RAND Education Assessment Finder

User Instructions

Laura S. Hamilton, Brian M. Stecher, Jonathan Schweig, Garrett Baker
Acknowledgments

The authors are grateful to the generous funders of this work: The William and Flora Hewlett Foundation, The Wallace Foundation, The Raikes Foundation, the Einhorn Family Charitable Trust, and the Overdeck Family Foundation. Our Advisory Committee provided useful feedback on the design and functioning of the Assessment Finder. The committee consisted of Lauren J. Bierbaum, Katie Buckley, Michelle Clark, Mari Jones, Sally Grubb Kingston, Clark McKown, Jim Soland, and Jeremy Taylor. We also want to acknowledge the helpful feedback we received from reviewers Cathy Stasz, Jeremy Taylor, and Chris Doss, and the support from Brittany Joseph, Kate Giglio, Monette Velasco, and Arwen Bicknell to prepare the materials for publication. Finally, we are grateful to Marissa Norris, Deanna Lee, Lee Floyd, and Willard Simmons for their excellent work on the website.
## Contents

Acknowledgments ................................................................................................................................. iii  
1. How to Use the RAND Education Assessment Finder ................................................................. 1  
2. Getting Started .............................................................................................................................. 3  
3. User Experiences .......................................................................................................................... 12
1. How to Use the RAND Education Assessment Finder

The RAND Education Assessment Finder helps users find the right tool to measure students’ interpersonal, intrapersonal, and higher-order cognitive competencies.

Introduction

Although U.S. schools have always emphasized a broad range of academic and nonacademic skills for students, skills that go beyond what is traditionally measured by standardized achievement tests have gained new prominence. Working on teams, managing emotions, and creatively solving problems are just a few of these skills. Extensive research shows how these and other interpersonal, intrapersonal, and higher-order cognitive competencies (IICC) are critical in promoting student success in and beyond the classroom.

As the evidence about the value of IICC grows, so does interest in learning more about these competencies. What are they? How do students develop them? How can educators tell whether students have developed IICC? Reviewing assessments that measure IICC can help educators develop their understanding of what IICC looks like and how these competencies are measured. As interest in IICC grows, so too does the number of programs and interventions designed to foster IICC in schools and through out-of-school-time programming. But implementing these programs sets off new kinds of questions: Are students benefiting from the program? What did they learn? Where did they start from? What should we continue to teach? Assessments help answer these questions and thus should be built into any plan to foster IICC. But with so many assessments to choose from, it is difficult to know which one can meet specific needs.

The RAND Education Assessment Finder is a free online tool that can help with this task. It is designed to help users learn more about IICC and identify useful assessments by having them think through specific details:

- What IICC are of interest?
- What grade/age are the students who will be assessed?
- How much time is available for assessment?
- Is computer testing an option?

The Finder uses these details to sort through its database of assessment options. The results show users what assessments are available that could meet the specified goals. It also enables users to examine assessment options more closely by presenting details related to further information on language availability, evidence for reliability and validity of scores, and where to obtain the assessment.

The database contains the most widely used IICC assessments and some that have been administered only as part of a single research study. There are links to most of the assessments
(although some might require special training, fees, or approvals to gain access). We hope to continue to add new assessments over time and to update the information on technical quality as new evidence becomes available.
2. Getting Started

There are two basic steps in using the RAND Education Assessment Finder:

1. **Select Criteria:** Users specify the IICC of interest and provide information about specific contexts (such as class size) related to their desired assessment.

2. **Review Search Results:** Users take a closer look at descriptions of assessments that might meet their criteria as defined in Step 1. The Finder can help users compare up to three assessments at once.

This user guide is organized according to these two steps. The first section presents definitions of key terms used in the criteria section. The second section explains the different elements found on the results page. These encompass an assessment’s general description and its administration, scoring, evidence of quality, and how interested users can obtain the assessment. In the section on User Experiences at the end of the document, three practical examples are provided to show how different users might use the Finder and implement changes based on the results in their schools and districts.

