The statute defining evidence-based practices requires that, when possible, a benefit-cost analysis be conducted. We use the benefit-cost analysis from Steps 2 and 3 to evaluate this criterion.[16] The WSIPP model uses Monte Carlo simulation to test the probability that benefits will exceed costs. Programs must have at least a 75% chance of a positive net present value to be classified as evidence-based.

When we locate no rigorous outcome evaluations for a program, we rely on the panel of experts assembled by OSPI to determine whether the program meets the criteria for promising. If a program is not listed on the inventory, we have not yet had the opportunity to review it or the program does not meet criteria to be promising.

[13] United States Census Bureau.
[14] We operationalize this as a weighted average across outcome evaluations included in the meta-analysis.
[15] In order to operationalize the benefit-cost criterion for a program to be classified as evidence-based, net benefits must exceed costs at least 75% of the time. After considerable analysis, we found that a typical program that WSIPP has analyzed may produce benefits that exceed costs roughly 75% of the time with a p-value cut-off of up to 0.20. Thus, we determined that programs with p-values < 0.20 on desired outcomes should be considered research-based in order to avoid classifying programs with desirable benefit-cost results as promising.
[16] WSIPP (2017).

Additionally, WSIPP has clarified classifications for programs that produce null or poor results since the last inventory update. In prior inventories, there was a single category for programs producing "null or poor outcomes." Programs with null effects on outcomes (i.e., p-value > 0.20) were inconsistently categorized as either "null or poor" or as "promising." For the current inventory, WSIPP has defined two separate categories to distinguish between programs producing null results (no significant effect on desired outcomes) and those producing poor (undesirable) outcomes and has standardized the application of these definitions. A program is designated as "null" if evaluations exist and indicate an effect on outcomes in the desired direction, but the current body of evidence also indicates the effect of the program cannot be statistically distinguished from zero (i.e., p-value > 0.20).[17] A program is designated as having "poor outcomes" if evaluations exist and indicate a statistically significant (p-value < 0.20) and undesirable effect on outcomes.[18]

The LAP inventory is displayed at the end of this report and is also available on our website.[19] Further information on the individual programs contained in the inventory can also be found on our website.[20]

Updates to the Inventory as of June 2018

Since the last inventory update in July 2016, WSIPP has reviewed previously examined programs that were due for updates and reviewed several programs-of-interest nominated by OSPI's panel of experts. Exhibit 3 provides an overview of programs with updated classifications and the reasons for classification changes. Exhibit 4 provides an overview of new programs and their classifications.

[17] We estimate results from a random-effects meta-analysis of multiple evaluations.
[18] We estimate results from a random-effects meta-analysis of multiple evaluations or one large multiple-site evaluation.
[19] http://wsipp.wa.gov/ReportFile/1687.
[20] http://www.wsipp.wa.gov/BenefitCost.

Exhibit 3

Classifications Revised Due to New "Null" Designation, New Research, or Correction to Prior Classification

Program name	Prior classification	Current classification	Reason for classification change
Tutoring support			
Tutoring: By adults for English language learner students	Promising	Null	Revised null definition
Tutoring: Supplemental computer-assisted instruction for struggling readers	Promising	Null	Revised null definition
Extended learning time			
Summer book programs: One-year, with additional support	Promising	Null	Revised null definition
Summer book programs: One-year intervention	Promising	Null	Revised null definition
Professional development			
Teacher professional development: Induction/mentoring	Promising	Null	Revised null definition
Teacher professional development: Not targeted	Null or poor outcomes	Null	Revised null definition
Educator professional development: Use of data to guide instruction ("train the trainers")	Null or poor outcomes	Null	Revised null definition
Parent outreach			
Families and Schools Together (FAST)	Research-based	Null	Included new research
Conjoint behavioral consultation	Promising	Null	Included new research
Behavior support			
Becoming a Man (BAM)	Research-based	Evidence-based	Corrected prior classification
School-Wide Positive Behavioral Interventions and Supports (SWPBIS)	Evidence-based	Research-based	Included new research
Responsive Classroom	Promising	Null	Revised null definition

