

## RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS

### What are the Relevance and Credibility Indices<sup>1</sup>?

*The Relevance Index and Credibility Index are rubrics designed to help decision-makers evaluate existing studies of educational programs, strategies or interventions to determine how seriously to take the findings and whether they can safely assume the findings might apply in their own context.\**

*The **Relevance Index** helps decision-makers determine how well a study's findings apply in their own context by assessing the extent to which various characteristics of the study population and context are similar to the decision-maker's own context.*

*The **Credibility Index** helps decision-makers assess how seriously to take the findings of a study by scoring it on the quality of study design.*

*\*Note that these indices are designed to be used to assess studies which focus on single investigations rather than for studies which summarize the results of many investigations, such as literature reviews or meta-analyses. While such summaries are useful for providing a general idea of "what works," it would be hard to assess the relevance of such overviews to your own context.*

### How to use this Worksheet to evaluate a research study:

- *Before reading a study you want to consider as evidence, look over the items in each index below so you know what to look out for.*
- *To find the information in the study that you will need for completing this worksheet, look for section headings in the study that use terms like "Methods" or "Study Design," "Measures," and "Sample" or "Setting." Search functions in a Word or PDF document can help with this (e.g., Ctrl F).*

### Key terms:

The **target population** of a study is the complete collection of students/teachers/schools or other units that we want to study or consider.<sup>2</sup> For example, the target population of a study on algebra skills of high school students in a particular school district may be all 150,000 students in every high school in the district.

An **outcome** is a change or impact caused by the program or strategy being evaluated, or it could be a characteristic of the respondents you want to measure.<sup>3</sup> An outcome of a reading program for third graders in an elementary school might be to raise reading skills among third grade students.

**Measures** are the items in a research study to which the participants respond<sup>4</sup>, which are used to assess performance on the outcomes of interest.

<sup>1</sup> The design of these indices was influenced by a number of sources: [Digital Promise's "Evaluating Studies of Ed Tech Products,"](#) [NESTA's "Standards of Evidence,"](#) [Alliance for Useful Evidence's "What Counts as Good Evidence,"](#) [David Gough's "Weight of Evidence: A Framework for the Appraisal of the Quality and Relevance of Evidence,"](#) and REL Central's Draft "Applicability of Evidence-Based Interventions" Tool.

<sup>2</sup> Lohr, S. L. (2010). *Sampling: Design and Analysis*. 2<sup>nd</sup> ed. Boston, MA: Brooks/Cole, Cengage Learning.

<sup>3</sup> <https://www.povertyactionlab.org/research-resources/measurement-and-data-collection>

<sup>4</sup> <http://www.uniteforsight.org/research-methodology/module4>

## A. Worksheet to Assess Relevance of Available Evidence to Your Purpose and Context:

### Step 1. First, establish if prior evidence exists:

Are there prior studies conducted on this educational program or strategy (Solution Option) which assess its effectiveness in improving student or teacher outcomes?

Yes                      No

- If *Yes*, go to Step 2.
- If *No*, you may need to plan to collect your own data. (See suggestions for resources to help with this in [DecisionMaker](#), Resources & Guidance, VIII. Evidence-Gathering to Evaluate Options).

**Step 2. If you find studies that provide information on effectiveness, follow the steps below (Steps 2.a – 2.e) to use the Relevance Index for establishing whether this evidence is relevant to your context.** Later, you will be guided to use the Credibility Index (Steps 3 – 14) to evaluate the credibility of studies you deem relevant to your context.

