

July 2020

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

Washington State provides funding to school districts to help provide supplemental services to underachieving students in the Learning Assistance Program (LAP).<sup>1</sup> In 2013, the Washington State Legislature directed 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.”<sup>2</sup>
- ✓ 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.”<sup>3</sup>
- ✓ 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.”<sup>4</sup>

Exhibit 1 summarizes when each of the legislative efforts started and the timeline for updates.

### Exhibit 1

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

Note:

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

<sup>1</sup> The Office of Superintendent of Public Instruction's website for the Learning Assistance Program.

<sup>2</sup> Engrossed Substitute Senate Bill 6002, Chapter 221, Sec. 609(3), Laws of 2014.

<sup>3</sup> Engrossed Substitute Senate Bill 5946, Chapter 18, Sec. 206(3), Laws of 2013.

<sup>4</sup> Ibid.

The documents created from the legislative direction—WSIPP’s inventory and OSPI’s menus of best practices<sup>5</sup>—are similar yet distinct. Staff members from WSIPP and OSPI coordinate efforts to ensure that results are consistent between the inventory and menus. WSIPP staff serve as non-voting members on the expert panels and provide research findings that inform the panel members’ deliberations about programs and classifications to include in the menus. OSPI’s LAP panel experts help WSIPP identify topics for analysis and review topics for inclusion on the inventory when no other research is available. The final classifications on WSIPP’s inventory and OSPI’s menus of best practices reflect each group’s independent judgment.

[Section I](#) of this report provides a high-level overview of Washington’s Learning Assistance Program and describes our approach for creating the inventory, including how we synthesize research evidence and conduct benefit-costs analyses. [Section II](#) describes how we determine program classifications, and [Section III](#) describes how classifications can change over time. [Section IV](#) lists updates to the current inventory.

The complete inventory is displayed at the end of this report and is available on our website.<sup>6</sup> Further information on individual programs in the inventory can also be found on our website.<sup>7</sup>

## I. Creating the LAP Inventory

---

Washington State’s Learning Assistance Program (LAP) was created by the legislature in 1987 to support students who score below grade-level standards in English Language Arts (ELA) and mathematics. In 2020, the legislature appropriated \$847.5 million for LAP for school years 2019-20 and 2020-21.<sup>8</sup> In 2018-19, the school year in which the most recent data is available, 174,535 students received LAP services (15% of the statewide population).<sup>9</sup>

In accordance with state statute,<sup>10</sup> funds are used for a variety of practices, strategies, and activities in K–12 schools, including the following:

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

---

<sup>5</sup> OSPI’s LAP expert panelists develop three menus focused on math, English Language Arts, and student behavior. Menus can be found on OSPI’s website [Menus of Best Practices and Strategies](#). Districts are required to select programs or practices from these menus in order to serve students in LAP. Districts can use alternative programs or practices not in the menu if they provide data that shows the program or practice increases student growth (RCW 28A.165.035).

<sup>6</sup> Cramer, J., & Wanner, P. (2020). *Updated inventory of evidence-based and research-based practices: Washington’s K–12 Learning Assistance Program Inventory*. Olympia: Washington State Institute for Public Policy.

<sup>7</sup> <http://www.wsipp.wa.gov/BenefitCost>

<sup>8</sup> Engrossed Substitute Senate Bill 6168, Chapter 357, Sec. 515, Laws of 2020.

<sup>9</sup> E. Seely, Data Analyst, Student Information, Office of Superintendent of Public Instruction (personal communication, May 21, 2020).

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 61 programs and practices.

### [WSIPP's Standard Approach to Meta-Analysis & Benefit-Cost Analysis](#)

Our approach to creating the inventory is the same approach we use for legislatively directed inventories in other policy areas.<sup>11</sup> We first use a rigorous, three-step research process to assess the evidence, economics, and risk for each program. Then, using information from this process, we classify programs according to standard definitions. WSIPP's three-step approach follows:

- 1) **Identify what works (and what does not).** We systematically review all rigorous research evidence and estimate the program's effect on the desired outcome or set of outcomes like high school graduation rates or student test scores.<sup>12</sup> 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 assess the riskiness of our conclusions by calculating the probability that a program will at least "break even" if critical factors—like the 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 2](#)) for each step. These standardized procedures support the rigor of our analyses and allow programs to be compared on an "apples-to-apples" basis. See WSIPP's [Technical Documentation](#) for additional information.<sup>13</sup>

---

<sup>11</sup> EBPI, & WSIPP. (2019). *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. E2SHB2536-10). Olympia: Washington State Institute for Public Policy.