Exhibit 4

New Program Classifications

Program name	Classification
Behavior support	
Mentoring: Community-based for children with disruptive behavior disorders	Research-based
Mentoring: School-based by teachers or staff	Research-based
Caring School Community (formerly Child Development Project)	Null
Other	
Growth mindset interventions	Research-based

Limitations

The benefit-cost analyses in this report reflect only those outcomes that were measured in the studies we reviewed. We focus primarily on outcomes that are "monetizable" with the current WSIPP benefit-cost model. "Monetizable" means that we can link the outcome to future economic consequences, such as labor market earnings, criminal justice involvement, or health care expenditures. At this time we are unable to monetize some relevant outcomes including suspensions/expulsions and attendance.[21]

Future Updates

The next update to this inventory will be published by July 1, 2020.

Acknowledgments

We would like to thank staff at the Office of Superintendent of Public Instruction and the Panels of Experts convened by OSPI for their guidance in identifying interventions to review and assistance in describing specific programs. We are especially grateful to Kristi Coe and Joshua Lynch for their input and collaboration throughout this process.

Suggested citation: Cramer, J., Bitney, K., & Wanner, P. (2018). *Updated inventory of evidence- and research-based practices: Washington's K–12 Learning Assistance Program.* (Document Number 18-06-2201). Olympia: Washington State Institute for Public Policy.

[21] We report meta-analytic results for non-monetizable outcomes like attendance on our website.

June 2018
Updated Inventory of Evidence- and Research-Based Practices:
Washington's Learning Assistance Program

Program/intervention	Level of evidence	Benefit-cost percentage	Reason program does not meet evidence-based criteria (see full definitions below)	Percent minority
Tutoring support				
Tutoring: By certificated teachers, small-group, structured	●	96%		63%
Tutoring: By adults, one-on-one, structured	●	94%		72%
Tutoring: By non-certificated adults, small-group, structured	●	78%		69%
Tutoring: By peers, cross-age[#]	⦿	81%	Heterogeneity	NR
Tutoring: By peers, same-age and classwide[#]	⦿	74%	Benefit-cost	62%
Tutoring: By adults, one-on-one, non-structured	⦿	71%	Benefit-cost	75%
Tutoring: Supplemental Educational Services (under Title I)	⦿	58%	Benefit-cost	95%
Tutoring: By adults, for English language learner students	**Null**	69%	Weight of evidence	91%
Tutoring: Supplemental computer-assisted instruction for struggling readers	**Null**	58%	Weight of evidence	91%
Extended learning time				
Double-dose classes	●	98%		91%
Out-of-school-time tutoring by adults	●	93%		75%
Summer learning programs: Academically focused	●	86%		85%
Summer book programs: One-year, with additional support	**Null**	57%	Weight of evidence	77%
Summer book programs: One-year intervention	**Null**	56%	Weight of evidence	86%
Summer book programs: Multi-year intervention	P		Weight of evidence	95%
Professional development				
Teacher professional development: Use of data to guide instruction	●	98%		54%
Teacher professional development: Targeted	●	79%		96%
Teacher professional development: Online, targeted	⦿	61%	Benefit-cost/heterogeneity	31%
Teacher professional development: Induction/mentoring	**Null**	64%	Weight of evidence	92%
Teacher professional development: Not targeted	**Null**	35%	Weight of evidence	51%
Educator professional development: Use of data to guide instruction ("train the trainers")	**Null**	31%	Weight of evidence	46%
Professional learning communities	P		No rigorous evaluation with outcome of interest	
Consultant teachers				
Consultant teachers: Online coaching	●	92%		53%
Consultant teachers: Coaching	●	81%		53%
Consultant teachers: Coaching: Literacy Collaborative	⦿	100%	Heterogeneity	29%
Consultant teachers: Content-Focused Coaching	⦿		Single evaluation	96%

● Evidence-based ⦿ Research-based **P** Promising ⦸ Poor outcomes **Null** Null outcomes **NR** Not reported See definitions and notes on page 13.