#### Overview of Step 2:

- a. Choose relevant factors
- b. Score each study on each important factor
- c. Compute score
- d. Determine relevance of each study
- e. Record results

#### To help you determine whether a study was conducted in a similar context to your school/ district/ state:

- a. **Choose Relevant Factors:** Think about which of the factors in the **Relevance Index** below are important to consider for your decision problem and mark the corresponding radio buttons. You may decide to choose all or just a few.
- b. **Score study on each important factor:** Read each study carefully and score it between 0 and 3 on each factor you chose. Your score should indicate how similar the study context is to your own context for this factor:  
*3 = very similar, 2 = moderately similar, 1 = slightly similar, 0 = not at all similar/information unavailable.*
- c. **Compute score:** Divide the total score earned by each study by the total maximum possible score on this index (the total maximum is simply 3 points multiplied by the number of factors you selected as important, e.g.,  $3 \times 7 = 21$ ). Then multiply by 100 to give you a Relevance Index between 0 and 100%. For example, if your ratings added up to 18 out of a total maximum score of 21, the study's Relevance Index will be  $(18/21 \times 100) = 86\%$ .

**RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS**

*Steps 2.a – c: Relevance Index*

<b>Contextual Factor</b>	<i>Things to look for and think about</i>	<i>Important consideration for me (fill in radio button or mark with an "x")</i>	<i>Details of the Study (If the information you are looking for is not provided in the study, you can either remove the factor from the list or note "Not available" in this column and score it 0)</i>	<i>Study context is similar to mine 3 = very 2 = moderately 1 = slightly 0 = not at all/unavailable</i>
Recency of study	Extent to which the educational activity, strategy or intervention (Solution Option) is still likely to be applicable in my context today	○		
Student demographics	Age of students on which the program/ intervention was tested	○		
	Baseline performance of students before implementing the program	○		
	Percent economically disadvantaged, e.g., as indicated by FRPL	○		
	Percent classified as minority	○		
	Percent ELL	○		
	Percent Special Education	○		
School context	Charter school vs. district school vs. private school	○		
	Selective admissions vs. open to all applicants	○		
	Qualifications of teachers (e.g., degree level, certification, tenure status, average experience)	○		
	Availability of necessary staff	○		
	Availability of necessary technology	○		
	Alignment with School Improvement Plan	○		
Relevance of measure used	Whether the outcome measure reported in the study is relevant for your goals, e.g., if you are trying to reduce behavior incidents, a study that reports suspension rates might be relevant, but one that reports attendance might be less relevant	○		
Scale	Is the scale at which the study was conducted similar to your context and needs (e.g., if you want to implement a reading program at large scale across several schools, did the study look at the effect of the intervention across many schools, or only at small scale in one school?)	○		
<b>Total possible score (Y) = Number of factors you selected x 3 points = _____</b>			<b>Total points you gave this study (X): _____</b>	
<b>Relevance Index = X/Y x 100 = _____ %</b>				

*Note. FRPL = Free and Reduced Price Lunch; ELL = English Language Learner.*

**RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS**

- d. Determine relevance of each study:** Compare the overall score for each study with the interpretation table below to reach a conclusion about which, if any, of the studies are relevant enough to your context.

**STEP 2.d: Relevance Index Interpretation Table**

<b>Relevance Index</b> ( = Total points for this study/Total possible score x 100)	<b>Relevance Rating</b>
<b>Less than 30%</b>	Low Relevance
<b>31 - 69%</b>	Moderate Relevance
<b>70% or higher</b>	High Relevance

- e. Record results:** Record your assessment of each study in the **Relevance Summary Table** below.

**STEP 2.e: Relevance Summary Table**

<b>Name of study</b>	<b>Authors</b>	<b>Year of study</b>	<b>(Column D) Relevance Score</b>	<b>(Column E) Total possible score</b>	<b>Relevance Index (Column D/Column E)</b>	<b>Relevance Rating (High, Moderate, Low)</b>	<b>Use this study as relevant evidence for this decision (Yes/No)</b>

## RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS

### Next Steps

- ❖ We recommend you move studies that earn a rating of High or Moderate relevance forward to determine whether they are credible.
- ❖ If you do decide to move forward with a study of low relevance to your context, you may want to be sure that it scores high on credibility in the next step.
- ❖ If none of the studies you review score better than Low Relevance, you may want to consider designing your own study and collecting your own data.