<sup>12</sup> 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.

<sup>13</sup> Washington State Institute for Public Policy (December 2019). *Benefit-cost technical documentation.* Olympia, WA: Author.

## Exhibit 2

### 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.

## 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.<sup>15</sup>

## II. Classifying Programs 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,<sup>16</sup> and WSIPP has 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 3](#)). Further, some programs are classified as “promising practices” when the OSPI-convened expert panel or the research evidence suggests the practice might improve student outcomes, but the topics did not meet the criteria for classification as evidence- or research-based.

Additionally, in the 2018 inventory update, WSIPP clarified classifications for programs that produce null or poor results. 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.” As of 2018, WSIPP defines 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 (see [Exhibit 3](#)).

---

<sup>15</sup> We report meta-analytic results for non-monetizable outcomes like attendance on our [website](#).

<sup>16</sup> [RCW 71.24.025](#). WSIPP’s adult behavioral health inventory can be found on our website; [EBPI & WSIPP \(2019\)](#).

### Exhibit 3

#### Definitions for Inventory Classification

##### Definitions in Statute (RCW 71.24.025)

###### *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 [. . .].*

##### WSIPP's Definitions for Null and Poor Results

###### *Null*

*A program or practice for which the results from a random-effects meta-analysis of multiple evaluations—or one large multiple-site evaluation—are not statistically significant for relevant outcomes.*

###### *Poor*

*A program or practice for which the results from a random-effects meta-analysis of multiple evaluations—or one large multiple-site evaluation—indicate undesirable (harmful) effects.*

###### Note:

The definitions of evidence-based, research-based, and promising outlined above are separate from the Tiers of Evidence defined by the federal Every Student Succeeds Act (ESSA). WSIPP and ESSA Tiers of Evidence share common criteria for determining classifications (e.g., programs must be tested in similar populations; study shows a statistically significant effect on a relevant outcome). While definitions sometimes overlap, they are not perfectly aligned. For more information on ESSA Tiers of Evidence, see the [U.S. Department of Education's guidance on using evidence to strengthen education investments](#).

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:

- [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 outcomes (p-value < 0.20).<sup>17</sup> To meet the research-based definition, one single-site evaluation must indicate the practice achieves desired outcomes (p-value < 0.20).

If results from a random-effects meta-analysis of multiple evaluations are not statistically significant (p-value > 0.20) for desired outcomes, the practice may be classified as “null.” If results from a random-effects meta-analysis of multiple evaluations or one large multiple-site evaluation indicate that a practice produces undesirable effects (p-value < 0.20), the practice may be classified as producing poor outcomes.

- [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.<sup>18</sup> The WSIPP model uses Monte Carlo simulation to test the probability that benefits will exceed costs. Programs that have at least a 75% chance of a positive net present value meet the “cost-beneficial” criterion.
- [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 who are children of color must be greater than or equal to the proportion of children of color aged 0 to 17 in Washington. From the 2010 Census, among Washington children aged 0 through 17, 68% were White, and 32% were children of color.<sup>19</sup> Therefore, if the weighted average of program participants in the outcome evaluations is at least 32% of students of color, 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 children of color (p-value < 0.20).

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

---

<sup>17</sup> Statisticians often rely on a metric, the p-value, to determine whether an effect is significant. The p-value is a measure of the likelihood that the difference could occur by chance—values range from 0 (highly significant) to 1 (no significant difference). For the purposes of WSIPP’s inventories, p-values <0.20 (a 20% likelihood that the difference could occur by chance) are considered statistically significant findings. We use a p-value of 0.20 (instead of the more conventional p-value of 0.05) in order to avoid classifying programs with desirable benefit-cost results as promising. 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.

<sup>18</sup> For information about WSIPP’s benefit-cost model see [WSIPP \(2019\)](#).