Organizing Principles of the Finder and More Information

The Finder is designed to help educators, policymakers, and researchers identify the IICC-related assessments for K–12 students that are appropriate for specific purposes in various contexts. The Finder is organized around three broad types of competencies: interpersonal, intrapersonal, and higher-order cognitive. This structure is based on a framework published in a National Research Council report.¹

Users interested in knowing more about the organizing framework, the selection of the included assessments, and the methodology used to develop the tool should consult the companion document, *Building a Repository of Assessments of Interpersonal, Intrapersonal, and Higher-Order Cognitive Competencies* (www.rand.org/t/RR2508). Users interested in further instructions about how to use assessments to guide their practice should refer to a second companion document, *Choosing and Using SEL Competency Assessments: What Schools and Districts Need to Know* (http://measuringsel.casel.org/pdf/practitioner-guidance.pdf).

Step 1: Identify Assessment Criteria

Identifying relevant assessments involves specifying both target learning goals and contexts in which assessments might be used. The context can affect the appropriateness of specific measures.

Three sections are presented to the user when they open the RAND Education Assessment Finder (Figure 2.1):

- **Section A** is an alphabetical list of the assessments in the database. The titles of the assessments are “live”—that is, any title can be clicked to review their details. This allows for “window shopping” if a user just wants to see what is available.
- **Section B** has check boxes that allow users to select up to three assessments for side-by-side comparison.
- **Section C** is where the user can enter broad criteria or select specific details and have the database filter results to match assessment goals and other details important to the user’s specific context.

When items are entered into section C filters, the list of measures in section A will increase or decrease automatically to match the criteria in the filters. If too few or too many assessments are presented, users can adjust their input to generate more or fewer results. The trick to initial filtering is to start with the criteria that are most meaningful and add other filters based on the initial list of measures.

![Figure 2.1. RAND Education Assessment Finder Opening Screen](image)
Using Selection Filters

Using filters early in the search process improves the likelihood of finding assessments that meet user needs early on in the search process. We provide brief descriptions of each filter for users to refer to as they fill in each field. More-detailed descriptions of the filters are in the next section of this guide, Step 2, for users wanting to interpret the search results.

Filter by Keyword

The first filter gives users a chance to search for assessments using terms that describe common IICC competencies. Examples include collaboration, confidence, optimism, motivation, self-efficacy, self-management, and decisionmaking.

When users enter a skill or competency here, the list of assessments is reduced to those that relate to that competency.

Grade Level

This filter allows the user to restrict the search to six grade spans. Check one or more of these boxes to find measures that are appropriate to use in certain grade ranges. This is likely one of the first considerations in choosing an appropriate assessment.

The database contains assessments designed for students in kindergarten through 12th grade. However, some of these assessments might also be designed for pre-kindergarten or postsecondary students in addition to K–12 students, which is why those filters exist. Assessments that are exclusively for use with preschool-aged children, postsecondary-aged students, or adults are excluded from the database.

Respondent

Although many assessments are completed by students, some measures are designed to be completed by teachers, parents, peers, and others. Each type of respondent offers a unique view of their own or another’s development in IICC. This filter allows the user to find assessments based on who will fill them out.

Method of Administration

This filter allows the user to restrict the search to assessments that are administered in a particular manner. Many assessments have standardized administration methods that keep the task identical and accessible for all respondents. Users should consider the method that is most appropriate for their respondents. For example, not all schools can offer access to digital mediums for an entire classroom at one time, or they might not have time to speak to every student individually.
Administration Time

This filter selects assessments based on the estimated length of time it takes to administer them. Taking an assessment requires time that is valuable in the course of a school day or out-of-school period. How much time can be dedicated to this assessment? The times of the assessments in the Finder range from ten minutes or less to more than 60 minutes. It should be noted that administration times are based on a typical setting and usually estimated by assessment developers.

Item Format

This filter refers to the type of responses that users make when completing the assessment. Here, users can choose what format best fits their intent.

- **Selected-response** assessments typically consist of multiple choice, true and false, or matching questions.
- **Free response** assessments usually consist of essay questions.
- **Performance task** assessments usually ask respondents to demonstrate their understanding and knowledge of a concept.
- **Oral responses** ask the respondent to demonstrate their understanding and knowledge in spoken form.

