IES Institute of Education Sciences

Regional Educational Laboratory Program

Developing and Validating Measures of Research Evidence Use, Research Capacity, and Partner Experiences with the Regional Educational Laboratories

August 2026
REL 2026-016
U.S. DEPARTMENT OF EDUCATION

A Publication of the National Center for Education Evaluation and Regional Assistance at IES

U.S. Department of Education
Linda McMahon
Secretary

Institute of Education Sciences
Matthew Soldner
Acting Director

National Center for Education Evaluation and Regional Assistance
Matthew Soldner — *Commissioner*
Sarah Brasiel — *Project Officer*

The Institute of Education Sciences (IES) is the independent, nonpartisan statistics, research, and evaluation arm of the U.S. Department of Education. The IES mission is to provide scientific evidence on which to ground education practice and policy and to share this information in formats that are useful and accessible to educators, parents, policymakers, researchers, and the public.

We strive to make our products available in a variety of formats and in language that is appropriate for a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other IES product or report, we would like to hear from you. Please direct your comments to ncee.feedback@ed.gov.

This report was prepared for the Institute of Education Sciences (IES) under Contract 91990023C0008 by Mathematica. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

August 2026

This report is in the public domain. Although permission to reprint this publication is not necessary, it should be cited as:

Bleeker, M., Crigler, F., Hurwitz, F., Carmichael, K., Poznyak, D., Neild, R. C. (2026). *Developing and validating measures of research evidence use, research capacity, and partner experiences with the Regional Educational Laboratories* (REL 2026-016). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/use-work/resource-library/reports

Authors' ORCID IDs:
Martha Bleeker 0000-0002-3213-3276
Forest Crigler 0009-0008-5509-0636
Felicia Hurwitz 0009-0006-8252-8467
Katherine Carmichael 0009-0000-4021-5709
Dmitriy Poznyak 0000-0002-9610-8424
Ruth Curran Neild 0000-0002-8696-3494

Developing and Validating Measures of Research Evidence Use, Research Capacity, and Partner Experiences with the Regional Educational Laboratories

Martha Bleeker, Forest Crigler, Felicia Hurwitz, Katherine Carmichael, Dmitriy Poznyak, and Ruth Curran Neild

August 2026

The 10 Regional Educational Laboratories (RELs), administered by the Institute of Education Sciences (IES), partner with state and local education agencies to support the use of research evidence in improving education outcomes. IES and the RELs have used surveys of their agency partners to assess REL performance and improve their products and services. To enhance data quality and strengthen the measures RELs can use to assess their services, IES commissioned a study during the 2022-27 REL cycle to develop reliable and valid survey measures. These measures assess REL partners' use of information in decisionmaking, their capacity to engage with and use research evidence, and their perceptions of REL services and products. The study team identified key constructs in each area and selected survey items to measure them. The team assessed reliability and validity using an online survey of REL partners and follow-up interviews with a subset of survey respondents.

The analysis confirmed sets of survey items that the RELs and other research partnerships can use to reliably measure 14 distinct constructs. The results showed that the items within each construct captured the same underlying dimension. Follow-up interviews with 10 REL partners further supported that the items reflected what they were intended to measure.

Findings based on these validated measures suggest that respondents draw on a mix of research evidence, professional experience, and local data when making professional decisions and addressing challenges. Although respondents expressed moderate confidence in their ability to use research evidence, they also noted opportunities to enhance its relevance and practical value for schools and districts. In addition, they highlighted the need for stronger organizational infrastructure and support to integrate research more effectively into decisionmaking processes. They also reported valuing their engagement with the RELs and the evidence and services the RELs provide.

The 10 Regional Educational Laboratories (RELs) partner with state and local education agencies to support the use of research evidence in education practice and decisionmaking. Together with their partners, the RELs design research studies, develop evidence-informed tools and professional development products, and provide engaging learning experiences and consultation about use of research evidence. The RELs collect information about their partners' use of research evidence through a Stakeholder Feedback Survey (SFS), which includes a set of required items and a bank of optional items administered to participants after the

The authors are grateful to the many people who contributed to this study and report. First and foremost, we thank the REL partners who participated in the study, along with REL staff who assisted with and provided information for the study. We also thank former IES staff who provided important and insightful technical input and study support. They include Chris Boccanfuso, Elizabeth Eisner, Amy Johnson, and Elizabeth Nolan. The study team also received important recommendations from content experts Elizabeth Farley-Ripple, Drew Gitomer, Caitlin Farrell, and Mark Rickinson. We also thank Maria Bartlett and Yange Xue for their contributions to the analysis, Jennifer Brown for editorial services, Dorothy Bellow and Stephanie Barna for helping format study documents and prepare the report for publication, and Lisbeth Goble for providing thoughtful feedback on data collection materials.

For additional information, including data collection methods and technical details about the analysis, see the appendices for the report on the Institute of Education Sciences website.

completion of REL training, coaching, and technical support activities and administered annually to REL partnership members. These data help REL staff understand the utility of activities, identify areas of partner need or concern, and inform future research and technical support activities. However, prior to this study, the items in the SFS had not been assessed for reliability and validity.

To improve the quality of these data and strengthen the measures the RELs can use to assess their services, the Institute of Education Sciences commissioned this study to develop and test survey measures that assess partners' use of information in decisionmaking, their capacity to engage with and use research evidence, and their perceptions of REL services and products. This study identified existing relevant constructs and survey items and evaluated their reliability and validity, and it produced descriptive findings in each of these three areas. The study also advances survey measures and empirical understanding of use of research evidence, which has been an emerging area of attention in education over the past decade (Farley-Ripple et al., 2020; Gitomer & Crouse, 2019; Penuel et al., 2017; Rickinson et al., 2022). Other education research partnerships can also use these measures to assess their partners' experiences with and use of research evidence.

Key terms used in the report, along with brief explanations of their use in the study, are defined in box 1.[2] The initial set of constructs assessed for reliability and validity is presented in box 2. The data sources, sample, and methods are summarized in box 3 and detailed in appendix A.