### Summary

**The output of Steps 2a-2e above on Relevance is to provide a Relevance rating for each study you review, and a decision on whether to use any of these studies as evidence for the Solution Option(s) (educational activities, strategies or interventions) you are considering, or to collect your own data.**

**If one or more studies pass the Relevance threshold (greater than 30%), then you can proceed with the relevant study to the next step to assess credibility. Our Relevance threshold is merely a guideline. You can follow our suggested threshold, or you can choose your own threshold by which to accept studies that makes sense given the context and decision problem at hand.**

**If no study passes the threshold for relevance, then we recommend you collect your own data to evaluate the Solution Option(s) you are considering.**

## B. Worksheet to Assess the Credibility of Available Evidence

*The Credibility Index can be used to assess how seriously you should take the findings of each study that met the Relevance threshold in Part A above.*

*Below is an overview of the Credibility Index:*

CREDIBILITY INDEX OVERVIEW	
<b>PART I: Source of Study</b>	<b>Scoring</b>
Section 1: Who conducted the study?	-1 to 2
Section 2: Who published the study?	-1 to 2
<b>Total Part I</b>	<b>Max 4 points</b>
<b>PART II: Implementation and Goals of the Program</b>	
Section 3: How long did participants engage in the program?	-1 to 2
Section 4: Are the components of the program clearly described?	-1 to 2
Section 5: Is there a clear description of how the program was implemented?	-1 to 2
Section 6: Are the outcomes being measured clear and meaningful?	-1 to 3
<b>Total Part II</b>	<b>Max 9 points</b>
<b>PART III: Study Sample</b>	
Section 7: Are the characteristics of the sample clear?	-1 to 2
Section 8: How many people are participating in the study?	-1 to 2
<b>Total Part III</b>	<b>Max 4 points</b>
<b>PART IV: Rigor of Methodology</b>	
Section 9: Is there a comparison group of any kind?	-1 to 1
Section 10: How was the comparison group selected?	1 to 6
Section 11: Do the authors discuss alternative possible explanations for the results?	0 to 1
Section 12: Measuring outcomes over time: At what time points were the data collected?	0 to 2
<b>Total Part IV</b>	<b>Max 10 points</b>
<b>TOTAL PARTS I, II, III and IV</b>	<b>XX/ 27</b>

RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS

*Credibility Index*

**Credibility Index Part I: Source of Study**

**Step 3. Select one relevant statement per section:** For each section (1-2) in Part I of the Credibility Index, select the statement that most accurately reflects the study you are reviewing and circle the relevant points. Select ONE statement only per section.

**Step 4. Calculate score:** Add up the points you awarded for these two sections.

PART I: SOURCE OF STUDY		MAX POINT: 4
<i>Details of the Study</i>	<i>What to look for and think about</i>	<i>Circle the relevant points for each section</i>
<b>SECTION 1</b>	<b>Who conducted the study?</b>	Max Points: 2
	It is not clear who conducted the study.	-1
	The study was conducted by the program vendor.	0
	The study was conducted by an external evaluator hired by the program vendor.	1
	The study was conducted by an external evaluator acting as an independent third party, i.e., not paid by the vendor. <sup>5</sup>	2
<b>SECTION 2:</b>	<b>Who published the study?</b>	Max Points: 2
	It is not clear where the study was published.	-1
	The study was published by the vendor.	0
	The study was published by a third party (i.e., other than the vendor) but not a peer-reviewed journal, e.g., a university, research organization, school district or other government agency. Keep in mind some technical reports are also later published in a peer-reviewed journal, so you may wish to check <a href="#">Google Scholar</a> to see if there is a later published version.	1
	The study was published in a peer-reviewed journal. <sup>6</sup>	2
<b>Total points for Part I: Sections 1-2</b>	<b>Max Possible: 4 points</b>	<b>Total points awarded: —</b>

<sup>5</sup> This may include collaborations between the evaluator and implementing partners such as school districts or other educational or research institutions, but excluding the program vendor. The key distinction is that those conducting and/or commissioning the evaluation should not stand to make a profit from the sale of the product/ intervention being tested. In situations where the developer is an educational institution like a university and another unit at the same institution is funded by a third party (e.g., a foundation or government funding agency) to conduct the evaluation, you may consider awarding 1.5 points to reflect the fact that the evaluator is not totally independent of the developer.