Fee for Use

With this filter, users can restrict their search based on the cost of the assessment. Users have three choices in this category.

- **Free and publicly available** indicates that users can obtain and administer the assessment without paying a fee; these are often available for download from the web.
- **Fee charged by developer** indicates that there is a cost for administration, and there might be additional charges for scoring. Because the fees change frequently and can vary according to such factors as the number of assessments purchased, we do not provide information about the fee other than whether one is charged.
- **No fee; limited access to instrument** indicates that there is no applicable fee but that there might not be a way to access the instrument without contacting developers directly or that access is limited in some way (e.g., the instrument is published in a journal that might require a subscription for access).

Filter by Competency

The last filter in the section enables users to restrict the search based on the three areas of competency developed by the National Research Council:

- **Cognitive**: critical thinking, problem solving, creativity, information literacy
- **Interpersonal**: communication, collaboration, empathy, trust, conflict resolution, leadership
- **Intrapersonal**: flexibility, adaptability, curiosity, perseverance, grit, integrity, self-regulation.

---

Step 2: Review Search Results

The second step enables users to take a closer look by reviewing detailed records for each assessment.

Information records for each assessment in the Finder are formatted in the same manner. There are five broad categories of information:

- general description
- administration
- scoring
- evidence of technical quality
- how to contact the developer.

To review the record for any single assessment in the Finder, simply click on its live link. Users can select up to three records at once to compare assessment information side by side. To do this, check the boxes to the right of each entry and then click “COMPARE SELECTED.” Figure 2.2 shows the record for the Active and Engaged Citizenship assessment.

Figure 2.2. Active and Engaged Citizenship Record
Categories of Information Provided

This section describes the information offered by each field in the Finder’s assessment records and offers definitions when appropriate. Remember that not all assessment records present the same amount of information. Users can contact developers or vendors for more information as needed.

General Information on the Measure

Purpose of the Measure

This part of the record displays the primary purpose of the assessment instrument as described by the vendor. This section also features what constructs are measured and the intended use of the assessment (if the vendor has provided the information).

Main Constructs Measured

This section notes whether interpersonal, intrapersonal, or cognitive competencies are measured by the assessment.

Applicable Grade Levels

This section tells users the age and/or education levels of the individuals whom the assessment is intended to measure.

Publication Year for Most Recent Version

This section displays the year in which the most recent version of the assessment was published.

Year Originally Developed

Assessments could be developed years before they are released for use, or they might be revised in the years that follow. This section tells users the year the assessment was first developed.

Notes

In this section are miscellaneous but relevant items of information related to the assessment and/or its development (e.g., whether the assessment was designed as a part of a specific intervention, or whether the assessment is a modification or update of another assessment).

Related Measures

This section lets users know whether there are any other assessments that are related to the one in the record. Related measures can include short forms of the assessment or versions intended to capture the responses of additional parties, such as parents. The link will be “live” if it is also in the Finder.
**Measure Administration**

**Respondent**
This section lets users know who will be responsible for completing the assessment. A respondent could be the student; their parent; or another reporter, such as a teacher.

**Method of Administration**
This section presents details about how an assessment is administered. Methods consist of pencil and paper, digital format, and interview questions.

**Number of Items and Format**
Here, users can see how many items are in the assessment and what format they are in. Formats consist of selected response (such as multiple choice), performance tasks, and extended response (such as an essay).

**Administration Time**
This section lets users know the approximate administration time of assessment as estimated by the developers.

**Available Languages**
This section lets users know whether the assessment is available in languages other than English.

**Fee for Use**
This section tells users whether they must pay a fee to access and use an assessment. Three categories are noted:

- **No**: There is no fee.
- **Yes**: There is a fee charged by the developer to access the instrument.
- **N/A**: There is no applicable fee, but there might not be a way to access the instrument without contacting developers, or access is limited in some other way.