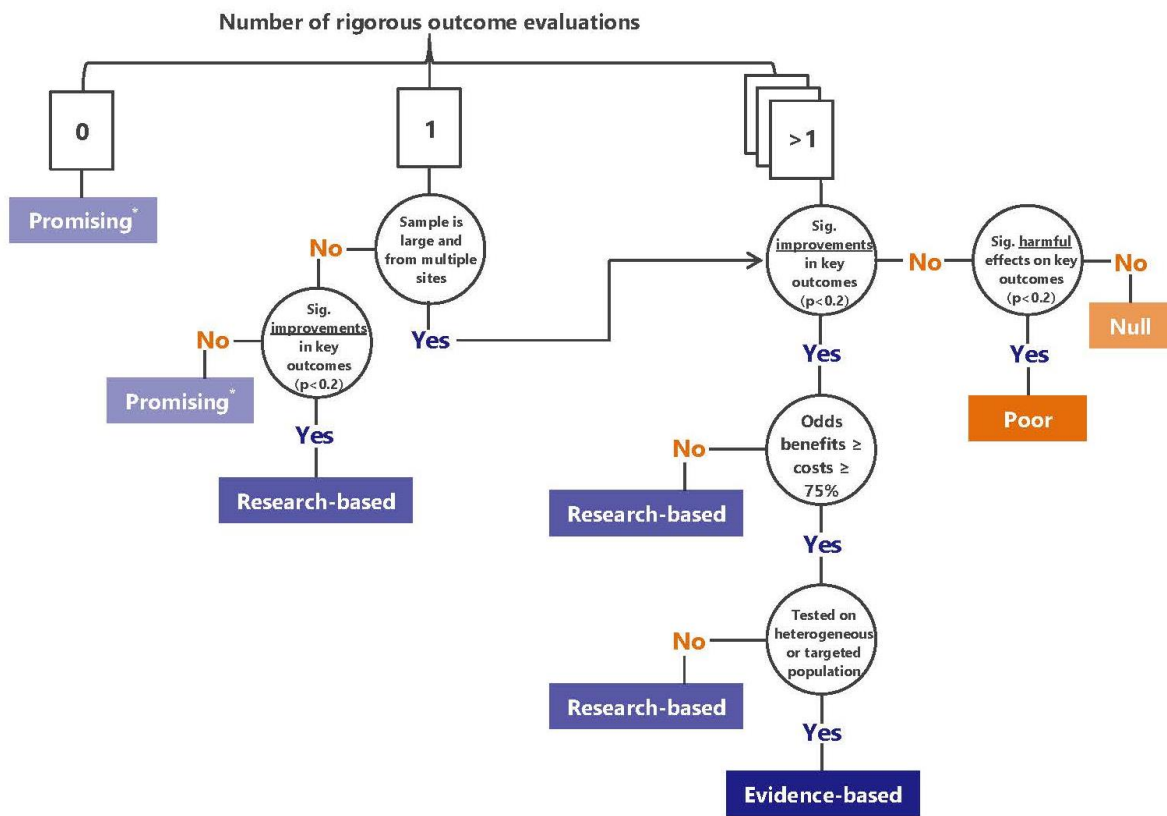
<sup>19</sup> [United States Census Bureau](#).

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 the criteria to be promising.

Exhibit 4 illustrates WSIPP’s process for implementing these criteria.

### Exhibit 4

Decision Tree for Program Classification  
For WSIPP’s Inventories of Evidence-Based, Research-Based, and Promising Practices



Note:

\*Considered “promising” if based on a logic model or well-established theory of change; RCW 71.24.025



### III. Why Classifications Change Across Iterations of the Inventory

---

The inventory is a snapshot that changes as new evidence and information are incorporated. While the definitions of evidence-based, research-based, and promising practices have not changed since the LAP Inventory was initially published in August 2014, programs may be classified differently with each update. The change in classifications could be due to changes in the meta-analyses, changes in the standard benefit-cost (BC) model, or both. The goal when implementing updates and revisions is to report rigorous, up-to-date, relevant information that addresses the needs of stakeholders.

- [Changes to program analyses](#). When WSIPP updates our review of a program or intervention (“program”), we conduct a complete literature search, update the meta-analyses, and construct new program costs. We may also make improvements to our meta-analytic methods to reflect current best practices.

We update our meta-analyses for specific programs when new research literature is available, OSPI’s expert panel members nominate a program for review, or when we receive legislative assignments or Board-approved projects that direct us to do so. Program updates are always contingent upon the capacity to execute these requests.