Research questions

The study addressed the following research questions:

1. What survey items provide reliable and valid measures of the survey constructs?
    a. To what extent do relevant items measure the same underlying construct? Which, if any, items should be removed to improve or streamline measurement of the underlying concept? To what extent do the items consistently assess what they were designed to measure?
    b. Based on follow-up interviews with survey respondents, to what extent do the items measuring instrumental use of information assess what they are intended to measure?
2. To what extent do REL partners who participated in the study report using different types of information to make decisions, report their capacity to engage with and use research evidence, and report valuing REL resources, based on the measures validated in this study?
    a. In general, what sources of information do study participants use when making decisions about policies and programs, developing new activities, or implementing programs?
    b. What are REL partners' perceptions of their capacity to engage with and use research evidence? Do their organizations support their use of research evidence?
    c. How do REL partners who participated in the study perceive their engagement with RELs and the evidence and services provided?

2. The study measures use two related but distinct terms—*research* and *research evidence*—to fully capture the experiences and perceptions of REL partners. The term *research* refers to the systematic collection and analysis of data driven by research questions, whereas *research evidence* refers to the information produced by research. To help simplify measure names, the study team used the term *research evidence* when naming the study constructs, even though the underlying survey items also touch on aspects of *research* production.

Box 1. Key terms

Content validity. The extent to which a measure covers the range of concepts it is intended to assess. The study team worked with content experts to confirm the items represented the scope of each construct being measured.

Construct validity. The extent to which a measure accurately assesses the theoretical concept it is intended to measure. The study team used confirmatory factor analysis to assess construct validity and conducted follow-up interviews with REL partners to further assess whether certain survey items measured what they were intended to measure.

Convergent validity. The extent to which survey items measure the same underlying dimension or construct. The study team used factor loadings from confirmatory factor analysis as a key indicator of convergent validity. Factor loadings indicate the strength of the relationship between an observed item and its hypothesized underlying factor. Measurement experts generally consider factor loadings of at least .50 as evidence of acceptable convergent validity, with loadings of .70 or higher indicating a strong association with the underlying construct.

Discriminant validity. Evidence that a construct is truly distinct from other constructs—both conceptually and statistically. In other words, it shows that a construct measures what it is supposed to measure and not something else.

Face validity. The extent to which a measure assesses what it is intended to measure, based on a subjective review of the measure. The study team worked with content experts to establish face validity of the survey constructs.

Factor. The statistical representation of a survey construct, derived from an analysis of the patterns of how respondents answer survey items related to that concept. Factor analysis is a statistical technique used to simplify complex data by grouping related variables (or survey items) into underlying factors. These factors are thought to represent the underlying constructs that influence responses to those questions. The study team used confirmatory factor analysis with Bayesian estimation methods to confirm whether responses for survey items for each construct grouped together as one factor.

Internal consistency reliability. A measure of reliability that assesses the extent to which the survey items within a survey construct are consistently measuring the same concept. This is determined by assessing the extent to which respondents rate items within a construct in a similar manner. The study team assessed internal consistency by calculating Cronbach's alpha and McDonald's omega for each construct, with higher scores (ranging from 0 to 1) indicating stronger reliability.

REL core partner. Participants in REL partnerships who are involved in planning the overall direction of the partnership with REL staff to address high-leverage needs. For example, a core partner may be an agency leader who is involved in planning partnership activities but does not necessarily attend the activities if they are designed for people in different roles (such as teachers or building leaders). Core partners often include district- or state-level staff and school-level administrators.

REL noncore partner. Participants in REL partnerships who are not highly involved in the overall planning, direction, or strategy of the partnership but receive REL services (such as training and coaching) and contribute their expertise to applied research and dissemination projects. They may have attended planning meetings for projects and activities under the partnership. Noncore partners often include district-level staff, school-level administrators, and teachers.

REL partnership. A structured collaboration between a REL and education agencies to develop research, tools, and learning supports that strengthen evidence-informed decision-making. Beginning in 2022, each REL developed several partnerships to improve long-term student success on a high-leverage topic within a specific state. Partnerships often include staff at the state, district, and school levels and have focused topics such as building the capacity for designing and evaluating professional learning for math coaches and teachers, strengthening the teacher pipeline, supporting the use of data to improve students' preparation for college and careers, and using data to foster supportive learning environments.

REL project. A specific activity or piece of work that aims to advance the goals of a REL partnership. Examples of projects include conducting research studies, developing professional learning materials and toolkits, and providing coaching or analytic support to local educators or district staff.

REL Program. A national network of 10 RELs that partner with regional educators and policymakers. REL work is change oriented and supports meaningful local, regional, and state decisions about education policies, programs, and practices

designed to improve learner outcomes. The RELs provide technical support, conduct research, and offer research-based learning opportunities to inform changes to policy and practice in an effort to improve students' education outcomes.

Research. Systematic data collection and analysis driven by research questions. The study team included this definition in the survey to ensure a common understanding among respondents.

Research evidence. The information produced by research. This includes evidence in reports, journal articles, and research summaries, as well as evidence embedded in curricula, practice tools, and assessments. The study team included this definition in the survey to ensure a common understanding among respondents.

Survey construct. A theoretical concept or factor that researchers measure using survey items. In this study, the team identified key constructs and selected survey items to measure each construct in three areas: partners' use of information in decisionmaking, their capacity to engage with and use research evidence, and their perceptions of REL services and products.

Box 2. Initial list of survey constructs

The study team worked with content experts and Institute of Education Sciences staff to identify an initial set of 13 survey constructs across three categories: use of information from a variety of sources, including research evidence, to make decisions (commonly described as "instrumental use" of information); partners' capacity to engage with and use research evidence; and partners' perceived value of Regional Educational Laboratory (REL) services and products. These constructs were based primarily on existing measures with prior evidence of reliability and validity (see appendix A for more detail). The study team refined and tested the measures for reliability and validity with REL partners, which resulted in the addition of one construct. The final set of 14 constructs reflect revisions made during the analysis and are described in the *Findings from the validation of survey measures* section and presented in table 5.

Use of information from a variety of sources to inform decisionmaking (or "instrumental use" of information)

These six constructs were based on items from Brennan et al. (2017), Lysenko et al. (2014), and Penuel et al. (2017).

- **Instrumental use of research evidence.** Extent to which research evidence is used when making decisions about policies and programs, developing new activities, or implementing programs.
- **Instrumental use of data.** Extent to which data from one's school, district, institution, or state are used when making decisions about policies and programs, developing new activities, or implementing programs.
- **Instrumental use of own expertise.** Extent to which one's own experience and expertise are used when making decisions about policies and programs, developing new activities, or implementing programs.
- **Instrumental use of colleagues' expertise.** Extent to which colleagues' input and feedback are used when making decisions about policies and programs, developing new activities, or implementing programs.
- **Instrumental use of feedback from students, parents, and community.** Extent to which feedback from students, parents, and the community is used when making decisions about policies and programs, developing new activities, or implementing programs.
- **Instrumental use of professional organizations and other resources.** Extent to which resources from professional organizations and other technical assistance providers are used when making decisions about policies and programs, developing new activities, or implementing programs.