**Credentials Required for Administration**
This section lets users know whether the assessment developer has indicated that the administrator of the assessment must possess specific credentials. Most assessments do not have such requirements, but some require that administrators meet specific criteria or participate in training. For example, some assessments require that administrators participate in interviewer training or receive training on survey administration techniques. Many commercially developed assessments have specific qualification criteria for users to purchase or administer the assessment. Most commonly, these criteria are referred to as “Qualification Level B,” which refers to an individual with one of the following:
• a master's degree in psychology, education, occupational therapy, social work, counseling, or a field closely related to the intended use of the assessment, plus formal training in the ethical administration, scoring, and interpretation of clinical assessments
• certification by or full active membership in a professional organization that requires training and experience in the relevant area of assessment
• a degree or license to practice in the health care or allied health care field
• formal, supervised mental health, speech/language, occupational therapy, social work, counseling, and/or educational training specific to assessing children or in the fields of infant and child development, plus formal training in the ethical administration, scoring, and interpretation of clinical assessments.

**Scoring**

**Overall Score Reporting**

This section tells users whether an overall score is reported for the assessment (as opposed to subscores that refer to parts of the assessment).

**Scoring Procedures**

This section describes the way that scores are calculated and who (or what) calculates the scores (e.g., hand-scored, computer-scored, online scoring).

**Subscore Reporting**

This section lists any subscores that are reported or calculated independently of the overall assessment score. Subscores are calculated using information from a portion of the assessment rather than the assessment as a whole and are typically intended to provide information about students’ performance on a specific competency as part of an assessment that measures multiple competencies.

**Interpretive Information**

This section lets users know whether the instrument developer provides any information to assist in score interpretation.

**Technical Quality**

**Populations for Which Technical Quality Evidence Has Been Collected**

This section presents information related to the demographic information about the sample that was used to collect evidence about score reliability and evidence about validity of inferences based on scores.

**Reliability Evidence**

This section relays any information related to the precision of the assessment—that is, the extent to which an assessment produces scores that are consistent and relatively free of...
measurement error. One way to think about reliability is this: A teacher administers an assessment to a student on one occasion, then administers it again on another occasion, and would expect the student to obtain roughly the same score each time, assuming the administration conditions were the same and nothing happened between administrations to cause the student’s skill or competency to change. An assessment that produces scores with low reliability is not particularly useful for decisionmaking. Typically, reliability indicators will take a value between 0 and 1, with values closer to 1 indicating higher reliability and lower measurement error. For more information on reliability evidence, see Choosing and Using SEL Competency Assessments: What Schools and Districts Need to Know (http://measuringsel.casel.org/pdf/practitioner-guidance.pdf).

Validity Evidence

This section relays any information about the extent to which evidence supports the intended inferences and uses of the assessment, including evidence related to how well the assessment measures what it was designed to measure. The evidence is grouped into four categories, and explanations of each category are provided in Choosing and Using SEL Competency Assessments: What Schools and Districts Need to Know (http://measuringsel.casel.org/pdf/practitioner-guidance.pdf). Users should not necessarily expect to see evidence related to each of the four categories; some sources are more relevant to certain assessment contexts than others. A key factor to keep in mind is that a single assessment is never demonstrated to be valid. Instead, users should look for evidence that supports an argument related to the assessment’s use for a specific purpose. Because interpreting validity evidence is complex and generally requires measurement expertise, users are encouraged to seek input from measurement experts to evaluate the adequacy and relevance of the available evidence for a particular assessment purpose. For more information on validity evidence, see the “Choosing and Using” guidance.

Locating the Measure

Obtaining a Copy of the Measure

If available, a link to the assessment is given to users in this section. If no website is available, the section offers other related information, such as the citation of the article in which the assessment is described.
3. User Experiences

These examples illustrate a few of the ways the assessment finder could be used by educators.

Example 1: Do My Elementary School Students Have a “Growth Mindset”?

1. Select Criteria

Ms. Steele is a third-grade teacher and has been reading a lot about the importance of students’ growth mindset over the summer. She wants to make sure that her students approach their work with the understanding that they are ultimately in control of improving their grades. An appropriate assessment will let her know whether intervention in this area is necessary and will support her request for assistance from her principal.