- [Changes in WSIPP’s standard benefit-cost model](#). WSIPP makes continuous improvements to our BC model. WSIPP uses a standard BC model across topic areas, including child welfare, juvenile justice, K-12 education, adult behavioral health, substance use, and more. When we make changes in our BC model, those changes are applied to all programs currently reported on WSIPP’s website and reflect the most up-to-date estimates of the valuation of programmatic benefits.

WSIPP makes updates to our BC model when we have legislative assignments or Board-approved projects that provide resources to do so.

[Exhibit 5](#) provides a list of changes that WSIPP made for the 2020 inventory update. The exhibit includes the type of change, the rationale for the change, and the program classifications impacted by the change.

Also note that the definitions for classification of poor, null, promising, and research-based programs all rely on unadjusted effect sizes from WSIPP’s meta-analyses. Therefore, any changes that affect unadjusted effect sizes may have implications for these program classifications. Changes to benefit-cost findings, however, affect only whether a program is classified as evidence-based.

## Exhibit 5

### Changes to WSIPP’s Meta-Analyses and Benefit-Cost Model and Implications for LAP Inventory Program Classifications in 2020

Change	Rationale for change	Meta/BC analysis elements affected <sup>^</sup>	Program classifications <sup>#</sup> potentially impacted
<b>Changes to program analyses</b>			
<i>Add new research literature</i>	New research is found in literature search; studies we could not include previously become usable due to improvements in statistical methods or ability to include new outcomes	Unadjusted effect sizes Adjusted effect sizes Placement of effects in time Program costs	All levels of program classification
<i>Remove research literature that was previously included</i>	Re-review indicates that a study does not meet criteria for rigor; studies pertain to populations or program implementations that are no longer included in the scope of the analysis; changes in our statistical methods mean we can no longer include certain measures of effect sizes	Unadjusted effect sizes Adjusted effect sizes Placement of effects in time Program costs	All levels of program classification
<i>Update meta-analytic methods</i>	Improvements to our statistical calculations; changes in best practices in the field of meta-analysis	Unadjusted effect sizes Adjusted effect sizes	All levels of program classification
<b>Changes to WSIPP's standard benefit-cost model</b>			
<i>Update economic parameters (inflation, discount rates, etc.)</i>	Updated data sources or new research becomes available that allows for more current parameters to be used in the model; changes in best practices in the field of benefit-cost analysis	Benefits associated with measured outcomes	Evidence-based classification only

**Notes:**

WSIPP may make other modifications, at researcher discretion, to ensure that our analyses represent the best evidence synthesis given the information we have available. For more detail on our approach, see WSIPP’s [Technical Documentation](#).

<sup>^</sup>This column lists the components of our meta/BC analyses that were affected by the relevant type of change. All of these elements have the potential to impact our benefit-cost findings.

<sup>#</sup>Classifications based on definitions described in [Exhibit 3](#).

## IV. Updates to the Inventory as of July 2020

Since the last inventory update in June 2018, WSIPP reviewed previously examined programs that were due for updates and reviewed programs-of-interest nominated by OSPI’s panel of experts. In total, WSIPP reviewed 18 programs for the 2020 inventory update. [Exhibit 6](#) provides an overview of new programs and their classifications. [Exhibit 7](#) provides an overview of programs in which classifications changed between 2018 and 2020 and the reasons for these changes. The exhibits in this section do not provide an exhaustive list of all programs in the inventory. The complete inventory begins on page 14 and contains 58 programs.

### Exhibit 6

#### New Program Classifications

Program/intervention name	Classification
<b>Tutoring support</b>	
By peers <sup>#</sup>	Evidence-based
<b>Behavior support</b>	
Restorative justice in schools	Research-based

Notes:

Classifications using definitions as described in [Exhibit 3](#).

<sup>#</sup>Peer tutoring was presented as two separate types of interventions in the 2018 inventory (tutoring by same-age peers and tutoring by cross-age peers). After re-reviewing the available research, we found there was not enough rigorous evidence to make a statement on cross-age tutoring alone. We decided to combine all of the evidence on these interventions into a general peer tutoring topic for the 2020 update.