Note:
This program is a special analysis for the purpose of this inventory and does not have a program-specific webpage on WSIPP's website.

June 2018
Updated Inventory of Evidence- and Research-Based Practices:
Washington's Learning Assistance Program

Program/intervention	Level of evidence	Benefit-cost percentage	Reason program does not meet evidence-based criteria (see full definitions below)	Percent minority
Parent outreach				
Parents as tutors with teacher oversight	⦿	56%	Benefit-cost	58%
Families and Schools Together (FAST)	**Null**	46%	Weight of evidence	83%
Conjoint behavioral consultation	**Null**	4%	Weight of evidence	21%
Parent and family engagement coordinators	P		No rigorous evaluation with outcome of interest	
Community partnerships				
Case management in schools	⦿	96%	Mixed results	61%
Mentoring: School-based (taxpayer costs only)#	⦿	16%	Benefit-cost	74%
Mentoring: School-based (with volunteer costs)#	⦿	13%	Benefit-cost	74%
Mentoring: Community-based (taxpayer costs only)#	⦿	70%	Benefit-cost	68%
Mentoring: Community-based (with volunteer costs)#	⦿	64%	Benefit-cost	68%
PROSPER	⦿	59%	Benefit-cost/heterogeneity	15%
Behavior support				
Positive Action	●	87%		63%
Becoming a Man (BAM)	●	79%		98%
Mentoring: Community-based for children with disruptive behavior disorders	⦿	78%	Heterogeneity	7%
Mentoring: School-based by teachers or staff	⦿	74%	Benefit-cost	86%
Becoming a Man (BAM) with high-dosage tutoring	⦿		Single evaluation	99%
Good Behavior Game	⦿	70%	Benefit-cost	50%
School-Wide Positive Behavioral Interventions and Supports (SWPBIS)	⦿	68%	Mixed results/Benefit-cost	50%
Behavioral Monitoring and Reinforcement Program (BMRP)	⦿	64%	Benefit-cost	41%
Coping Power Program	⦿	55%	Benefit-cost	80%
First Step to Success	⦿	53%	Benefit-cost	59%
"Check-in" behavior interventions	⦿	46%	Benefit-cost	72%
Second Step	⦿	30%	Benefit-cost	55%
Fast Track prevention program	⦿	0%	Benefit-cost	53%
Daily Behavior Report Cards	⦿		Single evaluation	13%
Caring School Community (formerly Child Development Project)	**Null**	61%	Weight of evidence	47%
Responsive Classroom	**Null**	49%	Weight of evidence	57%
Curriculum-based Support Group (CBSG)	P		Single evaluation	90%

● Evidence-based ⦿ Research-based **P** Promising ⦸ Poor outcomes **Null** Null outcomes **NR** Not reported See definitions and notes on page 13.

Note:
This program is a special analysis for the purpose of this inventory and does not have a program-specific webpage on WSIPP's website.

June 2018
Updated Inventory of Evidence- and Research-Based Practices:
Washington's Learning Assistance Program

Program/intervention	Level of evidence	Benefit-cost percentage	Reason program does not meet evidence-based criteria (see full definitions below)	Percent minority
Services for 8th, 11th, & 12th grades				
Credit retrieval	P		No rigorous evaluation with outcome of interest	
Other				
Special literacy instruction for English language learner students	●	80%		98%
Growth mindset interventions	◉	58%	Benefit-cost	71%
Academic vocabulary instruction	P		Weight of evidence	NR
Transition programs for incoming kindergarteners	P		Single evaluation	45%

● Evidence-based ◉ Research-based **P** Promising ⊘ Poor outcomes **Null** Null outcomes **NR** Not reported See definitions and notes on page 13.

Note:
This program is a special analysis for the purpose of this inventory and does not have a program-specific webpage on WSIPP's website.

