Survey of Evidence in Education for Schools (SEE-S) Technical Report

Henry May
Horatio Blackman
Sam Van Horne
Katherine Tilley
Elizabeth N. Farley-Ripple
Samantha Shewchuk
Darren Agboh
Deborah Amsden Micklos

October 2022

Center For Research Use In Education

University of Delaware
University of Minnesota
University of Pennsylvania

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C150017 to the University of Delaware. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Please cite this report as follows:

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.* The Center for Research Use in Education (CRUE) & the Center for Research in Education and Social Policy (CRESP), University of Delaware. Available from http://crue.cehd.udel.edu/

Acknowledgements

The authors would like to thank those who have contributed to this study.

This study would not have happened if it were not for the willingness of the educators and researchers who participated in the exploratory and cognitive interviews, the administrators of 64 schools who agreed to offer the pilot surveys to the staff and the 1,573 educators of the schools who participated in the pilot survey, the district administrators of 21 districts and 20 charter schools who agreed to offer schools the opportunity to participate in the field trial, the 134 school administrators within the districts and the 4,415 educators who participated in the field trial. Dr. James O'Toole was instrumental in recruiting schools to participate in the pilot trials and the districts and charter schools to participate in the field trial.

Dr. Lynn Okagaki of the University of Delaware, Dr. Rebecca Maynard of the University of Pennsylvania, Dr. Karen Seashore Louis of the University of Minnesota, and Dr. Ronald Gallimore have advised this project throughout. We greatly appreciated the guidance provided by the Advisory Committee composed of Atnre Alleyne, Dr. Robert Granger, Ben Herold, Dr. Karen Kolsky, Dr. Jonathan Supovitz, Dr. Lisa Thomas, Dr. Elliot Weinbaum, and Dr. Russell Whitehurst.

Dr. Allison Karpyn, Layne Humphrey, Kalyn McDonough, and Dreisen Heath provided support in the beginning phases and assisted with developing the conceptual framework and pilot surveys.

Dr. Akisha Jones Sarfo and Dr. Sara Grajeda assisted with data analysis of the pilot survey with educators and revisions to the *Survey of Evidence in Education for Schools.*

Graduate students who have supported the activities of this project are Rui Wang, Joseph Tise, Scott Sheridan, Liu Liu, Cara Cuccuini-Harmon, and Hilary Mead.

Undergraduate students who have assisted us in the project activities are Samantha Lester, Rosymar Magana, Nikita Patlolla, Nicole Mejia, Connor Morgan, Darnelle Evans, William Cavin, Mia Mazal, Mia Siebold, and Angelina Calandruccio

Lastly, we extend our sincere thanks to Drew Gitomer and Caitlin Farrell for their expert critique and helpful feedback as external peer reviewers of this report.

Table of Contents

Executive Summary

Funded by the Institute of Education Science (IES) at the U.S. Department of Education, the Center for Research Use in Education (CRUE) is tasked with expanding the study of knowledge utilization in the K-12 system. In this technical report, CRUE presents the methodological design of a large-scale quantitative investigation of research use by school-based practitioners through the *Survey of Evidence in Education for Schools (SEE-S)*. This report also provides basic descriptive results from the *SEE-S* Field Trial. Through the development and validation of multiple survey measures, this study aims to deepen the fields' understanding of the actions and activities that educators are involved in concerning the use of evidence in decision-making.

Survey Development and Administration

Development of the *SEE-S* began with interviews conducted with a small set of researchers, research brokers, and school-based practitioners including principals and classroom teachers. CRUE researchers used information gathered from the interviews to generate synthesis documents regarding key themes aligned to the study's framework. The documents were then used to draft an initial set of items. The items underwent refinement through multiple rounds of cognitive interviews with school-based practitioners, brokers, and researchers. Items were evaluated for construct and content validity. Two pilot tests of the *SEE-S* were also conducted. This process enabled the researchers to make improvements to the survey through item analysis and initial assessment of scale reliability. Sixty-four schools and 999 school-based professionals completed the survey during these two pilot studies (response rate = 34%).

A Field Trial of the *SEE-S* was conducted during the 2018-2019 and 2019-2020 school years. A total of 154 schools and 4,415 school-based practitioners participated in the survey (overall individual-level response rate = 51%). Hereafter, school-based practitioners will be referred to as practitioners. Results from the field trial are presented in this report.

Depth of Use Scaling and Reliability

A major component of the survey involves a battery of items intended to measure schools' *Depth of Use* (*DOU*) of research evidence. Motivated by Coburn's (2003) "depth of reform" which describes efforts to move "beyond surface structures and procedures" in reform implementation, and earlier conceptual work (Farley-Ripple, 2008a; Farley-Ripple, 2008b; Farley-Ripple & Cho, 2014) to frame "*Depth of Use*," which refers to the complex ways in which evidence use is meaningful, systematic, and likely to generate improvements in policy and practice (Farley-Ripple et al., 2018). The *DOU* measure is grounded in a school decision-making process identified by the respondent at the start of the survey, and therefore focuses on instrumental uses of research. By emphasizing *Depth of Use* as related primarily to instrumental use, we are able to attend to the specific role research evidence plays in decision-making, the types of decisions or problems research is most likely to inform, and the types of research decision-makers utilize in those processes.

Our *DOU* measures include six dimensions of practice that previous literature on organizational and evidence-based decision-making have suggested are important for generating meaningful systematic use; these are labeled as: (1) *evidence*, (2) *search*, (3) *interpretation*, (4) *participation*, (5) *frequency*, and (6) *stage of decision-making*. Multilevel models were used to produce aggregate school-level scores on each *Depth of Use* metric, as well as an overall *DOU* score with a possible range of 0-100 points. Cronbach's alpha for the *DOU* Total Score at the school level was .76, suggesting that the *DOU* Total Score has satisfactory reliability at the school level. The distribution of *DOU* Total Scores in our sample suggests that most schools score closer to the low end of the *research use scale*, with a mean of 15.6 points and a standard deviation of 5.9 points; however, there is considerable skew in the distribution, suggesting a small but significant number of schools that stand out as having exceptionally high scores on the *DOU* scale.

Perceptions and Assumptions Scaling and Reliability

In addition to the *DOU* metrics, several scales were produced to measure practitioners' perceptions and assumptions about research. Survey results indicate that a number of our scales met or exceeded standards for internal consistency (Lance et al., 2006). Appendix A provides the reliability coefficients (Cronbach's alpha) for each scale intended to represent a latent trait, including the fit statistics. Also included are the number of survey respondents, and number of items comprising each scale. Details on the concepts and components of each scale are included in the main sections of this technical report.

Introduction & Background

The Center for Research Use in Education (CRUE) seeks to expand the education community's understanding of how to improve the relationship between research and practice in the K-12 educational space. In addition, the center aims to describe and disseminate information on practices that improve knowledge mobilization; from the production of research to the uptake and application of strategies derived from research. This report documents the major technical aspects of the development of the *Survey of Evidence in Education for Schools* (*SEE-S*), including item development, sample selection, and reliability and validity assessment. Descriptive statistics for data collected during the survey field trial are also detailed in this report.

The Problem of Research Utilization

Foundational work in the field of evidence use focused on the under-utilization of social science research in social policy. Research explored barriers to use in policymaking and local decision-making processes, and sought to understand the nature and causes of weak ties between researchers and practitioners (Backer, 1993; Broekkamp & Hout-Walters, 2007; Davies & Nutley, 2008; Landry, 2001). In response to this work, throughout the late 1900s, federal and state governments sponsored initiatives in an attempt to bridge the gap between research-based knowledge and school practice. Among many other targeted initiatives was the development of the Education Resources Information Center (ERIC) and increases in funding for research focused on exploring and dismantling barriers to knowledge mobilization. Following this, the *No Child Left Behind Act* of 2001 and the *Education Sciences Reform Act* of 2002 (ESRA), established explicit expectations for the role of research in informing decisions about education programs, policies, and practice. From ESRA, the Institute of Education Sciences (IES) was created, pushing for more rigorous quantitative studies in education. IES established the What Works Clearinghouse (WWC), which reviews, critiques, and synthesizes evidence of impacts of education interventions. The WWC now includes hundreds of Intervention Reports and Practice Guides based on reviews of more than 11,000 studies.

Federal legislation has continued to include language intended to increase research use and implementation of evidence-based practices and programs. NCLB, reinforced by the *Every Student Succeeds Act* of 2015, called for significant changes in school and district policies and practices, establishing mandates for the use of evidence in informing decisions. Policy efforts to mandate research use are based on the premise that research can and should be used to support practice. In turn, those evidence-based decisions should lead to improved practices among school/district staff and educational outcomes for students; ultimately advancing our communities and the nation. Although this logic is clear, it is based on problematic assumptions about the nature of both research and decision-making. To date, we still have limited information about the degree to which educators utilize research evidence in their decision-making processes, especially relative to other forms of evidence.

The problem of knowledge utilization cannot be conceptualized as simply a problem of production and dissemination, nor a problem of merely increasing practitioner uptake. Rather, Lavis et al (2003) describes the challenge as developing a "decision-relevant culture" among researchers and a "research-attuned culture" among decision-makers. That is, the problem of knowledge utilization is dualistic in nature and must be addressed from two perspectives: that of researchers and that of practitioners. To realize the potential for education research to improve teaching and learning, we need a better understanding of the activities that constitute research use as well as the factors and conditions that influence the practices of and connections between researchers and practitioners. This report details the efforts of the Center for Research Use in Education to increase our understanding of how school-based educators access and use evidence to inform their practice, and the contextual influences on those practices.

The Center for Research Use in Education

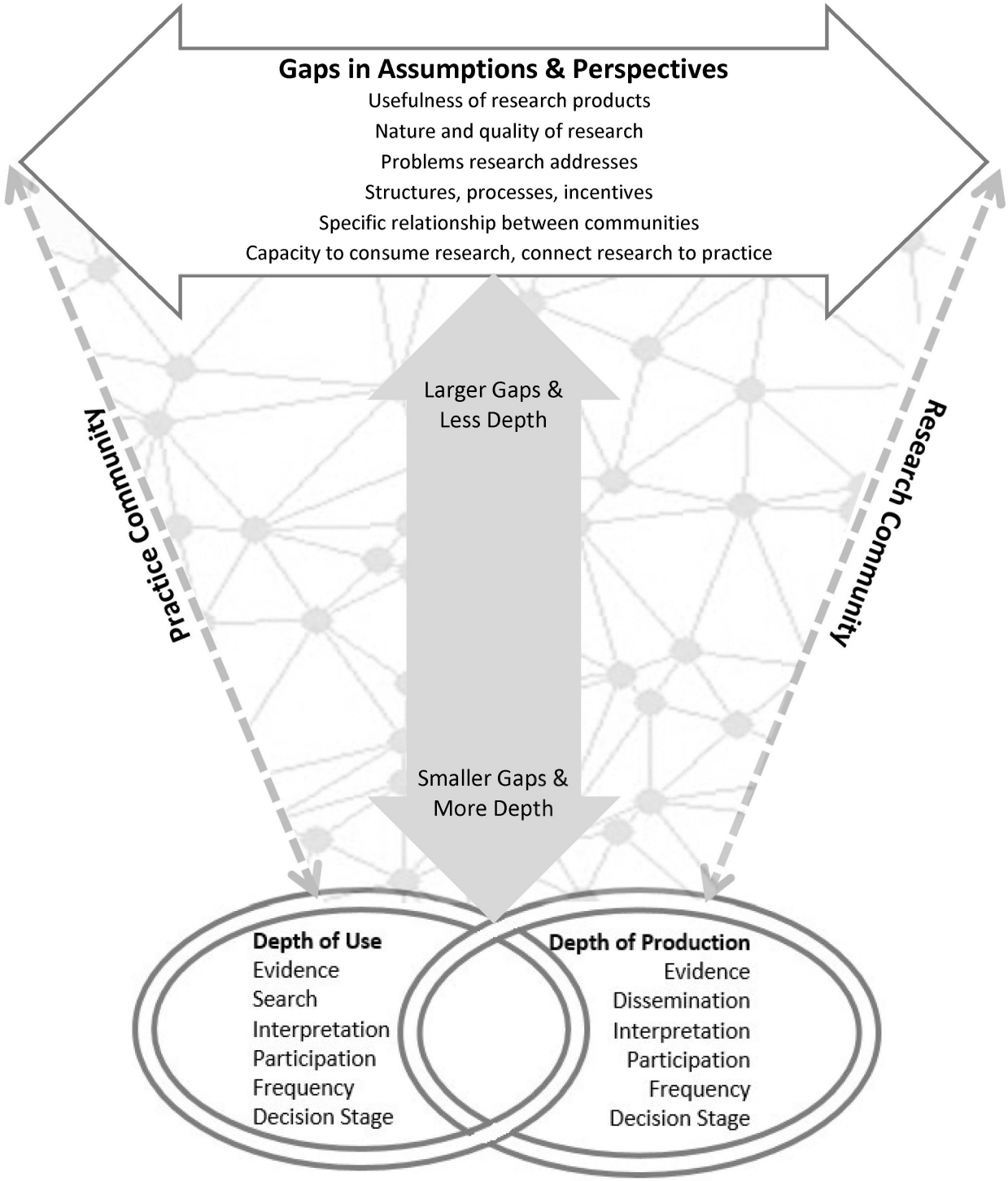
The ultimate goal of the Center for Research Use in Education (CRUE) is to expand the study of knowledge utilization and produce a more holistic picture of what drives research use, from the production of knowledge by researchers, to the sharing, or brokering of knowledge between

the two communities of research and practice, to the application of research knowledge in local decision-making processes.

Conceptual Framework

Figure 1 presents a detailed visualization of our conceptual framework. At the bottom of the figure are interrelated processes associated with use of research in decision-making and production of research. Each process features a parallel set of dimensions, and we describe variability in those processes in terms of depth, a concept we explore in our work. At the top of the image are five key assumptions and perspectives of the research and practice communities represented by the horizontal arrow. This arrow represents the size and scope of potential gaps in those assumptions and perspectives between the research and practice communities, larger at the top and becoming smaller towards the bottom. We hypothesize that where gaps between those communities are largest, we will see the least research use, or research use that lacks "depth" as indicated by the widening of the arrows as we move up. We believe these gaps to be driven by characteristics of individuals and their organizations. More information on the framework can be found in the *Educational Researcher* article Rethinking connections between research and practice in education: A conceptual framework (Farley-Ripple et al., 2018) and in the subsequent section of this report entitled Measurement Blueprint Development & Revision.

Figure 1. ***Conceptual Framework***



Farley-Ripple, E., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018). Rethinking connections between research and practice in education: A conceptual framework. *Educational Researcher, 47*(4), 235-245.

The Work of the Center

The Center has engaged in a series of three research studies that will: (a) investigate the nature and depth of research use in schools, (b) identify the factors in both the research and school communities and the relationships between them that hinder or facilitate research use, and (c) develop strategies to make more meaningful and impactful connections between research-based evidence and classroom practice.

Measurement Study

We designed, produced, piloted, and validated two parallel surveys for researchers and practitioners. Both surveys contain sets of items designed to assess three dimensions: (a) the different aspects of production, dissemination, and use of education research, (b) the gaps and assumptions that exist between the practitioner and research communities which support or hinder connections, and (c) the direct and indirect connections between research and practice communities. Additionally, the practitioner version of the survey includes a fourth dimension, which focuses on measuring the confidence that individual educators have in critically interpreting research.

Descriptive Studies

For our Descriptive Studies, we use our parallel surveys to test the hypotheses implied in our conceptual framework. Specifically, we aim to quantify and compare researchers' and practitioners' perspectives on research evidence and its utility.

Case Studies

In-depth case studies are being conducted to supplement the descriptive studies so that we can deepen our understanding of the factors that support connections between research and practice, as well as the use of research evidence in school decision-making.

Measurement Study: Development & Validation of the *SEE-S*

This report documents the school-based practitioner (heretofore, practitioner) side of the "Measurement Study," which is the foundational part of this project that produced the surveys used to collect the quantitative data used in the "Descriptive Studies," which are documented in our other Descriptive Reports, research presentations, and journal articles. We define "school-based practitioners" as teachers, principals, assistant principals, interventionists, and other instructional staff from schools and districts across the United States. Specifically, this Technical Report describes the process used to develop the *Survey of Evidence in Education for Schools (SEE-S*) and the primary results from analyses of data collected using the *SEE-S* during piloting and field trial phases of the study. It also includes information and requirements for those seeking to use the *SEE-S* (or parts thereof) in other projects and initiatives.

Survey Development

Development of the *SEE-S* began Fall 2015 and concluded in the summer of 2018. Our approach to instrument development followed a multi-phase approach, utilizing both qualitative and quantitative methods to produce reliable and valid survey measures. Our development process included drafting blueprints of various instruments utilizing our framework, exploratory interviews, and multiple rounds of cognitive interviews and pilot testing. Preliminary psychometric evaluation of items intended to represent specific constructs were also conducted after each stage of piloting.

Measurement Blueprint Development & Revision

Development of the *SEE-S* began with the creation and revision of a measurement blueprint based on our conceptual framework. The blueprint named and defined each construct, and it provided a target number of survey items for each. After drafting the blueprint, an advisory panel of experts from both the researcher and practitioner communities reviewed it. The purpose of this revision process was to ensure all relevant constructs were identified and defined. The team relied on these definitions to maintain shared understandings and to guide

future stages of development. The finalized constructs are presented in the following sub-sections.

Depth of Use

Our construct of "*Depth of Use*" describes the complex ways in which evidence use is meaningful, systematic, and likely to generate improvements in policy and practice. We focus primarily on instrumental uses of research, or situations in which research is specifically used to inform decision-making (Caplan, 1979). Instrumental research use is expected in current educational practice, particularly under the evidence-use expectations of federal accountability policy (e.g., NCLB, ESSA), but is found to be lacking. Other forms of research use, such as conceptual, political, and symbolic are less central in our work though important forms of research use in education.

Our conception of *Depth of Use* acknowledges the complexity and multidimensionality of evidence use as an organizational practice. Attending to this complexity, we focus on *evidence*, *search*, *interpretation*, *participation*, *frequency*, and the *stage of decision-making* where evidence is relied upon. Each dimension is understood as individual continua, rather than occurring or not occurring. As a larger construct, *Depth of Use* describes the degree to which research meaningfully and systematically informs decisions about education practice.

Evidence

Measuring the use of different types of evidence in school-based decision-making processes is a primary focus of this work. Decision-makers utilize a variety of evidence sources during the decision-making process (Coburn et al., 2007; Corcoran et al., 2003; Farley-Ripple, 2012; Penuel et al., 2016; Supovitz & Klein, 2003). Although different kinds of information may inform decisions, we are specifically interested in the use of evidence generated from research not conducted by the school or district (i.e., external research) compared with other forms of research (e.g., locally driven) and knowledge that influences decision-making. At CRUE, we do not consider one form of research to be better or more important than other forms of evidence. Rather, we are interested in understanding the relative use of external research compared to

other forms (i.e., locally generated) of evidence in school-based decision-making processes. The *evidence* dimension of *Depth of Use* ranges from no engagement with external research on one end of the spectrum to substantial inclusion of external research on the other, with the high end of this spectrum reflecting organization-wide engagement with external research (i.e., by many school staff and across many decisions).

Search

This dimension of *Depth of Use* focuses on the nature and extent of practitioners' search for evidence. The low end of this dimension is characteristized by a very limited search for research-based evidence, and a focus on local sources of evidence and sources of evidence which are already familiar to the decision-maker (David, 1981; Finnigan et al., 2012; Honig & Coburn, 2008; Williams & Cole, 2007). At the high-end, decision-makers utilize direct sources of research for an organizational decision.

Interpretation

For evidence to impact decision-making, it must be interpreted and transformed into usable knowledge (Bertrand & Marsh, 2015; Coburn et al., 2009; Davies & Nutley, 2008). This sensemaking process includes critical evaluation of information in order to determine its quality or rigor, applicability to one's context, and relevance to the problem (Davies, 1999; Williams & Cole, 2007). *Interpretation* as a construct is conceptualized as the extent to which actors assess critical aspects of research evidence, with higher scores for more activities related to assessing the relevance and quality of the information.

Participation

Participation relates to *who* participates in decision-making, and also the ways in which they are involved (i.e., collecting information, evaluating information, making the decision). Understanding who participates in research use during decision-making is important because individuals' positionality within the staffing hierarchy and their predispositions, knowledge, and goals influence what and how evidence is interpreted by group members (Coburn, 2001a;

Coburn et al., 2009; Finnigan & Daly, 2014; Kennedy, 1982; Weick, 1995; Weiss 1995). To capture variability in *participation*, the construct spans from no one or few individuals on the low end to collaborative teams representing high levels of participation among staff in decision-making.

Frequency

In describing the depth of evidence use, *frequency* is an indicator of the extent to which research informs decisions. *Frequency* also captures the extent to which evidence use practices are part of regular practice and whether evidence use is sustained. Studies typically document how often decision-makers use research as a way of assessing the centrality of research evidence in decision-making—Does it play a role sporadically, or is it institutionalized in decision-making practices? Although no research exists in this specific domain, the regularity with which research evidence is brought to bear on decisions may be an indicator of greater or lesser systematic use.

Stage of decision-making

This dimension of *Depth of Use* describes the stages of decision-making where research may play a role. These stages of decision making are:

1. Identifying a problem in a school or district,
2. Determining reasons behind a particular problem,
3. Identifying multiple potential strategies for addressing the problem,
4. Selecting which strategies to implement,
5. Informing how best to implement the decision, and
6. Adjusting the solution to improve implementation or outcomes.

Theories on research use suggest that different types of evidence may be preferred at different points (Bass, 1983; Farley-Ripple, 2008b), such as problem framing or identification of a potential solution (Coburn et al., 2009). Here, we measure practitioners' use of research evidence at different stages of decision-making.

Depth of Use *as an Overall Construct*

As described above, each dimension of depth is supported in the literature as a substantive component of research use in practice. However, examining any one dimension provides a partial picture of schools' engagement with research but risks reducing evidence use to an administrative task rather than multiple activities constituting a complex practice. We therefore conceptualize these dimensions as fitting together to help us describe research use as a multi-faceted practice in terms of "depth". Depth characterizes schools' practice across dimensions, with deeper engagement across dimensions promoting more meaningful and systematic use of research, and with lesser engagement suggesting more emergent research use practice. *Depth of Use* as a construct also helps us to differentiate among different schools' practices and to understand the relationships among different dimensions of use. Lastly, as an organizational measure, and one which spans multiple decisions as reported by different people within the organization, we hypothesize that the individual dimensions of *Depth of Use* will form a coherent and reliable overall scale that can be used to differentiate schools.

Depth of Production

Our conceptual framework also implies a parallel set of dimensions for the production and dissemination of research. These dimensions are defined and discussed in a separate report on the Survey of Evidence in Education for Researchers (SEE-R). Notably, the SEE-R is not expected to produce a single overarching scale for depth of production, as the focus of that survey is on a single research study (i.e., characteristics of studies may be more like a checklist, and not exhibit correlations across dimensions), and responses come from only one person (i.e., it is not an organizational measure).

Assumptions & Perspectives about Research

The horizontal arrow in Figure 1 represents our approach to understanding and describing the factors that influence the use of research in school-based decisions. Dunn's (1980) five categories of culture grounds our work. Here, we look to describe the differences in assumptions and perspectives between the research and practice community across five

dimensions: (1) the usefulness of research products, (2) the nature and quality of research, (3) problems that research addresses, (4) the structures, processes, and incentives surrounding research production and use, and (5) the relationships between communities.

Usefulness of research products

One potential gap between research and practice relates to the usefulness of research products. Prior research suggests that the characteristics and type of research products influence their use in schools (Corcoran et al., 2001; West & Rhoton, 1994). The usefulness dimension of the gap represents the degree to which the characteristics valued by practitioners are incorporated in research products.

Nature and quality of research

Findings from prior work suggest that practitioners may value characteristics of research that differ from the primary concerns of researchers. For example, while the research community often places greater emphasis on internal validity for causal inference, practitioners often prefer evidence based on work from organizations similar to their own; suggesting greater weight on external validity (i.e., other schools and districts; Corcoran et al., 2001; Finnigan et al., 2013; Supovitz & Klein, 2003). This dimension explores features of research valued by practitioners.

Problems that research addresses

The extent to which the evidence produced by the research community is considered timely and relevant to the problems confronting practitioners is an indicator of this dimension of the gap. Published research directly investigating the types of problems faced by school-based practitioners is extremely limited (Leithwood & Steinbach, 1995; Neeleman, 2019). Additionally, no standardized mechanism exists for practitioners to communicate their needs to the research community in a systematic manner (National Research Council, 2012). Even when research is conducted on relevant problems, the research and practice communities operate on very different timelines, where the pace of research is often slow compared to the fast-paced nature of working in schools (Penuel et al., 2016).

Structures, processes, and incentives

Theories around the causes of the research-practice gap also point to structures, processes, and between communities incentives that influence the use of research (Supovitz & Klein, 2003). Practitioners do not operate in isolation. Rather, their actions are influenced by the contexts in which they work including organizational structure and culture (Coburn & Talbert, 2006; Corcoran et al., 2001; Finnigan et al., 2013; Honig, 2003; Massel et al., 2012; Spillane, 1998; Weiss, 1995; West & Rhoton, 1994). Therefore, participants' opportunity to use research evidence may be related to the presence of and use of particular structures, procedures, routines, and supports.

Relationships and Connections Between Communities

Depicted as network ties in our framework (see Figure 1), research use may also be considered a function of the relationship between education decision-making by practitioners and the production of research (Coburn & Stein, 2010; Honig & Venkateswaran, 2012; Huberman, 1990; Landry et al., 2001). Relationships are conceptualized as user-pulled (e.g., active search by users), producer-pushed (e.g., dissemination), or exchanges (e.g., interaction during key decision-making or research processes). The nature and extent of connections and interactions between individuals across the two communities is an indicator of the relationship dimension—specifically whether practitioners know how to connect with researchers, and desire those connections, or whether researchers know how to connect with practitioners, and desire those connections. Additionally, both researchers and practitioners may have direct connections and interact with not just individuals in the other community, but organizations and media outlets that support and encourage connections between the research and practice communities.

Confidence to Critically Evaluate Research Evidence

Our framework also centers on educators' capacity to critically evaluate research. Prior research suggests practitioners may lack confidence in their research use abilities (Hill & Briggs, 2020; Williams & Coles, 2007) and may lack capacity to critically interpret research (Coburn & Talbert,

2006; Supovitz & Klein, 2003). Educators' training and experiences related to research and data analysis may impact their capacity to critically evaluate research evidence (Supovitz & Klein, 2003).

Research Brokerage

We also recognize the importance of research brokerage in facilitating connections between research and practice. Research brokers are seen as "linking mechanisms," and include a variety of actors such as funding organizations, advocacy groups, professional associations, and individuals who operate in brokerage roles (Malin et al., 2018; Neal et al., 2015; Scott et al., 2014). This dimension of the framework includes a focus on individual educators' brokerage practices, including understanding mechanisms for accessing research as well as whether and how they share research within their schools.

Exploratory Interviews

The instrument development process included 44 exploratory interviews with members of the communities of interest (14 practitioners, 14 researchers, and 16 brokers). At the time of the exploratory interviews, practitioners who participated were actively employed at a public K-12 institution. Researchers who were interviewed were employed at a research university. Brokers were employed at organizations, outside of universities or public K-12 institutions, focused on providing services for schools and educators (e.g., Rodel Foundation, Friday Institute, Search Institute). The study's principal investigators used the conceptual framework to inform the creation of a series of semi-structured interview protocols for each community of interest. Interview protocols were piloted with another researcher prior to their use. Interviews were approximately one hour in length and were conducted over the phone. Audio was recorded to allow for transcription. The exploratory interviews were intended to gather information about the relevance of the framework, and also to identify aspects that were missing. The purpose of these interviews was to gather information that led to the development of synthesis documents that would serve as the basis for item development.

Practitioner Interviews

The practitioner interviews covered three primary domains: 1) decisions and problems, 2) resources used to address problems, and 3) role of research in decision processes. Practitioners were first asked to describe some of the problems their school was working to address. Interviewers also asked about the kinds of decisions made to address those problems. Following this, interviewees were asked to describe the decision-making process; who was involved, what specific process was followed, and their personal involvement in those processes and decisions.

Interviewees were also asked about their understanding and use of research; both in their own practice and related to the problems and decisions they referenced earlier. Last, practitioners were asked about their experiences searching for research and their connections with a researcher or research organization.

Researcher Interviews

Interviews of researchers focused on 1) research problems and 2) dissemination and connection to practice. Researchers were first asked about their current work, including issues they are engaged in researching, the factors that shaped their choices about which work to pursue, and goals for their work. They were then asked about their dissemination practices. Interviewees also responded to questions about how their work impacts practice, and aspirations for further impact. Last, they were asked about their relationship to the practitioner community and their ideal views of the nature of that relationship.

Broker Interviews

The broker interviews focused on understanding the role the broker and their organization plays in connecting research and practice. The interview protocol focused on 1) organization mission and problem focus, 2) conceptions and use of research, and 3) efforts to connect research and practice. Brokers were first asked to describe the focus of their work and the strategies employed to address specific problems of practice. They were then asked to describe their

general beliefs about research and how research should connect to practice in an ideal situation.

Interview Analysis

Interview transcripts were analyzed using the qualitative analysis software Dedoose Version 7.0.23 (SocioCultural Research Consultants, 2016). The survey blueprints served as the basis for the development of an a priori coding framework. Transcripts were coded iteratively. Summaries of information relevant to the development of the survey were created. These summaries then led to a report for each construct that described 1) the important issues derived from the data, 2) what was necessary to capture on the survey, and 3) the vocabulary to use when surveying practitioners (and researchers).

Survey Item Development & Revision

Item development began with the creation of a blueprint for each survey, derived from the conceptual framework guiding this study. Survey topics (i.e., introduced as constructs in the previous sections) were defined based on a review of the literature and project aims. The construct blueprint and the feedback generated from the exploratory interviews were then used to create the first draft of the *SEE-S*.

Cognitive Interviews

Following the initial development of survey items, cognitive interviews were conducted with educators to further refine the survey. The cognitive interviews were carried out as follows. First, draft *SEE-S* survey items were created in Qualtrics, an online survey platform. Convenience sampling was used to recruit practitioners (n=35). Interviewees completed the survey while discussing their reactions to and interpretations of the survey items with an interviewer. The purpose of cognitive interviews is to aid in the design of survey questions that accurately reflect the intent of the question (Campanelli, 1997; Willis & Artino, 2013).

During each interview, the interviewer instructed the interviewee to think aloud as they moved through the survey, sharing their understanding of the survey items and responses to them. The

interviewers also had a short protocol with specific questions to ask the participant. The survey was divided into four sections based on topic (e.g., *Depth of Use*). Three rounds of cognitive interviews were conducted for each topic area, with five interviews per round. Transcribed data was synthesized by members of the research team, after which, revisions to the survey were made. Subsequent rounds of interviewing and refinement were carried out until the research team deemed that the feedback no longer suggested substantial changes to the instrument.

Pilot Testing and Psychometric Evaluation

The first pilot of the *Survey of Evidence in Education for Schools* was administered in the Winter and Spring of 2017. The second pilot was administered late Spring and early Fall 2017. Potential participating schools were contacted either through the district office or directly via conversations with the center's recruitment coordinator. Districts/schools were selected by the center's recruitment coordinator based on criteria set by the research team for locale (urban, suburban, rural) and district size. The goal for the second pilot was to have a sample from 30 schools representing a diverse set of localities.