Capacity to engage with and use research evidence

- **Confidence in ability to use research evidence.** Confidence and capacity to conduct and evaluate research and use the evidence produced by research. Items in this construct were based on items from Brennan et al. (2017).
- **Usefulness of research evidence.** Perceptions about the usefulness and value of research and the evidence produced by research. Items in this construct were based on items from May et al. (2022).
- **Conceptual use of research evidence.** Extent to which research evidence is used to understand issues and challenges and discuss possible solutions. Items in this construct were based on items from Penuel et al. (2017).
- **Organizational support for use of research evidence.** Extent to which one's organization expects and encourages the generation and use of research evidence. Items in this construct were based on items from Brennan et al. (2017) and Rickinson et al. (2021).

- **Organizational infrastructure for use of research evidence.** Perceptions about the resources and processes one's organization provides to support use of research evidence. Items in this construct were based on items from Rickinson et al. (2021).

Perceived value of REL services and products

- **Value of REL engagement.** Extent to which REL activities and engagement are valued. Items in this construct were drawn from existing survey items in the REL Stakeholder Feedback Survey.
- **Value of REL evidence and services.** Extent to which REL research evidence and services are relevant and accessible. Items in this construct were based on items from Lysenko et al. (2014). Based on findings from the current study, the team divided this initial construct into two distinct constructs: relevance of REL evidence and services and accessibility of REL evidence and services. See the *Findings from the validation of survey measures section* and table 5 for additional details.

Box 3. Data sources, study sample, and methods

Data sources. This study used two data sources, and all data were collected in 2024.

- **Online survey.** The survey measured partners' use of information from a variety of sources to make decisions, their capacity to engage with and use research evidence, and their perceived value of REL services and products.
- **Follow-up interviews.** The study team conducted follow-up interviews with a subset of survey respondents. Interviews assessed the construct validity of the six "instrumental use" constructs by examining whether survey responses aligned with partners' actual practices over time. The study team also gathered feedback on the usability and format of all survey items.

Study sample. This study involves partners who were working with the Regional Educational Laboratories (RELs) operating during the 2022-27 cycle. The study team used background information on REL partnerships to identify a purposive sample for the survey and interviews.

- **Survey sample.** The study team invited 166 partners across nine RELs to participate in the study, representing a typical mix of core and noncore partners in roles such as teachers, administrators, and staff at the district, state, and postsecondary levels. A total of 103 partners completed the survey (64 percent response rate[1]), providing sufficient power[2] for reliability and validity analyses. Respondents represented a mix of core and noncore REL partners from 11 REL partnerships, nine RELs, and 11 REL projects. One of the 10 RELs was not included in the study because its partnerships' activities fell outside of the targeted timeline.
- **Interview sample.** The study team invited 38 survey respondents to participate in follow-up interviews based on their survey responses; interviews were conducted approximately three to four months after survey completion. The purposive sample included partners who frequently reported using research evidence and other sources of information in decisionmaking, those with high or low ratings of REL evidence and services, and respondents with survey patterns that warranted further examination. The team ensured representation across roles and experience levels and included all nine RELs. Ten partners ultimately completed interviews, representing seven RELs and seven REL partnerships.

Methodology. The study team used the following methods to address the research questions.

- *Research Question 1a. To what extent do relevant items measure the same underlying construct? Which, if any, items should be removed to improve or streamline measurement of the underlying concept? To what extent do the items consistently assess what they were designed to measure?*
  - The study team used confirmatory factor analysis (CFA) with Bayesian estimation methods to examine whether items measured the same underlying construct (convergent validity). CFA was used to determine whether responses to items within each construct grouped together on a single factor, indicating that they capture the same concept. The team selected Bayesian estimation because it produces valid parameter estimates and standard errors even with sample sizes below 100. See appendix B for more details. The team reviewed the CFA factor loading estimates, which show the strength of the relationship between each survey item and the underlying construct, and identified for possible removal any items that were insufficiently aligned with their construct. Loadings between .50 and .95 indicate

a good association, with values of .70 or higher reflecting strong alignment (Kline, 2016). Loadings below .40 suggest weaker relationships with the underlying construct (Stevens, 2009).
    - To further assess whether items were statistically associated with a factor or underlying construct, the study team also calculated posterior predictive p-values. Nonsignificant values above .05 indicate good model fit, with values near .50 representing particularly strong fit (Meng, 1994). See appendix B for more details.
    - The study team also reviewed the magnitude of factor loadings to identify potentially redundant survey items.[3] Items with loadings within ±.03 of one another (e.g., .92, .95) were considered closely similar. Because small differences among high loadings have a small effect on reliability (calculated as the squared loading), the team considered whether items with similar loadings provided unique value or relevance to the REL context before recommending removal.
    - To assess the reliability of items within each construct, the study team calculated internal consistency using Cronbach's alpha and McDonald's omega, both of which range from 0 to 1, with higher values indicating greater reliability. In general, a Cronbach's alpha or McDonald's omega of .7 is considered acceptable, and .8 or higher is desirable. See appendix B for more details.
- *Research question 1b. Based on follow-up interviews with survey respondents, to what extent do the items measuring instrumental use of information assess what they are intended to measure?*
    - To test construct validity for the items measuring instrumental use of information, the study team conducted follow-up interviews with survey respondents. Interview questions focused on whether participants described behaviors that were aligned with their survey responses, helping to determine whether survey items accurately represented how respondents used information in practice.
- *Research question 2. To what extent do REL partners who participated in the study report using different types of information to make decisions, report their capacity to engage with and use research evidence, and report valuing REL resources, based on the measures validated in this study?*
    - The study team calculated descriptive statistics, including means and standard deviations, based on the final set of survey items for each construct. These values represent respondents' average scores across the full sample. No statistical significance tests were conducted because the study aimed to describe patterns in use of information sources, research capacity, and perceived value of REL services and products rather than compare their relative frequency.

Additional details about the data collection methods are described in appendix A, and additional technical details about the analysis are provided in appendix B.