<sup>6</sup> It is not always easy to know whether a journal is truly peer-reviewed even if it appears in Google Scholar. We suggest you Google the journal’s name to find its website and skim the content for confirmation that it is peer-reviewed. If in doubt, you can score the item 1 or 1.5.

**RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS**

**Credibility Index Part II: Implementation and Goals of the Intervention**

**Step 5: Select one relevant statement per section:** For each section (3-6) in Part II of the Credibility Index, select the statement that most accurately reflects the study you are reviewing and circle the relevant points. Select ONE statement only per section.

**Step 6: Calculate score:** Add up the points you awarded for these four sections.

<b>PART II: IMPLEMENTATION AND GOALS OF THE INTERVENTION</b>		<b>MAX POINTS: 9</b>
<i>Details of the Study</i>	<i>Things to look for and think about</i>	<i>Circle the relevant points for each section</i>
<b>SECTION 3:</b>	<b>Length of time participants are exposed to the educational program/strategy being studied</b>	<b>MAX POINTS 2</b>
	Length of exposure is not clear from the study.	-1
	Length of exposure is too short to make a difference.	0
	Length of exposure is too long to reflect likely effect in regular practice.	1
	Length of exposure is about right.	2
<b>SECTION 4:</b>	<b>Components of the program:</b> Is it clear what program participants were expected to do compared with business as usual conditions?	<b>MAX POINTS: 2</b>
	The components of the program/treatment are not clear.	-1
	The components of the program/treatment are partially clear.	1
	The components of the program/treatment are completely clear.	2
<b>SECTION 5:</b>	<b>Implementation of the program:</b> Is it clear how the program was implemented in the study, e.g., how much time participants actually spent engaging in the program activities, how often, whether it was implemented in/out of the classroom, before/ after school hours etc.?	<b>MAX POINTS: 2</b>
	The implementation details of the program are not clear.	-1
	The implementation details of the program are partially clear.	1
	The implementation details of the program are completely clear.	2



**RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS**

<b>SECTION 6:</b>	<b>Meaningful outcomes</b>	<b>MAX POINTS 3</b>
	It is not clear which outcomes are being measured, e.g., it is not clear whether the study is evaluating geometry skills or algebra skills.	-1
	The outcomes measured are not at all aligned with the ultimate goal for implementing the intervention, e.g., the study investigates whether an after-school supplemental math program improves geometry skills, despite the fact that the program aims to improve algebra skills.	0
	The outcomes measured only capture short-term behavioral changes but not longer-term educational outcomes, e.g., the study only documents whether students are attending an after-school math program, but does not measure whether their math skills are improving.	1
	The outcomes measured are aligned with some but not all of the stated goals for implementing the intervention, e.g., the study is measuring algebra skills when the primary goal of the program is to improve both algebra and geometry skills.	2
	The outcomes measured are aligned with all of the stated goals for implementing the intervention, e.g., the study is measuring algebra skills when improving algebra skills is the primary goal of the program.	3
<b>Total points for Part II: Sections 3-6</b>		<b>Max Possible: 9 points</b>
		<b>Total points awarded:</b>

### Credibility Index Part III: Study Sample

*What is a sample? Imagine that you are the superintendent of a large, diverse school district, and want to investigate the social and emotional competencies of all high school students in your district. The school district contains 80 high schools with 150,000 students. This is the “target” population that you are interested in. You decide to ask a research team to collect the data for you but have a limited budget. It might not be feasible to go into all 80 high schools in the district and have all 150,000 students take an assessment of social and emotional competencies, so the research team will likely select a subset of students - a “sample”- to make the data collection process more feasible.*

#### Key terms

A **sample** is a subset of a population, for example, a subset of the students/teachers/schools that make up a target population for an evaluation.<sup>7</sup>

A sample is **representative** if the sample is similar to the target population on all important characteristics.