Thirty-three schools participated in the first pilot and 31 in the second, for a total of 64 schools. The schools were located in Connecticut, Delaware, and Pennsylvania. Participants from each school were given three months to complete the survey which was administered online via the Qualtrics online survey platform. Nine-hundred ninety-nine out of 2,921 participants completed the full battery of surveys (response rate=34%). An additional 574 participants completed about half of the survey. There was no minimum school-level response rate.

After each pilot round, items intended to measure latent constructs were preliminarily assessed for internal consistency using reliability analysis and exploratory factor analysis.

Pilot Sample Demographics

Teachers were the primary staff role surveyed for the *SEE-S* pilot studies, although all instructional staff and school administrators were asked to complete the pilot survey. Additionally, at least one district administrator agreed to complete the survey at the time our

sample recruiter confirmed a school's participation in the survey. Nearly 74% of the sample was teachers, 7% instructional coaches, and 4% school administrators. Table 1 details the survey sample. We expected special education teachers to be common in the sample, and we wanted to be able to distinguish them from classroom teachers and coaches or specialists, which is why the special education teacher role was listed as a separate response option for this item.

Table 1. ***Pilot Survey Participants by Role***

Staff Role	Frequency	Percent
Teacher	945	60.0%
Special Education Teacher	216	13.8%
Instructional Coach or Specialist	113	7.2%
School Administrator	60	3.8%
District Administrator	20	1.3%
Other School Instructional Staff	206	13.1%
Other District Staff	13	0.8%

N=1,573

The majority of schools included in the two pilot rounds were elementary schools (69%). Several middle and high schools were also surveyed. Table 2 describes the distribution of schools by grade span.

Table 2. ***School Distribution by Grade Level***

School-Level	Frequency	Percent
Elementary School	44	68.8%
Middle School	8	12.5%
Middle/High School (Grades 7-12)	3	4.7%
9th Grade Academy	1	1.6%
High School	8	12.5%

N=64

Survey Revisions Based on Pilot Results

The pilot survey contained 413 items. After the pilot rounds, the survey was revised to have 338 items, having deleted 41.6% (n=172) of the original items. Of the items in the field trial, 38.5% were items that were presented the same way in both the pilots and field trial, 32.8% were items that were revised based on the experience of the pilot trials. Of the items in the field trail, 28.7% were new questions added to the survey to provide information for the constructs being assessed.

Field Trial Study Design

Here, we describe the field trial to validate the *Survey of the Evidence in Education for Schools* (*SEE-S*) as a reliable and valid measure of school-based professionals' use of and perspectives about research. Using the version of the *SEE-S* following the development and piloting stages described previously, we aimed to administer the *SEE-S* survey to a nationally representative sample of school-based professionals (educators, administrators, staff) across the United States. The sections that follow include information about the target population, our sampling procedures and sample, data collection process, and the procedure used for statistical analyses of our survey items and scales.

Population

The target population for the survey includes school-based practitioners (heretofore, practitioner) responsible for instructional practice or policy, defined as teachers, principals, assistant principals, interventionists, and other instructional staff from schools and districts across the United States. A small number of district staff were also surveyed, if nominated by a participating district or school as a source for research or assistance connecting research to practice (i.e., if a school relied on a specific district staff person to help lead their use of research, we wanted to survey them too). However, the focus of our study is school-based professionals as our primary goal is to expand the study of knowledge utilization at the school level, as opposed to the district or state level. This focus is intended to motivate new approaches to increasing research use in schools and classrooms through a deeper understanding of the current practices related to, and beliefs about, the use of research in decision-making.

Sampling Procedures

To yield the most representative data for the *SEE-S,* we aimed to achieve a nationally representative sample of districts and charter schools across the U.S. A two-stage stratified random sample was used to select public school districts using the Common Core of Data as the sampling frame. Districts were stratified on urbanicity (i.e., suburban, rural, urban) and then sampled with probability proportional to size (i.e., total enrollment). Schools within districts were stratified by school level (elementary, middle, high) and then randomly sampled within strata with probability equal to the proportion of schools from that district in that strata. Charter schools that were part of a larger network were stratified on network size (i.e., 8% of charters were in networks with <5 schools, 1% of charters were in networks with ≥5 schools, 91% of charters were independent) and then charter networks and independent charter schools were sampled with probability proportional to size (i.e., total enrollment) within strata, with a restriction of no more than 2 large networks included in the sample. Public and charter schools with fewer than 10 teachers were excluded from the sampling frame because the goal was to

obtain responses from multiple respondents within schools to understand how research is used in decision making. Sampling was done using SAS 9.4 statistical software with the PROC SURVEYSELECT procedure.

Recruitment

We partnered with a professional recruiter who has extensive experience recruiting large nationally representative samples of schools for research. The recruitment strategy involved multiple modes of communication (e.g., email, print mail, telephone) to make initial contact, with follow-up as necessary to reach the desired sample size. Districts and schools that declined to participate were replaced with randomly sampled districts or schools from the same strata. A total of 99 districts and 44 charter school networks were contacted during our recruitment phase. Of those contacted, 21 districts (containing 1,033 schools total) and 10 charter school networks (containing 25 schools total) agreed to participate. Within the participating districts, respondents from 134 public schools and 20 charter schools agreed to participate and eventually completed the *SEE-S*.

Participation Incentives

Schools were offered a monetary incentive of up to $1,000 for each school if the school's final response rate was ≥ 90%. Lower response rates yielded smaller incentives (e.g., $900 for 89%-80%, $800 for 79%-70%, etc. with minimum = $100).

Field Trial Analyses & Results

Data Collection and Response Rate

The *SEE-S* was administered to school-based practitioners and administrators during the 2018-2019 and 2019-2020 school years. Names and email addresses of participants were collected from each district or school prior to administration. Surveys were administered via the Qualtrics online survey platform. Each participant was given four weeks to complete the survey. Weekly reminders were sent to non-respondents and partial completers. Additionally, a weekly response rate report was emailed to each principal at participating schools. After the

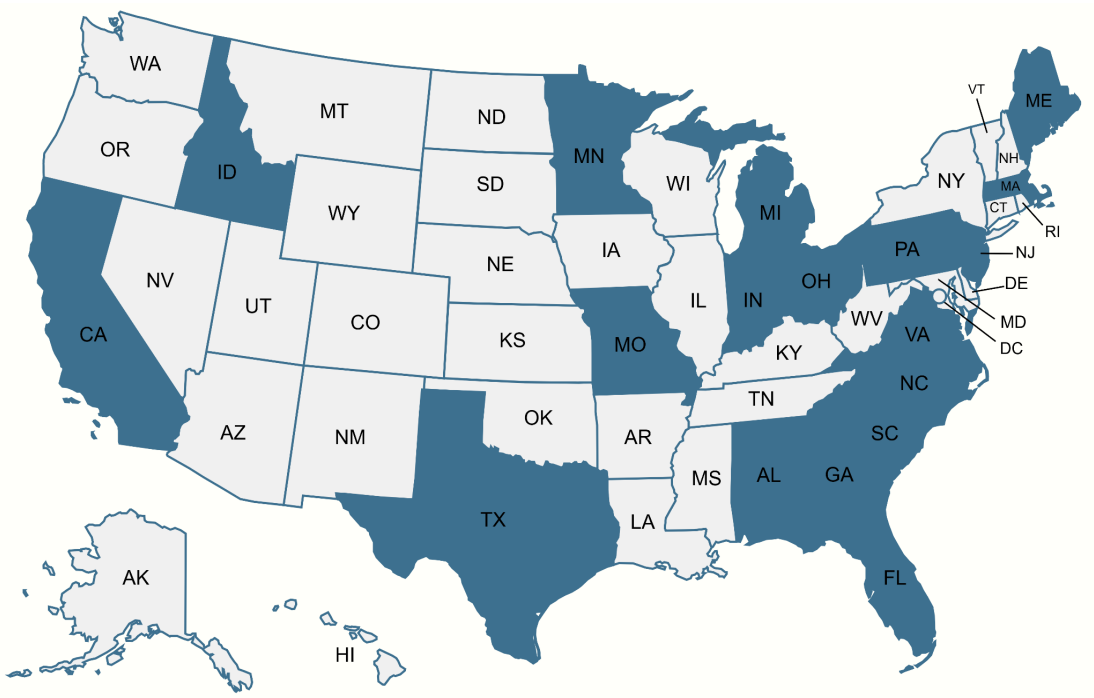
administration window for a school closed, the school was emailed a school-level descriptive report of their responses. Respondents could not return to previous sections that determined the sets of items in the survey (e.g., after a respondent indicated they were familiar with an organizational decision, they could not return to change those answers), but within sections respondents could return to responses and they did not need to complete the survey in one session.

The overall individual-level response rate for the *SEE-S* was 53%. The response rates by school varied from a low of 0% to a high of 100%. The average school response rate was 56%. The average time it took for an individual to complete the survey was 28 minutes.

Field Trial Sample Demographics

A total of 134 traditional public schools from 21 districts, as well as 20 charter schools, were successfully recruited into the sample for the field trial administration of the *SEE-S.* Schools from 18 different states were represented in the sample. Figure 2 depicts the geographic distribution of participating districts, by state.

Figure 2. ***Survey of Evidence in Education Field Trial: Participating Districts by State***



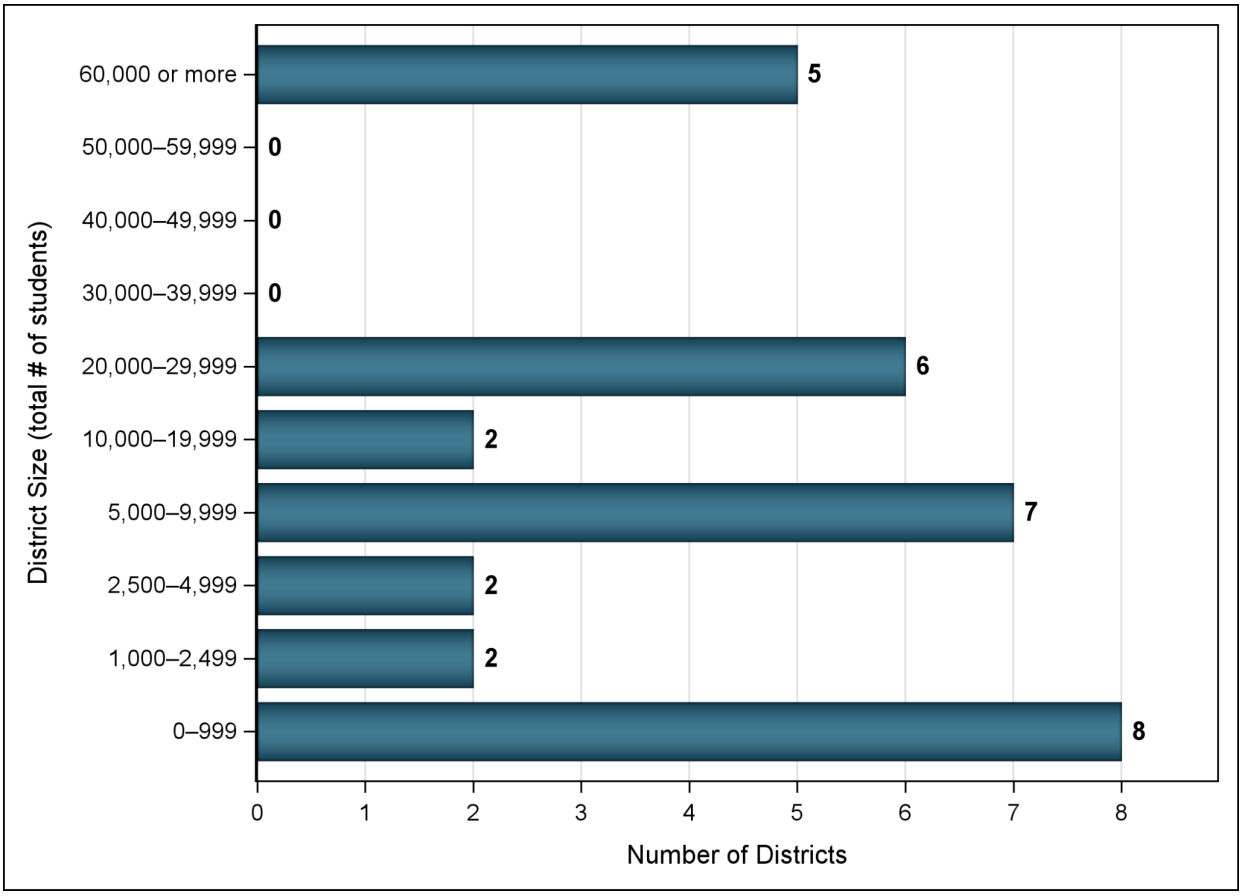
The respondents consisted of 4,390 school-based practitioners and 25 district staff. Respondents from an equal number of 'Suburban' and 'Urban' districts participated in the study. Table 3 details district-level sociodemographic characteristics.

Table 3. ***Sociodemographic Characteristics of Participating Districts***

Characteristic	*n*	*%*
Public Status		
Public	21	67.7
Charter	10	32.3
Urbanicity		
Rural	5	16.2
Suburb	13	42.9
Urban	13	42.9

Figure 3 illustrates the distribution of student enrollment by district. The majority of districts enrolled between 0 and 9,999 students. Five districts enrolled 60,000 or more students.

Figure 3. ***Distribution of Participating Districts by Student Enrollment***



Our final sample included 134 traditional public and 20 charter schools. With regard to the distribution of schools by grade level served, our sample mirrored national statistics from the 2018–19 school year (NCES, 2020). More specifically, 55% of schools in our sample fit the definition for an elementary school (e.g., K–5 or K–6), which matches the national percentage of 55%. Sixteen percent of our sample were middle schools (e.g., typically grades 6-8), which matches the national percentage of 16%. Twenty-one percent of our sample were high schools (e.g., typically grades 9–12), whereas the national percentage is 22%. The remaining 9% of schools spanned wider grade ranges (e.g., K–8 or K–12).

Table 4. ***Characteristics of Participating Schools***

Characteristic	*n*	*%*
Status		
Public	134	87.0
Charter	20	13.0
Urbanicity		
Charter	20	13.0
Rural	28	18.2
Suburban	75	48.7
Urban	31	20.1
Grade Level		
Elementary	84	54.5
Middle	24	15.6
High	32	20.8
Elem./Mid	6	3.9
Mid/High	2	1.3
Elem./Mid/High	6	3.9

Sociodemographic Characteristics of Participating Schools

A number of different school personnel were eligible to take the *SEE-S*. Among others, this included teachers, principals, assistant principals, and instructional coaches. Table 5 depicts the distribution of participants by role for the *SEE-S*.

Table 5. ***SEE-S Participants by Role***

Staff Role	*n*	*%*
Classroom teacher	2818	63.8%
Special education teacher	420	9.5%
Arts or electives teacher (e.g., CTE/STS, music, art, P.E., etc.)	298	6.7%
School administrator (e.g., principal or assistant principal)	181	4.1%
Paraprofessional or teacher assistant	152	3.4%
Instructional coach	129	2.9%
Interventionist	126	2.8%
Guidance counselor	108	2.4%
Librarian	44	1.0%
Speech language therapists	43	1.0%
Health professional (e.g., school nurse)	32	0.7%
District administrator/staff	25	0.6%
Technology/Media specialist	25	0.6%
School psychologist	8	0.2%
Occupational therapist	5	0.1%
Physical therapist	1	0.02%

N=4,415 Individuals who completed at least one full portion of the survey (e.g., *Depth of Use*)

Methods for Establishing Validity and Reliability of the *SEE-S*

Several techniques were used for item analysis, scaling, and calculation of derived variables. The first stage of analyses involved calculation of simple frequencies for the response categories of each closed-response item. These were used to confirm the absence of ceiling or floor effects and to confirm the appropriateness of response categories. Several items were also linked to conditional follow-up items, with open-ended responses used to confirm the validity of responses for the fixed-choice item. For example, a respondent who reported that external research influenced a decision was then required to provide information that could be used to identify what research they were referring to (e.g., an author name, a title, a web URL). After coding open-ended items and imposing validity filters on linked items, the next stage of

analyses involved exploratory and confirmatory factor analysis, reliability analysis, and calculation of survey scales for sets of items believed to measure a latent construct.

To examine the properties of survey scales in the Assumptions and Perspectives About Research, Capacity to Critically Consume Research, and Brokerage sections, we first conducted exploratory factor analysis with individual-level data using the PROC FACTOR procedure with an oblimin rotation with a weight of zero given our assumption that the sub-factors were correlated. The number of factors extracted for each set of items was determined based on examination of scree plots and interpretation of subsets of items relative to our measurement blueprint and conceptual framework, then verified via confirmatory factor analysis (CFA). The CFA of survey scales was conducted with the PROC CALIS procedure in SAS 9.4. Results indicated that a number of our scales met or exceeded standards for internal consistency (Hu & Bentler 1999). Table A1 in Appendix A provides the reliability coefficients (Alphas) for each construct intended to represent a scale (Cronbach, 1951). Also included are the number of survey respondents, and number of items comprising each scale, and basic descriptive statistics.

For constructs related to *Depth of Use* (which are focused on the organizational conditions and efforts related to research use), validity and reliability analyses were performed using school-level data after aggregation using multilevel linear or non-linear mixed models (a.k.a, hierarchical models, or HLM). These multilevel models were used to produce empirical-Bayes estimates of school-level intercepts (see Raudenbush & Bryk, 2002, p. 45-51), which serve as precision-weighted estimates of the school-level averages (see the subsequent section entitled "Aggregate School-Level Scores for Depth of Research Use" for details). This technique addresses the issue of varying numbers of respondents from each school, reducing the likelihood of erroneous identification of outliers due to imprecision. Thus, there was no minimum school-level response rate for the analysis of *Depth of Use* scores. Factor analysis and reliability calculations for school-level constructs were performed using these aggregated estimates.

Depth of Research Use

The first section (i.e., Section "1S") of the *SEE-S* focuses on the six dimensions of *Depth of Use* of education research by school-based practitioners (see Figure 1). We focus on two different types of decisions, organizational decisions and personal practice decisions, recognizing that not all practitioners will be involved in or familiar with organizational decisions. We define organizational decisions as "decisions about policy and practice made at the school or district level that affect a significant number of teachers and/or students." We define personal practice decisions for individual respondents as "a decision you have made about your own practice. For example, using a new instructional strategy, changing your classroom organization, or implementing what you learned from PD training." We began this section of the survey by asking participants to describe an organizational decision made by their school community in the last year and then to rate their familiarity with the decision-making process. We call this component of the survey "Path A." If the participants could not name an organizational decision or had little to no familiarity with the decision listed, they were redirected in the survey to a parallel set of items (i.e., we call this "Path B") in which they were asked to describe a personal practice decision. Adequate familiarity with an organizational decision was defined as an average score of at least 1.5 on a 0-3 scale ('not familiar'=0, 'somewhat familiar'=1, 'mostly familiar'=2, 'very familiar'=3) across the following three items:

> *Although you may or may not have been involved in the decision you described above, how familiar are you with each of the following aspects of the decision process?*
> *1) The information used to inform the decision*
> *2) The process for gathering the information*
> *3) Who was involved in gathering information and/or making the decision*

Some respondents (n=412 or 9.3% of our sample) were unable or unwilling to describe either an organizational or a personal decision and were allowed to bypass the 1S section of the survey after confirming that they were unable to describe any decisions. Other respondents who were unfamiliar with an organizational decision but were able to describe a decision they made "in

their personal practice" were routed to Path B (n=2,660 or 66.5% of our sample). Following this dynamic routing, practitioners responded to a nearly identical set of questions focused on either an organizational or a personal decision.

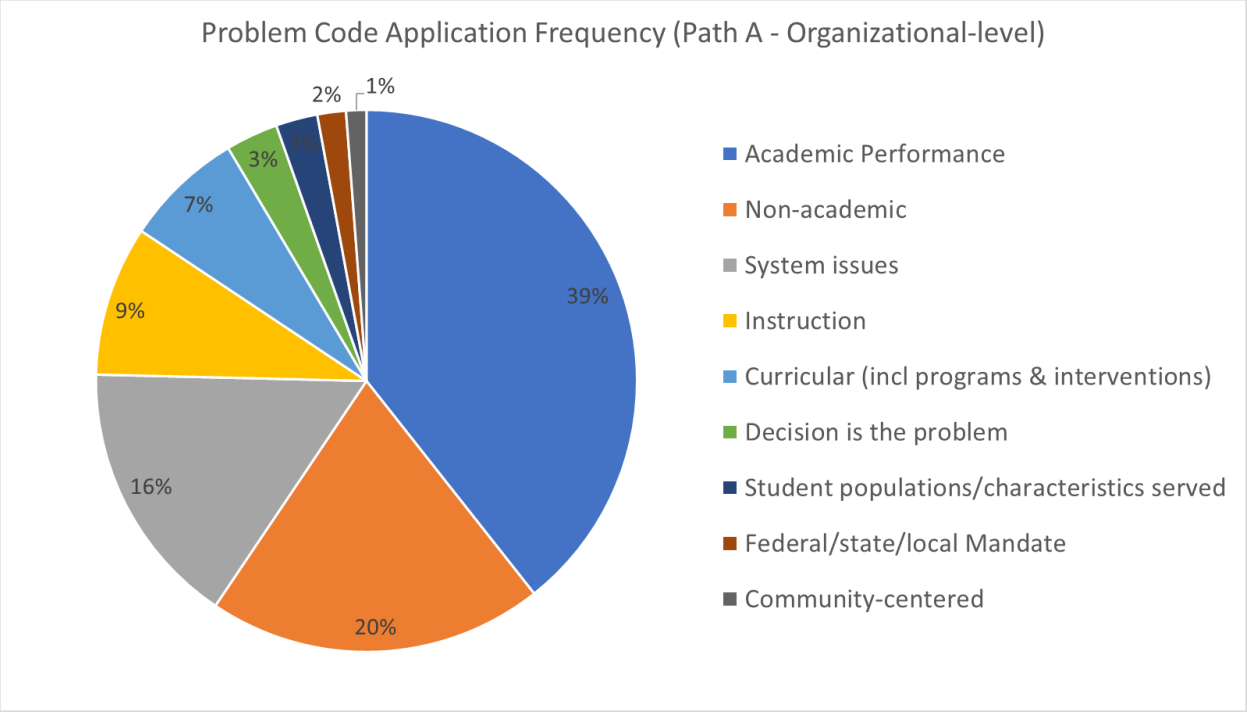
Problems & Decisions - Qualitative data

Types of Organizational Problems Faced & Decisions Made in Schools

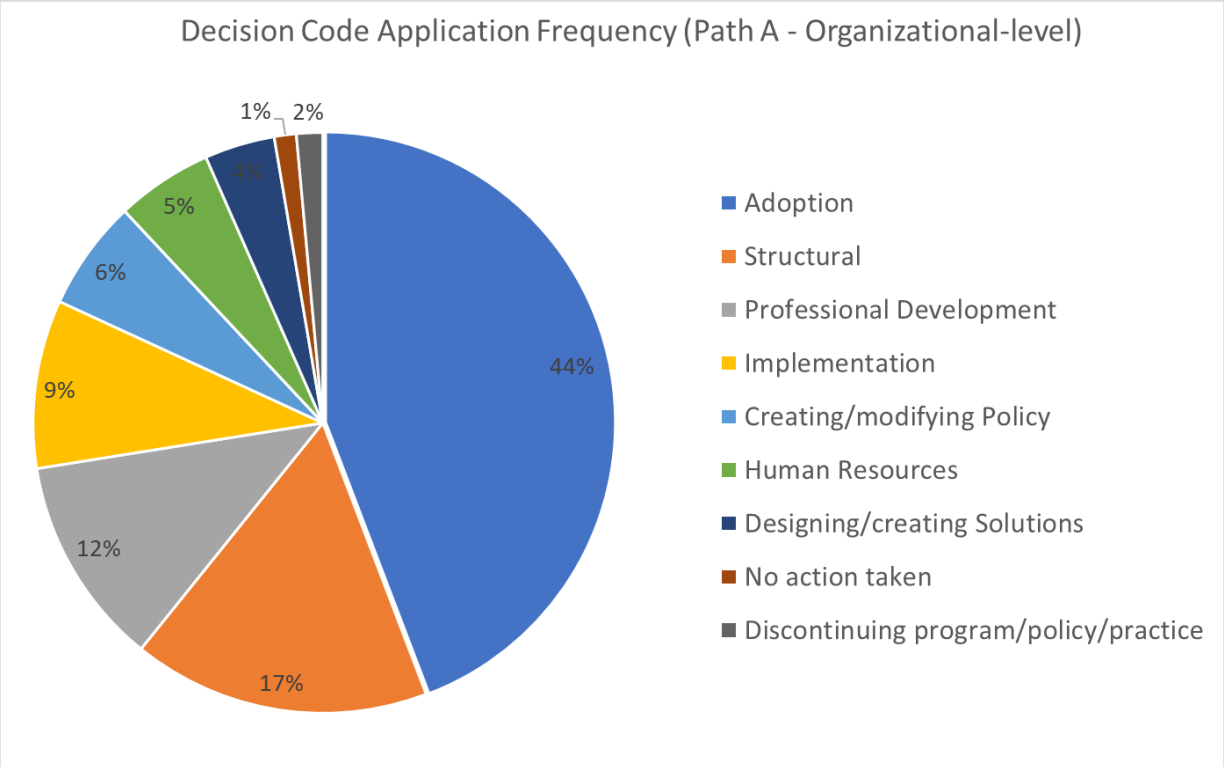
The qualitative data gathered by the open-ended survey items about 1) what organizational decision was made and 2) the challenge or problem that the decision addressed was analyzed using an a priori coding framework developed and refined using two rounds of pilot data from the *SEE-S* survey. To establish this coding framework the team established an iterative process for coding the data using an emergent thematic approach. Content and types were categorized through an iterative discussion in which different themes were created, tested with sample responses, and modified. Through this process multiple categories of problems and decisions were identified. Organizational problems could be classified as pertaining to "academic performance," "non-academic issues," "instruction," "curriculum & programs," "systemic issues," "community-centered," "student characteristics/populations," and "federal/state/local mandates." Organizational decisions observed fell into ten categories: "adoption," "implementation," "human resources," "structural," "external," "professional development," "no action taken," "discontinuing a policy or practice," "creating or modifying a policy," and "designing or creating a solution."

After refining the coding framework over the course of two pilot administrations of the *SEE-S*, it was applied to the data collected in the field trial. The most frequently reported organizational problems were related to academic performance (39%, n=553), followed by non-academic issues (20%, n=282) and systemic issues (16%, 224). The least commonly reported organizational problems were those related to community-centered issues (1%, n=17). The most frequently reported organizational decisions were decisions to adopt something new (44%, n=575), structural changes (17%, n=216), and professional development (12%, n=152).

Figure 4. ***Types of Organizational Problems and Decisions in Participating Districts***



Note. N= 1,405

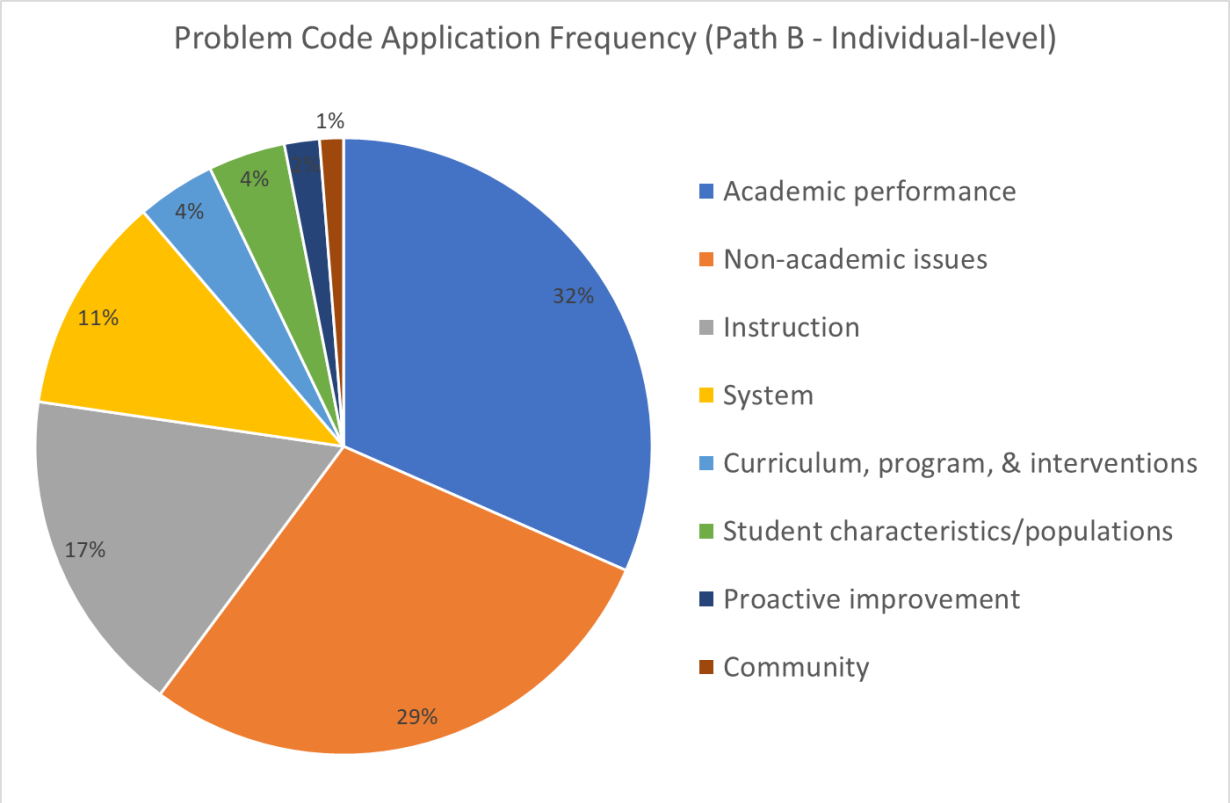


Note. N= 1,313

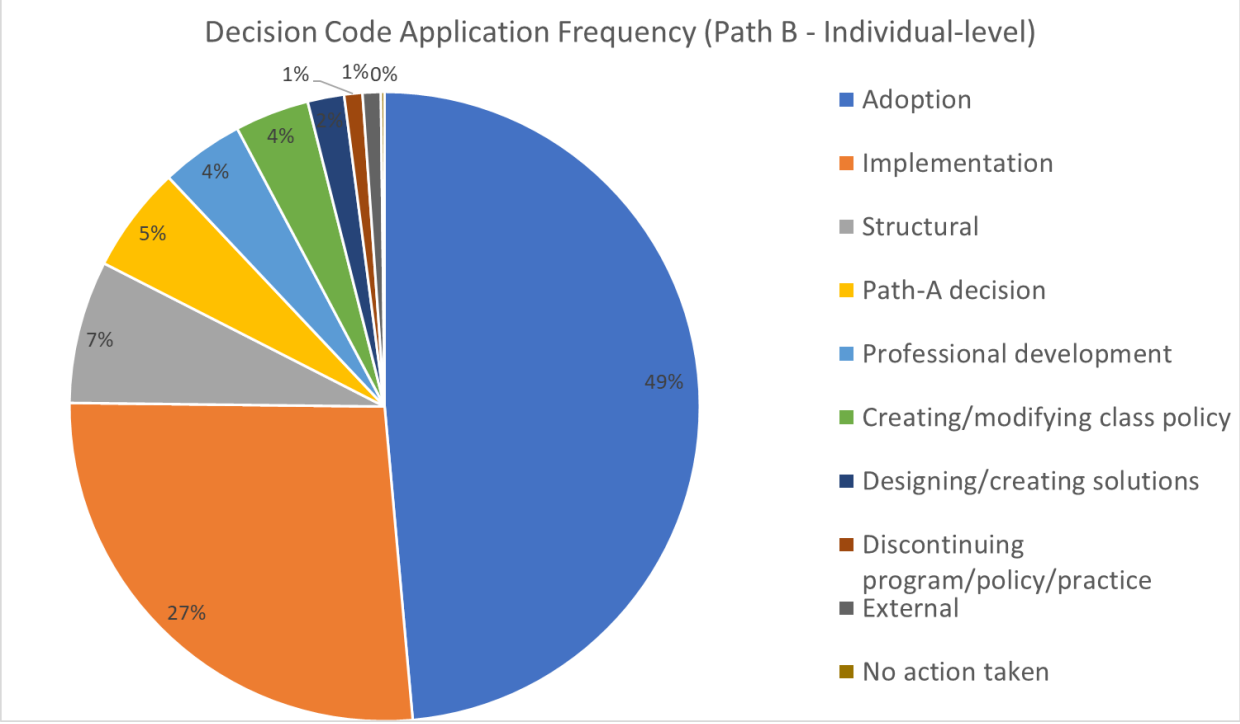
The coding framework developed for the Path-A data (organizational-level decision) was applied to a subset of Path-B data (individual-level decision) from the two pilot administrations of the *SEE-S* to test whether or not modifications to the existing coding framework were necessary. Through testing and team discussions, it was determined that the same coding framework could be applied to the Path-B data with three modifications: the addition of a new code, the re-parenting of an existing code, and the elimination of an existing code. The problem code "Proactive Improvement" was added, the decision code "Human Resources" was eliminated, and the problem code "Federal/state/local mandate" was nested under the parent code "Systemic Issue" (for more details on these coding decisions see Appendix B).