Notes

1. Of the 166 REL partners invited to participate in the survey, 5 were removed from the final sample and response rate calculation because the study team received email bounce backs and confirmation from the RELs that these partners had left their district or organization before the partnership sessions ended. The 58 REL partners who did not respond to the survey represented eight RELs, 10 REL partnerships, and nine REL projects. Like respondents, nonrespondents included a mix of core and noncore partners from a variety of education roles.

2. The study team conducted power calculations (Arifin, 2017; Bonett, 2002) and Monte Carlo simulations (Lee, 2015) to determine what sample sizes were needed to replicate previous types of analyses conducted on existing items. Results from the power calculations suggested that at least 30 participants were needed to calculate Cronbach's alpha for each construct and at least 50 participants were needed to calculate CFA.

3. The study team used a categorical CFA model, which is mathematically equivalent to a graded-response item response theory model (Muthén, 1984). In this framework, factor loadings correspond to discrimination parameters, indicating how strongly each item relates to the latent construct. Items with similar loadings—especially when paired with high residual correlations—may be redundant because they contribute little unique information. The team also examined item difficulty parameters (or thresholds), which mark points on the latent trait where respondents are likely to shift from one response category to the next (e.g., from "not at all agree" to "somewhat agree"). Nonsignificant thresholds suggest weak separation between adjacent response options, whereas unusually wide spacing between thresholds indicates that some response categories are rarely or inefficiently used. These patterns may imply that the item is not effectively distinguishing respondents along the scale.

Findings from the validation of survey measures

Across all constructs, items generally measured the same underlying concept, with strong model fit and high factor loadings. To streamline the measures, the study team removed several overlapping or weakly aligned items, and reliability indicators showed that the remaining items assessed intended concepts consistently.

All six constructs measuring instrumental use of information to make decisions functioned well without requiring revisions

Statistical tests provided evidence that each set of items measured a distinct construct related to instrumental use of different types of information (research evidence, data, the respondent's own expertise, colleagues' expertise, community feedback, and resources from professional organizations and other providers). Posterior predictive p-values were nonsignificant and ranged from .43 to .51 (table 1), indicating good fit. In addition, individual survey items were closely related to the construct they were designed to measure, with all but one of the statistical factor loadings above .70 and none below .60 (see table 5 later in the report). Although a few items within the same construct had similar factor loadings, suggesting some overlap, the study team determined that each item captured a distinct and important dimension of the associated construct. For example, two items in the instrumental use of colleagues' expertise construct–which asked how often respondents used colleagues' input, feedback, or consultation "when developing new activities, programs, or guidelines" and "when implementing activities and programs"–had factor loadings of .98; however, these separate items help measured the range of concepts this construct was intended to assess (content validity), so the study team decided to retain both items.

The survey constructs also demonstrated good reliability, meaning the questions within each survey construct worked well together to assess the underlying concept. Across the six constructs, Cronbach's alpha ranged from .85 to .90, and McDonald's omega ranged from .92 to .97 (see table 1).

All five constructs measuring capacity to engage with and use research evidence functioned well; revisions were made to three constructs to reduce respondent burden and remove redundant survey items

For two constructs–confidence in ability to use research evidence and usefulness of research evidence–statistical tests indicated good model fit, with nonsignificant posterior predictive *p*-values of .37 and .17, respectively. All factor loadings were above .72, suggesting a strong association between the individual survey items and their intended latent constructs (see the second section of table 5 later in the report). These two constructs also demonstrated strong reliability, with Cronbach's alpha values of .93 and .89 and McDonald's omega values of .97 and .93, respectively (table 2).

Initial analyses indicated good model fit for the constructs measuring conceptual use of research evidence and organizational infrastructure for use of research evidence and acceptable model fit for the construct measuring organizational support for use of research evidence. Items within each construct loaded together on one factor as expected, all posterior predictive *p*-values were greater than .05, and the initial factor loadings for items

Table 1. Final reliability metrics for the constructs assessing use of information to inform decisionmaking

Construct	Number of items	Posterior predictive *p*-value	Cronbach's alpha	McDonald's omega
Instrumental use of research evidence	4	.51	.89	.92
Instrumental use of data	4	.43	.85	.92
Instrumental use of own expertise	4	.43	.89	.97
Instrumental use of colleagues' expertise	4	.47	.90	.96
Instrumental use of feedback from students, parents, and community	4	.50	.85	.94
Instrumental use of professional organizations and other resources	4	.50	.86	.94

Note: Posterior predictive *p*-values with nonsignificant values above .05 indicate good model fit, with values near .50 representing particularly strong fit. Cronbach's alpha and McDonald's omega both range from 0 to 1, with higher values showing stronger reliability. Reliability coefficients of approximately .70 are generally acceptable, and values of .80 or above are viewed as strong. See appendix B for more details.

Source: Authors' analysis of data from the online survey conducted in 2024.

Table 2. Final reliability metrics for the constructs assessing capacity to engage with and use research evidence

Construct	Number of items	Posterior predictive *p*-value	Cronbach's alpha	McDonald's omega
Confidence in ability to use research evidence	5	.37	.93	.97
Usefulness of research evidence	5	.17	.89	.93
Conceptual use of research evidence	5	.43	.90	.94
Organizational support for use of research evidence	5	.46	.87	.91
Organizational infrastructure for use of research evidence	3	.60	.83	.88

Note: Posterior predictive *p*-values with nonsignificant values above .05 indicate good model fit, with values near .50 representing particularly strong fit. Cronbach's alpha and McDonald's omega both range from 0 to 1, with higher values showing stronger reliability. Reliability coefficients of approximately .70 are generally acceptable, and values of .80 or above are viewed as strong. See appendix B for more details.

Source: Authors' analysis of data from the online survey conducted in 2024.

were above .60. However, the study team identified several items that were either redundant or insufficiently aligned with their constructs. The team removed these items to reduce respondent burden while maintaining reliability.

For conceptual use of research evidence, the six-item construct initially demonstrated good model fit (posterior predictive *p*-value = .44), with all factor loadings above .82. However, the study team removed one item ("Change the way you look at problems facing your school, district, institution, or state") because its content overlapped with other items in the construct and had similar factor loadings as those items. In addition, there was a wide gap between item difficulty parameters (or thresholds), suggesting that respondents did not use the response scale as consistently as other items.