She begins her search by selecting appropriate filters. She knows she wants an assessment that measures “growth mindset,” and searches what is available for third- to fifth-grade students.

2. Review Search Results

The first step resulted in 27 assessments, illustrated partially in Figure 3.1. Ms. Steele reads the short descriptions of the assessments and selects two to examine and compare more closely. She chooses the How I Feel About Myself assessment because it measures students’ well-being specifically in a school context, along with the Implicit Theories of Intelligence Scale for Children because the description directly mentions “growth mindset.”

![Figure 3.1. Assessment Results for Example 1](image-url)
She compares the two assessment records side by side (Figure 3.2). She notes that the *Implicit Theories* assessment can be taken online by students in less than ten minutes and offers an overall score that will give her a quick understanding of her students’ growth mindset.

Ms. Steele clicks on the live link to the assessment and initiates contact with the developer to learn more.

---

**Example 2: Should I Try to Develop an Innovative Program to Support Students’ “Persistence” in Writing?**

1. **Select Criteria**

   Mr. Torres is the head of the English language arts program at a large high school. He has noted that too many students seem to give up on essays after the second—and even first—draft. He has heard that *persistence* is an important IICC and would like to know whether it is possible...
to foster that quality in ninth- and tenth-grade student writers. He is not sure how to measure persistence, how much of it his students already have, or whether they are more or less persistent than students in other schools. He would like to discuss this with the teachers in the English Department and is looking for relevant information. They are not looking for a specific measure now, but he would like to get some information to share.

The user begins his search by selecting appropriate filters. In the “Filter Measures” section, he types “writing” in the keyword search box. He also selects assessments for grades 9–12 because something for younger students will likely not be appropriate (Figure 3.3).

Figure 3.3. Assessment Results for Example 2
2. Review Search Results

The first step resulted in three assessments. Mr. Torres chooses to compare all three at one time. He notes several similarities and differences, including how persistence is defined and what students are asked to do to measure it.

Mr. Torres sends the record to other English language arts teachers. At the meeting, he reviews the three measures with the other teachers. They agree that the assessment will help them understand students’ unwillingness to work on second and third drafts of their writing. They decide to eliminate one measure because it relies on behavioral observation, and they would prefer a traditional written online assessment (Figure 3.4). Mr. Torres contacts the vendors of the remaining two assessments with plans to present the information to the group in the next meeting.

![Figure 3.4. Assessment Comparison for Example 2](image-url)
Example 3: Are Students in Our After-School Program Developing Creativity Skills?

1. Select Criteria

Mrs. Yang and Mr. Taylor developed an after-school program to foster creative thinking through a hands-on approach to everyday science-related problems. After the first year, the team wants to know if the program is having an effect on students’ thinking. An assessment of creativity could be part of their efforts to judge program impact and think about making changes in the program if needed. Without a comparison group, they cannot draw strong conclusions about the effects of the program. However, if they can find an appropriate measure that is sensitive to changes over a year’s time, they could report changes to school and district administrators. This could offer some relevant evidence.

The users begin their assessment search by typing “creativity” in the keyword search box. They limit the search to “pencil and paper” methods of administration (there are no computers available during after-school hours) and know they cannot spend more than 30 minutes on this task. They also choose to limit selections by their students’ middle-school grade level (6–8) (Figure 3.5).

Figure 3.5. Assessment Results for Example 3
2. Review Search Results

The first step resulted in five assessments. Mrs. Yang and Mr. Taylor note that three of the five directly call out “creativity” in their descriptions. They choose to compare two: the Creativity Assessment Packet (The Williams Scale) and the Creativity Assessment Packet (Test of Divergent Thinking), which are two related assessments that come from the same bundle of assessments (Figure 3.6).

The team is immediately drawn to the second of these—the Test of Divergent Thinking—because it offers subscores on curiosity, imagination, complexity, and risk-taking. These factors, if correctly measured, could inform specific program changes if necessary, which could add to the robustness of reports to different committees by including such detail. The team immediately contacts the developer for more information, including whether there is validity evidence to support the use of the assessment as a measure of change in students’ creativity from one year to the next.

Figure 3.6. Assessment Comparison for Example 3