The most frequently reported individual problems were related to academic performance (32%, n=859), followed by non-academic issues (29%, n=777) and instructional issues (17%, n=467). The least commonly reported individual problems were those related to community-centered issues (1%, n=34). The most frequently reported individual decisions were decisions to adopt something new (49%, n=1,298), implementation modifications (27%, n=711), and structural changes (7%, n=196).

Figure 5. ***Types of Individual Problems and Decisions***



Note. N= 2,720



Note. N= 2,672

Evidence

The first section of the survey following the identification of an organizational or an individual decision included one item measuring the influence of fourteen different forms of evidence on those decisions (see Figure 6). Clickable links for the term 'research' were included that provided a pop-up definition for research: "*Systematic data collection and analysis driven by research questions*."

Figure 6. *SEE-S Item Capturing Influence of Different Forms of Evidence*

To what extent did the following types of information influence the decision? Please check a response for every row.

Note: Click on underlined text for a definition of the word or phrase.

	Had **little or no** influence	Had **some** influence	**Heavily** influenced	I Don't Know/Remember
Articles, reports, books, or summaries based on external research or program evaluation (paper or web-based)	○	○	○	○
Research or program evaluation conducted by central office staff	○	○	○	○
Research or program evaluation conducted by teacher(s) or principal(s)	○	○	○	○
Research or program evaluation led by students or local youth	○	○	○	○
Other formal analysis of school-wide or district-wide data	○	○	○	○
Informal data collected by school/district staff	○	○	○	○
Materials from a program developer, publisher or from professional development training	○	○	○	○
Opinion of national expert(s)	○	○	○	○
Guidance from Federal or State Departments of Education	○	○	○	○
Advice from local education leader(s) (e.g., district superintendent)	○	○	○	○
Other practitioners' experiences/advice	○	○	○	○
Opinions of parents or other community members	○	○	○	○
My own professional experience	○	○	○	○
Other (Please Specify)	○	○	○	○

The Influence-of-Evidence item was followed by four conditional items. The first conditional item was presented to respondents who reported that external research influenced the decision. This item stated, "*You indicated that articles, reports, books, or summaries based on independent research or program evaluation influenced the decision: Please tell us about a research or evaluation study that influenced the decision, including as much information as possible about the study author, title, or web URL,*" followed by an open-ended text response box. The other three conditional items were presented to respondents who reported that local research and/or data analyses influenced the decision, and the items inquired about the analyses performed, who conducted them, and what was produced. These conditional items were used to create validity filters for the responses to the first five evidence items covering the influence of external research, local research, and formal analyses of local data. The goal of these validity filters was to reduce social-desirability response bias in our *Depth of Use* measures (e.g., respondents claiming that research was used, even though they have no idea what that research was). Unless an individual respondent could cite basic details about the evidence they claimed was influential, we did not count their response to that influence item as a valid indicator of *Depth of Use*. In other words, *Depth* is a function not only of actual use, but also the degree to which the broader school staff are aware of the evidence used to inform the decision.

Validity Filter - External Research

Of the 1,343 respondents routed to Path A (i.e., organizational decision), 857 respondents (64%) reported that external research had some or heavy influence on the decision. These 857 respondents were then routed to the first validity check item, "*Please tell us about a research or evaluation study that influenced the decision, including as much information as possible about the study author, title, or web URL.*" This item was left completely blank by 165 respondents (19%). The remaining responses were coded using a rubric (see Appendix C) designed to (a) confirm the existence of research cited by the respondent, (b) classify the citation as direct versus indirect (e.g., providing information about the research source versus the name or entity who provided the research), and (c) classify the type of citation (e.g., a book, an author's name,

a research article, etc.). Only 34% of responses claiming influence by external research were confirmed as research, and the proportion of Path A decisions that were validated as influenced by external research dropped to 22% (i.e., 292 out of 1,343). As stated previously, *Depth of Use* is a function not only of actual use, but also the degree to which the broader school staff are aware of the evidence used to inform the decision. Invalidated responses regarding use of external research do not necessarily suggest that research was not used. Rather, given that respondents familiar with a decision could not provide information that could be used to confirm what research was used, depth of use is simply more shallow.

Of the 2,660 respondents routed to Path B (i.e., individual decision), 1,703 respondents (64%) reported that external research had some or heavy influence on the decision. These 1,703 respondents were then routed to the first validity check item, "*Please tell us about a research or evaluation study that influenced the decision, including as much information as possible about the study author, title, or web URL.*" This item was left completely blank by 337 respondents (20%). The remaining responses were coded using the same rubric that was used for organizational decisions described above (see Appendix C). Only 33% of responses claiming influence by external research were confirmed as research, and the proportion of Path B decisions that were validated as influenced by external research dropped to 15% (i.e., 395 out of 2,660). As stated previously, these invalidated responses suggest that depth of use is more shallow because respondents familiar with a decision could not identify what research was used.

Validity Filter - Local Research

The influence of evidence item (see Figure 6) included three forms of local research: *Research or Program Evaluation Conducted by Central Office Staff*, *Research or Program Evaluation Conducted by Principals or Teachers*, and *Research or Program Evaluation Led by Students or Local Youth*. Of the 1,343 respondents routed to Path A (i.e., organizational decision), 1,116 respondents (83%) reported that at least one form of local research had some or heavy influence on the decision. These 1,116 respondents were then routed to the second validity check item. This conditional item simply asked what product(s) were produced from the

research (i.e., *A summary report, An academic journal article, A summary presentation, Customized reports for individual schools or teachers, Data visualizations (e.g., charts, plots, figures), Other,* or *I don't know*). Forty percent (n=448) of the 1,116 respondents selected "I don't know", and this was used as the validity filter—affirmative responses to the influence of local research items were invalidated if the respondent did not know what was produced from the research. After implementing this validity filter, the proportion of Path A decisions that were validated as influenced by local research dropped to 50% (i.e., 668 out of 1,343) after combining data across the three forms of local research. As stated previously, these invalidated responses suggest that depth of use is more shallow because respondents familiar with a decision could not identify what research was used.

Of the 2,660 respondents routed to Path B (i.e., individual decision), 1,902 respondents (72%) reported that at least one form of local research had some or heavy influence on the decision. These 1,902 respondents were then routed to the second validity check item. This conditional item simply asked what product(s) were produced from the research (i.e., *A summary report, An academic journal article, A summary presentation, Customized reports for individual schools or teachers, Data visualizations (e.g., charts, plots, figures), Other,* or *I don't know*). Forty-six percent (n=871) of the 1,902 respondents selected "I don't know", and this was used as the validity filter—affirmative responses to the influence of local research items were invalidated if the respondent did not know what was produced from the research. After implementing this validity filter, the proportion of Path B decisions that were validated as influenced by local research dropped to 39% (i.e., 1,031 out of 2,660). Again, these invalidated responses suggest that depth of use is more shallow because respondents familiar with a decision could not identify what research was used.

Validity Filter - Local Data Analyses

The Influence-of-Evidence item (see Figure 6) captured analyses of local data as "*Other formal analyses of school-wide or district-wide data*" in order to distinguish it from local research. Such analyses may not have been guided by specific research questions and they may not have resulted in a formal product; however, other details can be used to confirm the validity of

reports from survey participants. More specifically, those respondents who reported that formal analyses of local data influenced a decision were presented with five follow-up questions shown in Figure 7.

Of the 1,343 respondents routed to Path A (i.e., organizational decision), 1,043 respondents (78%) reported that formal analyses of local data had some or heavy influence on the decision. Of these respondents, 182 (17% of the 1,043) did not know what was produced (i.e., they selected "I don't know" for products, and "No" or "I don't know" for the interactive dashboard) and they selected "I don't know" for at least 2 of the remaining 3 items: *Who did the analyses?, What data were included?,* and *What kinds of analyses were done?* The remaining 861 (83% of the 1043) respondents were able to report what was produced from the analyses and/or at least two of the following pieces of information: who did the analyses, what data were used, and what kinds of analyses were performed. After implementing this validity filter, the proportion of Path A decisions that were validated as influenced by local data analyses dropped to 64% (861 out of 1,343). As stated previously, the invalidated responses suggest that depth of use is more shallow because respondents familiar with a decision could not describe basic details of the local data analyses used.

Of the 2,660 respondents routed to Path B (i.e., individual decision), 1,548 respondents (58%) reported that formal analyses of local data had some or heavy influence on the decision. Of these respondents, 356 (23% of the 1,548) did not know what was produced (i.e., they selected "I don't know" for products, and "No" or "I don't know" for the interactive dashboard) and they selected "I don't know" for at least 2 of the remaining 3 items: *Who did the analyses?, What data were included?,* and *What kinds of analyses were done?* The remaining 1,192 (77% of the 1,548) respondents were able to report what was produced from the analyses and/or at least two of the following pieces of information: who did the analyses, what data were used, and what kinds of analyses were performed. After implementing this validity filter, the proportion of Path B decisions that were validated as influenced by local data analyses dropped to 45% (1,192 out of 2,660). Again, invalidated responses suggest that depth of use is more shallow because respondents familiar with a decision could not describe basic details of the local data analyses

used. Additional details on responses to the local data analyses follow-up items are presented in a subsequent section entitled, “Item Analyses - Local Research and Data Analyses.”

Figures 8 and 9 show the relative influence of all forms of evidence included in the Influence-of-Evidence item (see Figure 6) for organizational decisions and individual decisions, respectively. For organizational decisions, *informal data collected by school or district staff* is most often influential, followed by *advice from local education leaders*, *other practitioners experiences/advice*, and *formal analyses of school or district data.* School research (i.e., *local research conducted by school or district staff*) is also influential in a substantial number of decisions—however, other forms of research appear to be far less likely to influence decisions. Results for individual decisions suggest that the professional experiences of individual educators and their colleagues are most influential, and that informal and formal analyses of local data is again more often influential than other forms of research. As noted in the previous sections describing the validation filters for external research, local research, and local data analyses, many respondents (i.e., up to 49%) were unable to provide details supporting their claims that formal data analyses or research influenced a decision.

Figure 7. ***SEE-S Validity Filter Items for Influence of Local Research and/or Local Data Analyses***

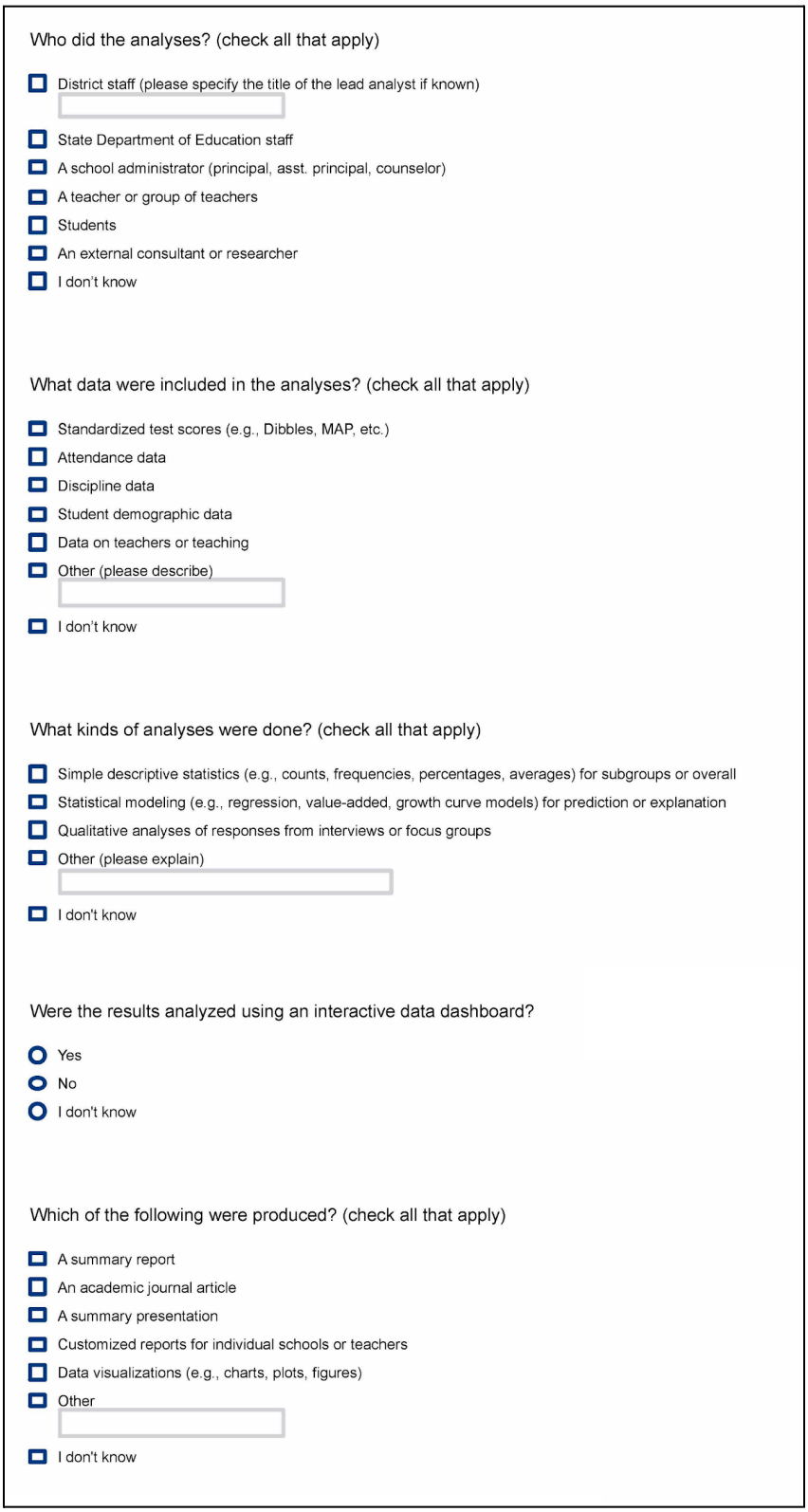
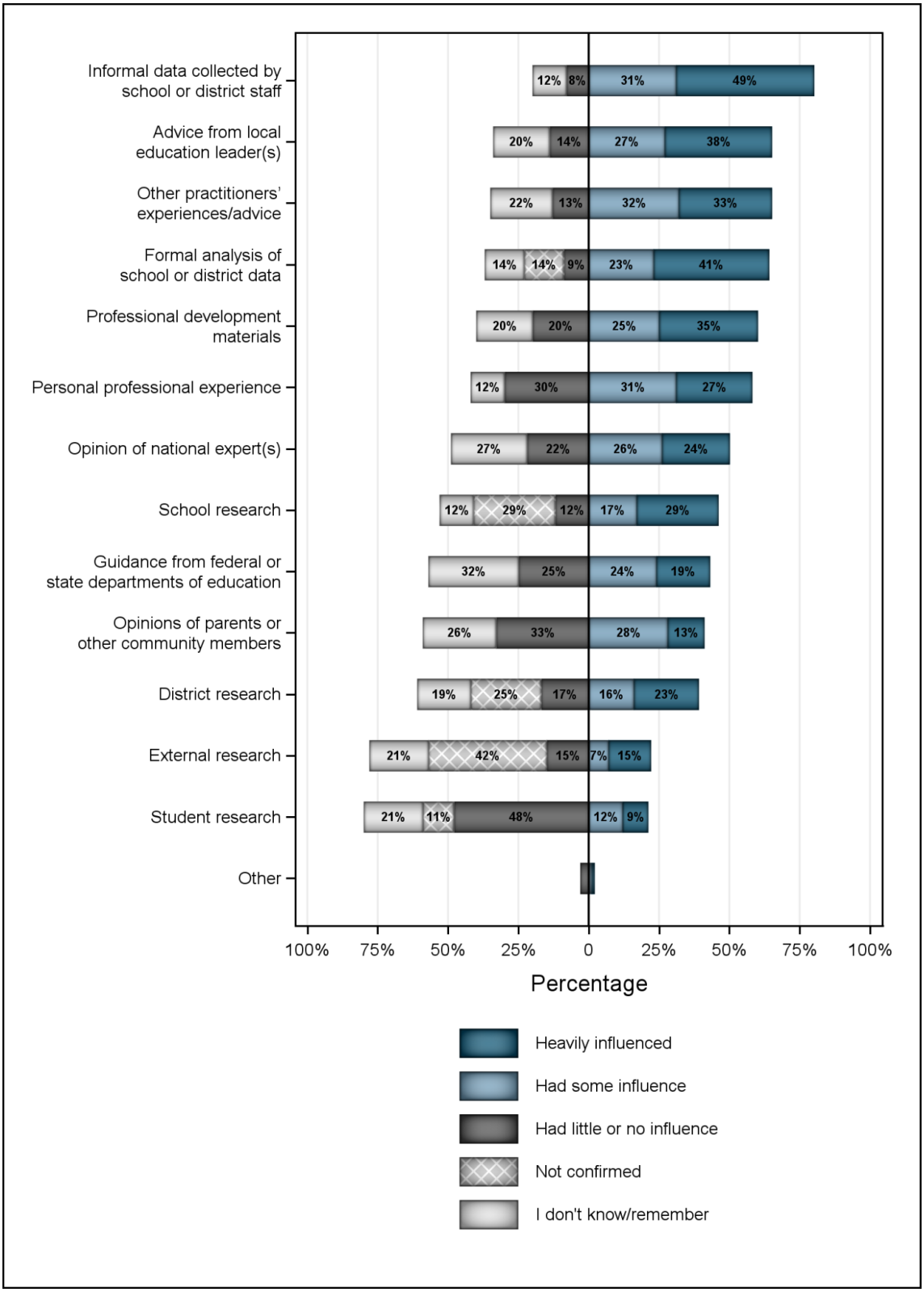
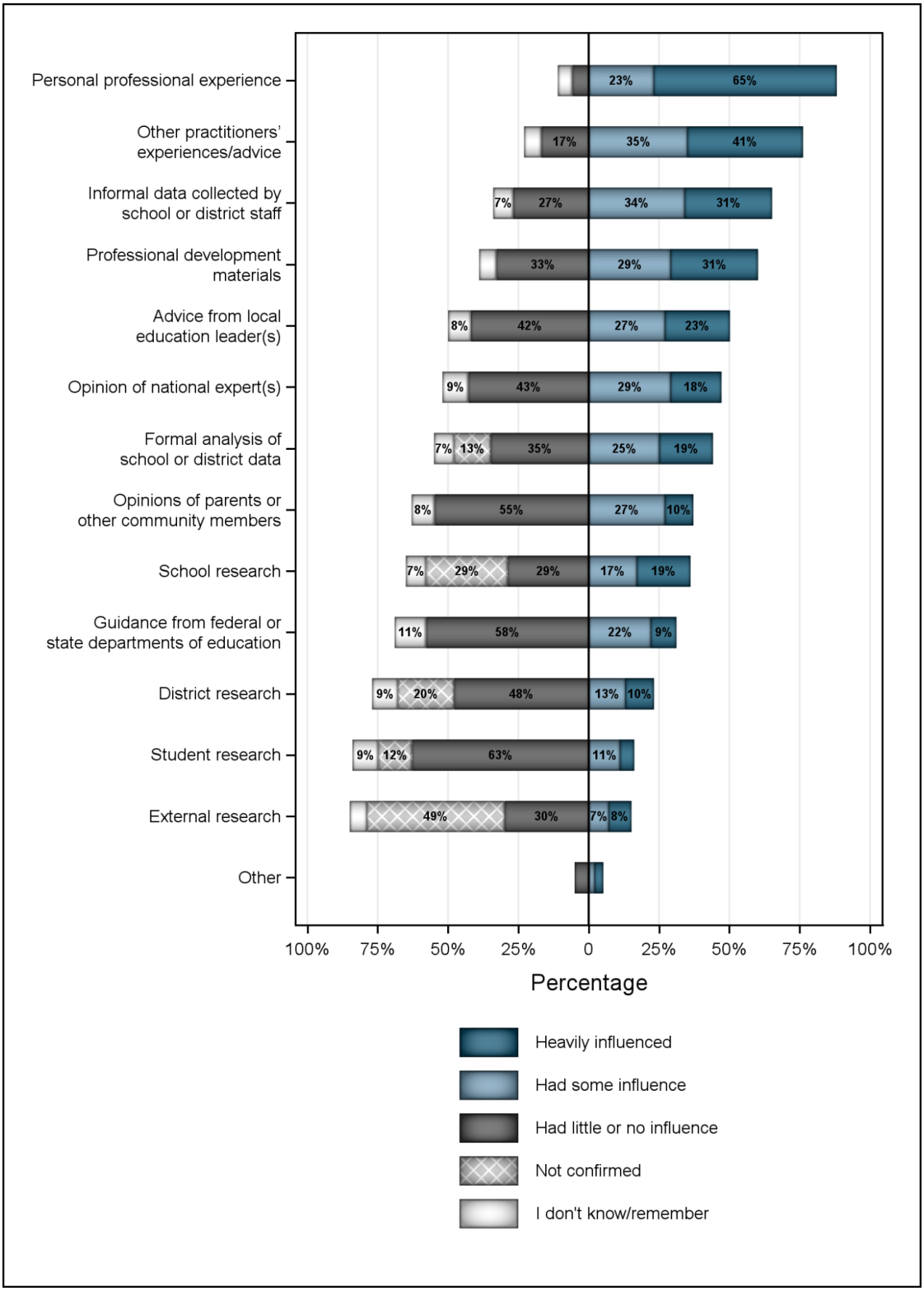


Figure 8. ***Influence of Different Forms of Evidence, Organizational Decisions***



Note. Item response N ranges from 982 to 1,343

Figure 9. ***Influence of Different Forms of Evidence, Personal Practice Decisions***

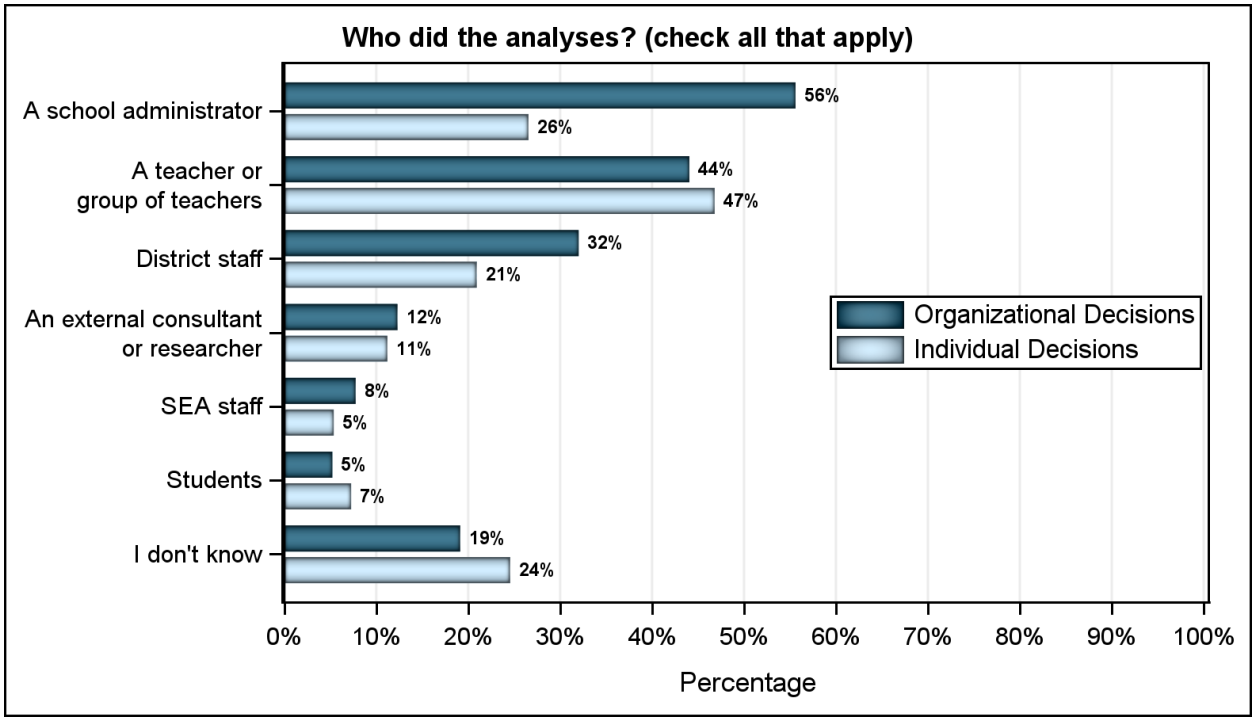


Note. Item response N ranges from 2,380 to 2,660

Item Analyses - Local Research and Data Analyses

When asking about the evidence used to inform the decision, several options categorized as local research and data analyses were presented. Participants who responded that 'District Research', 'School Research', 'Student Research', or 'Formal Analyses of Local Data' were used during the decision-making process were asked follow-up questions about the type of data used in the analysis, the analysis methods, and who conducted the analysis. These results are depicted in Figure 10.

Figure 10. ***Involvement in Analysis of Local Data to Inform Decisions***



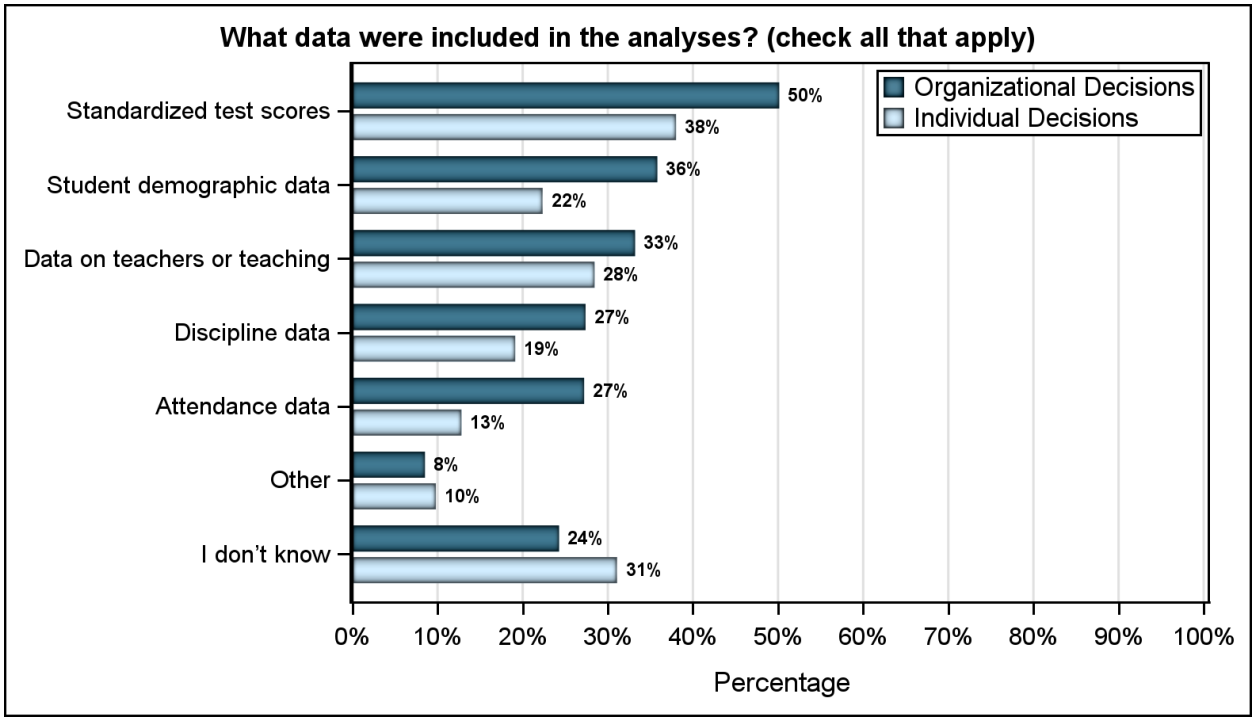
Note. Item response N = 1,187 for organizational decisions; N = 2,059 for individual decisions.

For organizational decisions, school administrators were most often involved in analyses of local data (i.e., noted by 56% of respondents). Teachers were involved 44% of the time, district staff were involved in analysis 32% of the time, while external researchers were involved 12% of the time. About 20% of respondents were unsure who was involved in the analysis. Results were similar for individual decisions, except that school administrators conducted analyses for only

27% of the decisions reported. For individual decisions, teachers were most often involved in data analysis (47%, i.e. using their own or their colleagues' data). School administrators were involved in analysis 27% of the time, district staff were involved 21%, and others were involved far less often. Nearly a quarter of respondents were unsure about who engaged in analysis.

Respondents also described the type of local data that was used. Standardized test score data were analyzed for 50% of organizational decisions. Student demographic data was cited in 36% of responses, and just over 30% of decisions involved data on teachers or teaching. Individual decisions tended to use fewer types of data for any one decision, but the relative prevalence of each type was similar, except for student demographics, which was less prevalent for individual decisions. Figure 11 provides complete information on local data analyzed for organizational and individual decisions.