For organizational support for use of research evidence, the original seven-item construct demonstrated acceptable model fit (posterior predictive *p*-value = .13), although a value closer to .50 would indicate better model fit (Meng, 1994). All factor loadings were above .64, but the similarity across loadings suggested some items were measuring similar aspects of the same construct. The study team removed two items. One item ("Leaders generally convey that it is important to use research evidence in policy or program development") was removed because it was created by the study team rather than drawn from an established measure and because it was conceptually similar to other items addressing leadership support for research use; in addition, there were wide gaps between item difficulty parameters (or thresholds), indicating that response categories were not functioning as expected, potentially leading to poor measurement precision. Another item ("We are encouraged to engage in informed risk-tasking in educational practices") was removed because it had the lowest factor loading (.64) and did not explicitly reference research evidence; in addition, the item difficulty parameter for one of the thresholds was nonsignificant, suggesting that the respondents did not reliably distinguish between adjacent response categories. After these changes, model fit improved (posterior predictive *p*-value = .46; see table 2).

For organizational infrastructure for use of research evidence, the initial four-item construct showed good model fit (posterior predictive *p*-value = .53), with all factor loadings above .65. However, the study team removed one item ("There are processes in place to support collaborative learning") because it had the lowest factor loading (.65) and was not clearly tied to use of research evidence. In addition, there was a wide gap between item difficulty parameters (or thresholds), suggesting that respondents did not use the response scale consistently and that the item did not discriminate effectively across levels of the latent factor.

After changes to these three constructs, posterior predictive *p*-values ranged from .43 to .60 (see table 2), and all factor loadings were .77 or above (see the second section of table 5 later in the report), suggesting that the items are closely related to the underlying constructs. The three revised constructs also demonstrated good

reliability, with Cronbach's alpha values ranging from .83 to .90 and McDonald's omega values ranging from .88 to .94 (see table 2).

Constructs measuring the perceived value of REL services and products functioned well; one construct was refined by removing redundant survey items and another was divided into two separate constructs to reflect distinct underlying dimensions

For the value of REL engagement construct, items loaded together on one factor as expected, the posterior predictive *p*-value was nonsignificant (p = .20; table 3), and the factor loadings were above .70. These results indicate that the survey items measure the underlying construct. However, several items had highly similar factor loadings, suggesting conceptual overlap among them. To reduce respondent burden without compromising internal consistency reliability or content validity, the study team removed two survey items ("REL engagement is relevant to my work" and "REL engagement increased my desire to support and participate in change in my school, district, institution, or state"), which were created by the study team rather than drawn from an established measure and whose content was already represented by other items in the construct. After these changes, factor loadings were .72 or above, suggesting that the items are closely related to the underlying construct (see the final section of table 5 later in the report).

For the value of REL evidence and services construct, the posterior predictive p-value was statistically significant (p = .03), which indicates possible model misfit and could suggest that the items measured more than one underlying construct. Based on these findings, the study team divided the seven items into two separate constructs: one focused on relevance of REL evidence and services and one focused on accessibility of REL evidence and services. The posterior predictive p-values for the two separate factors were not statistically significant, suggesting that these items may indeed represent two distinct constructs. In addition, the factor loadings for both constructs exceeded .70, indicating strong associations between the items and the two underlying constructs that capture distinct dimensions of relevance and accessibility. Finally, although the two factors are correlated (r = .81), the value falls below common thresholds for discriminant validity, implying that the constructs overlap but are not interchangeable.

After the items were divided into two separate constructs, the factor loadings for two items from the construct measuring accessibility of REL evidence and services were similar and suggested potential redundancy. The study team removed the item that was less reflective of the accessibility construct (e.g., "Are reliable"). After these changes, factor loadings for the two constructs were .74 or above, suggesting that the items are closely related to the underlying constructs (see the final section of table 5 later in the report).

Across all three constructs measuring the perceived value of REL services and products, internal consistency reliability was strong. Cronbach's alpha values ranged from .84 to .92, and McDonald's omega values ranged from .91 to .96 (see table 3).

Table 3. Final reliability metrics for constructs measuring perceived value of Regional Educational Laboratory (REL) services and products

Construct	Number of items	Posterior predictive *p*-value	Cronbach's alpha	McDonald's omega
Value of REL engagement	7	.20	.92	.96
Relevance of REL evidence and services	3	.43	.84	.91
Accessibility of REL evidence and services	3	.45	.85	.91

Note: Posterior predictive *p*-values with nonsignificant values above .05 indicate good model fit, with values near .50 representing particularly strong fit. Cronbach's alpha and McDonald's omega both range from 0 to 1, with higher values showing stronger reliability. Reliability coefficients of approximately .70 are generally acceptable, and values of .80 or above are viewed as strong. See appendix B for more details.

Source: Authors' analysis of data from the online survey conducted in 2024.

Follow-up interviews confirmed that the six constructs measuring use of information to make decisions accurately assessed the intended concepts

The study team conducted follow-up interviews to assess how well survey responses aligned with REL partners' actual practices in the months after completing the survey. The interviews also served to confirm with participants that the study team's interpretation of each construct accurately reflected their experiences.

These interviews focused on validating survey items related to the instrumental use of research evidence, data, personal expertise, colleagues' expertise, community feedback, and professional organizations when making decisions. To validate these items, the study team identified respondents who reported "often" or "always" using research evidence (or other types of information) in their decisionmaking. The study team then asked respondents to describe specific examples of how they used this type of information to make decisions over the past three months. Respondents were given a category (e.g., research evidence, data, colleagues' expertise) and asked to provide an example of the information they used, so no additional coding of responses was required.

The study team transcribed all interviews and held peer debriefing sessions after each one. During these sessions, team members reviewed participants' responses related to each category of information and critically examined their own interpretations to minimize personal bias and strengthen the credibility of the findings. After completing the interviews, the team also examined responses across participants to assess the consistency of evidence supporting each instrumental use construct. These discussions ensured that responses were mapped to constructs in a consistent way.

In general, based on responses from all 10 interview respondents, the study team was able to confirm that the items worked as expected. Interview respondents were consistently able to provide specific examples of what they did in practice over the past several months, validating their survey responses. Their ability to describe recent real-world examples also helped rule out that survey respondents were not just providing socially desirable survey responses (May et al., 2022). Examples of activities described by respondents in the followup interviews, mapped to each construct measuring use of information, are presented in table 4.

After completing the follow-up interviews and conducting the previously described analyses, the study team finalized 14 survey constructs. These constructs, along with their corresponding items, response scales, and factor loadings, are presented in table 5.