### Exhibit 7

#### Classifications Revised Due to Updated Meta-Analyses or Benefit-Cost Modeling

Program/intervention name	Prior (2018) classification	Current (2020) classification	Primary reason for classification change
<b>Tutoring support</b>			
By non-certificated adults, small-group, structured	Evidence-based	Research-based	Removed research literature that was previously included
<b>Behavior support</b>			
First Step to Success	Research-based	Null	Added new research literature
Good Behavior Game	Research-based	Evidence-based	Revisions to standard benefit-cost model
Becoming a Man (BAM)	Evidence-based	Research-based	Revisions to standard benefit-cost model
Second Step	Research-based	Evidence-based	Added new research literature

Note:

Classifications using definitions as described in [Exhibit 3](#).

On March 13, 2020, Governor Inslee announced the closure of all Washington State public and private schools in response to the outbreak of the disease caused by the novel coronavirus

(COVID-19).<sup>20</sup> On April 9<sup>th</sup>, the directive was extended through the rest of the 2019-20 school year. Although they are at various stages of preparedness, districts are working to meet student needs by providing “instruction using printed learning materials, phone contact, email, technology-based instruction, or a combination of these.”<sup>21</sup>

At the time of publication (July 2020), it is unknown to what extent regular instruction will resume in the 2020-21 school year. To date, WSIPP’s analyses of interventions relevant to teaching and learning in a remote environment are limited to supplemental computer-assisted instruction for students with various types of learning needs. To be clear, the research evidence in our analyses described below compare supplemental computer instruction to regular classroom instruction or another form of supplemental assistance. To date, we have not evaluated the effect of remote learning compared to regular classroom instruction and cannot make an empirical statement on whether remote learning is more or less effective than regular classroom instruction.

Overall, we found that supplemental computer-assisted instruction does not reliably improve outcomes beyond regular classroom instruction, or work better, on average than other types of specialized assistance except among English language learners.

### [Supplemental Computer-assisted Instruction Compared to Regular Classroom Instruction](#)

WSIPP identified one program on the inventory—supplemental computer-assisted instruction for struggling readers (vs. regular classroom instruction)—that educators may consider relevant to teaching in a remote environment. This type of approach assists students who test below grade-level standards in reading and provides self-guided practice as a supplement above and beyond regular classroom instruction. In our analysis, we found that the weight of the research evidence indicates this type of approach has no reliable effect on student outcomes (null results) above and beyond regular classroom instruction.

### [Supplemental Computer-assisted Instruction Compared to Other Types of Supplemental Assistance](#)

We also found studies comparing supplemental computer-assisted instruction with other specific types of assistance for students struggling in mathematics and for English language learners. On our website, we display findings for three additional supplemental computer-assisted interventions.<sup>22</sup> These programs are not included on the LAP inventory and are not classified because the effects in analyses reflect the impact of supplemental instruction compared to other types of assistance, rather than compared to regular classroom instruction. Again, we identify these programs as relevant for remote teaching and learning situations but encourage educators

---

<sup>20</sup> Office of the Superintendent of Public Instruction website [Novel Coronavirus \(COVID-19\) Guidance & Resources](#) accessed June, 2020.

<sup>21</sup> OSPI. (2020). [Continuous Learning 2020: Resources and Planning Tools](#).

<sup>22</sup> See WSIPP’s website for detail on supplemental computer-assisted instruction for [struggling readers \(vs. other assistance\)](#), for [students struggling in math](#), and for [English language learners \(ELL\)](#).

to consider program results in the context of their instructional environments and student populations. The weight of the research evidence indicates that supplemental computer-assisted instruction for students struggling in math and reading is no better or worse than other types of assistance (null results). For English language learners, we found one rigorous study with a significant and positive effect on students’ reading test scores, indicating that supplemental computer-assisted instruction for English language learners may be superior to other types of supplemental approaches.

Finally, for some interventions on the LAP inventory, WSIPP compiles multiple brand-name and non-brand name programs into general categories. For example, a school-based mentoring topic on the inventory may include a brand-name program like Big Brothers Big Sisters and other non-brand name programs that share similar components like the use of volunteer mentors. Some programs are combined into general categories on the inventory but reported separately on our website (see Exhibit 8).