Figure 11. ***Types of Local Data Analyzed to Inform Decisions***

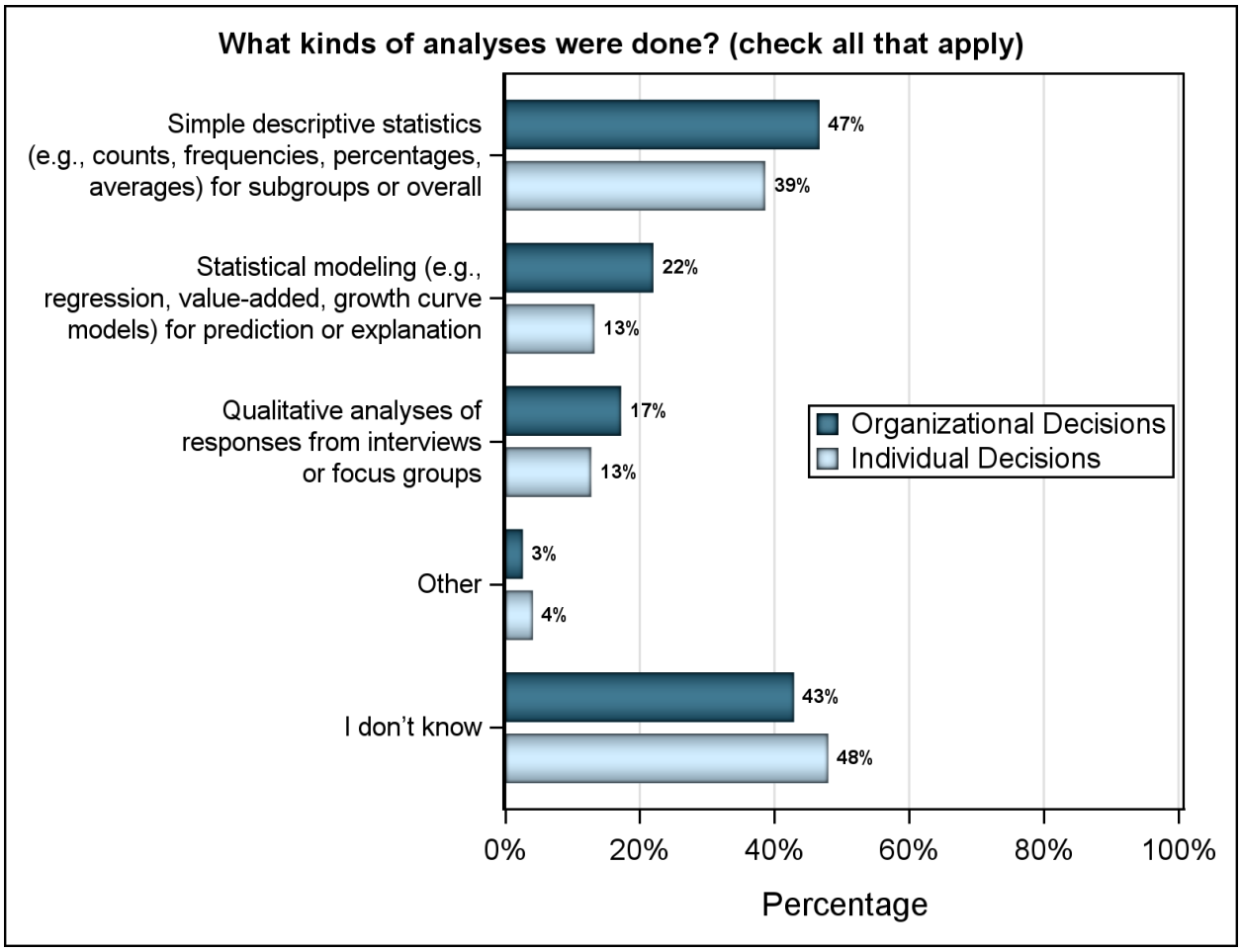


Note. Item response N = 1,187 for organizational decisions; N = 2,059 for individual decisions.

Survey participants reported on the types of analyses conducted using local data. Primarily, simple descriptive statistics were calculated, though 22% of Path-A respondents indicated that

statistical modeling (e.g., regression) was used to inform an organizational decision. Almost half of Path-A respondents (43%) indicated that they were unsure of the type of analysis that was conducted. Analyses supporting individual decisions also tended to rely on descriptive analyses, and about half of Path-B respondents (48%) did not know what kinds of analyses were performed. Complete data for this question is detailed in Figure 12.

Figure 12. ***Types of Analyses Applied to Local Data to Inform Decisions***

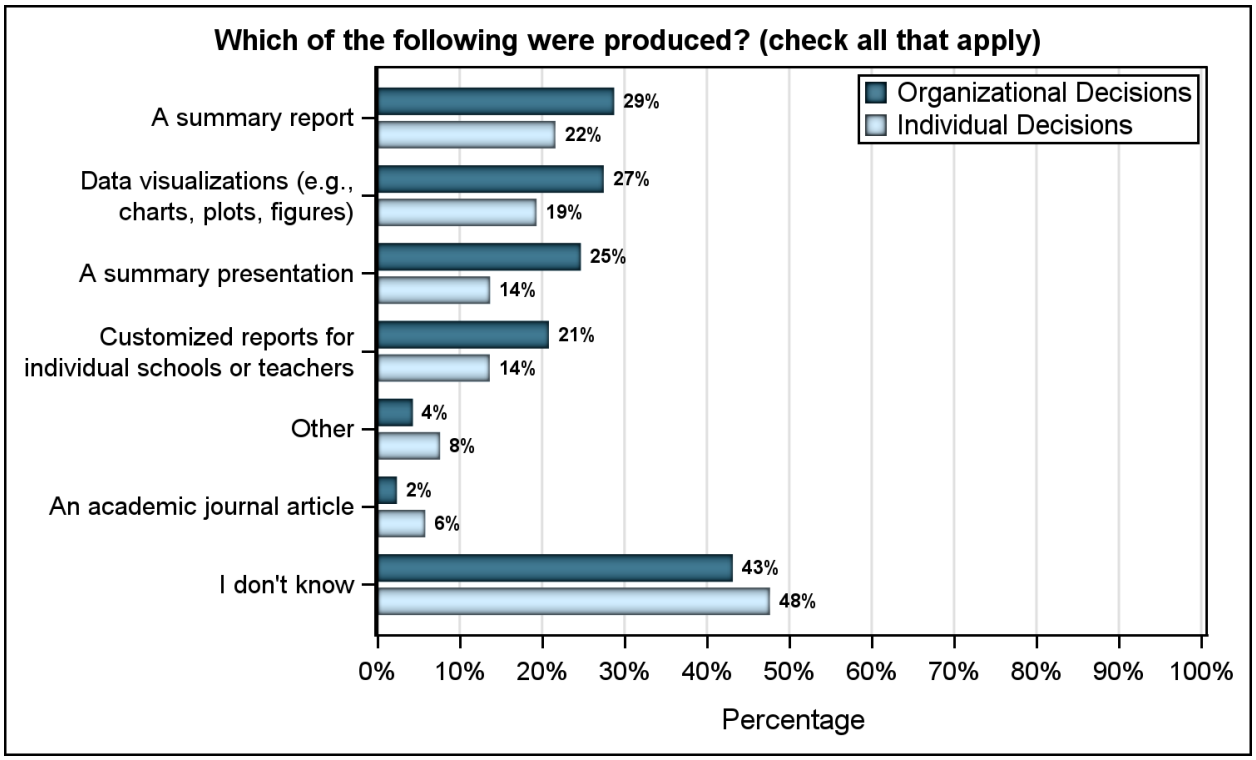


Note. Item response N = 1,187 for organizational decisions; N = 2,059 for individual decisions.

Twenty-nine percent (29%) of respondents indicated that a summary report was produced from the analysis of local data used to inform organizational decisions. Data visualizations (e.g., charts/graphs) were created in 27% of cases where local data was analyzed and informed

organizational decisions. Respondents also indicated that a summary presentation of the data analysis results was not common, with only 25% of Path-A respondents reporting that this occurred. For individual decisions, the prevalence of most formal products was lower, except for academic journal articles, which were cited for 6% percent of individual decisions (versus only 2% for organizational decisions). Detailed results can be found in Figure 13.

Figure 13. ***Products Produced from Local Data Used to Inform Decisions***



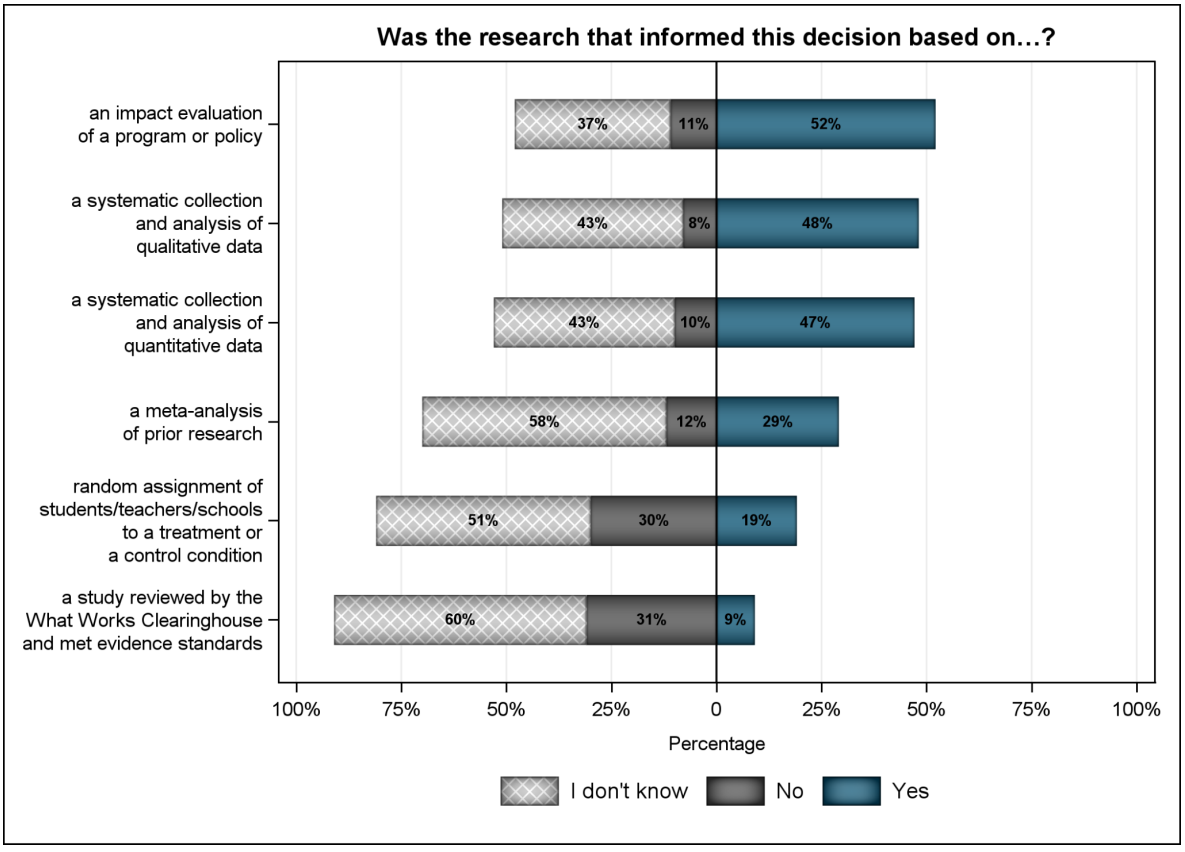
Note. Item response N = 1,187 for organizational decisions; N = 2,059 for individual decisions.

Respondents were also asked an additional survey question, "*Were the analyses intended to determine whether a strategy worked better than an alternative strategy (or what's been done in the past)?*". Forty-three percent of Path-A participants and 43% of Path-B respondents reported that the analyses were conducted for this purpose. However, 46% of Path-A respondents and 43% of Path-B respondents were unsure. A final item in this section asked, "*Were the results analyzed using an interactive data dashboard?*" Only 12% of Path-A respondents and 9% of Path-B respondents indicated that an interactive data dashboard was used to analyze the data.

Causality, Generalizability, and Precision of Research Evidence

Survey respondents answered items related to respondents' perceptions of the causal inference and research design underlying the evidence that informed the decision, the generalizability of the evidence, and the precision and level of inference of the research. Survey respondents tended to perceive that organizational decisions were based on systematic collections of qualitative or quantitative data and results of program evaluation. They were less likely to perceive that the evidence was from research that employed random assignment of subjects to study conditions (see Figure 14).

Figure 14. ***Respondents' Perceptions of Causality of Evidence for Organizational Decisions***

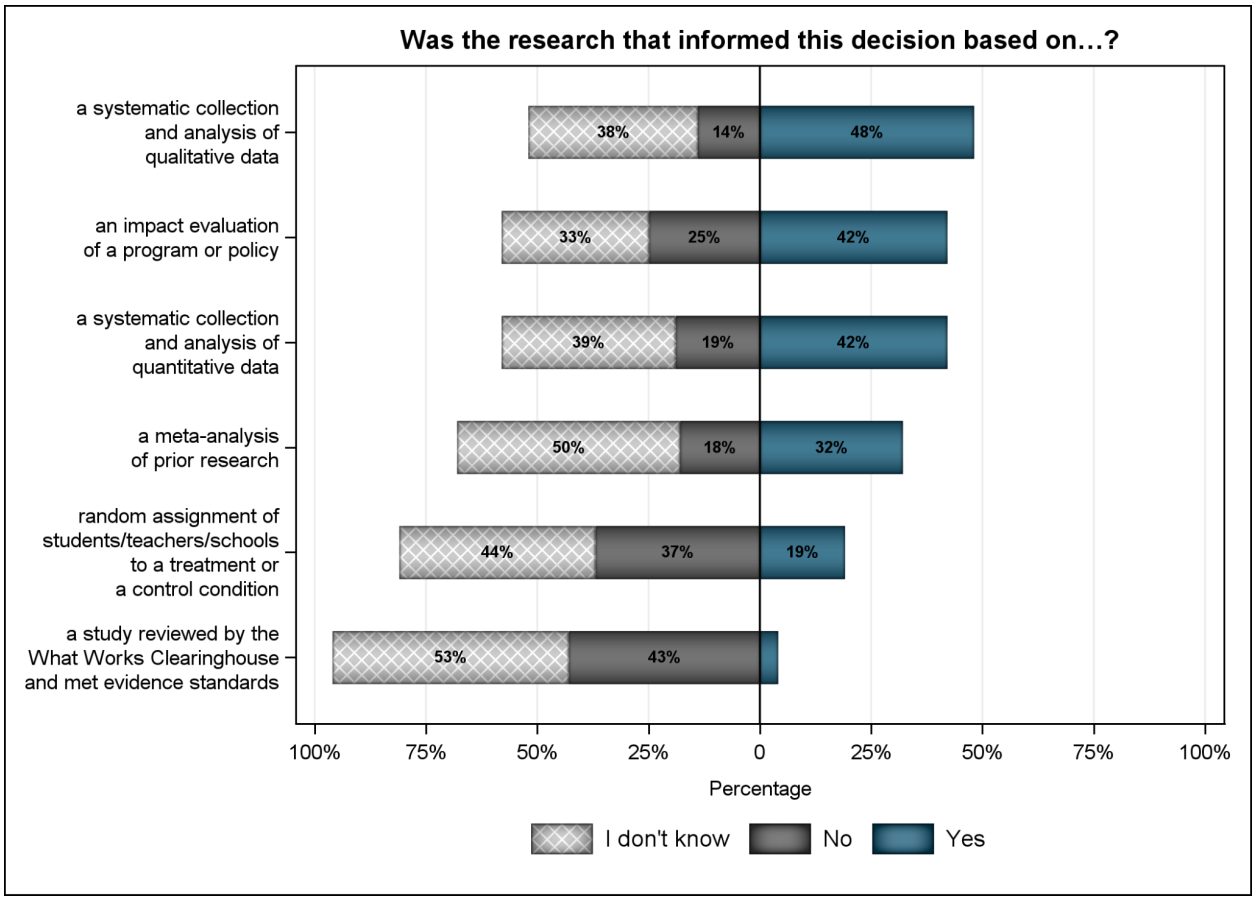


Note. Item response N = 289
Includes only individuals whose Path-A responses were validated as influenced by external research.

Similarly, respondents who reported about the causality of evidence that informed individual decisions (see Figure 15) tended to believe that evidence was based on systematic collection of

qualitative or quantitative data. These respondents were also less likely to perceive that the evidence was based on research that involved random assignment of participants to treatment or control conditions.

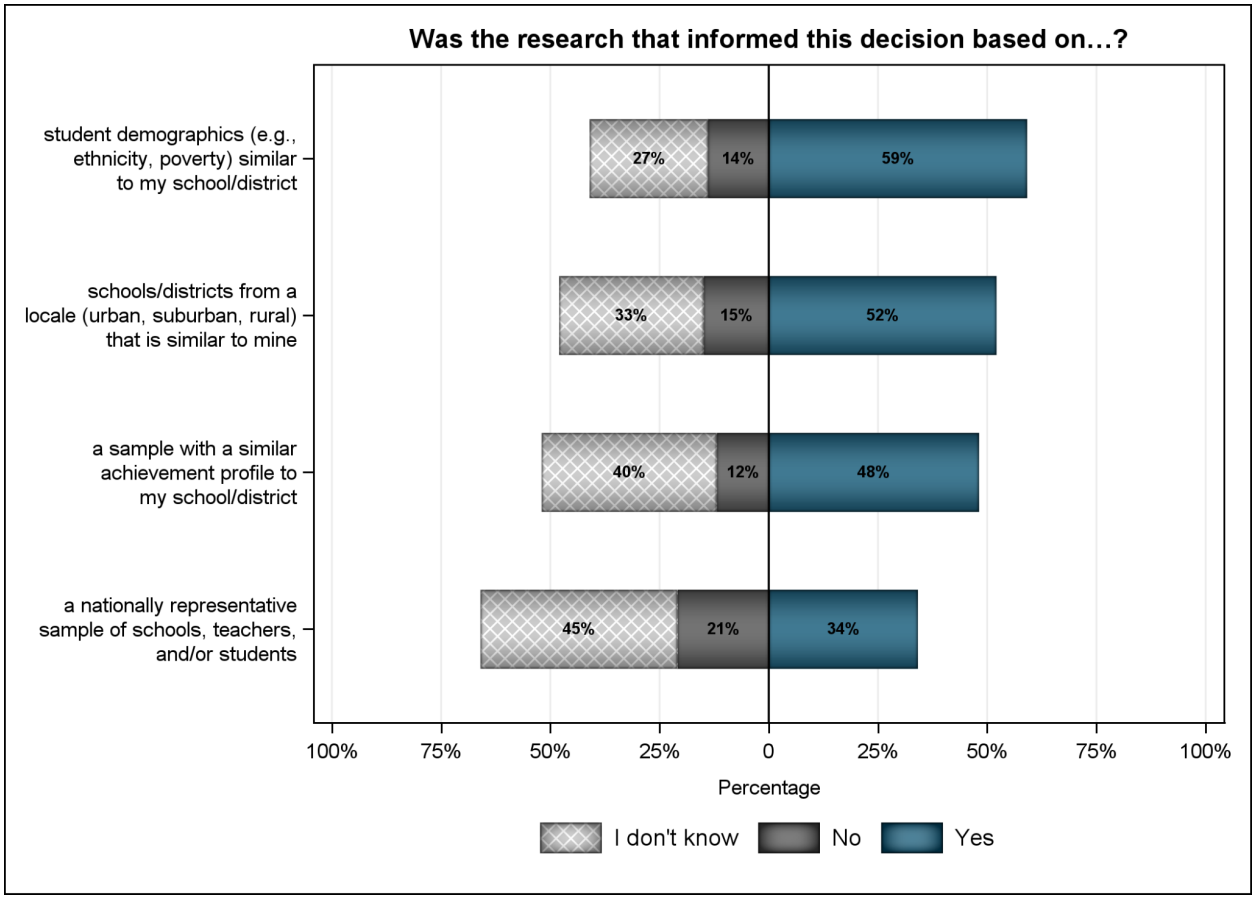
Figure 15. ***Respondents' Perceptions of Causality of Evidence for Individual Decisions***



Note. Item response N = 390
Includes only individuals whose Path-B responses were validated as influenced by external research.

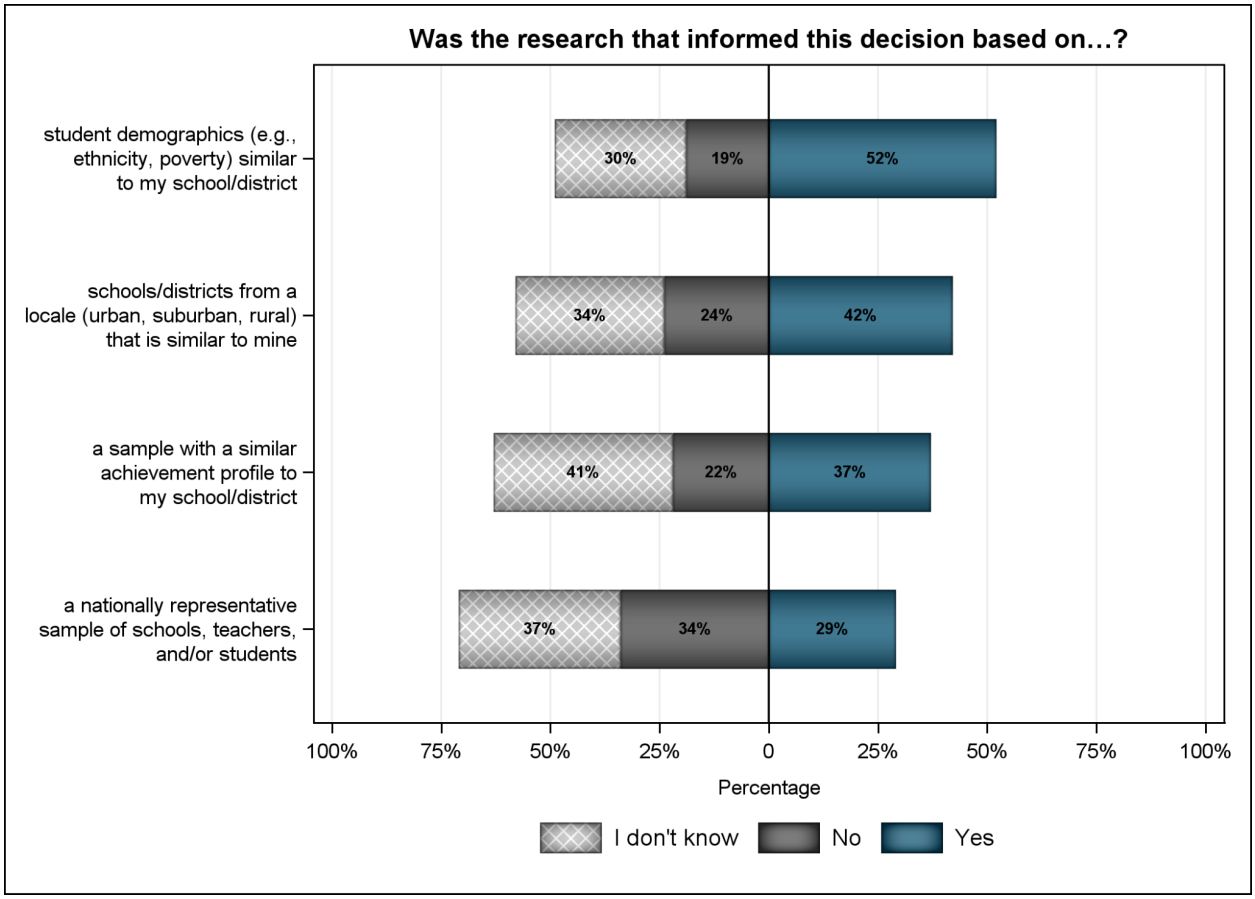
With regards to generalizability, survey respondents tended to perceive that both organizational and personal decisions were based on samples of students and schools that had characteristics that were similar to students in their own institutions. Figures 16 and 17 show detailed results.

Figure 16. ***Respondents' Perceptions of Generalizability of Evidence for Organizational Decisions***



Note. Item response N = 289
Includes only individuals whose Path-A responses were validated as influenced by external research.

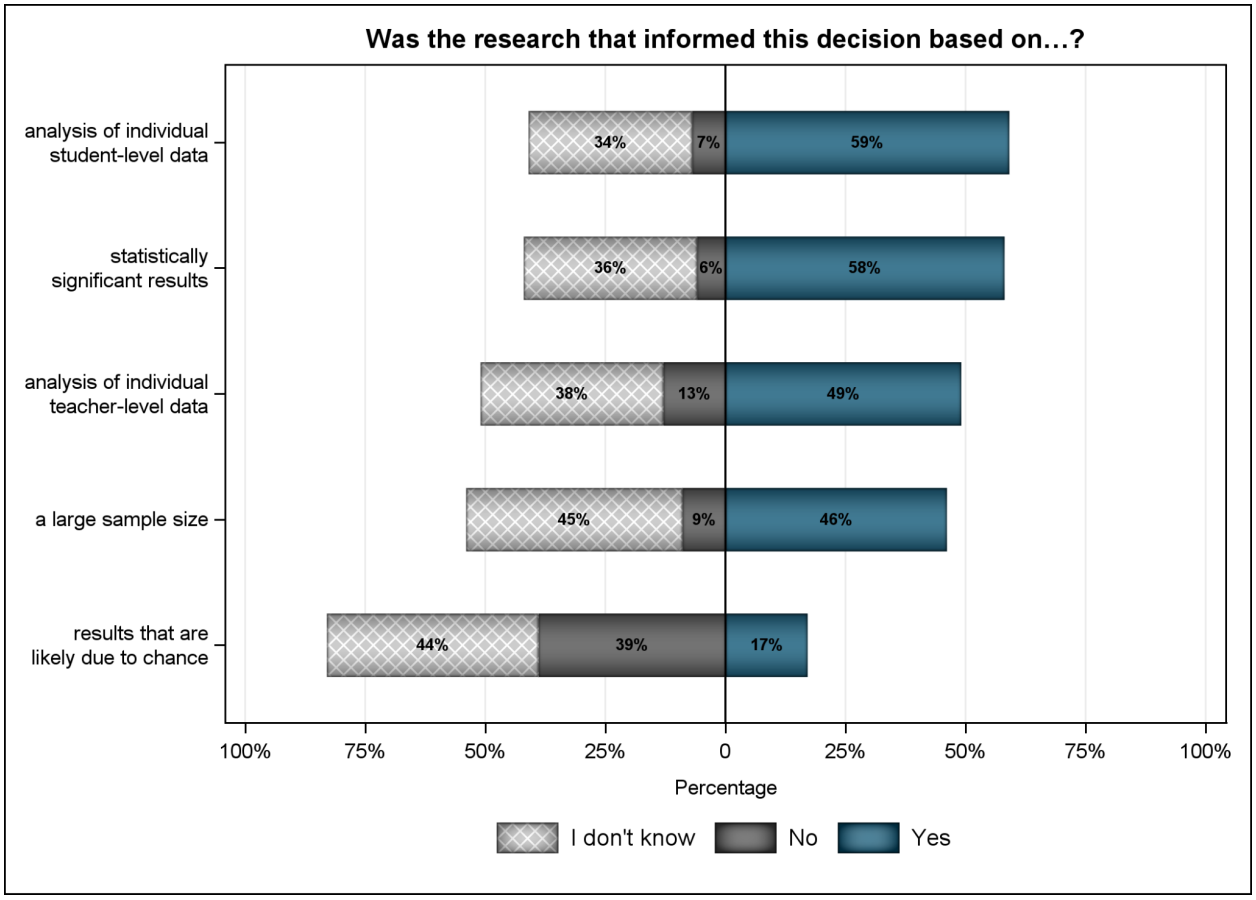
Figure 17. ***Respondents' Perceptions of Generalizability of Evidence for Individual Decisions***



Note. Item response N = 390
Includes only individuals whose Path-B responses were validated as influenced by external research.

Regarding the precision of the evidence used, survey respondents tended to perceive that organizational decisions were based on statistically significant results from the analysis of individual student-level and/or teacher-level data (see Figure 18). Very few respondents perceive that organizational decisions were based on results that are likely due to chance, although 44% are unsure.

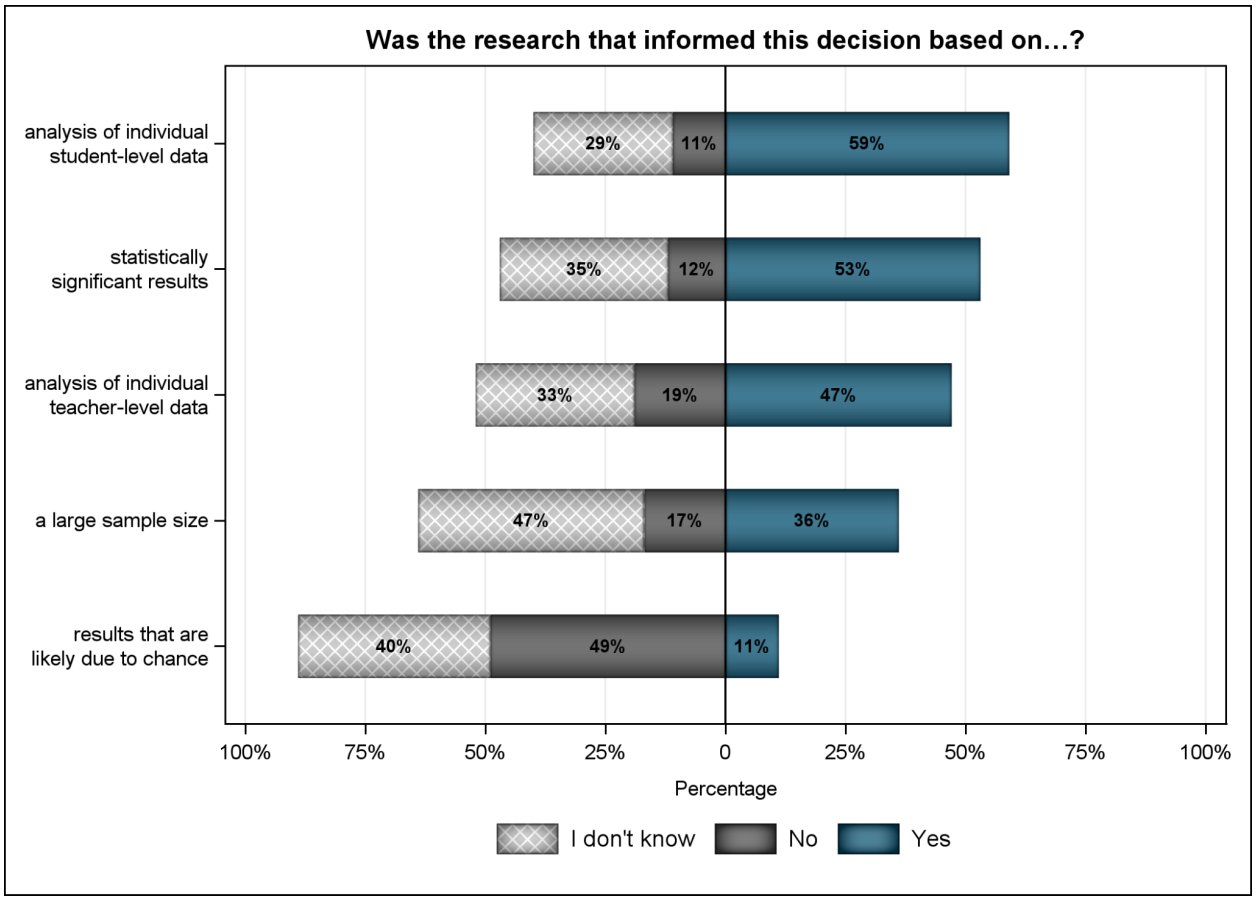
Figure 18. ***Respondents' Perceptions of Precision of Evidence for Organizational Decisions***



Note. Item response N = 289
Includes only individuals whose Path-A responses were validated as influenced by external research.

Similarly, survey respondents tended to perceive that individual decisions were based on statistically significant results from the analysis of individual student-level and/or teacher-level data (see Figure 19). Compared to organizational decisions, even fewer respondents perceive that individual decisions were based on results that are likely due to chance, although 40% are unsure.

Figure 19. ***Respondents' Perceptions of Precision of Evidence for Individual Decisions***



Note. Item response N = 390
Includes only individuals whose Path-B responses were validated as influenced by external research.