Table 4. Examples of instrumental use of information used to make decisions, shared by Regional Educational Laboratory partners

Construct	Examples provided in interviews
Instrumental use of research evidence	• Reviewed evidence-based practices identified in research when considering new interventions • Reviewed best practices for teacher recruitment and retention • Used research to create professional development plans
Instrumental use of data	• Collected outcomes on programs administered over the course of the year and used data to make program modifications • Reviewed comparable data from the region about staff compensation and retention • Used survey data from other departments to create professional development plans
Instrumental use of own expertise	• Relied on own experiences when no other staff members in the building had the same role • Incorporated past experiences working in other districts when planning a program presentation
Instrumental use of colleagues' expertise	• Consistently bounced ideas off colleagues • Gathered teacher feedback before implementing new programs
Instrumental use of feedback from students, parents, and community	• Considered feedback from students/parents when developing action steps and goals • Considered student/parent feedback provided on surveys when making decisions
Instrumental use of professional organizations and other resources	• Collaborated closely with partners at local department of education office • Attended professional development sessions with other organizations and brought takeaways back to school

Source: Authors' summary of data from the follow-up interviews conducted in 2024.

Table 5. Final list of survey constructs and survey items, with factor loadings

Construct and item	Final factor loading
Constructs assessing use of information to inform decisionmaking	
Instrumental use of research evidence Response scale: Never (1), Sometimes (2), Often (3), Always (4), Not applicable (This is not part of my role) *How often do you use research evidence, including evidence provided by RELs, when …*	
Deciding about the content or direction of a policy or program?	.92
Developing new activities, programs, or guidelines in your role or agency?	.94
Implementing activities and programs?	.92
Resolving problems in your daily professional practice?	.65
Instrumental use of data Response scale: Never (1), Sometimes (2), Often (3), Always (4), Not applicable (This is not part of my role) *How often do you use data from your school, district, institution, or state, when …*	
Deciding about the content or direction of a policy or program?	.81
Developing new activities, programs, or guidelines in your role or agency?	.90
Implementing activities and programs?	.96
Resolving problems in your daily professional practice?	.79
Instrumental use of own expertise Response scale: Never (1), Sometimes (2), Often (3), Always (4), Not applicable (This is not part of my role) *How often do you use your own experience and expertise when …*	
Deciding about the content or direction of a policy or program?	.98
Developing new activities, programs, or guidelines in your role or agency?	.95
Implementing activities and programs?	.91
Resolving problems in your daily professional practice?	.91
Instrumental use of colleague's expertise Response scale: Never (1), Sometimes (2), Often (3), Always (4), Not applicable (This is not part of my role) *How often do you use colleague input, feedback, or consultation when …*	
Deciding about the content or direction of a policy or program?	.92
Developing new activities, programs, or guidelines in your role or agency?	.98
Implementing activities and programs?	.98
Resolving problems in your daily professional practice?	.81
Instrumental use of feedback from students, parents, and community Response scale: Never (1), Sometimes (2), Often (3), Always (4), Not applicable (This is not part of my role) *How often do you use feedback and concerns from students, parents, or community members when …*	
Deciding about the content or direction of a policy or program?	.87
Developing new activities, programs, or guidelines in your role or agency?	.97
Implementing activities and programs?	.95
Resolving problems in your daily professional practice?	.77
Instrumental use of professional organizations and other resources Response scale: Never (1), Sometimes (2), Often (3), Always (4), Not applicable (This is not part of my role) *How often do you use professional organization or technical assistance resources when …*	
Deciding about the content or direction of a policy or program?	.87
Developing new activities, programs, or guidelines in your role or agency?	.91
Implementing activities and programs?	.95
Resolving problems in your daily professional practice?	.82

(continued)

Construct and item	Final factor loading
Constructs assessing capacity to engage with and use research evidence	
Confidence in ability to use research evidence Response scale: Not at all confident (1), Somewhat confident (2), Mostly confident (3), Very confident (4) *I feel confident that I have the knowledge and skills to …*	
Find research evidence to inform practice, policy, or program development	.91
Evaluate the quality of research	.92
Interpret the results of research	.94
Apply research evidence to practice, policy, or program development	.93
Design evaluations of practices, policies, or programs	.90
Usefulness of research evidence Response scale: Not at all agree (1), Somewhat agree (2), Mostly agree (3), Strongly agree (4) *Please rate your level of agreement with the following statements about research. Please think about research evidence broadly and not limited to what REL provides.*	
Most research evidence suggests actionable steps to take in practice.	.72
Researchers have a solid grasp on evolving problems in schools and districts.	.89
Research evidence informs the most important issues schools and districts face.	.89
Research evidence is produced quickly enough for me to make use of it.	.91
Research takes into consideration the varying levels of resources available to schools and districts to implement research findings.	.86
Conceptual use of research evidence Response scale: Never (1), Sometimes (2), Frequently (3), All the time (4) *Thinking about research evidence broadly (and not limited to what REL provides), how often do you use research evidence to …*	
Expand your understanding of an issue?	.87
Suggest alternative solutions to problems facing your school, district, institution, or state?	.91
Bring attention to an issue that your school, district, institution, or state has not yet faced?	.82
Use a research-based framework when making improvements?	.88
Provide a common language and set of ideas for you and your colleagues?	.90
Organizational support for use of research evidence Response scale: Not at all agree (1), Somewhat agree (2), Mostly agree (3), Strongly agree (4) *In my organization …*	
It is expected that research evidence will be used in policy or program development.	.80
We refer to evidence of what works when deciding which programs or initiatives to implement.	.77
Generating new research evidence to inform policy or program development is encouraged.	.89
It is expected that policies and programs will be evaluated.	.86
Interaction or collaboration with researchers or research organizations is encouraged.	.81
Organizational infrastructure for use of research evidence Response scale: Not at all agree (1), Somewhat agree (2), Mostly agree (3), Strongly agree (4) *To what extent do you agree with the following statements about the current environment within your organization?*	
We have adequate time to access and review research evidence.	.86
We have sufficient access to research evidence.	.86
There are formal processes to help staff engage critically with different information sources.	.82
Constructs assessing perceived value of REL services and products	
Value of REL engagement Response scale: Not at all agree (1), Somewhat agree (2), Mostly agree (3), Strongly agree (4) *To what extent do you agree with the following statements?*	
I would like to participate in additional support activities offered by the REL.	.78
The format of REL engagement provides ample opportunity for participants to meaningfully interact with each other.	.72
REL engagement is relevant to the issues faced by my school, district, institution, or state.	.87
The benefits from my REL engagement were worth the time and resources invested to participate.	.94
REL engagement improved my ability to support change in my school, district, institution, or state and make sure the change is sustained.	.93
REL engagement is helping my school, district, institution, or state achieve its goals.	.85
I expect to apply information from my REL engagement in my work.	.95