June 2018
Updated Inventory of Evidence- and Research-Based Practices:
Washington's Learning Assistance Program

Definitions and Notes:

Reasons Programs May Not Meet Suggested Evidence-Based Criteria:

Benefit-cost: The proposed definition of evidence-based practices requires that, when possible, a benefit-cost analysis be conducted. We use WSIPP's benefit-cost model to determine whether a program meets this criterion. Programs that do not have at least a 75% chance of a positive net present value do not meet the benefit-cost test. The WSIPP model uses Monte Carlo simulation to test the probability that benefits exceed costs. The 75% standard was deemed an appropriate measure of risk aversion.

Heterogeneity: To be designated as evidence-based, the state statute requires that a program has been tested on a "heterogeneous" population. We operationalize heterogeneity in two ways. First, the proportion of program participants belonging to ethnic/racial minority groups must be greater than or equal to the proportion of minority children aged 0 to 17 in Washington. From the 2010 Census, for children aged 0 through 17 in Washington, 68% were white and 32% were from minority backgrounds. Thus, if the weighted average of program participants in the outcome evaluations of the program is at least 32% ethnic/racial minority, then the program is considered to have been tested in a heterogeneous population.

Second, the heterogeneity criterion can also be achieved if at least one of the studies has been conducted on youth in Washington and a subgroup analysis demonstrates the program is effective for minorities ($p < 0.20$). Programs passing the second test are marked with a ^. Programs that do not meet either of these two criteria do not meet the heterogeneity definition.

Programs whose evaluations do not meet either of these two criteria do not meet the heterogeneity definition.

Mixed results: If findings are mixed from different measures (e.g., undesirable outcomes for behavior measures and desirable outcomes for test scores), the program does not meet evidence-based criteria.

No rigorous evaluation with outcome of interest: The program has not yet been tested with a rigorous outcome evaluation.

Single evaluation: The program does not meet the minimum standard of multiple evaluations or one large multiple-site evaluation contained in the current or proposed definitions.

Weight of evidence: To meet the evidence-based definition, results from a random effects meta-analysis (p-value < 0.20) of multiple evaluations or one large multiple-site evaluation must indicate the practice achieves the desired outcome(s). To meet the research-based definition, one single-site evaluation must indicate the practice achieves the desired outcomes (p-value < 0.20).

Level of Evidence:

Evidence-based: A program or practice that has been tested in heterogeneous or intended populations with multiple randomized and/or statistically-controlled evaluations, or one large multiple-site randomized and/or statistically-controlled evaluation, where the weight of the evidence from a systematic review demonstrates sustained improvements in at least one outcome. Further, "evidence-based" means a program or practice that can be implemented with a set of procedures to allow successful replication in Washington and, when possible, has been determined to be cost-beneficial.

Research-based: A program or practice that has been tested with a single randomized and/or statistically-controlled evaluation demonstrating sustained desirable outcomes; or where the weight of the evidence from a systematic review supports sustained outcomes as identified in the term "evidence-based" in RCW (the above definition) but does not meet the full criteria for "evidence-based."

Promising practice: A program or practice that, based on statistical analyses or a well-established theory of change, shows potential for meeting the "evidence-based" or "research-based" criteria, which could include the use of a program that is evidence-based for outcomes other than the alternative use.

Other Definitions:

Benefit-cost percentage: The percent of the time where the monetary benefits exceed costs, according to a Monte Carlo simulation run by WSIPP.

Null outcome(s): If desired results from multiple evaluations are not statistically significant ($p > 0.20$), a program is classified as "Null".

Poor outcome(s): If results from multiple evaluations or one large multiple-site evaluation indicate that a program produces undesirable effects ($p < 0.20$), a program is classified as producing "poor outcomes."

For further information, contact:
Julia Cramer at 360.664.9073, julia.cramer@wsipp.wa.gov

Document No. 18-06-2201

Washington State Institute for Public Policy

The Washington State Legislature created the Washington State Institute for Public Policy in 1983. A Board of Directors—representing the legislature, the governor, and public universities—governs WSIPP and guides the development of all activities. WSIPP's mission is to carry out practical research, at legislative direction, on issues of importance to Washington State.