Search

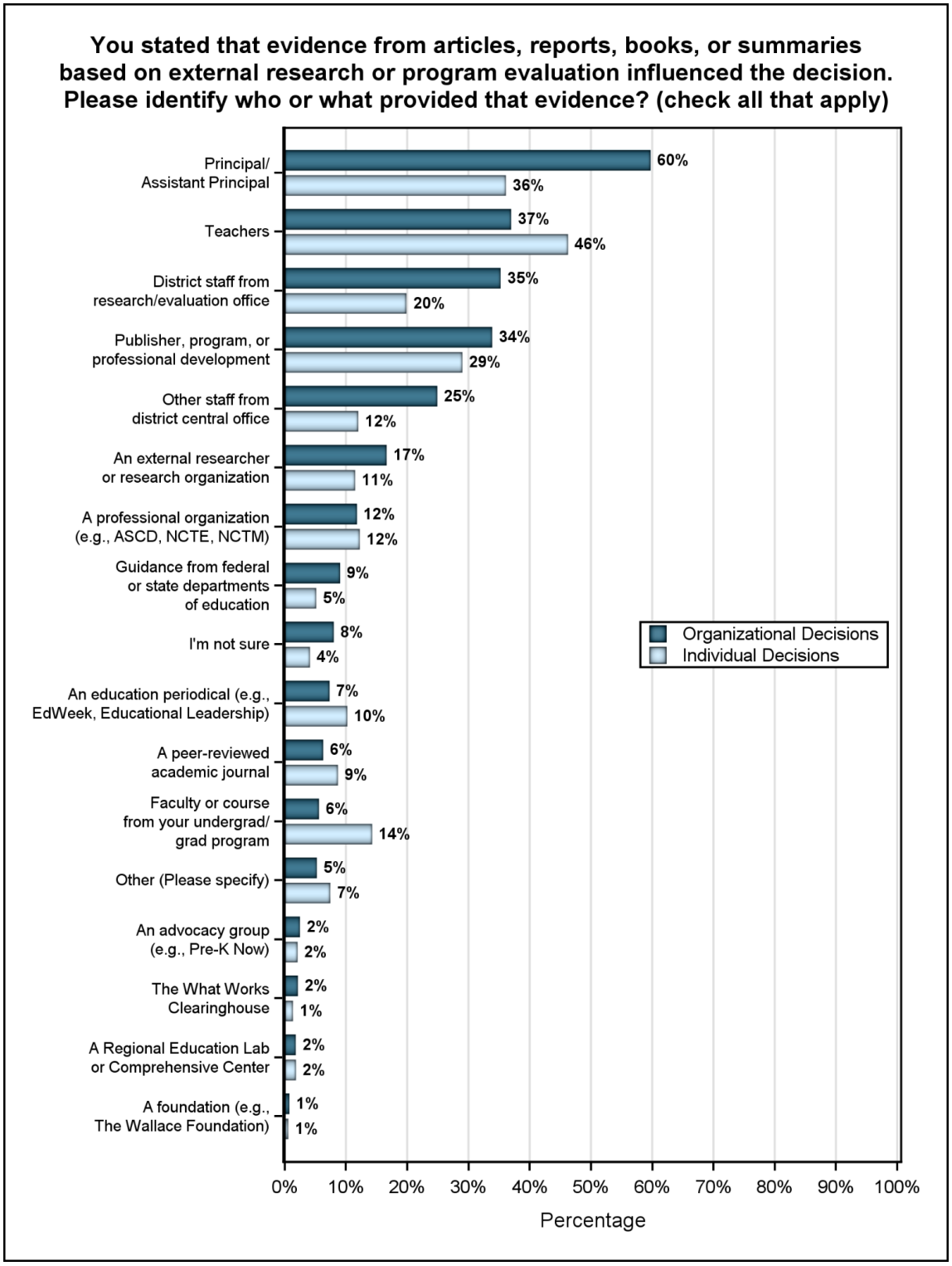
After indicating the types of evidence that were used to inform the cited decision, survey respondents reported about the nature of the search for that evidence. Respondents completed items that asked how information was found, what costs were associated with acquiring the information, and if the evidence was research-based, what specific source provided the information.

Source of Evidence (Research Evidence only)

Participants who reported that evidence from research products was used were asked to answer questions about the source of that evidence. For organizational decisions, 60% of respondents reported that their principal or assistant principal provided that information.[1] Thirty-seven percent reported that a teacher provided that evidence. Four percent of respondents indicated that the evidence came from either the What Works Clearinghouse (WWC) or a Regional Education Laboratory (REL). Reported sources of research evidence were similar for individual decisions. The primary source of evidence cited by respondents were teachers (46%), followed by principals/assistant principals (36%). Three percent of respondents indicated that evidence from a REL or the WWC influenced their individual decisions. Full results for organizational decisions and individual decisions can be found in Figure 20.

[1] Note that of the 173 respondents who indicated that the principal supplied the information, 94% reported that they were mostly or very familiar with the information used to make the decision. As with the evidence validity filters, the prevalence of "I don't know" responses in Figures 16 through 19 suggest that *Depth of Use* is more shallow because respondents familiar with a decision could not describe details of the research that was used.

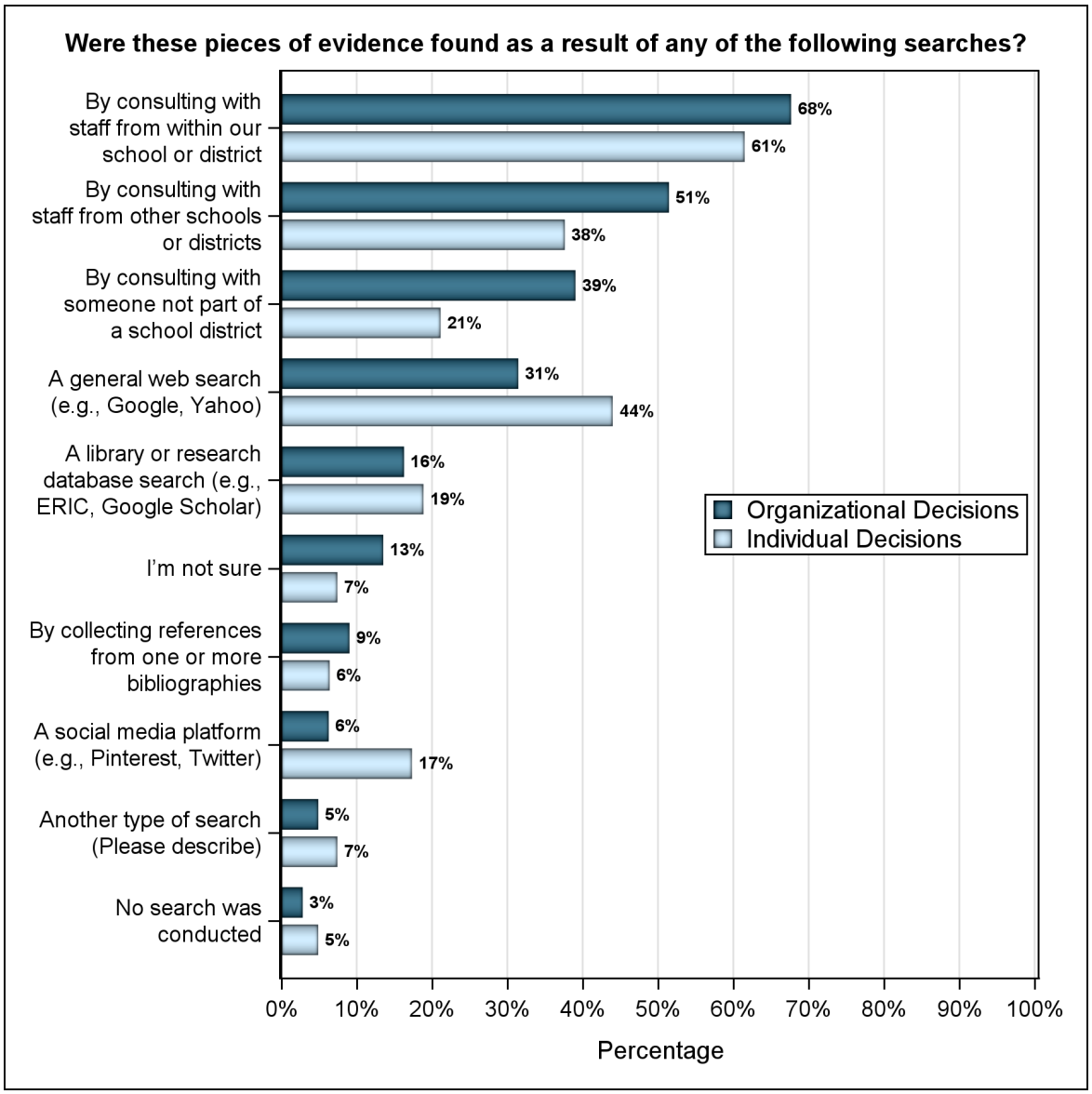
Figure 20. ***Source of Evidence, Organizational Decisions and Individual Decisions***



Note. Item response N = 290 for organizational decisions; N = 394 for individual decisions
Includes only individuals whose responses were validated as influenced by external research.

SEE-S respondents also reported on the extent of their search for evidence. Results are shown in Figure 21. Search was primarily constrained to personal networks, with respondents overwhelmingly relying on staff from their own or other schools and districts. For organizational decisions, searches using web search engines (e.g., Google) and research databases such as ERIC were much less frequent. Conversely, for personal practice decisions, a general web search was the second most frequently cited search.

Figure 21. ***Extent of Search for Evidence***



Note. Item response N = 290 for organizational decisions; N = 394 for individual decisions
Includes only individuals whose responses were validated as influenced by external research.

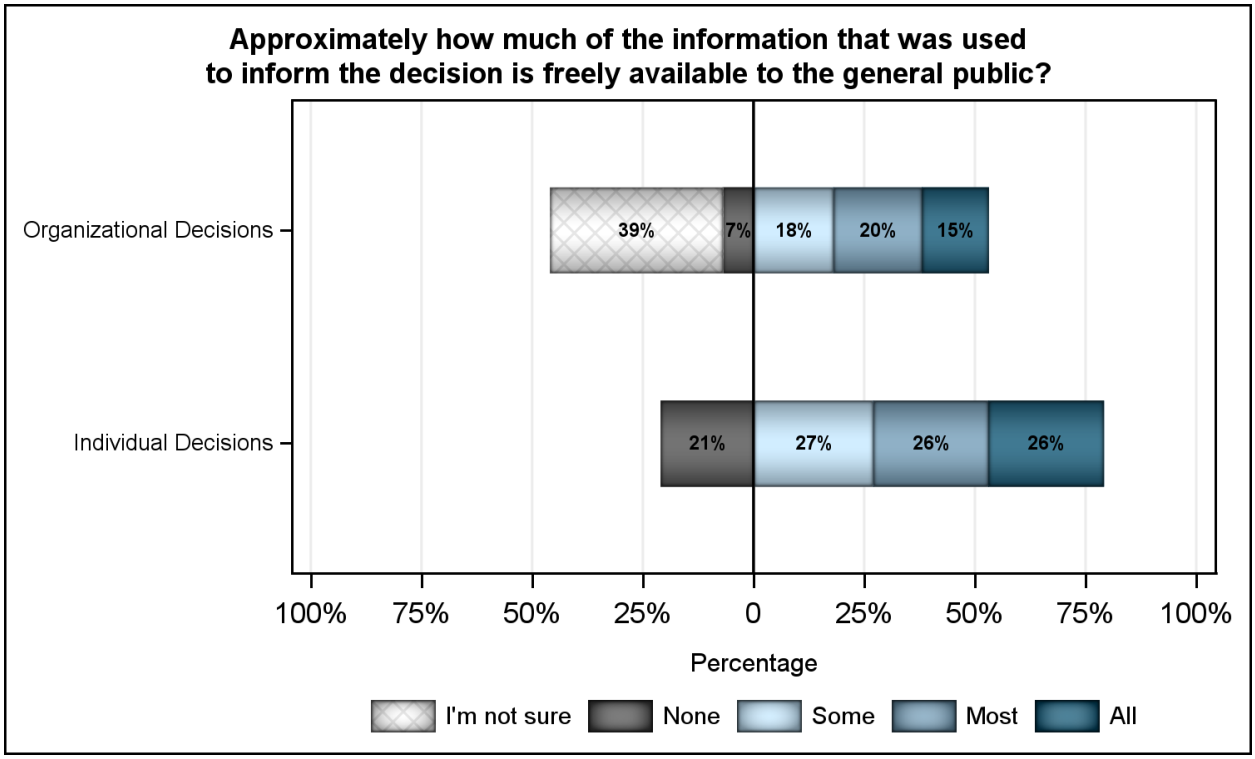
Search: Direct to Research

We also derived an indicator of whether a direct search for research was conducted. This binary indicator, "Search: Direct to Research," was coded 1 if the respondent reported that a search was conducted via an active search of a "library or research database search (e.g., ERIC, Google Scholar)," or "by collecting references from one or more bibliographies," or via an active (i.e., looking for research) or passive search (i.e., being provided with research) via "an external researcher or research organization," "a Regional Education Lab or Comprehensive Center," "the What Works Clearinghouse," "a peer-reviewed academic journal," or "a faculty person or a course from your undergrad/grad program." If none of these active or passive search mechanisms were used, the "Search: Direct to Research" indicator was coded zero.

Search Cost

Participants answered several questions about the monetary cost of acquiring the evidence used to inform the decision (i.e., all evidence, not just research evidence), and the time spent to collect the information and make the decision. Thirty-five percent of participants who referenced an organizational decision indicated that 'most' or 'all' of the information was freely available to the public. Thirty-nine percent were unsure about the public nature of the information. Fifty-two percent of participants who referenced an individual decision indicated that 'most' or 'all' of the information was freely available to the public. Twenty-one percent were unsure about the public nature of the information.

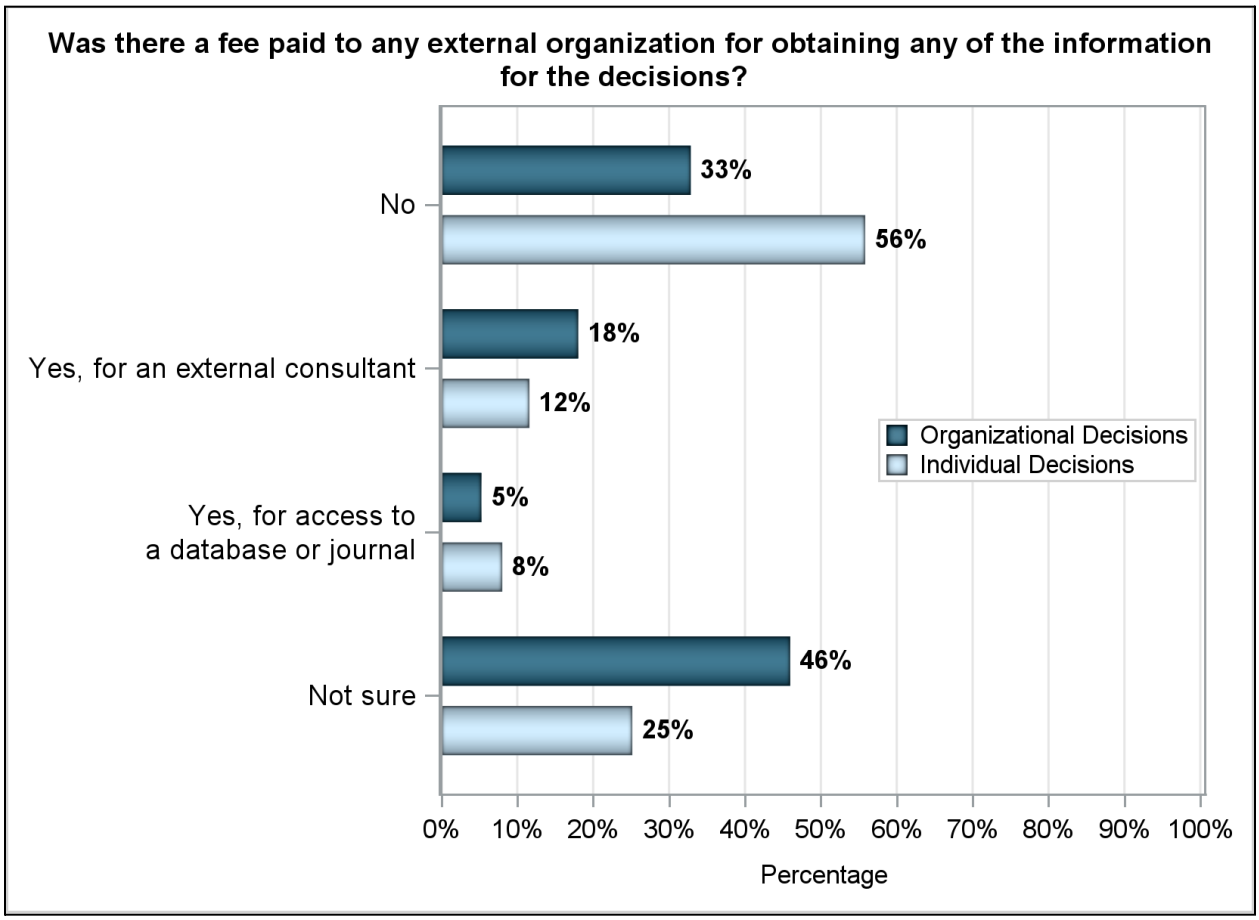
Figure 22. ***Amount of Information That Was Freely Available to the General Public***



Note. Item response N = 1,317 for organizational decisions; N = 2,573 for individual decisions

As to fees paid for obtaining information for organizational decisions, 23% of respondents indicated a fee was paid for the information, but nearly half of the respondents did not know whether a fee was paid. Twenty percent of the respondents who referenced an individual decision indicated that a fee was paid for the information, 25% were unsure whether any fee was paid.

Figure 23. ***Fees Paid for Information for Organizational and Individual Decisions***

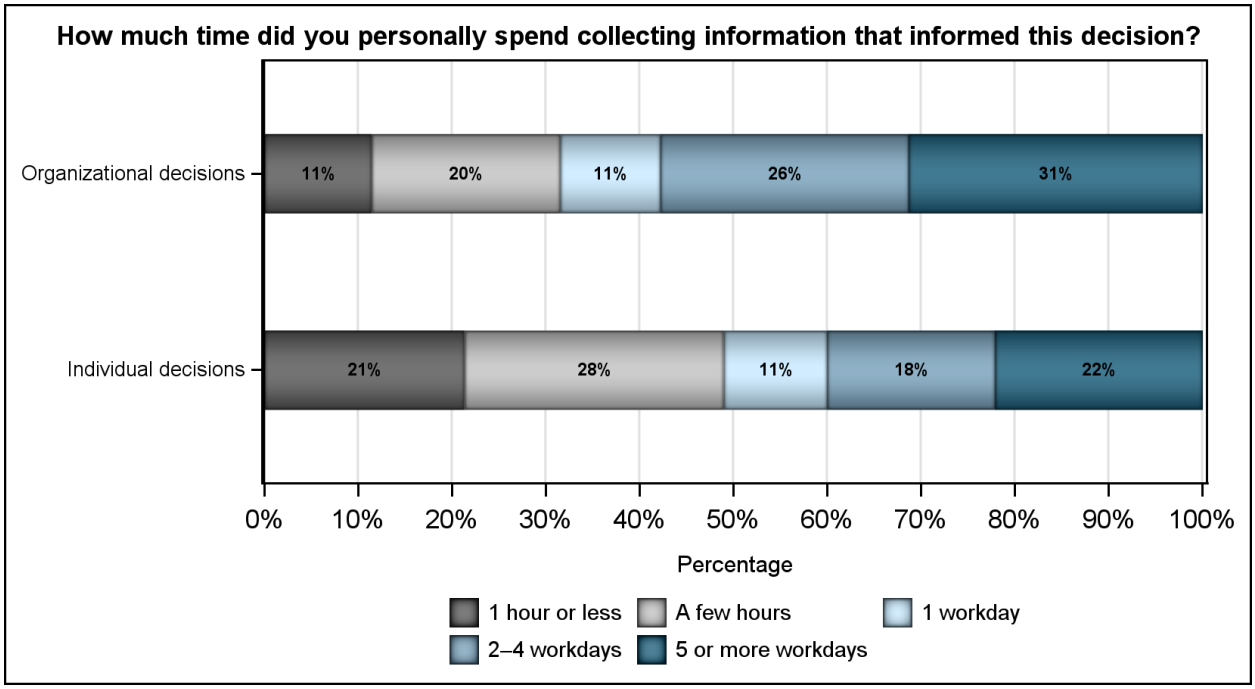


Note. Item response N = 290 for organizational decisions; N = 391 for individual decisions
Includes only individuals whose responses were validated as influenced by external research.

Fifty-seven percent of participants who were involved in an organizational decision reported that they personally spent two or more workdays to collect information that informed the decision. Thirty-seven percent reported that it took fewer than six months to complete the decision-making process. However, 39% indicated that they were unsure of the time the entire process took. Fewer than 40% of participants who referenced personal decisions indicated that it took them at least two workdays to collect information that informed the decision. Approximately 20% reported spending one hour or less. Forty-seven percent of respondents

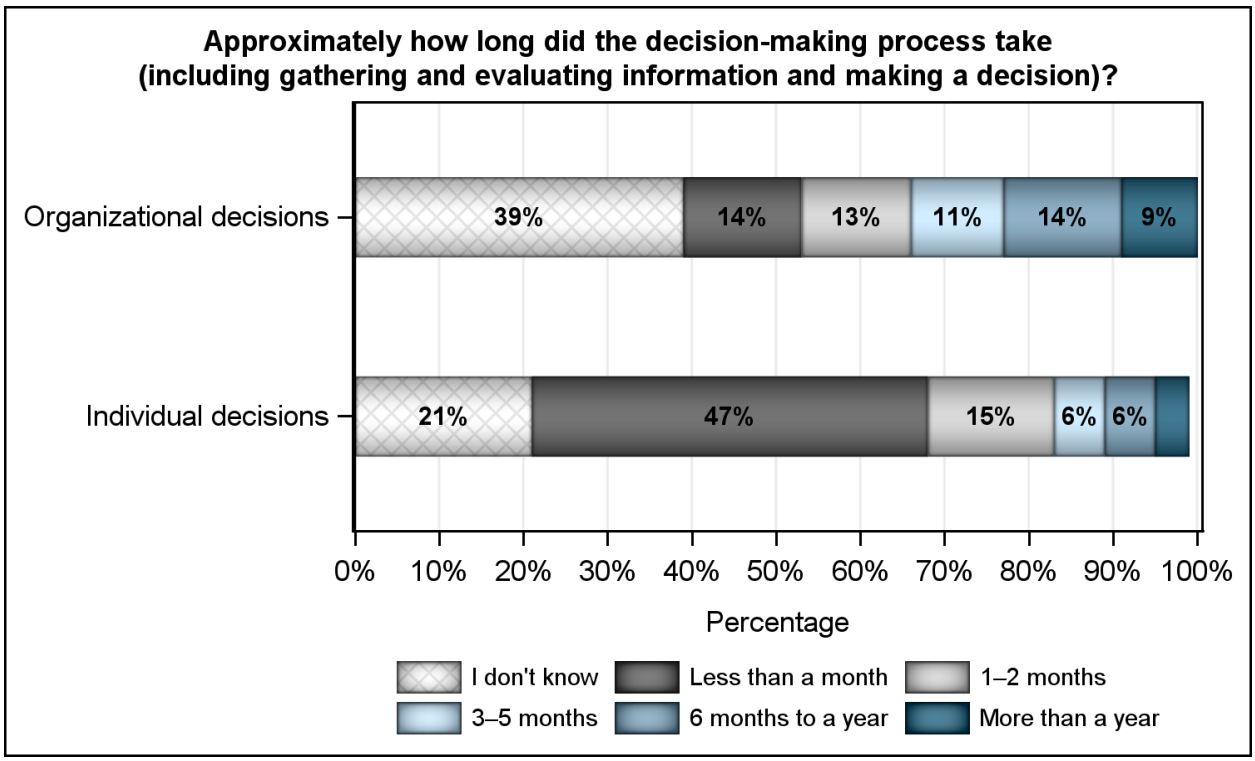
who cited a personal decision spent less than one month to move through the decision-making process.

Figure 24. ***Personal Time Spent Collecting Information That Informed the Decision***



Note. Item response N = 393 for organizational decisions; N = 2,576 for individual decisions
Includes only individuals who were personally "involved in gathering evidence".

Figure 25. ***Total Time Spent on the Decision-Making Process***

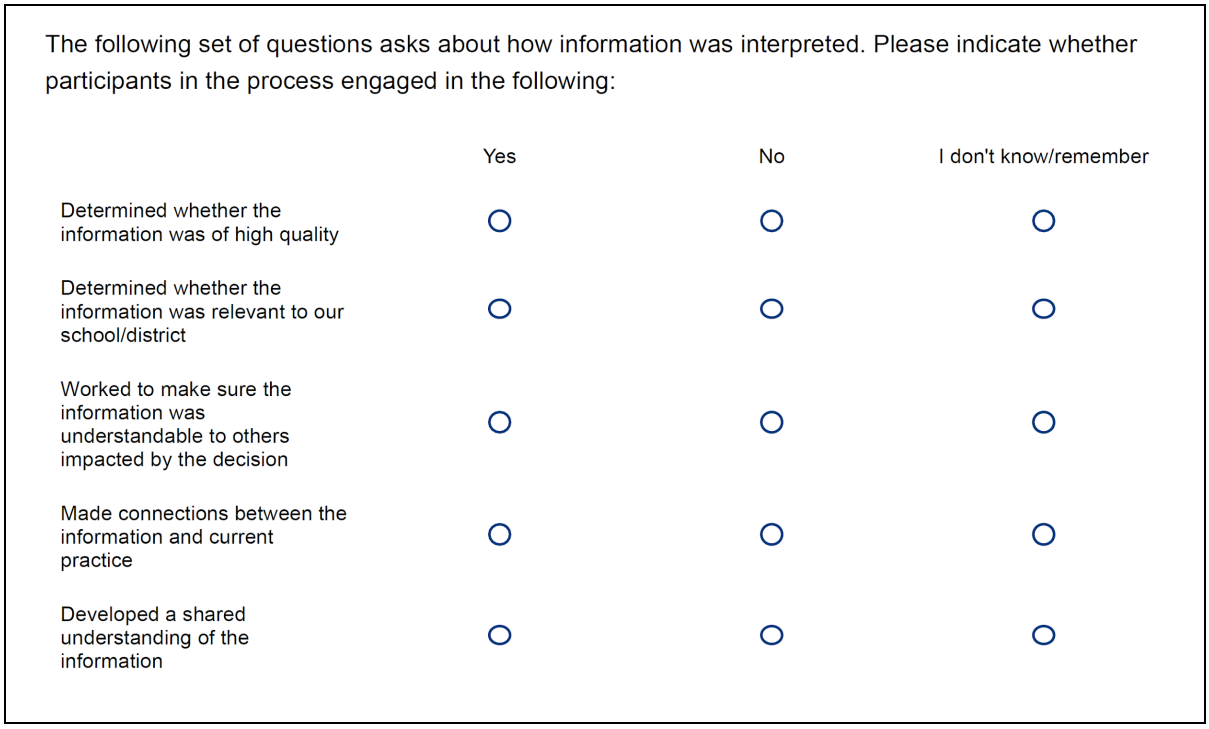


Note. Item response N = 1,317 for organizational decisions; N = 2,610 for individual decisions

Interpretation

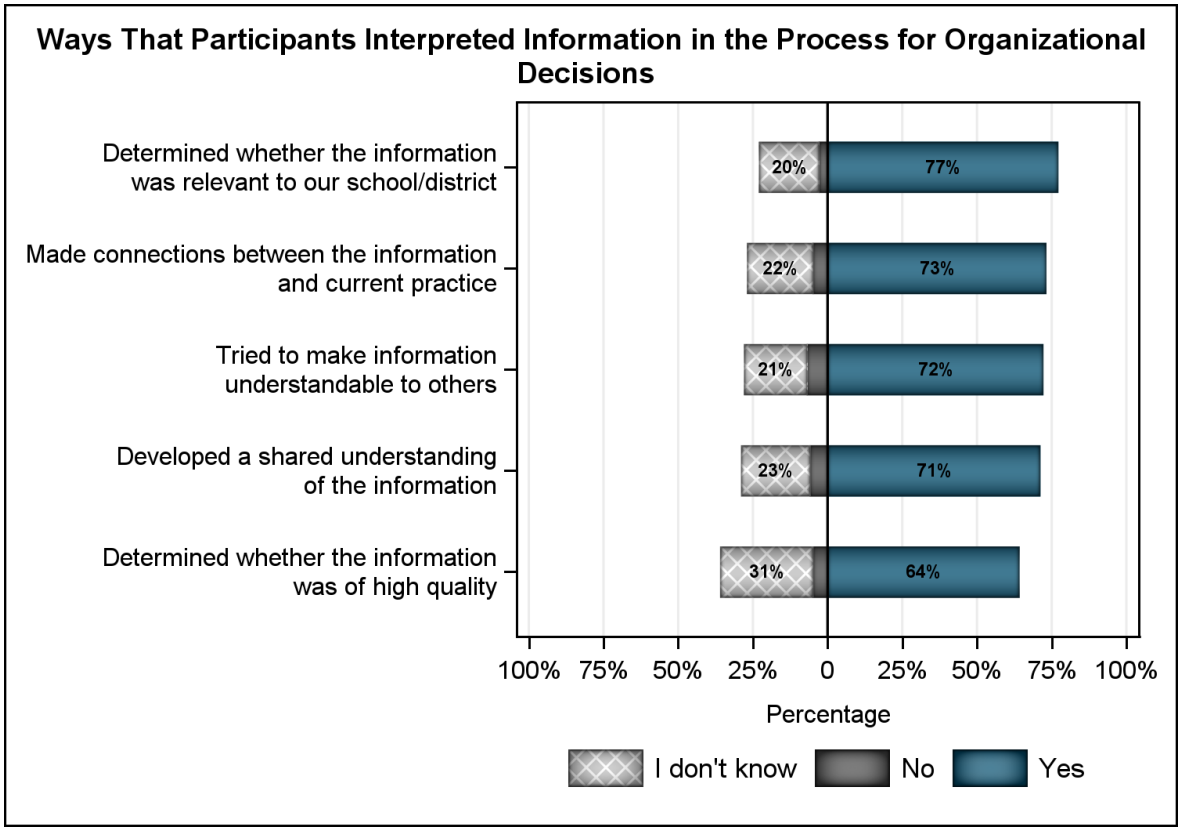
An overall score for *Interpretation* within the *Depth of Use* framework was calculated as a sum of the number of checked items within the following interpretation item. For both organizational decisions and individual decisions, scores ranged from 0 to 5. The mean for organizational decisions (with standard deviation in parentheses) was 3.57 (1.95). For individual decisions, the mean was 3.34 (1.76).