**Sample Size:** The number of units (e.g., students/teachers/schools) in a sample.

**Statistical power:** The probability that the estimate of the program effect will be found statistically significant if a given effect size is determined to have occurred.<sup>8</sup>

There are two concerns with drawing a sample in order to get trustworthy results in an effectiveness study:

- Are the characteristics of the sample clear?
  - For example, the researchers should clearly explain characteristics of the sample, such as % FRPL, % minority and % ELL.
- Is the sample large enough to detect an effect when indeed there is one?
  - For example, if the sample only has 10 participants and the study aims to measure outcomes for social and emotional competencies, this would probably be too small a sample. However, if the study aims to measure ease of implementation across 10 different classrooms, this sample size would be more reasonable.

---

<sup>7</sup> Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation: A systematic approach*. 8<sup>th</sup> Edition. Sage Publications.

<sup>8</sup> Ibid.

**RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS**

**Step 7: Select one relevant statement per section:** For each section (7-8) in Part III of the Credibility Index, select the statement that most accurately reflects the study you are reviewing and circle the relevant points. Select ONE statement only per section.

**Step 8: Calculate score:** Add up the points you awarded for the two sections.

<b>PART III: STUDY SAMPLE</b>		<b>MAX POINTS: 4</b>
<i>Details of the Study</i>	<i>Things to look for and think about</i>	<i>Circle the relevant points for each section</i>
<b>SECTION 7:</b>	<b>Sample characteristics: Are the characteristics of the study participants clear? (e.g., grade, gender, FRPL, ELL and Special Education status).</b>	<b>MAX POINTS: 2</b>
	The characteristics of the study participants are not clear.	-1
	The characteristics of the study participants are moderately clear.	1
	The characteristics of the study participants are very clear.	2
<b>SECTION 8:</b>	<b>Sample size: Does the size of the sample, i.e., the number of participants in the study, seem adequate? (For researchers: do you think there is enough power to detect an effect if indeed there is one?)</b>	<b>MAX POINTS 2</b>
	The number of participants in the study is unclear.	-1
	The study does not have an adequate number of participants.	0
	The study had a fairly adequate number of participants.	1
	The study has a very adequate number of participants.	2
<b>Total points for Part III: Section 7 and 8</b>	<b>Max Possible: 4 points</b>	<b>Total points awarded:</b>

## Credibility Index Part IV: Rigor of Methodology

### Key terms

A **comparison group** is a group that did not participate in a program and whose results on the outcome of interest can be compared with those of the group that *did* participate in the program. The key challenge in a good efficacy study is to find a comparison group that closely resembles program participants, meaning both groups should be similar on average across important characteristics, such as student demographics or school characteristics.

To assess the effects of a social program, researchers conducting effectiveness studies need to measure the outcomes for the individuals exposed to the program (the “treatment group”) and find a credible way to estimate the outcomes that would have occurred in the absence of the program.<sup>9</sup> To do so, researchers must identify a comparison or “control” group that is similar to the group exposed to the program except for participation in the program.

There are two important considerations with respect to comparison groups:

- The most important consideration is **whether the study includes a credible comparison group**. A credible comparison group is one that is similar to the group that received the program on characteristics that are relevant for the goals of program. To assess how credible a comparison group is, ask **how** program participants were selected to participate in the program: were program participants selected because of certain student, teacher or school characteristics?
  - For example, were students performing below grade level in literacy chosen to participate in a reading program? If so, does the comparison group perform at the same baseline reading levels as program participants?
  - Or were schools with motivated principals and strong infrastructure selected to participate in the program? If so, does the comparison group have equally motivated principals and similar infrastructure to the program schools?
- The second consideration is **whether data on outcomes are collected multiple times** for the treatment and the comparison group, e.g., before and after the program, and on subsequent occasions. This can be useful if you want to account for baseline differences between the two groups, or if you want to measure longer-term outcomes.