(continued)

Construct and item	Final factor loading
Relevance of REL evidence and services Response scale: Not at all agree (1), Somewhat agree (2), Mostly agree (3), Strongly agree (4) *Please rate the extent to which you agree with the following statements about REL research evidence and services. REL research evidence and services …*	
Are relevant to the learners and communities I serve	.74
Are easy to transfer into my own professional practice	.93
Are useful to guide my own professional practice	.93
Accessibility of REL evidence and services Response scale: Not at all agree (1), Somewhat agree (2), Mostly agree (3), Strongly agree (4) *Please rate the extent to which you agree with the following statements about REL research evidence and services. REL research evidence and services …*	
Offer timely information	.80
Are easy to understand	.91
Are easy to find	.92

REL is Regional Educational Laboratories.

Note: Factor loadings of .50 or higher indicate acceptable convergent validity, whereas loadings of .70 or higher indicate a strong association with the underlying construct.

Source: Authors' analysis of data from the online survey conducted in 2024.

Descriptive findings from the validated survey measures

REL partners reported regularly drawing on multiple information sources when making decisions and described engaging with research evidence, although with varying levels of confidence and organizational support for doing so. They also expressed that their experiences with the RELs were valuable and that REL evidence and services were relevant to and accessible in their work. Although the study sample does not represent all REL partners, responses across these constructs provide a snapshot of how REL partners use information when making decisions, their capacity to engage with and use research evidence, and their perceived value of REL services and products. The descriptive findings that follow summarize patterns in respondents' decisionmaking, research engagement, and experiences with the RELs.

Respondents indicated that their professional decisionmaking is commonly guided by varied sources of information, including their own expertise; data from their school, district, institution, or state; and input from colleagues

Respondents rated how frequently they use six sources of information when making decisions about policies, programs, new activities, or implementation, using a four-point frequency scale. They reported often relying on their own expertise (mean = 3.50) and on data from their school, district, institution, or state (mean = 3.37). They also often consider input from colleagues (mean = 3.30) and the broader community (mean = 3.01). Respondents also reported sometimes using research evidence (mean = 2.62) and information from professional organizations (mean = 2.88) as part of their decisionmaking processes. The means for each of the six instrumental use constructs, which were all measured with the same response scale, are displayed in table 6.

Respondents reported using research evidence in problem solving and indicated moderate levels of perceived usefulness and confidence in using research evidence

Using a four-point frequency scale, respondents reported frequently using research evidence in conceptual problem solving (mean = 3.02). However, using a four-point confidence scale, respondents reported feeling only somewhat to mostly confident in using research evidence (mean = 2.87), and using a four-point agreement scale, they reported that they somewhat or mostly agreed that research evidence is useful (mean = 2.62). Additional details about these constructs are in table 7.

Table 6. Means for constructs measuring instrumental use of information to inform decisionmaking about policies and programs, developing new activities, or implementing programs

Construct	Response scale	*N*	Mean	Standard deviation
Instrumental use of research evidence	(1) Never, (2) Sometimes, (3) Often, (4) Always	67	2.62	0.67
Instrumental use of data from school, district, institution, or state	(1) Never, (2) Sometimes, (3) Often, (4) Always	67	3.37	0.54
Instrumental use of own expertise	(1) Never, (2) Sometimes, (3) Often, (4) Always	67	3.50	0.57
Instrumental use of colleagues' expertise	(1) Never, (2) Sometimes, (3) Often, (4) Always	66	3.30	0.55
Instrumental use of feedback from students, parents, and community	(1) Never, (2) Sometimes, (3) Often, (4) Always	67	3.01	0.60
Instrumental use of professional organizations and other resources	(1) Never, (2) Sometimes, (3) Often, (4) Always	67	2.88	0.61

Note: Sample sizes for the constructs in this table are smaller than those in tables 7 and 8 because only respondents involved in decisionmaking about policies and programs, new activities, or implementation in their district or organization completed these items. Respondents could skip any item that did not apply to their role. To be included in the analysis, respondents had to answer at least three of the four items in a given construct.

Source: Authors' analysis of data from the online survey conducted in 2024.

Respondents noted the need for stronger organizational support and infrastructure to integrate research more effectively into decisionmaking

Using a four-point agreement scale, respondents reported moderate organizational support for research use (mean = 2.89) and moderate agreement that their organization provides the infrastructure needed to fully integrate research into decisionmaking (mean = 2.28). These findings suggest opportunities for the RELs to provide tailored support to REL partners on using research evidence and to partner organizations on how to create and support an infrastructure for research use. Additional details about these constructs are in table 7.

Respondents value their engagement with the RELs and the evidence and services provided

Using a four-point agreement scale, most respondents agreed that engagement with their REL was valuable for themselves and for their school, district, or state (mean = 3.14). They also reported positive perceptions of the relevance (mean = 3.19) and accessibility (mean = 3.12) of REL evidence and services. Together, these findings suggest that the RELs provide useful and accessible information that meets their partners' needs. Additional details about these constructs are in table 8.

Table 7. Means for constructs measuring capacity to engage with and use research evidence

Construct	Response scale	*N*	Mean	Standard deviation
Confidence in ability to use research evidence	(1) Not at all confident, (2) Somewhat confident, (3) Mostly confident, (4) Very confident	100	2.87	0.71
Usefulness of research evidence	(1) Not at all agree, (2) Somewhat agree, (3) Mostly agree, (4) Strongly agree	99	2.62	0.62
Conceptual use of research evidence	(1) Never, (2) Sometimes, (3) Frequently, (4) All the time	100	3.02	0.65
Organizational support for use of research evidence	(1) Not at all agree, (2) Somewhat agree, (3) Mostly agree, (4) Strongly agree	100	2.89	0.70
Organizational infrastructure for use of research evidence	(1) Not at all agree, (2) Somewhat agree, (3) Mostly agree, (4) Strongly agree	100	2.28	0.74

Source: Authors' analysis of data from the online survey conducted in 2024.