Figure 26. ***SEE-S Interpretation Item for Depth of Use***



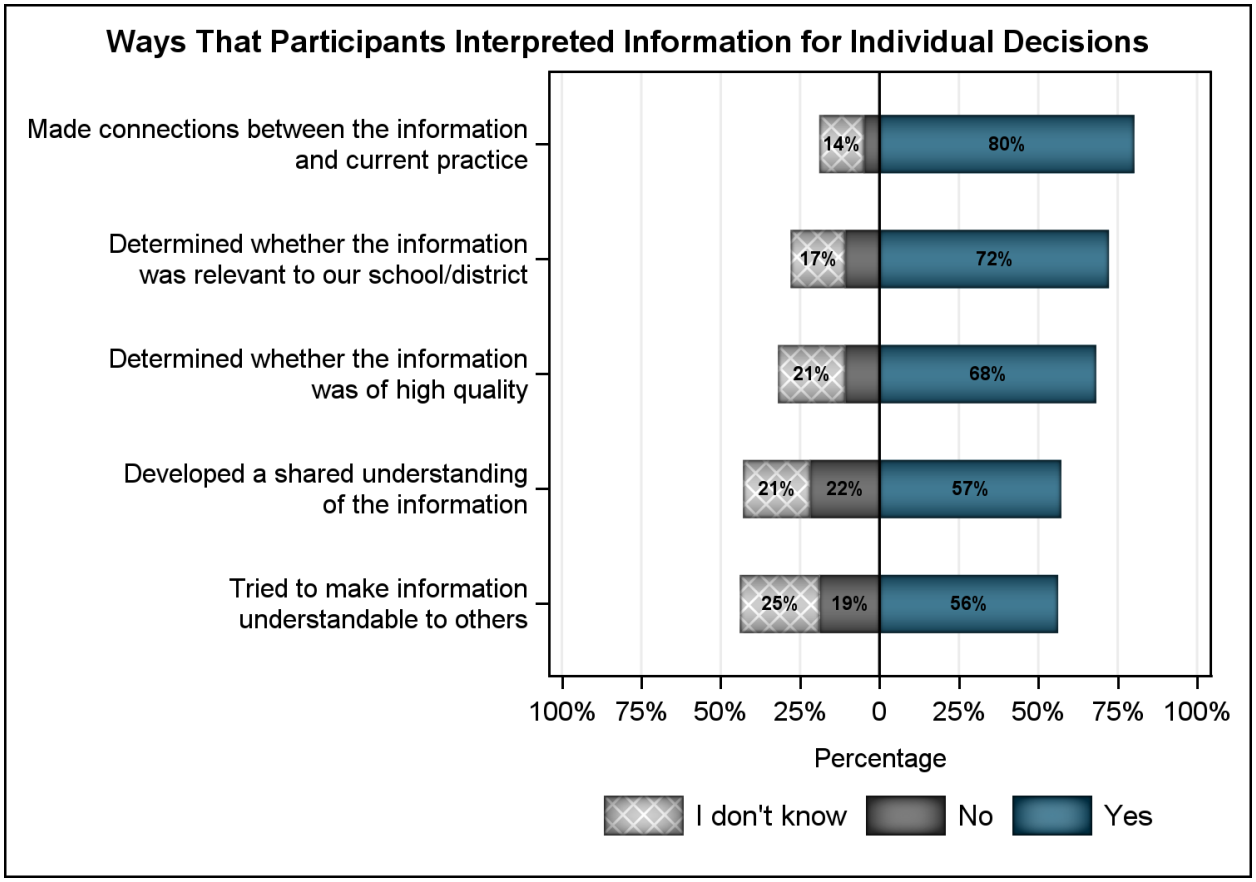
Respondents who referenced an organizational decision made it a priority to determine if the information was relevant to their local context, and were less often engaged in determining whether the information was of high quality (see Figure 27). For both organizational and individual decisions, participants more often determined whether the information was relevant to their situational context and tried to make connections between information and practice. Respondents who referenced an individual decision were more engaged in determining whether the information was of high quality, and were less often concerned with developing shared understandings.

Figure 27. ***Ways That Participants Interpreted Information in the Process for Organizational Decisions***



Note. Item response N = 1,312

Figure 28. ***Ways That Participants Interpreted Information for Individual Decisions***



Note. Item response N = 2,599

Participation

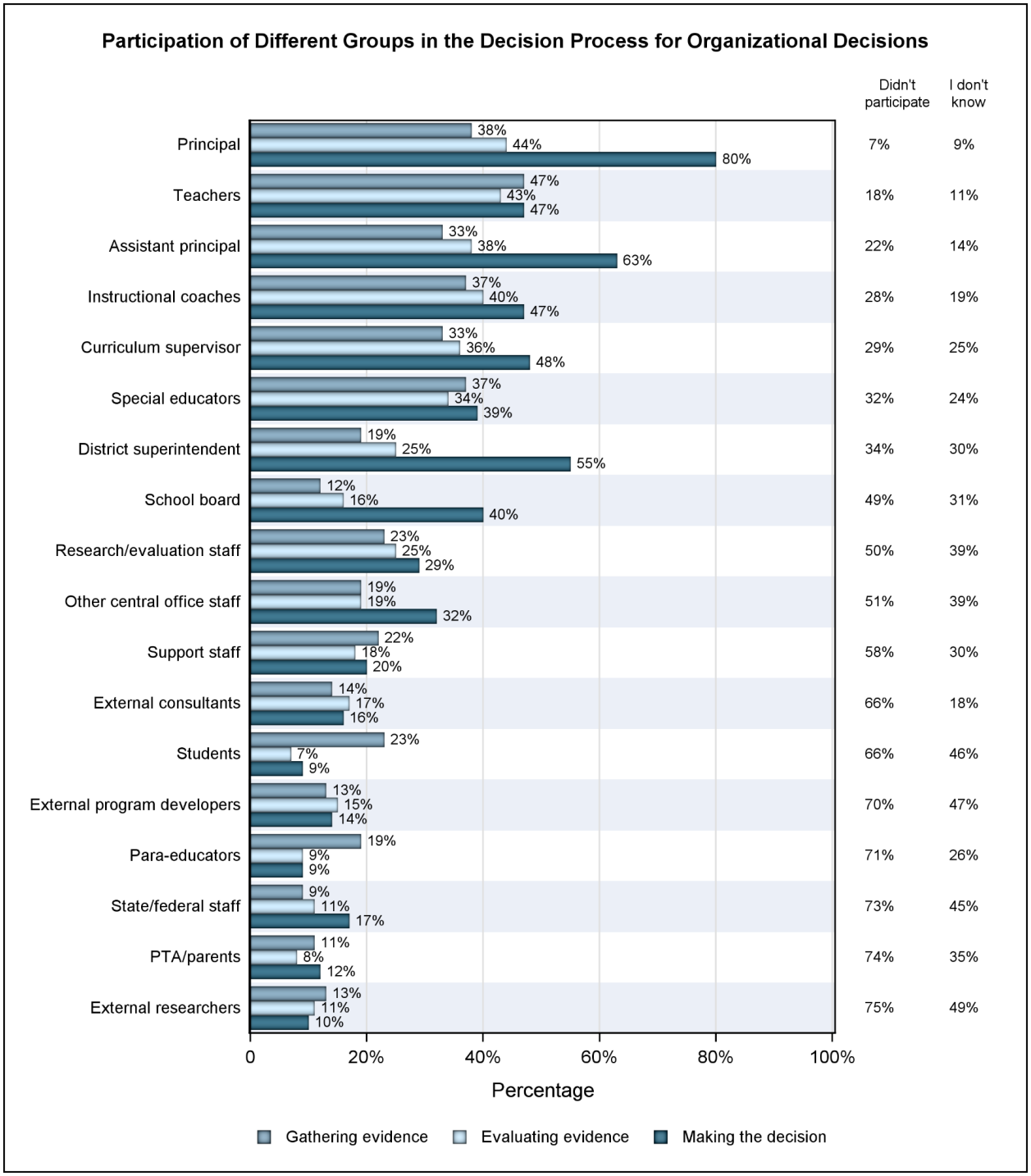
Figure 29 presents results for the *Participation* item in our *Depth of Use* framework regarding organizational decisions. Results show that school-based organizational decisions most often involve school administrators, teachers, and other educators at the school. The next group most often involved include the district superintendent and other district office staff. School and district administrators are most likely to be involved in making the decision. Other groups that are not employees of the school or district are less likely to be involved in collecting evidence, evaluating evidence, or making the decision.

An overall score for *Participation* within the *Depth of Use* framework was calculated as a count of the number of groups, among the following eighteen, that were reported by the respondent to be "involved in gathering and evaluating evidence to inform the decision," and/or "involved in making the decision."

1. School Board
2. Principal
3. Asst. Principal
4. Teachers
5. Students
6. Instructional Coaches
7. Special Educators
8. Para-educators/Teaching Assistants
9. Curriculum/Instructional Supervisor
10. Support Services Staff
11. District Superintendent
12. Research/Evaluation Staff
13. Other Central Office Staff
14. PTA / Parents
15. State/Federal Staff
16. External Program Developer(s)
17. External Researcher(s)
18. External Consultant(s)

Although we focus on organizational decisions for the development of *Depth of Use* scores, the resultant scores for the *Participation* metric for both organizational and individual decisions ranged from 0 to 18. For organizational decisions, 1285 responded to the *Participation* items and the mean (standard deviation) of the *Participation* metric was 6.97 (3.94). For individual decisions, a subset of respondents (n=1,376, or 52% of respondents to Path B) indicated that others were involved in the decision. For this subset of respondents in Path B, the mean (standard deviation) for the number of respondents with usable responses for the *Participation* metric (n=1,338) was 5.36 (3.39).

Figure 29. ***Participation of Different Groups in the Decision Process (Organizational Decisions)***



Note. Item response N ranges from 1,133 to 1,311

Frequency

An overall score for the *Frequency* measure within the *Depth of Use* framework was calculated as the expected proportion of decisions, within a school, for which external research, or local research, or local data analyses influenced the decision. For each analysis, the outcome variable was coded 1 if a respondent indicated that that type of evidence influenced the decision or 0 if it did not influence the decision. See the section entitled "Aggregate School-Level Scores for Depth of Research Use" for more details on this school-level measure. Note that this scale is not a direct measure of frequency as defined in our conceptual framework. Such a direct measure would require collection of longitudinal data.

Decision Stage

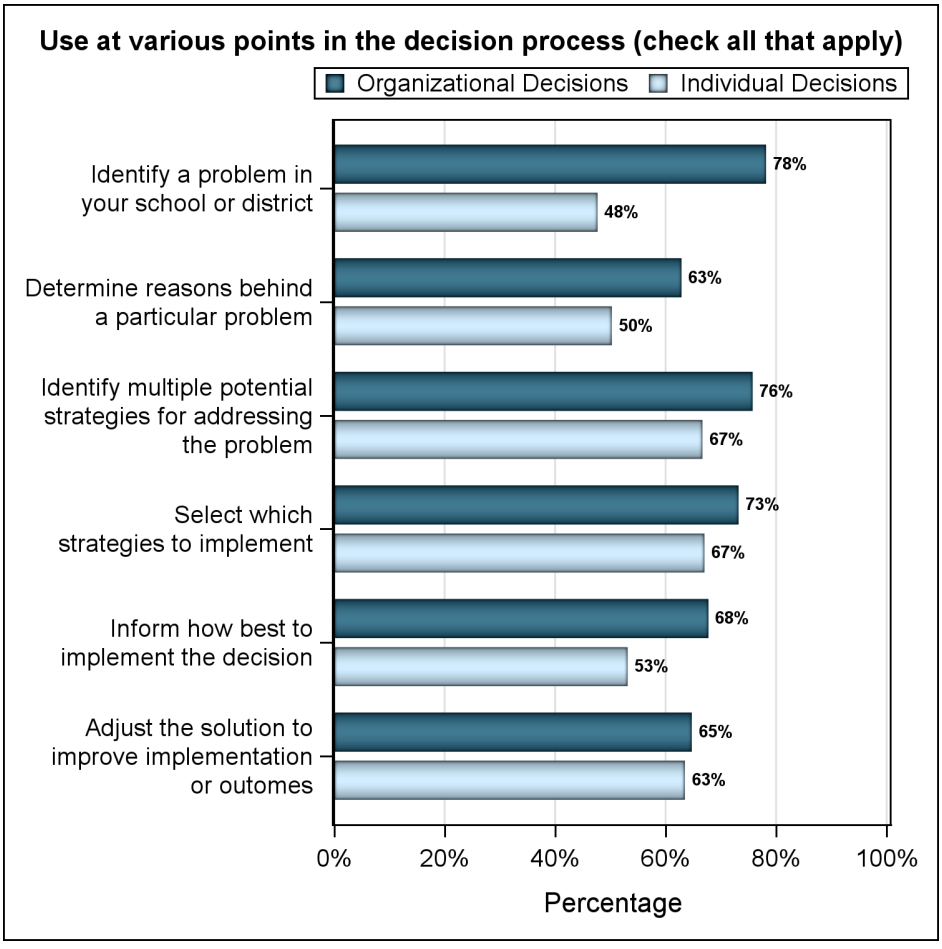
Research and other forms of evidence may be utilized at different points in the decision-making process. Participants were asked to indicate when the information they referenced earlier in the survey was used in the decision-making process. Participants were asked to select all options that applied. This set of items gives us insight into specific uses of evidence and allows for identification of instrumental, conceptual, strategic or symbolic uses of research. Instrumental uses of research refer to situations where research is used to inform a specific decision (Caplan, 1979; Rich, 1977). Symbolic use refers to attempts to meet normative expectations for engaging in evidence-based decisions, but often in inauthentic ways, for example with "the research says" (Weiss, 1979; Farley-Ripple, 2012). Conceptual use refers to gradual shifts in awareness and perspectives as new knowledge is incorporated into their thinking (Caplan, 1979; Rich, 1977). Research use is considered strategic when evidence is used to attain specific power or profit goals (Huberman, 1990). Figure 30 depicts responses to the decision stage items for both organizational and individual decisions.

Participants who indicated familiarity with an organizational decision responded that evidence was most frequently used to identify a problem in their school or district (78%), and to identify the range of potential strategies to address that problem (76%). Evidence use was least frequent

when determining the reasons behind the problem (63%), and when adjusting the solution to improve implementation or outcomes (65%).

Respondents who answered items about a specific individual practice decision indicated using evidence the most when identifying potential strategies to address the problem (67%) and selecting which strategies to implement (67%). Evidence was least used to identify the problem 48%). Results are shown in Figure 30.

Figure 30. ***School-Based Practitioners' Evidence Use in Different Stages of Decision-Making***



Note. N = 862 for organizational decisions; N = 1,633 for individual decisions

Although correlations between the six stages were relatively low, this is not surprising given that this item constitutes more of a checklist rather than a latent trait. As such, a composite score

that is simply a count is still justified. The overall score for *Decision Stages* within the *Depth of Use* framework was calculated as a count of the number of decision stages for which the evidence was used. Resultant scores ranged from 0 to 6 with a mean of 2.69 and a standard deviation of 2.49.

Aggregate School-Level Scores for Depth of Research Use

For each school, we calculated an overall score of *Depth of Use* of Research, as well as subscale scores on eight metrics: (1) *Influence of External Research*, (2) *Influence of Local Research*, (3) *Influence of Formal Analyses of Local Data*, (4) *Search: Direct to Research*, (5) *Interpretation*, (6) *Participation*, (7) *Frequency*, and (8) *Decision Stage*. See previous section for each of these eight indicators for details on the construction of each indicator used to create school-level scores. This section details the statistical models used to estimate the school-level scores.

Because survey data were collected from multiple respondents in each school, with different sample sizes in each school, and because respondents from the same school could report on the same organizational decision, multilevel models were used to produce aggregate school-level scores on each *Depth of Use* metric. For continuous (or semi-continuous) metrics based on responses to items measured with a rating scale or checklist (i.e., *Influence of External Research*, *Influence of Local Research*, *Influence of Formal Analyses of Local Data*, *Interpretation*, *Participation*, and *Decision Stage*), a three-level Hierarchical Linear Model (HLM) was estimated with individuals at level 1, decisions at level 2, and schools at level 3. Notably, only about 20% of organizational decisions within a school included data from more than one respondent, so the models were formulated as partially-nested HLMs, with level-2 random effects present only for those decisions with more than one respondent. The mathematical form of the HLM is as follows.

$$Y_{ids} = \alpha_s + \sum_{d=1}^{D} \gamma_d + \varepsilon_{ids}$$

Where Y_{ids} is the value of the outcome variable (e.g., influence of external research), α_s is a random effect for school *s* from a $N(0,\tau^2)$ distribution, γ_d is a random effect for decision *d* from

a $N(0,\varphi^2)$ distribution (where D equals the total number of decisions with more than one respondent), and ε_{ids} is a residual for respondent *i*, regarding decision *d*, for school *s* from a $N(0,\sigma^2)$ distribution. Note that there are no fixed coefficients in this model (e.g., no overall intercept, β_0), so the estimates for α_s are empirical-Bayes estimates of school-level means, with precision-weighting given within-school sample sizes and between-within school variances. Therefore, each school-level score is an aggregation across all decisions in a school described by all respondents who reported being mostly or very familiar with the decision (i.e., it is a measure of depth as an organizational measure), and higher scores on the external research, local research, and local data analysis subdimensions require not only that research was used, but that a school have several respondents with basic familiarity of the decision made and the research evidence used (e.g., they can name the author, title, or provide other information that points to the research used or describes the local analyses performed). This is a critical part of our definition of deep use of research.

For responses to binary metrics (i.e., *Search*: *Direct to Research*, *Frequency*), a multilevel logistic regression was conducted using a three-level Hierarchical Generalized Linear Model (HGLM). Given the discrete dependent variable and very small sample sizes within decisions, estimating a random effect for decisions was not possible. Therefore, an alternative approach using observation weights was used to account for multiple respondents reporting on the same decision. The weight for each respondent was calculated as 1 divided by the number of respondents who reported on that decision. The mathematical form of the logistic HGLM is as follows.

$$ln\left[\frac{\pi}{1-\pi}\right] = \alpha_s \text{ where } Y_{is} \sim Bernoulli(\pi)$$

Where Y_{is} is the binary value of the outcome variable (e.g., direct search for research: yes=1, no=0) assumed to follow a Bernoulli distribution with probability *Y*=1 equal to π, and α_s is a random effect for school *s* from a $N(0,\tau^2)$ distribution. Note that there is no respondent-level residual because this is a logistic model, and the residual term is implied by the logistic link function and Bernoulli distribution. Also, as with the HLM, there are no fixed coefficients in this

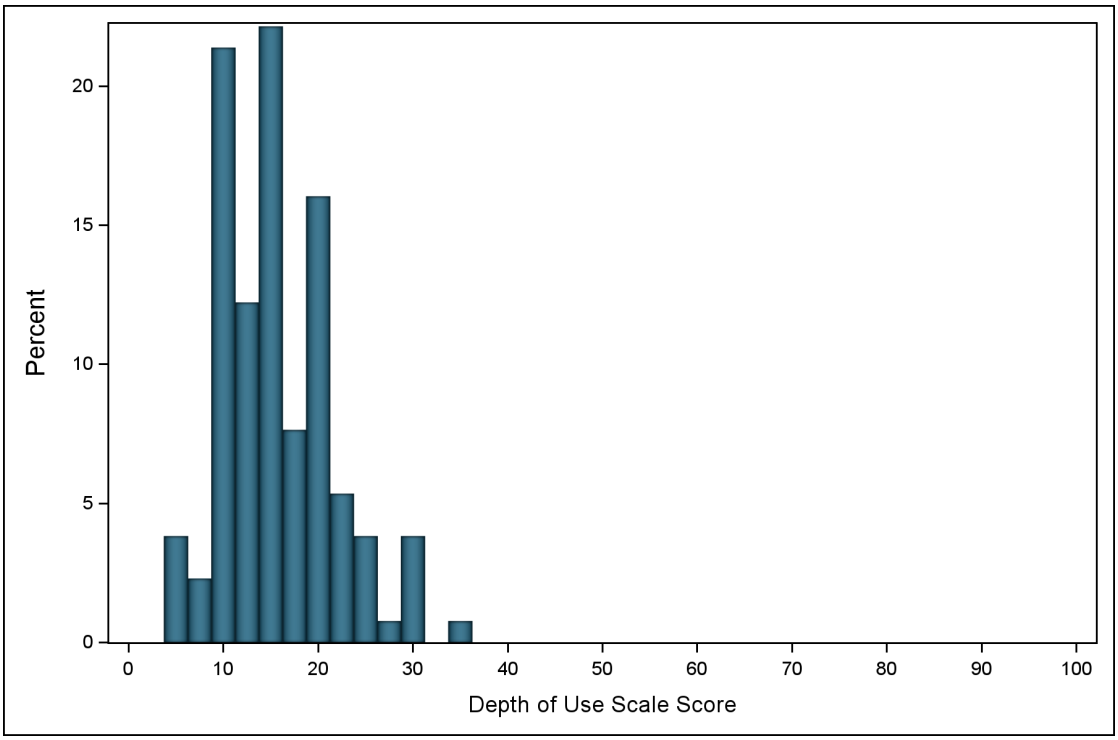
model (e.g., no overall intercept, β_0), so the estimates for α_s are empirical-Bayes estimates of school-level log-odds, with precision-weighting given within-school sample sizes and between-within school variances. Exponentiating α_s produces predicted probability values for each school, which serve as school-level scores for binary metrics. For *Search: Direct to Research*, the scores can be interpreted as estimates of the proportion of decisions that involve active or passive search for research involving research databases, bibliographies, research organizations, academic journals, or university faculty or courses. For *Frequency*, these scores can be interpreted as estimates of the proportion of school organizational decisions that our data and analyses could confirm as influenced by external or local research or local data analyses.

Depth of Use (*DOU*) of Research Total Scores for Schools

Overall *Depth of Use* (*DOU*) scores for each school, and associated *DOU* plots, were produced using the eight school-level scores estimated above. An eight-spoke radar chart was produced for each school with each spoke corresponding to one of the eight *Depth* metrics, and scaled to range from the minimum score (i.e., zero) through the maximum score for that metric (e.g., 1 for *Influence of Evidence*, *Search*, and *Frequency* metrics; 5 for *Interpretation*; 18 for *Participation*; and, 6 for *Decision Stage*). Each radar plot forms eight triangles, with the area of each triangle determined by the position of the school's scores on adjacent metrics within the total ranges of those spokes. Overall *Depth of Use* scores for each school are calculated as the percentage of the total possible area in the radar chart that is covered by the eight triangles for that school. Although the total possible range of *DOU* scores runs from 0 to 100, the resultant *DOU* scores for the sample of schools participating in the R4S field trial ranged from 5.0 to 36.0, with a mean of 15.6 and a standard deviation of 5.9. Figure 31 shows a histogram of school-level scale scores for the *DOU* measure, Figure 32 shows histograms for each of the eight subdimensions, and Table 6 presents descriptive statistics for the overall scores and each subscale score. Most dimensions of *Depth of Use* show considerable variability across schools, but the use of *Local Research* and *Local Data Analyses* is clearly more prevalent than use of *External Research*. Relatedly, direct searches for research appear to be rare. The histogram for

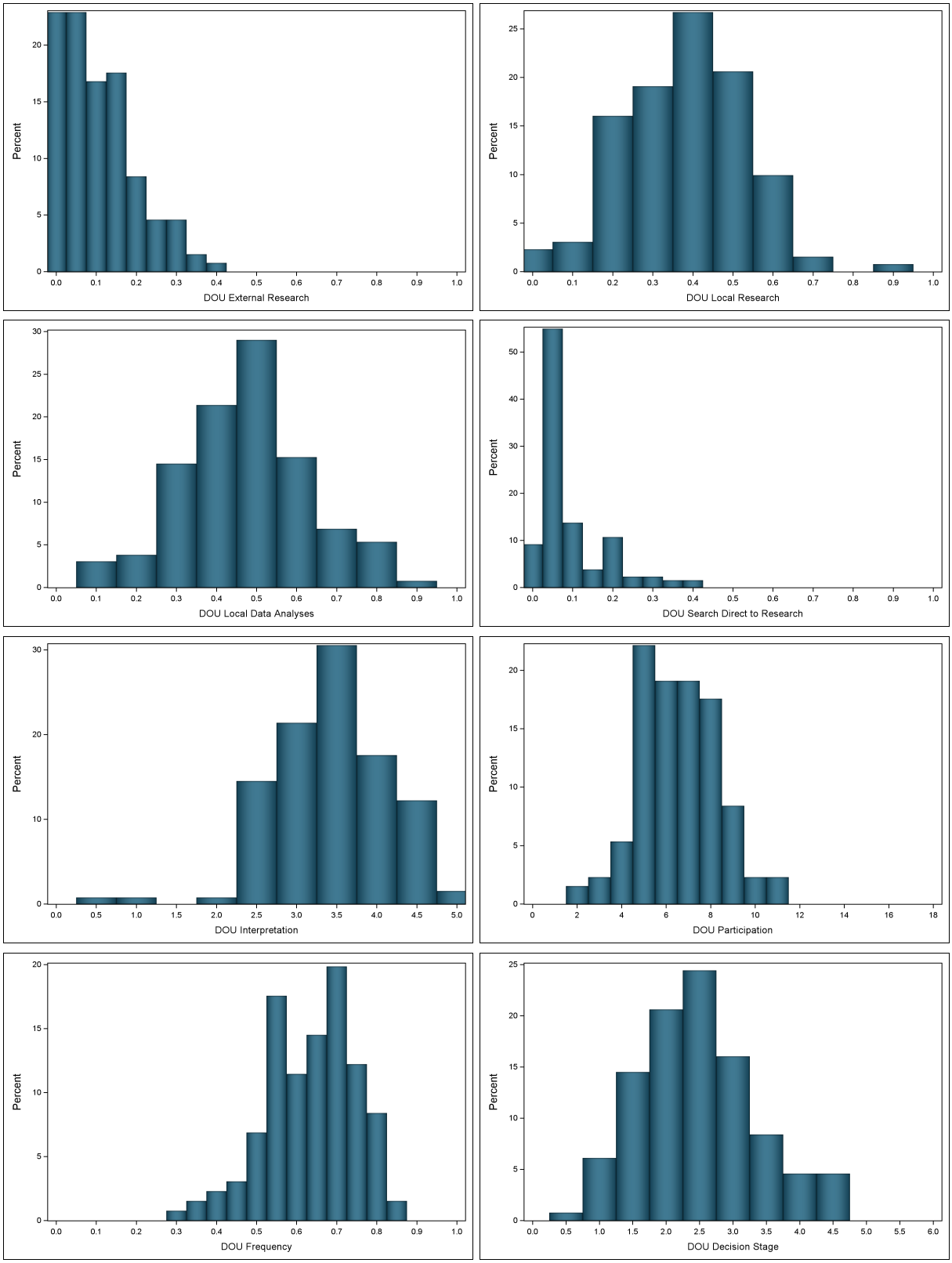
Participation makes evident that decision processes tend to involve multiple groups of constituents, and the histogram for *Interpretations* suggests that those people interpreting evidence sought to ensure that it was high quality, relevant, and understandable. Lastly, the histogram for *Decision Stage* suggests that empirical evidence (including external research as well as local research and analyses) came to bear on the majority of decisions reported, and at multiple stages during the decision process.

Figure 31. ***Histogram of School-Level Scale Scores for Depth of Use of Research Evidence***



Note. N = 131 schools. Individual-level item responses ranged from 1,285 to 1,344

Figure 32. ***Histograms of School-Level Scale Scores for DOU Subscales***



Note. N = 131 schools. Individual-level item responses ranged from 1,285 to 1,344

Table 6. ***Descriptive Statistics for Depth-of-Use Scale Scores***

Scale	N	Possible Range	Mean	Standard Deviation	Minimum	Maximum
Total *Depth of Use*	131	0-100	15.6	5.91	5.00	36.00
External Research	131	0-1	0.10	0.09	0.00	0.39
Local Research	131	0-1	0.38	0.15	0.00	0.87
Local Data Analysis	131	0-1	0.47	0.16	0.07	0.86
Direct to Research	131	0-1	0.09	0.09	0.02	0.40
Interpretation	131	0-5	3.44	0.72	0.75	4.79
Participation	131	0-18	6.52	1.77	1.93	10.60
Frequency	131	0-1	0.64	0.11	0.30	0.87
Decision Stage	131	0-6	2.50	0.87	0.54	4.67

Practitioner Assumptions and Perspectives about Research

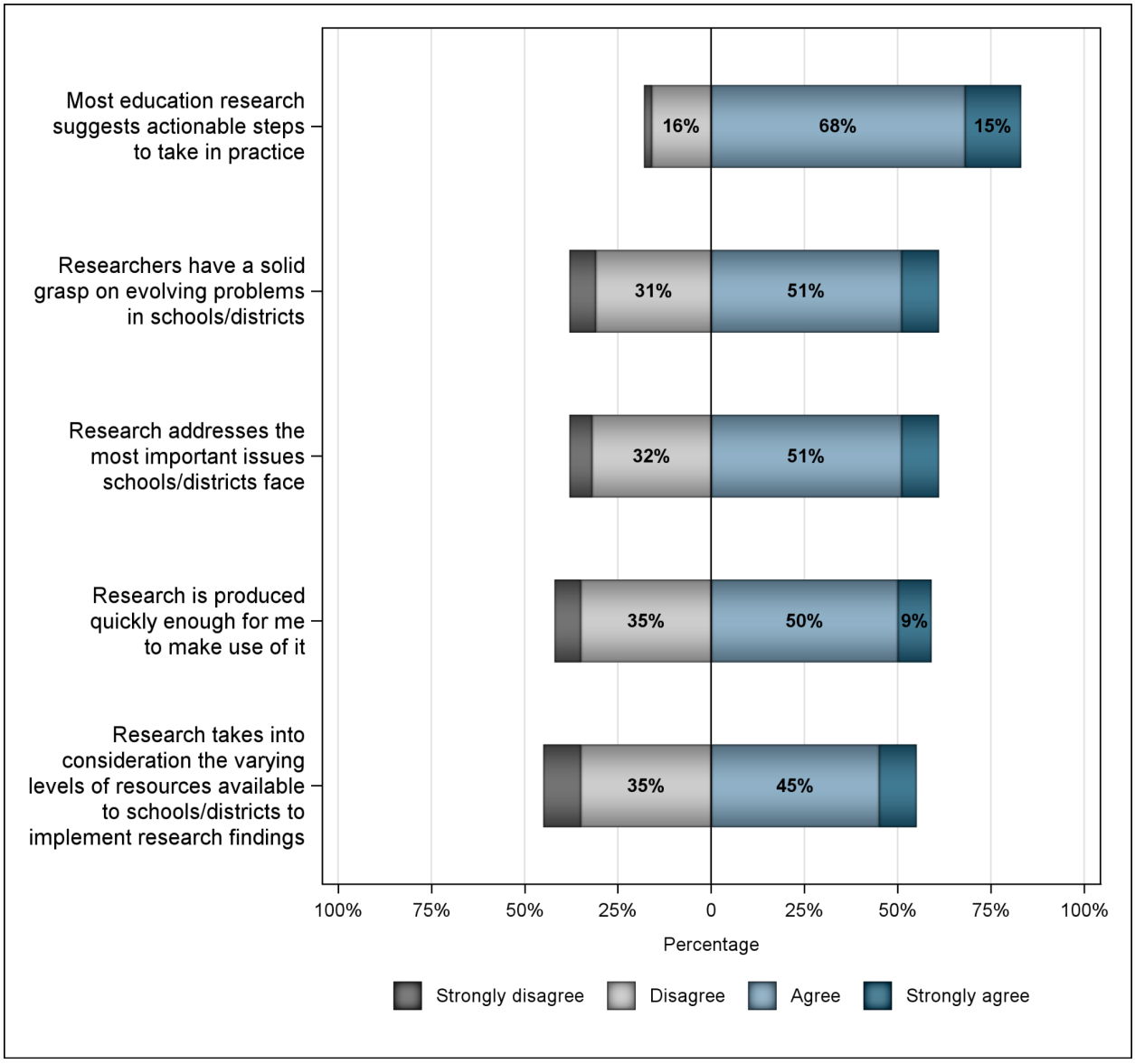
This portion of the *SEE-S* focuses on school-based practitioners' perceptions and assumptions related to the five potential gaps between the research and practice communities specified in our conceptual framework (see Figure 1). Covered in this section are items about the relevance and usefulness of research, the characteristics of products that practitioners find useful, and questions about the quality of research concerning internal and external validity. Participants were also asked to respond to questions about the structures, processes, and incentives that may facilitate or constrain research use in their schools and districts, and their relationships with the research community. (Thus, these responses are not about the problems and decisions reported in the previous section, but rather are about the general practices and assumptions of school-based practitioners who responded to the *SEE-S*.)