**Step 9: Select one relevant statement per section:** For each section in Part IV of the Credibility Index, select the statement that most accurately reflects the study you are reviewing and circle the relevant points. Select ONE statement only per section.

**Step 10: Calculate Score:** Add up the points you awarded in the last column for Section 9-12.

---

<sup>9</sup> Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation: A systematic approach*. 8<sup>th</sup> Edition. Sage Publications.

RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS

<b>PART IV: RIGOR OF METHODOLOGY</b>		<b>MAX POINTS 10</b>
<i>Details of the Study</i>	<i>Things to look for and think about</i>	<i>Circle the relevant points for each section</i>
<b>SECTION 9: First determine whether there is a comparison group of any kind.</b>		
	It is not clear whether there is a comparison group.	-1
	There is no comparison group.	-1
	Study includes a comparison group which does not participate in the program being studied.	1
<b>SECTION 10. Then, identify how that comparison group was selected. (only answer 10a. and 10b. if the study includes a comparison group)</b>		
	<b>10a. Clarity of selection process for program v. comparison groups.</b>	
	Unclear how study participants were selected to be in the program or the comparison group.	-1
	The selection process for program versus comparison group participants is clearly described.	0
	<b>10b. Strength of study design to ensure the comparison group is similar to the program participants</b>	
	The study compares outcomes for students/teachers/schools who are receiving the program with outcomes for counterparts who are <u>not</u> participating in the program, but this comparison group is not similar on relevant characteristics, or it is unclear whether they are similar.	0
	The study compares outcomes for students/teachers/schools who are receiving the program with outcomes for counterparts who have similar characteristics but are <u>not</u> participating in the program. It may do so either by identifying a comparison group that shares several known characteristics with the program participants, e.g., same grade, gender, SES (statistical matching), or by first matching program participants with non-participants who <u>could</u> have been just as likely to participate in the program, as predicted by known characteristics such as age and gender, and then comparing outcomes for the matched pairs (propensity score matching techniques). <sup>10</sup>	1

<sup>10</sup> <https://www.povertyactionlab.org/sites/default/files/resources/2016.08.31-Impact-Evaluation-Methods.pdf>

**RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS**

	The intervention is provided to students/teachers/schools who are above a cutoff point for eligibility. The study compares participants who are <u>just above</u> the eligibility cutoff, and therefore receive the intervention, with students/teachers/schools who are just below the eligibility cutoff, and therefore do not receive the intervention. This design should ensure the two groups are highly comparable.	3
	The study uses a randomized controlled trial (RCT) in which students/teachers/schools are chosen at random to either <u>participate in the program</u> or to serve in a <u>comparison group</u> .	5
	The study uses a randomized controlled trial (RCT) and also uses propensity score matching or statistical matching techniques.	6
<b>SECTION 11. Addressing other possible explanations for the results: If the study does not assign students/teachers/schools at random to participate in the program or in the comparison group, do the authors attempt to account for, or at least discuss the possibility of, other factors besides the treatment explaining the difference in outcomes between treatment and control groups?</b>		
	The study does not address other possible explanations for the results.	0
	The study discusses other possible explanations for the results OR does not need to because participants were assigned at random to treatment or comparison groups and were comparable at the beginning and end of the treatment period (i.e., were comparable at baseline and did not suffer high attrition).	1
<b>SECTION 12. Measuring outcomes over time</b>		
	The study only assesses student outcomes after they have participated in the program, i.e., it includes <u>only post-test</u> measures.	0
	The study includes before and after measures, e.g., a pre-test/survey/observation before the intervention and a post-test/survey/observation after the intervention.	1
	The study assesses outcomes multiple times before/ during and after the intervention, which may include a second post-test several months after the intervention has ended.	2
<b>Total for Part IV: Section 9</b>	<b>Max Possible: 10 points</b>	<b>Total points awarded:</b>

RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS

**Step 11: Add up points in each part:** Now add up the points earned by this study for Parts I, II, III and IV of the Credibility Index

**Summary Score for Credibility Index**

Totals Points for Parts I-IV:	Max Possible	Points earned
Part I	4	
Part II	9	
Part III	4	
Part IV	10	
<b>GRAND TOTAL</b>	<b>27</b>	Total points awarded: —

**Step 12: Assess credibility:** Use the Credibility Index Interpretation Table below to find the credibility band your score falls into.

**Credibility Index Interpretation Table**

Credibility Index (Total Points for Parts I, II, III, IV)	Credibility Rating	Credibility Parameter
Less than 10	Low credibility	0.2
10-19	Moderate credibility	0.6
20-27	High credibility	1

**Step 13: Adjust effect size:** For low or moderate credibility studies, you can multiply the effect size found in the study you reviewed by the relevant Credibility Parameter to adjust it downwards. Note that this adjustment should be viewed as a reflection of your professional judgement rather than as scientific evidence!

For example, if you are trying to assess a computer-assisted learning program for impact on standardized test scores and a study you reviewed of “Program A” reported an effect size of 0.3 but received a Credibility rating of “Moderately credible,” this is how you would proceed:

Credibility Parameter: Moderately credible = 0.6

Impact on standardized test scores (as taken from the evaluation study) = 0.3

Multiply the effect size reported in the study by the Credibility Parameter:  $[0.3] \times [0.6] = [0.18]$

Use the new effect size, 0.18, as the expected effectiveness for Program A. In [DecisionMaker](#) you would enter 0.18 in the evaluation measures table.

Hollands, F. M., Escueta, M.M., Pan, Y., Levin, H. M., Corter, J., & Wang, K. (2019). *Relevance and credibility indices for studies of educational interventions*. Teachers College, Columbia University.

RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS

**Step 14: Record your assessments of each study:** Use the Relevance and Credibility Summary Table below to document your assessments of each study you reviewed.

**Relevance and Credibility Summary Table**

Name of study	Authors	Yr of study	Relevance Index	Relevance Rating (High, Moderate, Low)	Use study as relevant evidence for decision (Yes/No)	Credibility Index (Total Score adding Parts I, II, III, IV)	Credibility Rating (High, Moderate, Low)	Effect size reported in study	Adjusted effect size

Summary

- ❖ The output of Step B on Credibility is to assign a low, medium or high Credibility Parameter to a study that evaluates the effectiveness of an educational program or strategy. This parameter will function as a weight between 0 and 1.
- ❖ You can multiply the effect size found in a study by this weight to adjust it for credibility. Effect sizes from high credibility studies would remain unchanged while the effect sizes for low and moderate credibility studies would become smaller.
- ❖ You may also simply consider not relying at all on studies that have very low credibility. Instead, you could either plan your own evaluation of the educational activity or find alternative educational activities for which better evidence of effectiveness is available.



RELEVANCE AND CREDIBILITY INDICES FOR STUDIES OF EDUCATIONAL INTERVENTIONS

❖ Optional Summary Table for the Relevance Index

<b>RELEVANCE INDEX SUMMARY TABLE</b>			
<b>Factor Selected</b>	<b>Details About My Context</b>	<b>Details About Study</b>	<b>Relevance (Scale of 0 to 3)</b>
<i>Factor 1:</i>			
<i>Factor 2:</i>			
<i>Factor 3:</i>			
<i>Factor 4:</i>			
<i>Factor 5:</i>			
<i>Factor 6:</i>			
<b>Total Number of Factors Selected * 3 = XX possible points</b>			<b>Total relevance score: Y</b>
Score on Relevance Index			Y/100= --