### Exhibit 8

Programs Combined on the LAP Inventory but Reported Separately on the WSIPP Website

Programs displayed on the inventory	Programs displayed on the website
Mentoring: School-based (taxpayer costs only)	Mentoring: School-based by volunteers (taxpayer costs only) Mentoring: Big Brothers Big Sisters School-Based (taxpayer costs only)
Mentoring: School-based (with volunteer costs)	Mentoring: School-based by volunteers (including volunteer costs) Mentoring: Big Brothers Big Sisters School-Based (including volunteer costs)
Mentoring: Community-based (taxpayer costs only)	Mentoring: Community-based (taxpayer costs only) Mentoring: Big Brothers Big Sisters Community-Based (taxpayer costs only)
Mentoring: Community-based (including volunteer costs)	Mentoring: Community-based by volunteers (including volunteer costs) Mentoring: Big Brothers Big Sisters Community-Based (including volunteer costs)
Case management in schools	Communities in Schools City Connects

## Future Updates

---

The next update to this inventory will be published by July 1, 2022.

## 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 assist in describing specific programs. We are especially grateful to Sheila Gerrish, Joshua Lynch, and Annie Pennell for their input and collaboration throughout this process.

**July 2020**  
**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 suggested evidence-based criteria	Percent children of color
<b>Tutoring support</b>				
Tutoring: By certificated teachers, small-group, structured	●	97%		63%
Tutoring: By adults, one-on-one, structured	●	92%		66%
Tutoring: By peers	●	81%		46%
Tutoring: By non-certificated adults, small-group, structured	⊙	69%	Benefit-cost	65%
Tutoring: By adults, one-on-one, non-structured	⊙	69%	Benefit-cost	77%
Tutoring: Supplemental Educational Services (under Title I)	⊙	62%	Benefit-cost	95%
Tutoring: By adults, for English language learner students	Null	60%	Weight of evidence	91%
Tutoring: Supplemental computer-assisted instruction for struggling readers (vs. regular classroom instruction)	Null	64%	Weight of evidence	91%
<b>Extended learning time</b>				
Double dose classes	●	98%		91%
Out-of-school-time tutoring by adults	●	93%		75%
Summer learning programs: Academically focused	●	87%		85%
Summer book programs: One-year, with additional support	Null	58%	Weight of evidence	77%
Summer book programs: One-year intervention	Null	57%	Weight of evidence	86%
Summer book programs: Multi-year intervention	P	71%	Weight of evidence	95%
<b>Professional development</b>				
Teacher professional development: Use of data to guide instruction	●	98%		54%
Teacher professional development: Targeted	●	79%		96%
Teacher professional development: Online, targeted	⊙	60%	Benefit-cost/heterogeneity	31%
Teacher professional development: Induction/mentoring	Null	64%	Weight of evidence	92%
Teacher professional development: Not targeted	Null	38%	Weight of evidence	51%
Educator professional development: Use of data to guide instruction ("train the trainers")	Null	29%	Weight of evidence	46%
Professional learning communities	P		No rigorous evaluation with outcome of interest	
<b>Consultant teachers</b>				
Consultant teachers: Online coaching	●	93%		53%
Consultant teachers: Coaching	●	81%		53%
Consultant teachers: 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 pages 18 and 19.

Note:

# This program is a special analysis for this inventory and does not have a program-specific webpage on WSIPP’s website.

*The classifications in this document are current as of July 2020.*  
*For the most up-to-date results, please visit the program’s page on our website <http://www.wsipp.wa.gov/BenefitCost>*