Relevance and Usefulness of Research

The first dimension of the gap between research and practice suggests that research may lack relevance for school-based practitioners. Among others, these issues include the actionability, and the timeliness of research, and attention to current problems faced by practitioners (Hemsley-Brown, et al., 2009; Maynard, 2006). The intent of our survey scale is to measure the degree to which practitioners find research to be relevant and useful for informing their work. Figure 33 details practitioners' responses to our items about the relevance and usefulness of education research. Generally, survey respondents had mildly positive perceptions about the relevance and usefulness of research. The vast majority of respondents (83%) 'agreed' or 'strongly agreed' that most education research offered actionable steps to take in practice. Responses were slightly less positive regarding researchers' understanding and attention to the problems and issues schools face, the timeliness of research, and the degree to which research acknowledges limited school resources. For example, 38% of respondents reported perceiving that researchers do not have a solid understanding of evolving problems in schools and districts. We conducted exploratory and confirmatory factor analyses of the individual items and the

results suggested that there was one latent variable: *Problems of Practice.* The factor loadings for this scale are included in Appendix D.

Figure 33. ***School-Based Practitioners' Perceptions of the Relevance and Usefulness of Research***



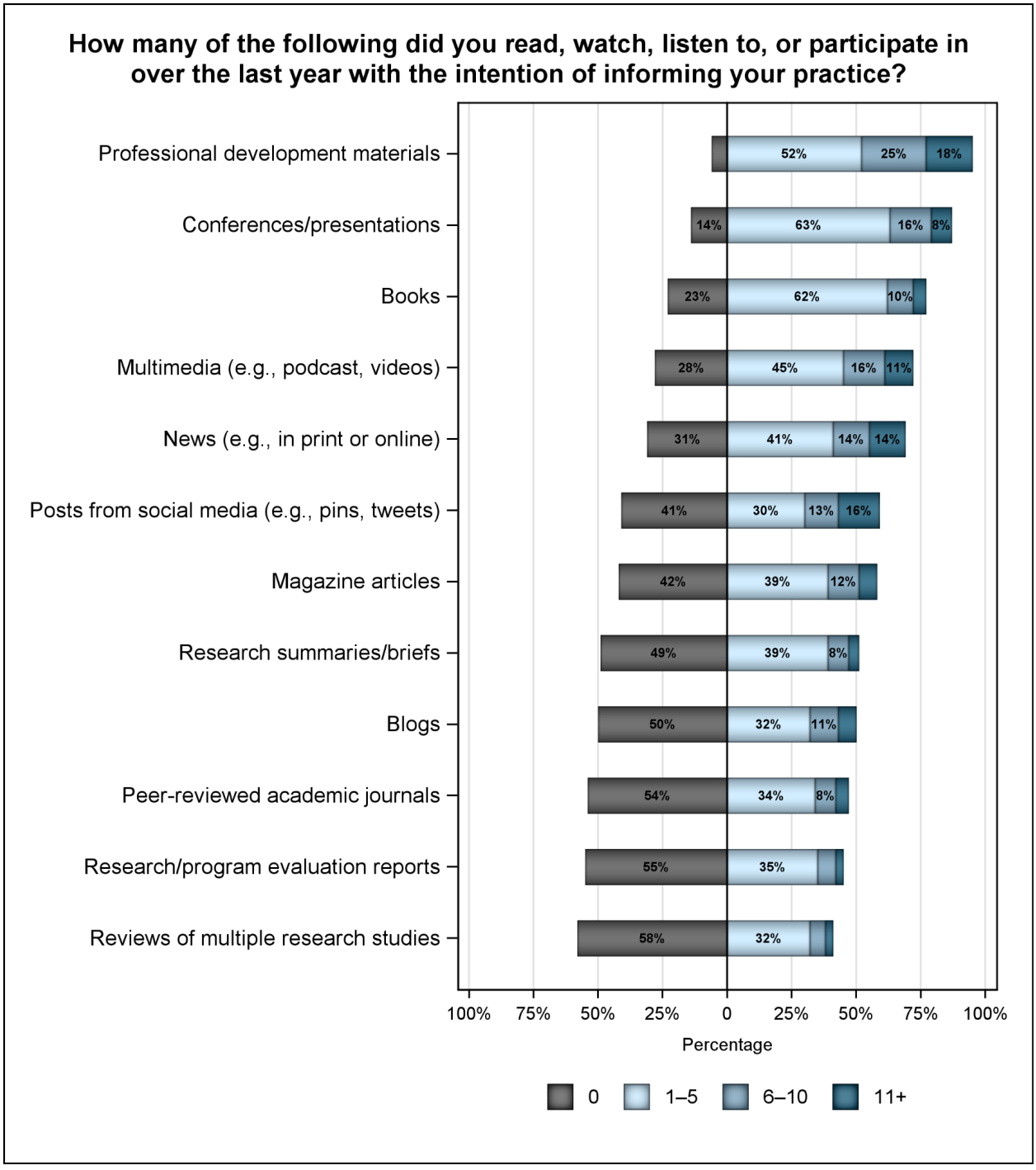
Note. Item response N ranges from 4,155 to 4,174

Products Consumed by School-Based Practitioners

Our research is also concerned with identifying the types of research products, and other types of products practitioners use to inform their practice. Prior research suggests that practitioners consume information from a variety of sources, including materials from formal professional learning events, books, social media, and more traditional research products (Behrstock-Sherratt, 2011; Farley-Ripple, 2012; Penuel et al., 2017). Figure 34 presents frequencies on the number of different products consumed by survey respondents over the past year, with the intention of informing their practice. Among the most consumed products were posts from social media and materials from professional development. Least consumed were research products such as reviews of multiple research studies, research/evaluation reports, and academic journal articles.

We conducted exploratory and confirmatory factor analyses of the individual items related to products consumed by practitioners, and the results suggested three latent variables for the types of products consumed by practitioners: research products, professional development materials, and media products. The subscales and the factor loadings of individual survey items are included in Appendix D.

Figure 34. ***Type of Products and Quantity Consumed by School-Based Practitioners***



Note. Item response N ranges from 4,085 to 4,198

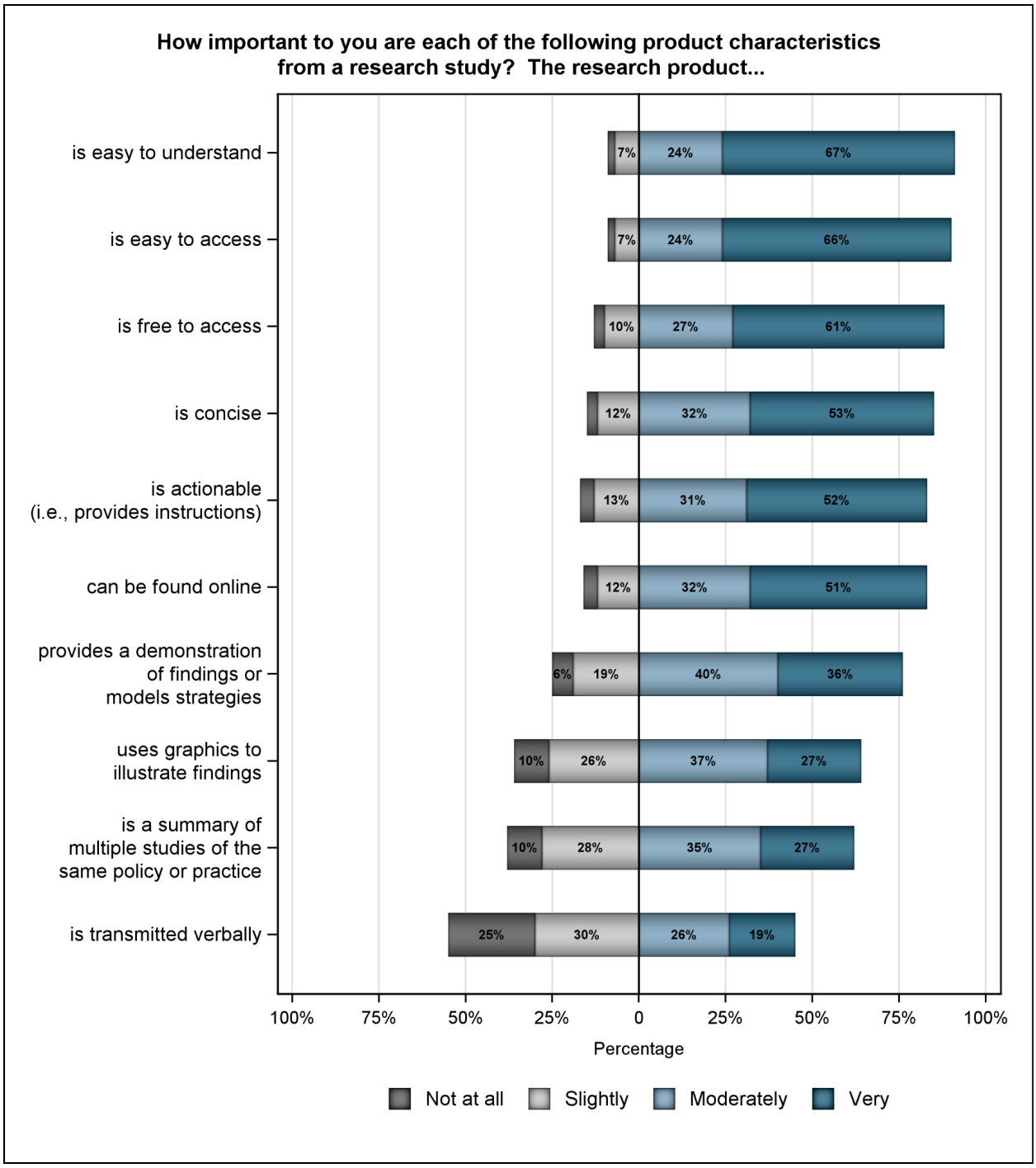
Educators Preferences for Characteristics of Research

Another set of items in this section focuses on practitioners' preferences regarding the characteristics of the products they consume. Prior research suggests that practitioners place greater weight on different characteristics of products, such as product accessibility (Gross, et al, 2005; Corcoran et al. 2001), and format and complexity (Reichardt, 2000; West & Rhoton, 1994). Figure 35 presents respondents' preferences regarding product characteristics. Practitioners felt most strongly that the product from a research study be easy to understand (67% 'strongly agree') and easy to access (66% 'strongly agree'). Of least importance to practitioners was that the information be presented verbally (19% 'strongly agree').

We conducted exploratory and confirmatory factor analyses of the individual items related to preferences for different characteristics of research, and the results suggested two latent variables: *Helpfulness of Features* and *Ease of Access*. The subscales and the factor loadings of individual survey items are included in Appendix D

.

Figure 35. ***School-Based Practitioners' Perceptions of Research Product Characteristics***



Note. Item response N ranges from 4,167 to 4,194

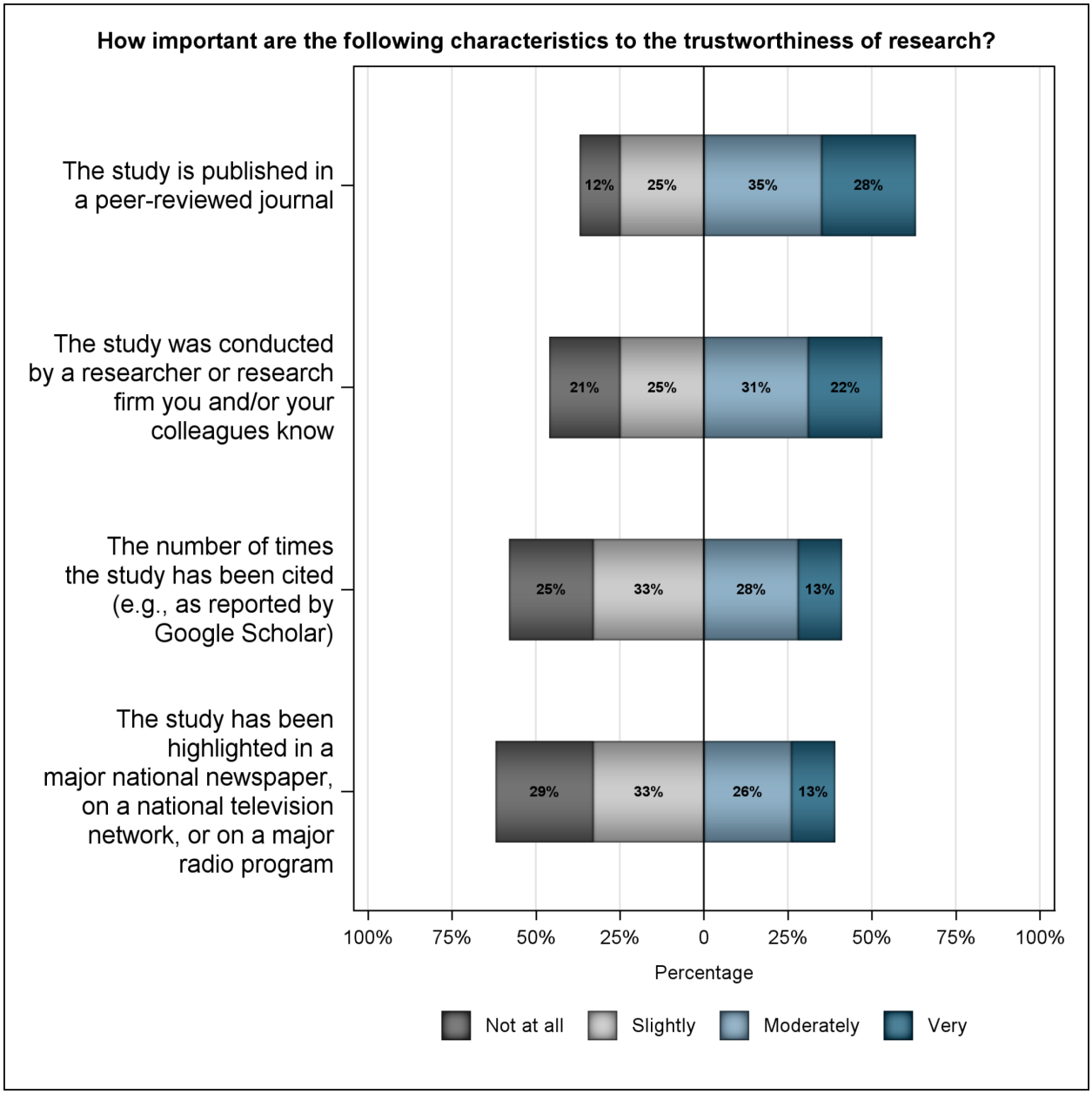
Determining the Trustworthiness and Credibility of Research

This dimension of the framework pertains to the practitioner community's valuing of different qualities of research, including internal and external validity. Research suggests that practitioners place more weight on external validity (Finnigan et al., 2013; Supovitz & Klein, 2003). Figure 36 presents practitioner's responses about the degree of importance of a number of characteristics of research. Regarding the trustworthiness of a research study, 63% of respondents indicated they place a high degree of importance on the research study being published in a peer-reviewed journal. Slightly over half of respondents (53%) reported that it was important that they or their colleagues know the researcher or research firm who conducted the study.

Respondents also responded to items about the characteristics of a research study and whether they were important for the study's credibility. Figure 37 depicts survey responses to this set of items. Of greatest importance to the credibility of a study was whether the study produced statistically significant results and reported findings for all outcomes, positive or negative. Of the 4,080 practitioners who responded to these items, almost half (48%-49%) 'strongly agreed' that these were important in their judgment of the credibility of a research study. As shown in Figure 37, few respondents (14%) 'strongly agreed' that whether the study had met What Works Clearinghouse evidence standards was important to the credibility of the study.

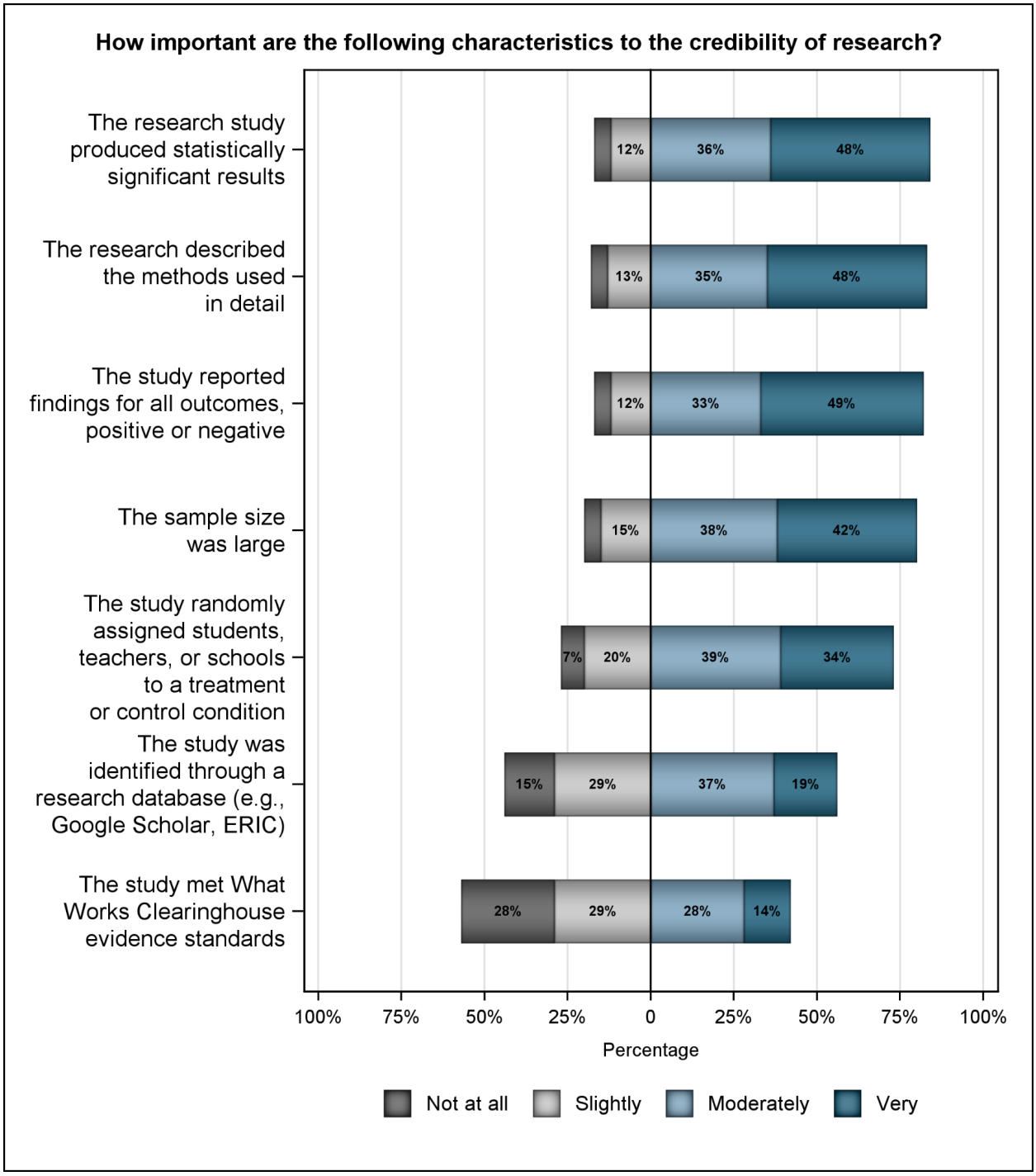
After conducting exploratory and confirmatory factor analyses, there was evidence of one latent variable we have labeled *Reputation-Based Trustworthiness* and another we have labeled *Methods-Based Credibility*. The details of the factor loadings for individual items for these latent variables are included in Appendix D.

Figure 36. ***School-Based Practitioners' Views About What Makes Research Trustworthy***



Note. Item response N ranges from 4,170 to 4,186

Figure 37. ***School-Based Practitioners' Views About What Makes Research Credible***



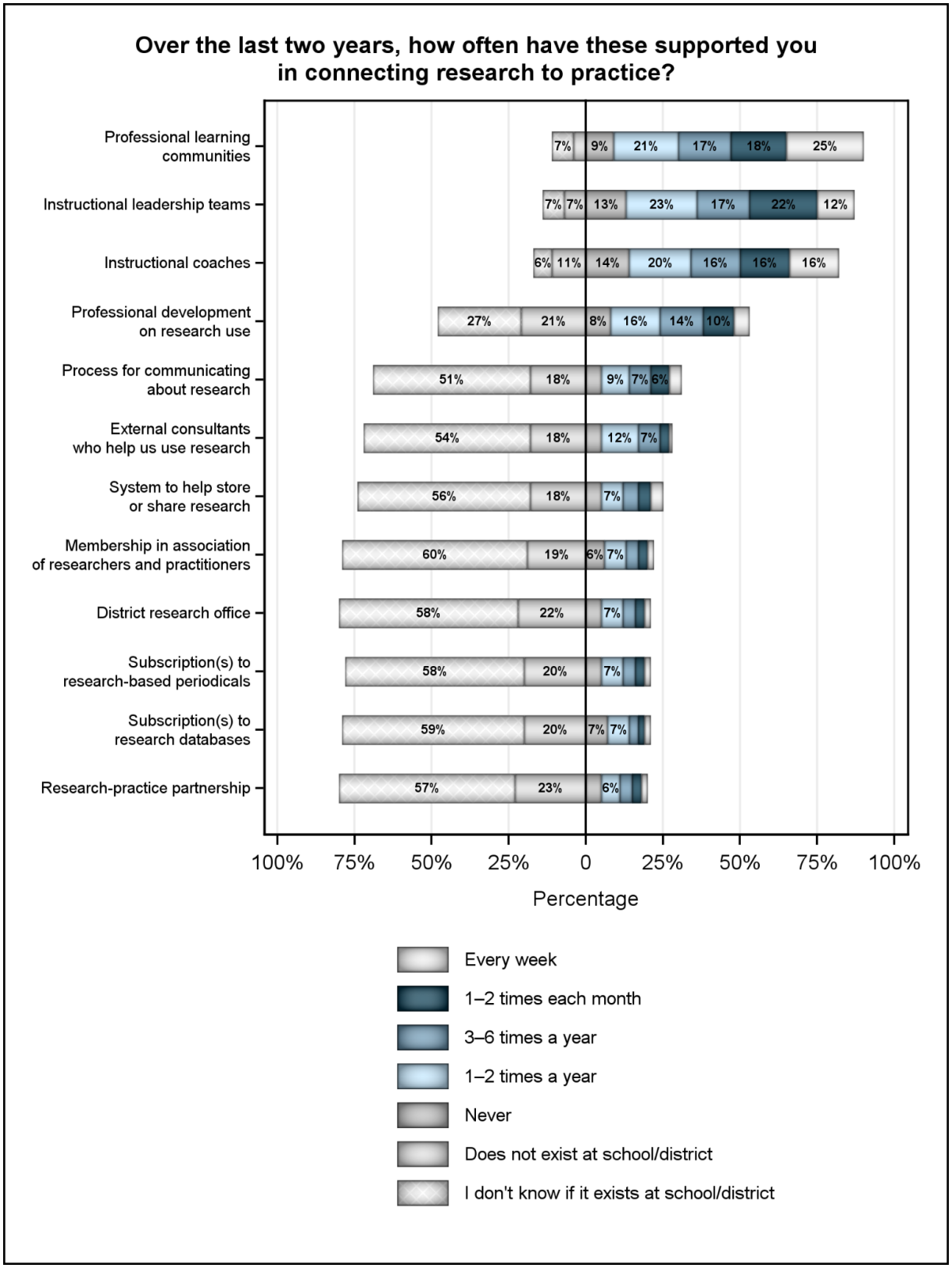
Note. Item response N ranges from 4,150 to 4,173

School and District Structures, Processes, and Incentives

The data in Figure 38 details survey takers' responses about the presence of a number of different school or district structures, as well as the frequency of support that structure provided them in connecting research to practice over the past two years. We consider certain roles (e.g., instructional coaches) to be school structures because they may serve as gatekeepers to research and other forms of evidence. The vast majority of respondents reported that professional learning communities or PLCs (90%), instructional leadership teams (87%), and instructional coach(es) (82%), were present in their school or district. Conversely, few respondents indicated that their school or district had a system or tool to help store and/or share research (26%), or a process for sharing/communicating about research (31%).

Regarding the frequency of support provided by elements present in schools, PLCs were cited as providing the most frequent support for connecting research to practice with 25% of respondents indicating receiving support "every week." Sixteen (16%) of respondents reported that instructional coaches provided this type of support "every week." For those elements that were rarely present, including district research offices and subscriptions to research-based periodicals, respondents most commonly cited these as providing support either "never" or "1-2 times per year."

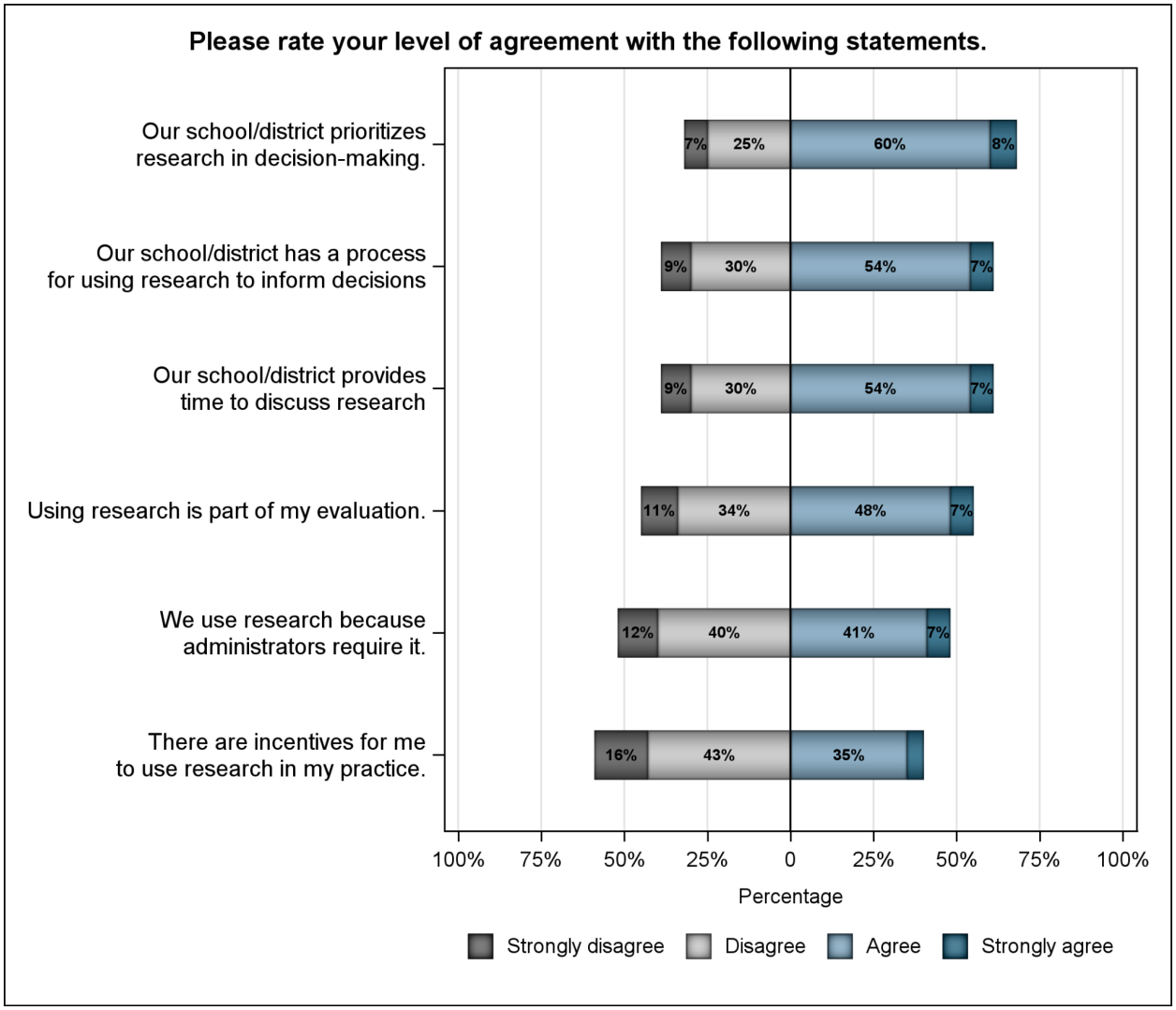
Figure 38. ***School and District Structures***



Note. Item response N ranges from 4,094 to 4,190

Practitioners were also asked to respond to items about school and district norms and processes that may facilitate or impede efforts to use research, as shown in Figure 39. Sixty-eight percent (68%) of respondents reported that they 'agree' or 'strongly agree' that their school/district prioritizes research in decision-making. Most respondents (60%) stated that they 'agree' or 'strongly agree' that their school/district provides time to discuss research. We conducted exploratory and confirmatory factor analyses for the *Processes and Incentives* items and the results provided evidence for one latent variable that included all items, as well as individual latent subscales for *Processes for Using Research* and *Incentives to Use Research.* The factor loadings of individual survey items for these latent variables are included in Appendix D.

Figure 39. ***Processes and Incentives for Using Research***

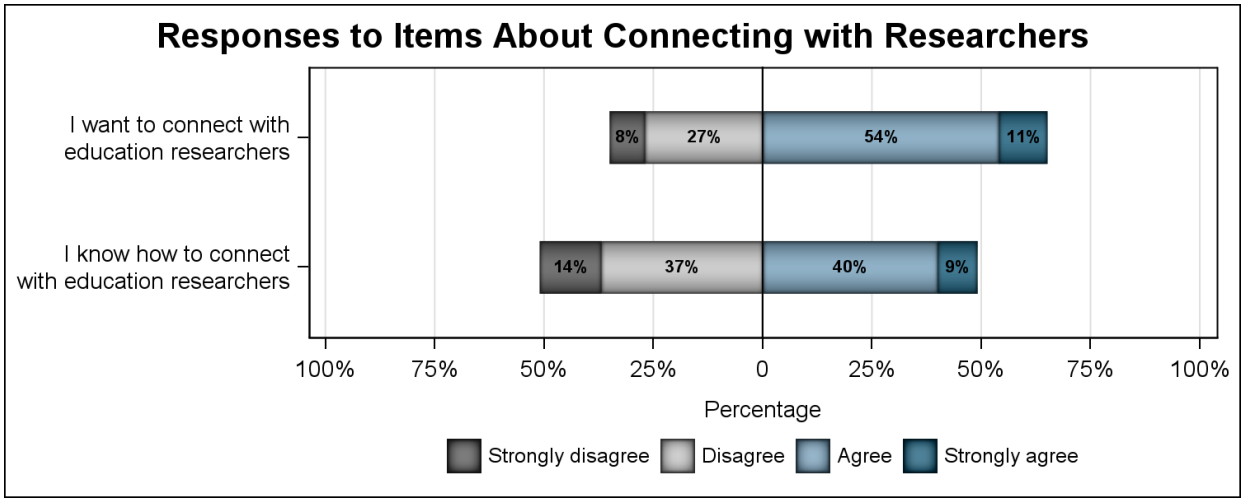


Note. Item response N ranges from 4,095 to 4,107

Relationships Between Research and Practice Communities

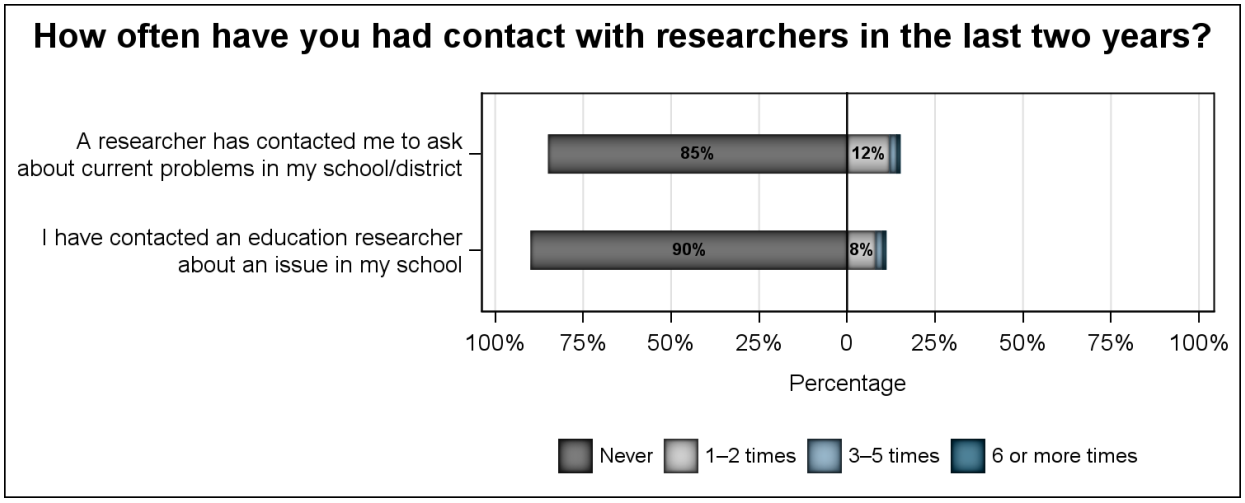
As shown in Figure 40, respondents to the survey generally agreed that they wanted to connect with education researchers--approximately 65% of respondents agreed that they wanted to connect with education researchers. But a smaller percentage indicated that they agreed with the statement about knowing how to connect with education researchers. Furthermore, despite their interest in connecting with education researchers, only about 10% of respondents reported that they contacted an education researcher within the last two years (see Figure 41).