Table 8. Means for constructs measuring perceptions of Regional Educational Laboratory (REL) services and products

Construct	Response scale	*N*	Mean	Standard deviation
Value of REL engagement	(1) Not at all agree, (2) Somewhat agree, (3) Mostly agree, (4) Strongly agree	99	3.14	0.66
Relevance of REL evidence and services	(1) Not at all agree, (2) Somewhat agree, (3) Mostly agree, (4) Strongly agree	100	3.19	0.63
Accessibility of REL evidence and services	(1) Not at all agree, (2) Somewhat agree, (3) Mostly agree, (4) Strongly agree	100	3.12	0.67

Source: Authors' analysis of data from the online survey conducted in 2024.

Implications

The study results provide initial support for the reliability and validity of the survey constructs capturing REL partners' use of information from a variety of sources, including research evidence, to make decisions; perceptions of their capacity to engage with and use research evidence; and perceived value of existing REL services and products. Findings provide evidence that supports the RELs' use of these measures to better understand how their partners perceive and use research evidence and to tailor professional learning, tools, and supports accordingly. For example, if survey responses indicate that partners do not feel confident using research evidence to inform program development, RELs could provide targeted training on how to interpret and apply research evidence in their work. Such efforts may strengthen partners' capacity to use research evidence effectively and promote more informed decisionmaking.

These constructs may also be useful to other educators seeking to understand use of research evidence in their own contexts. Users may select and administer the individual constructs that are most relevant to their needs, rather than using the full set of measures. For instance, school districts could use the instrumental use constructs to understand how school leaders and teachers use research evidence and data when making decisions about content or developing new activities. Similarly, school districts could use the confidence in ability to use research evidence constructs to identify gaps in staff confidence and determine where additional professional development may be helpful. School districts or school leaders may also use the constructs focused on organizational support and infrastructure for research use to reflect on their current organization's practices and consider what barriers may influence use of research evidence and whether any resources might be helpful to encourage use of research evidence.

There are several limitations to this study. The study used a purposive sample, so participants do not fully represent all individuals who engage with the REL Program. Although the survey response rate was relatively strong and respondents reflected a typical mix of REL partners—including teachers, administrators, and staff at the district, state, and postsecondary levels—some sampled individuals did not participate. Nonresponse was especially prevalent among district and school staff, who were difficult to reach because of competing responsibilities. As a result, the findings should be interpreted as descriptive of the respondents in this study and should not be generalized to any specific REL or to the overall REL Program. Additional validity testing would be beneficial to further examine whether the updated items and constructs are supported in a different sample of REL partners.

In addition, it is important to note that these items were not designed or intended to function as longitudinal indicators of growth. For example, in the REL context, the individuals who work with the RELs and complete the survey may differ at each time point, and responses are anonymous. As a result, changes over time can be assessed only by comparing group-level scores rather than by tracking individuals, and it should not be assumed that the same respondents are represented at each time point. In addition, the nature of the technical assistance provided by the RELs can shift over time based on partner needs, project priorities, and staffing. Consequently, observed changes in scores may reflect variation in respondents or in ongoing work rather than true changes in the value of

REL support or partners' use of research evidence. Moreover, as partners deepen their engagement with research evidence, they may become more aware of their own gaps or limitations. This increased awareness could lead to declines in certain measures even as actual capacity improves. Given these considerations, these items are best suited for identifying patterns, areas of strength, and areas where additional support may be needed.

References

Arifin, W. N. (2017). *Sample size calculator* (Version 2.0) [Spreadsheet file]. https://wnarifin.github.io/ssc.html

Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics, 27*(4), 335–340. https://doi.org/10.3102/10769986027004335

Brennan, S. E., McKenzie, J. E., Turner, T., Redman, S., Makkar, S., Williamson, A., Haynes, A., & Green, S. E. (2017). Development and validation of SEER (Seeking, Engaging with and Evaluating Research): A measure of policymakers' capacity to engage with and use research. *Health Research and Policy Systems, 15,* Article 1. https://doi.org/10.1186/s12961-016-0162-8

Farley-Ripple, E. N., Oliver, K., & Boaz, A. (2020). Mapping the community: Use of research evidence in policy and practice. *Humanities and Social Sciences Communications, 7,* Article 83. https://doi.org/10.1057/s41599-020-00571-2

Gitomer, D. H., & Crouse, K. (2019). Studying the use of research evidence: A review of methods. William T. Grant Foundation. https://wtgrantfoundation.org/wp-content/uploads/2019/05/A-Review-of-Methods-FINAL003.pdf

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Lee, S. (2015). Implementing a simulation study using multiple software packages for structural equation modeling. *SAGE Open*, *5*(3). https://doi.org/10.1177/2158244015591823

Lysenko, L. V., Abrami, P. C., Bernard, R. M., Dagenais, C., & Janosz, M. (2014). Educational research in educational practice: Predictors of use. *Canadian Journal of Education/Revue Canadienne De l'éducation*, *37*(2), 1–26. https://journals.sfu.ca/cje/index.php/cje-rce/article/view/1477

May, H., Blackman, H., Van Horne, S., Tilley, K., Farley-Ripple, E. N., Shewchuk, S., Agboh, D., & Micklos, D. A. (2022). *Survey of Evidence in Education for Schools (SEE-S) technical report*. Center for Research Use in Education and Center for Research in Education and Social Policy, University of Delaware. https://eric.ed.gov/?id=ED628010

Meng, X.-L. (1994). Posterior predictive pvalues. *The Annals of Statistics, 22*(3), 1142–1160. https://doi.org/10.1214/aos/1176325622

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49*(1), 115–132. https://doi.org/10.1007/BF02294210

Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., Farrell, C., & Allen, A. R. (2017). How school and district leaders access, perceive, and use research. *AERA Open, 3*(2), 1–17. https://eric.ed.gov/?id=EJ1194150

Rickinson, M., Cirkony, C., Walsh, L., Gleeson, J., Cutler, B., & Salisbury, M. (2022). A framework for understanding the quality of evidence use in education. *Educational Research, 64*(2), 133–158. https://doi.org/10.1080/00131881.2022.2054452

Rickinson, M., Gleeson, J., Walsh, L., Cutler, B., Cirkony, C., & Salisbury, M. (2021). *Research and evidence use in Australian schools: Survey, analysis and key findings*. Q Project, Monash University. https://bridges.monash.edu/articles/report/Research_and_Evidence_Use_in_Australian_Schools_Q_Project_Survey_analysis_and_key_findings/14445663

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). Routledge.