**July 2020**  
**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 suggested evidence-based criteria	Percent children of color
<b>Parent outreach</b>				
Parents as tutors with teacher oversight	⊙	56%	Benefit-cost	58%
Families and Schools Together (FAST)	Null	50%	Weight of evidence	83%
Conjoint behavioral consultation	Null	23%	Weight of evidence	21%
Parent and family engagement coordinators	P		No rigorous evaluation with outcome of interest	
<b>Community partnerships</b>				
Case management in schools <sup>#</sup>	⊙	68%	Mixed results/benefit-cost	61%
Mentoring: School-based (taxpayer costs only) <sup>#</sup>	⊙	19%	Benefit-cost	74%
Mentoring: School-based (with volunteer costs) <sup>#</sup>	⊙	16%	Benefit-cost	74%
Mentoring: Community-based (taxpayer costs only) <sup>#</sup>	⊙	66%	Benefit-cost	68%
Mentoring: Community-based (including volunteer costs) <sup>#</sup>	⊙	57%	Benefit-cost	68%
PROSPER (PROmoting School-community-university Partnerships to Enhance Resilience)	⊙	39%	Benefit-cost/heterogeneity	15%
<b>Behavior support</b>				
Positive Action	●	94%		63%
Good Behavior Game	●	76%		50%
Second Step	●	85%		53%
Becoming a Man (BAM)	⊙	74%	Benefit-cost	98%
Becoming a Man (BAM) with high-dosage tutoring	⊙		Single evaluation	99%
Mentoring: Community-based for children with disruptive behavior disorders	⊙	67%	Benefit-cost/heterogeneity	7%
Mentoring: School-based by teachers or staff	⊙	71%	Benefit-cost	86%
School-Wide Positive Behavioral Interventions and Supports	⊙	74%	Benefit-cost	50%
Behavioral Monitoring and Reinforcement Program (BMRP)	⊙	64%	Benefit-cost	81%
Coping Power Program	⊙	57%	Benefit-cost	80%
"Check-in" behavior interventions	⊙	57%	Benefit-cost	72%
Restorative justice in schools	⊙	11%	Mixed results/benefit-cost	65%
Fast Track prevention program	⊙	0%	Benefit-cost	53%
Daily Behavior Report Cards	⊙		Single evaluation	13%
First Step to Success	Null	47%	Weight of evidence	59%
Caring School Community (formerly Child Development Project)	Null	60%	Weight of evidence	47%
Responsive classroom	Null	4%	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 pages 18 and 19.

**Note:**

<sup>#</sup> This program is a special analysis for this inventory and does not have a program-specific webpage on WSIPP’s website.

*The classifications in this document are current as of July 2020.*  
*For the most up-to-date results, please visit the program’s page on our website <http://www.wsipp.wa.gov/BenefitCost>*



**July 2020**  
**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 suggested evidence-based criteria	Percent children of color
<b>Services for 8<sup>th</sup>, 11<sup>th</sup>, &amp; 12<sup>th</sup> grades</b>				
Credit retrieval	P		No rigorous evaluation with outcome of interest	
<b>Other</b>				
Special literacy instruction for English language learner students	●	81%		98%
Growth mindset interventions	⊙	56%	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 pages 18 and 19.

Note:

# This program is a special analysis for this inventory and does not have a program-specific webpage on WSIPP’s website.

**July 2020**  
**Updated Inventory of Evidence- and Research-Based Practices:**  
**Washington's Learning Assistance Program**

---

**Definitions and Notes**

---

**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:**

Cost-beneficial: Benefit-cost estimation is repeated many times to account for uncertainty in the model. This represents the percentage of repetitions producing overall benefits that exceed costs. Programs with a benefit-cost percentage of at least 75% are considered to meet the "cost-beneficial" criterion in the "evidence-based" definition above.

**July 2020**  
**Updated Inventory of Evidence- and Research-Based Practices:**  
**Washington's Learning Assistance Program**

**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, a program must have been tested on a "heterogeneous" population. We operationalized heterogeneity in two ways. First, the proportion of program participants must be greater than or equal to the proportion of children of color in Washington State aged 0 to 17. From the 2010 Census, among children aged 0 through 17 in Washington, 68% were White and 32% were children of color. Thus, if the weighted average of program participants were at least 32% children of color then the program was considered to have been tested on 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 children of color ( $p < 0.20$ ). 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.
- 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 measuring outcome of interest:** The program has not yet been tested with a rigorous outcome evaluation.
- Null outcomes:** If results from multiple evaluations or one large multiple-site evaluation indicate that a program has no significant effect on outcomes of interest ( $p > 0.20$ ), a program is classified as producing "null outcomes."
- 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."
- 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:** Results from a random-effects meta-analysis ( $p > 0.20$ ) indicate that the weight of the evidence does not support desired outcomes, or results from a single large study indicate the program is not effective.

Suggested citation: Cramer, J., & Wanner, P. (2020). *Updated inventory of evidence- and research-based practices: Washington's K–12 Learning Assistance Program*. (Document Number 20-07-2202). Olympia: Washington State Institute for Public Policy.

For further information, contact:

Julia Cramer at 360.664.9073, [julia.cramer@wsipp.wa.gov](mailto:julia.cramer@wsipp.wa.gov)

Document No. 20-07-2202



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.