Figure 40. ***Items About Connecting with Researchers***



Note. Item response N ranges from 4,159 to 4,167

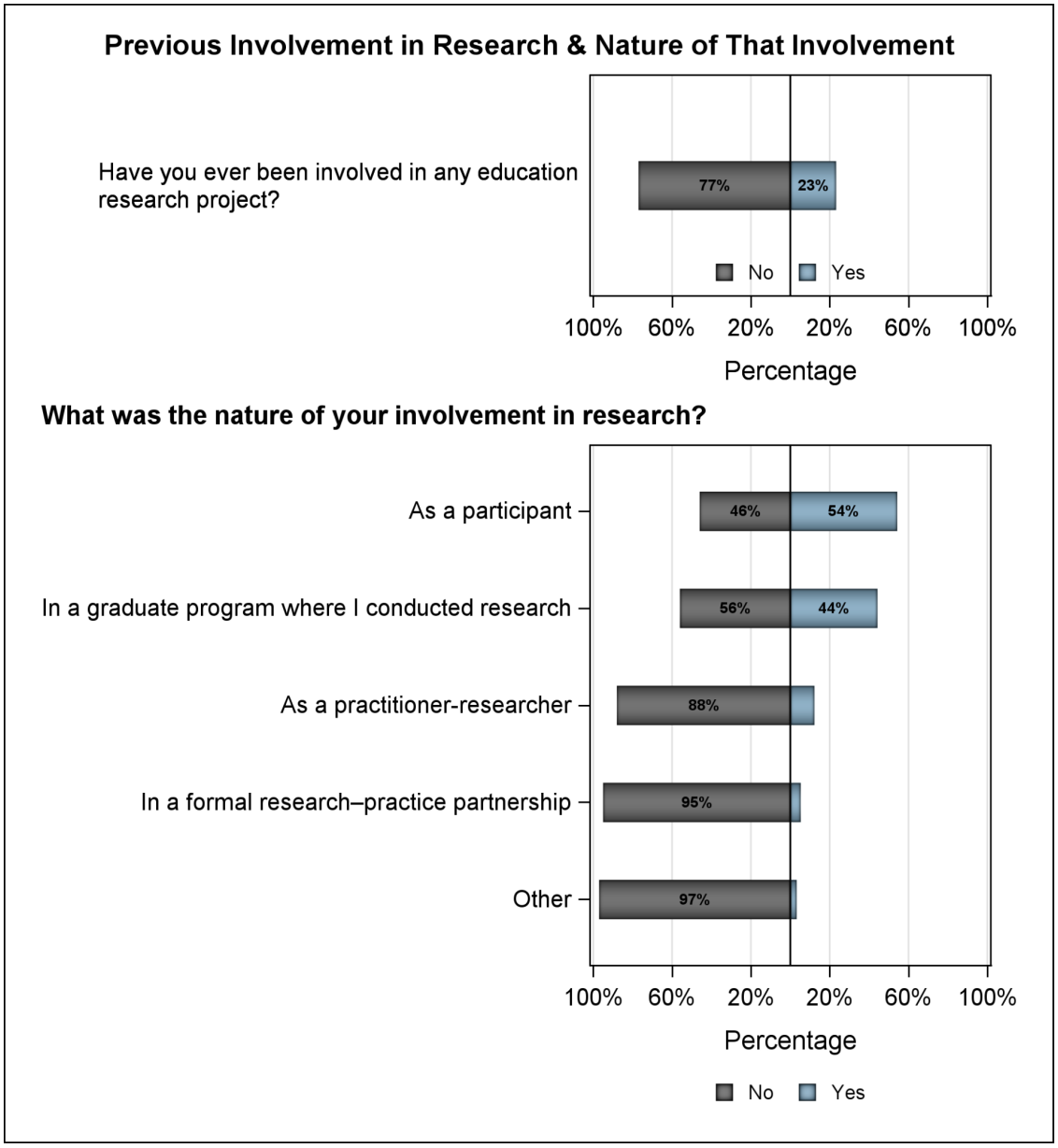
Figure 41. ***Contact with Education Researchers***



Note. N = 4,197

Respondents were also asked about prior involvement with research as another means of gauging the relationship between communities (Figure 42). Only 23% of respondents reported prior participation in research, and of those, most (54%) indicated they were a participant in a research study, and 44% were in a graduate program where they conducted research. Few reported doing their own research or participating in a research-practice partnership.

Figure 42. ***Previous Involvement in Research***



Note. Item response N for top panel = 4,147; for bottom panel N = 955

Reliability of Individual-Level Survey Scales on Practitioner Assumptions and Perspectives about Research

As described in the previous section entitled, "*Methods for Establishing Validity and Reliability of the SEE-S*," we conducted exploratory and confirmatory factor analyses (EFA and CFA) on sets of items from the *Assumptions and Perspectives* section of the survey that were not checklist items or "yes/no" survey items. Detailed results from these factor models (i.e., final factor loadings) are included in Appendix D of this report. Several constructs include only a subset of items from a survey section, as items that did not have a substantial factor loading were dropped from a scale. Once the final set of items for a scale were identified, internal consistency reliability was estimated using Cronbach's Alpha, and scale scores were calculated as a simple average across items using raw item responses (i.e., coded 1, 2, 3, 4). Descriptive statistics and reliability estimates for the survey scales for the *Assumptions and Perceptions* portion of the survey are included in Table 7.

Table 7. ***Descriptive Statistics and Reliability of Assumptions & Perceptions Survey Scales***

Construct	N	# of Items	Alpha Reliability	Range	Mean	Standard Deviation
Problems of Practice	4098	5	0.88	1-4	2.68	0.61
Processes and Incentives	4046	6	0.89	1-4	2.51	0.61
Processes	4085	3	0.86	1-4	2.62	0.66
Incentives	4072	3	0.82	1-4	2.41	0.67
Products						
Media	4028	5	0.79	1-4	1.96	0.74
Research	4049	4	0.91	1-4	1.60	0.68
Professional Resources	4132	3	0.71	1-4	2.23	0.62
Products - Characteristics						
Ease of Access	3939	6	0.90	1-4	3.42	0.65
Helpful Features	3954	4	0.84	1-4	2.76	0.79
Trustworthiness and Credibility						
Reputation-Based Trustworthiness	3994	4	0.80	1-4	2.46	0.79
Methods-Based Credibility	3957	5	0.89	1-4	3.19	0.72

Social Network Survey of Practitioners' Connections to Research

The social network survey portion of the *SEE-S* asked survey respondents to identify up to 10 individuals, 10 organizations, and 10 media sources they rely on to connect with educational research in general (as opposed to in the context of a specific decision, as in earlier sections of the survey). Respondents entered open ended names for resources and were asked to describe them in terms of predetermined categories. Here we describe some basic analyses of the SNA data generated from the *SEE-S* Field Trial.

Data were entered into UCINet (Borgatti et al., 2002) to generate network statistics for individuals and schools. UCINet offers a range of analytical tools for ego-networks that can address a wide range of questions related to individuals' networks. In this report, we utilize two simple descriptors of ego-networks that provide insight into the ties that link educators to educational research: size and composition. Network size is often conceptualized as a measure of social capital. In other words, the larger the network, the more resources available. In the context of educators' networks, more ties may indicate greater access to educational research. However, larger networks often come at a cost, of either time or effort, to maintain those ties, limiting educators' ability to make use of the resources available through them. Network size is ascertained from the total number of resources reported by educators in the sample, and then the total number of resources reported by educators within each school. In terms of composition, individual and school networks include many types of resources. Composition is useful in differentiating the nature and quality of ties for accessing research. Composition statistics were generated for each category of resource based on the proportion of the network constituted by each. Proportions eliminate the bias of comparisons based on network size and permit comparisons among educators' networks.

Respondents reported a mean of 7.69 resources (SD=6.57) they use, and the school level mean was 138.42 resources (SD=98.34). Table 8 below summarizes the total frequency of responses by category, the proportion of individuals and schools with at least one tie to a particular category of resources, and the proportion of individuals' and schools' networks composed of resources from each category.

Table 8. ***Descriptive Statistics from the Social Network Survey Component of the SEE-S Survey***

	Proportion of all nominations	**Percent of individuals with at least one tie**	**Percent of schools with at least one tie**	**Composition of Networks (Individual)**		**Composition of Networks (School)**	
N	21,275	2,772	153	2,772		153	
Individuals	**%**	**%**	**%**	**Mean**	**SD***	**Mean**	**SD***
Teacher	14.7%	40.3%	95.4%	0.134	0.225	0.038	0.034
Principal/Assistant Principal	13.1%	59.7%	98.0%	0.164	0.215	0.132	0.062
Instructional Coach	7.3%	41.9%	93.5%	0.094	0.163	0.072	0.050
Other School Staff	2.9%	14.5%	77.1%	0.024	0.078	0.026	0.027
District Administrator or Staff	6.3%	30.7%	94.1%	0.066	0.141	0.067	0.048
Other (please specify)	2.2%	9.5%	74.5%	0.021	0.090	0.020	0.022
Interventionists (e.g., math or reading specialist, etc)	1.8%	10.2%	73.2%	0.017	0.069	0.018	0.022
Professor	1.7%	8.9%	68.0%	0.017	0.076	0.016	0.019
External PD Provider, Program Developer, or Publisher	1.6%	9.1%	68.6%	0.025	0.076	0.031	0.031
External Researcher	1.3%	6.2%	58.8%	0.013	0.072	0.015	0.032
Organizations							
Professional Association	7.7%	32.7%	92.2%	0.081	0.161	0.074	0.051
PD Provider, Program Developer, or Publisher	3.0%	15.3%	85.6%	0.018	0.077	0.019	0.039
University-based Research Organization	2.1%	10.7%	79.7%	0.017	0.067	0.022	0.027

*Note: SD in this table represents the standard deviation from the distribution of individual means. It is not equal to $\sqrt{(p/1 - p)}$.

Table 8 (continued). ***Descriptive Statistics from the Social Network Survey Component of the SEE-S Survey***

	Proportion of all nominations	Percent of individuals with at least one tie	Percent of schools with at least one tie	Composition of Networks (Individual)		Composition of Networks (School)	
Organizations (continued)	**%**	**%**	**%**	**Mean**	**SD***	**Mean**	**SD***
Independent Research Center	0.9%	5.6%	54.2%	0.008	0.049	0.010	0.022
School District	4.4%	20.6%	85.6%	0.037	0.095	0.042	0.043
Foundation	0.6%	3.8%	45.1%	0.005	0.035	0.006	0.010
Advocacy Group	0.4%	3.1%	39.9%	0.005	0.046	0.006	0.027
Government Agency (e.g., State Department of Education)	1.8%	11.5%	72.5%	0.017	0.070	0.017	0.017
Other (please specify)	1.4%	5.1%	49.7%	0.012	0.066	0.011	0.020
Media Sources							
Other Resources (e.g., YouTube, Teachers-Pay-Teachers)	4.6%	23.0%	70.6%	0.045	0.116	0.047	0.038
Social media	3.9%	19.2%	86.3%	0.038	0.104	0.150	0.081
Magazine (Online, Print)	2.8%	16.4%	86.3%	0.026	0.083	0.029	0.028
News Source (Online, Print, TV)	2.7%	13.4%	78.4%	0.021	0.073	0.029	0.036
Book	2.6%	12.7%	75.2%	0.018	0.059	0.023	0.023
Research Database (e.g., Google Scholar)	2.3%	12.5%	77.1%	0.020	0.075	0.022	0.025
Other website (please specify)	2.3%	11.9%	41.2%	0.022	0.082	0.022	0.026
Blog	1.7%	9.5%	70.6%	0.015	0.064	0.017	0.024
Peer-Reviewed Journal	1.3%	8.0%	64.1%	0.013	0.062	0.012	0.014
Other (please specify)	1.0%	3.5%	91.5%	0.009	0.063	0.009	0.022

*Note: SD in this table represents the standard deviation from the distribution of individual means. It is not equal to $\sqrt{(p/1 - p)}$.

Individual respondents reported most frequently turning to other individuals to find research-based information, most notably others in their school community: other teachers comprise 13.4% of the average individual network and 3.8% of the average school network, while administrators comprise 16.4% of the average individual network, 13.2% of the average school network. Among organizations that provide access to research-based information, professional associations were most often noted, comprising 8.1% and 7.4% of individual and school networks respectively. Media sources were also used widely, with "other" web-based resources (including sites like Teachers Pay Teachers) as frequently reported linking mechanisms to research, with 23% of educators and 70.6% of schools reporting use of at least one source of this type. Overall, connections to traditional sources of research, such as researchers, research organizations, and peer-reviewed journals were among the least frequently reported sources, though when aggregated to the school level, at least 50% of schools report at least one direct connection to research.

Confidence to Critique Research

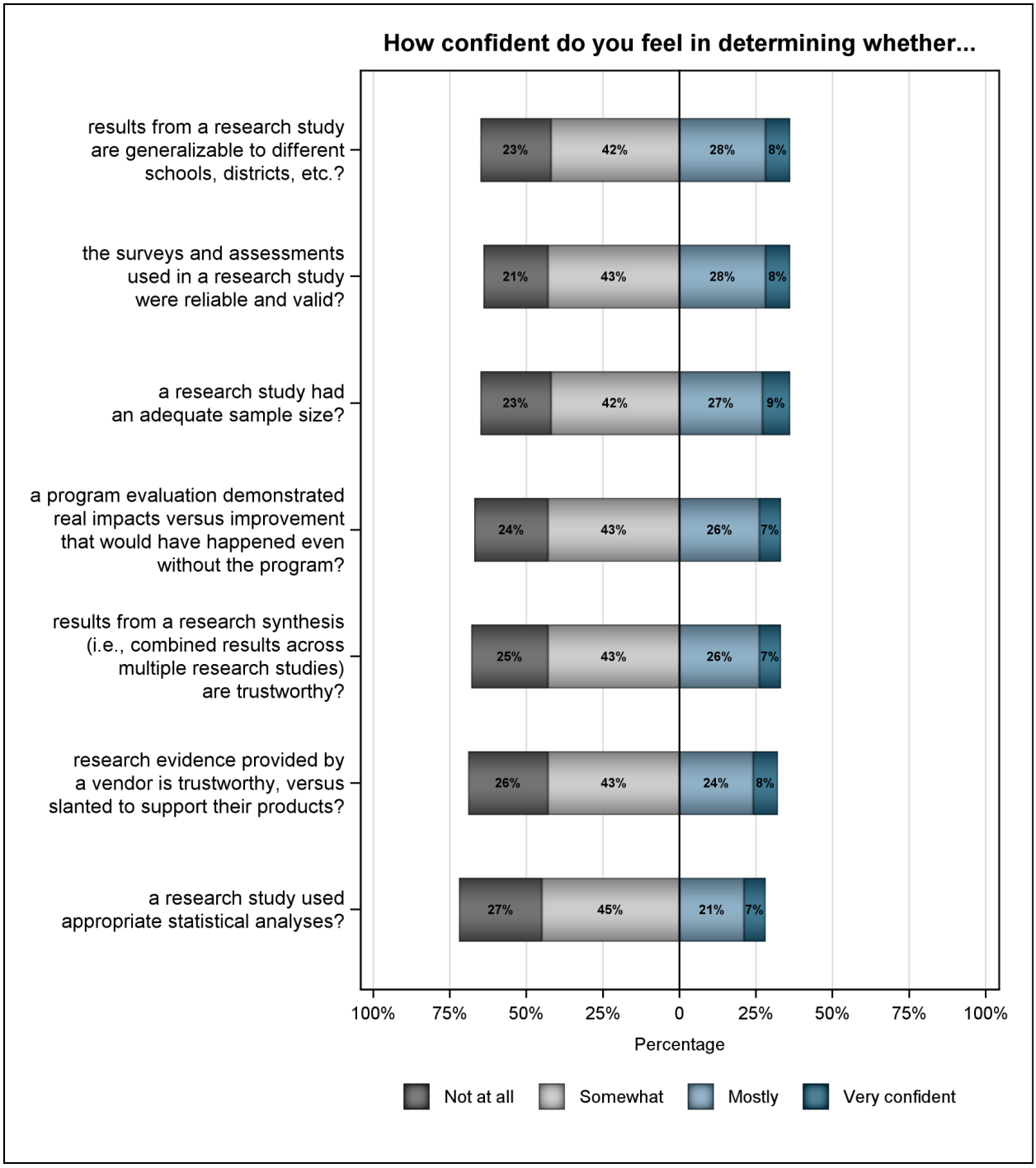
Covered in this section of the *SEE-S* are survey items that measure individual educators' confidence to critically interpret research. Included are items that 1) ask about their confidence in critiquing research, and 2) their research-related training and experiences.

Confidence to Critique Research

This set of items focuses on educators' confidence in their capacity to critically interpret or evaluate research. Figure 43 details practitioners' responses. Generally, survey respondents were not confident in their ability to critically interpret or evaluate research. For example, 64% of respondents reported they were 'not at all' or only 'somewhat' confident in determining whether surveys and assessments used in research were reliable and valid. Additionally, only 32% reported being 'mostly' or 'very' confident in their ability to determine if research evidence provided by a vendor was trustworthy. A confirmatory factor analysis conducted with PROC CALIS in SAS 9.4 indicated strong fit of a unidimensional model, and the coefficient alpha for the

seven items in the *Confidence to Critique Research* scale was very high (α = .97). The factor loadings are included in Appendix D.

Figure 43. ***School-Based Practitioners' Confidence in Critiquing Research***

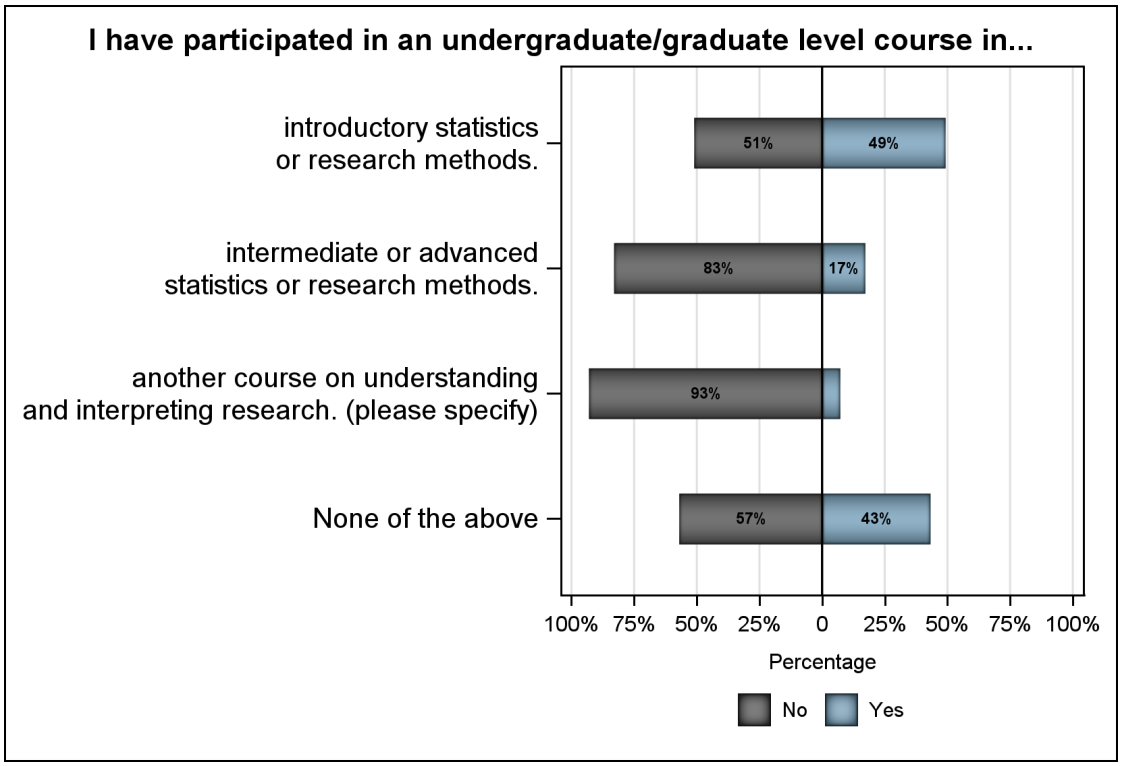


Note. Item response N = 4,082

Research-Related Training and Experiences

Educators were also asked to respond to a series of questions about their training in research methods and statistics (Figure 44). Fewer than half of the respondents (49%) had taken an introductory statistics or research methods course as an undergraduate or graduate student. The vast majority of respondents (83%) had not taken an advanced statistics or research methods course.

Figure 44. ***School-Based Practitioners' Research-Related Training***



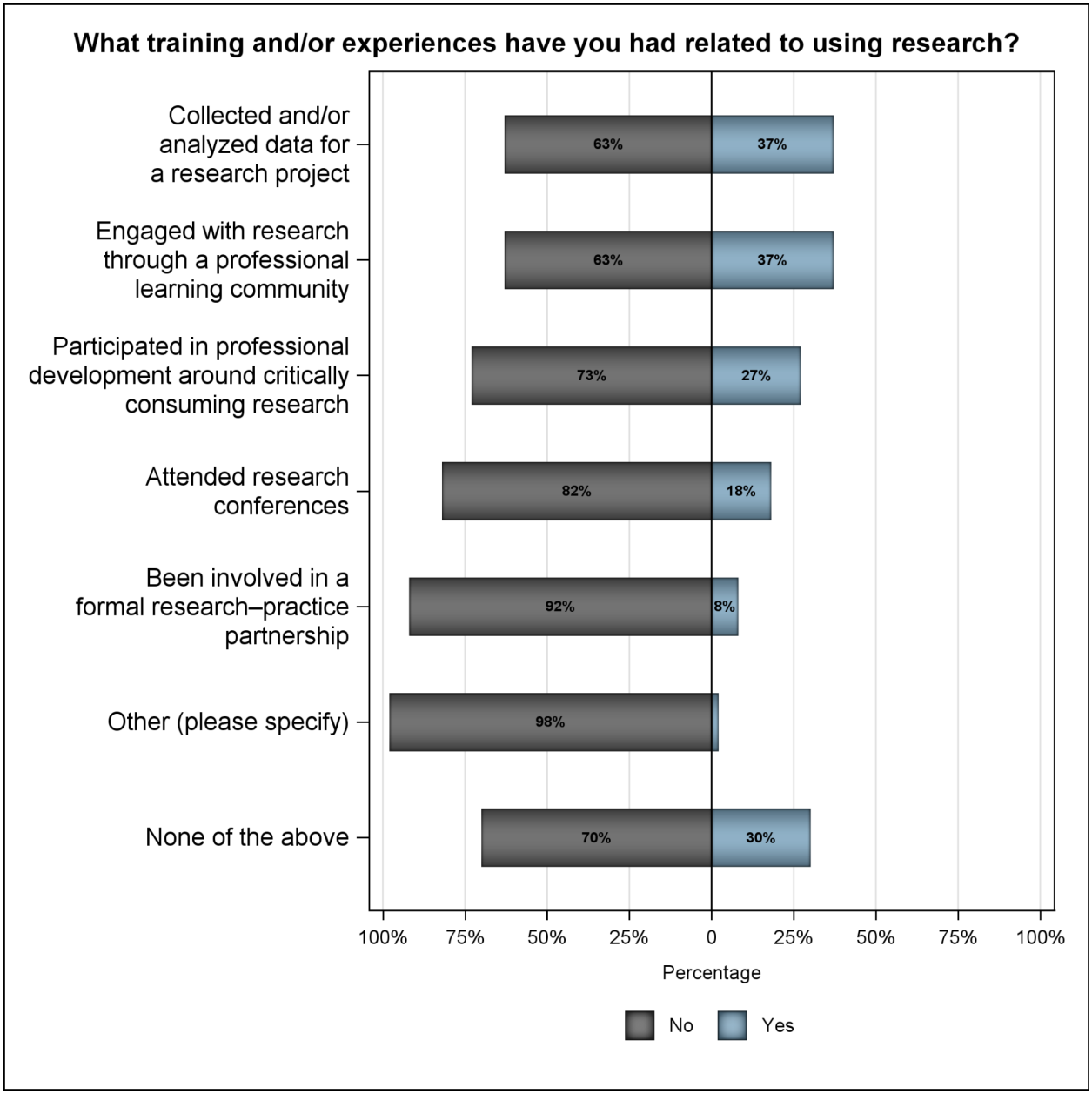
Note. Item response N = 4,082

Training & Experiences Related to Using Research

The *SEE-S* included items intended to describe educators' training and experiences related to research (see Figure 45). Research related experiences may be related to practitioners' use of research in practice and their broader connection to the research community. Responses to

these questions are detailed in Figure 45. Over 80% of survey respondents reported having never attended a research conference. Seventy-three percent indicate that they have never participated in professional learning on critically consuming research. Overall, respondents tended to have limited experience with activities related to using research. Less than 10%, for example, reported being part of a research–practice partnership.

Figure 45. ***Training and Experiences Related to Using Research***



Note. Item response N = 4,082

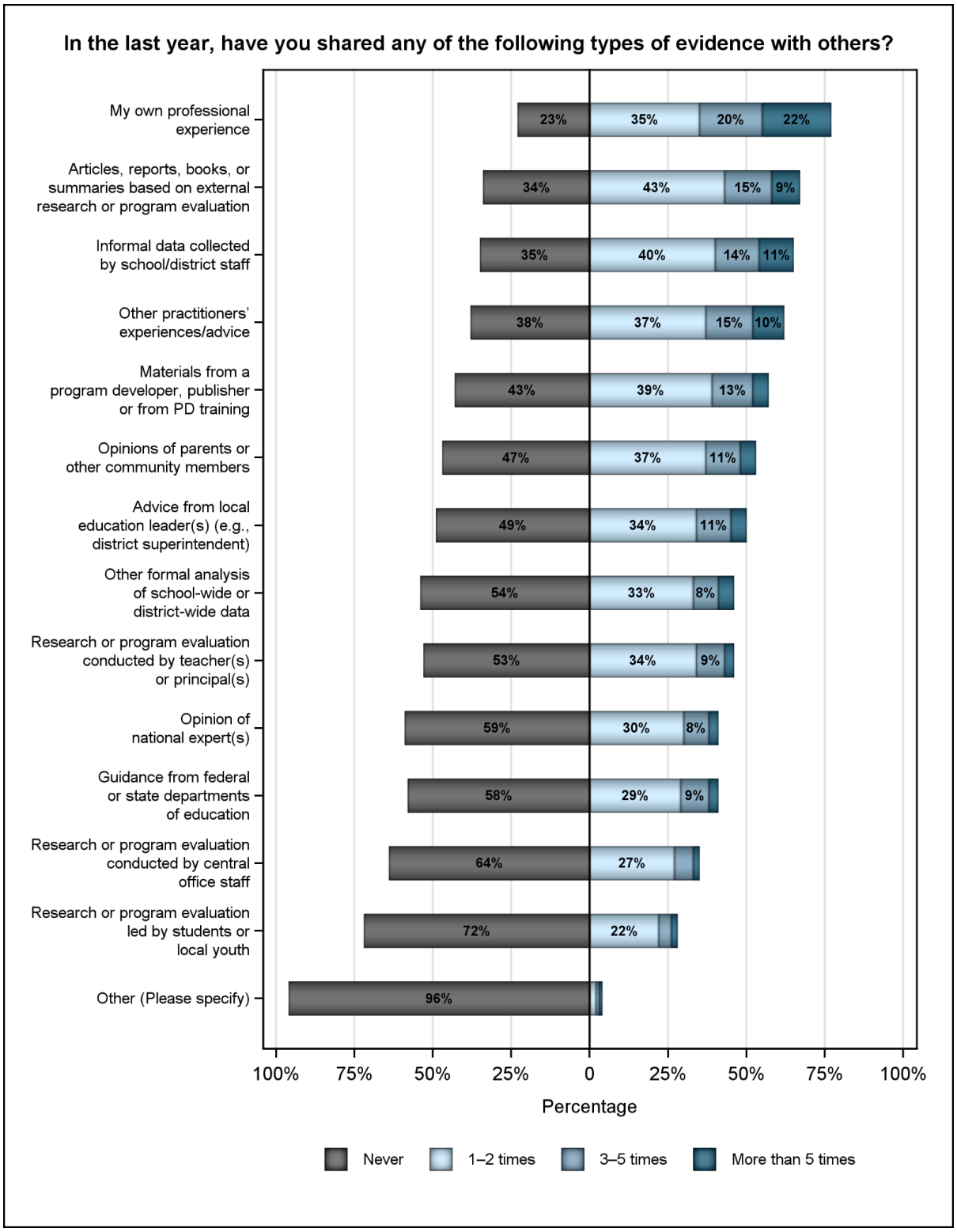
Brokerage of Research and Other Information

Brokering Different Types of Information

To further our understanding of the dissemination and use of research compared to other forms of evidence, respondents were asked a series of questions about their sharing of evidence. Educators were first asked how many times they shared different types of evidence with others in the last year. Among the list of items, the vast majority of practitioners (77%) reported sharing their own professional experiences. Sixty-seven percent (67%) of respondents reported sharing *articles, reports, books, or summaries based on external research or program evaluation* at least once during the previous year. Although *research or program evaluation conducted by central office staff* was among those least frequently shared (by 28% of respondents), recall that only 20% of respondents reported that their district had a research office (see Figure 38). As can be seen in Figure 46, *informal data collected by school/district staff*, *formal analyses of school-wide or district-wide data*, and *research conducted by school staff*, were reportedly shared by 65%, 46%, and 47% of respondents. Figure 46 presents detailed results for fourteen types of information shared by educators responding to the *SEE-S*.

We conducted exploratory and confirmatory factor analyses for the *Brokering Types of Information* items and the results provided evidence for three latent variables: *Research Products*, *Local Knowledge*, and *Experts/State and Federal Sources.* The factor loadings of individual survey items for these latent variables are included in Appendix D.

Figure 46. ***Brokering of Different forms of Evidence as Reported by School-Based Practitioners***

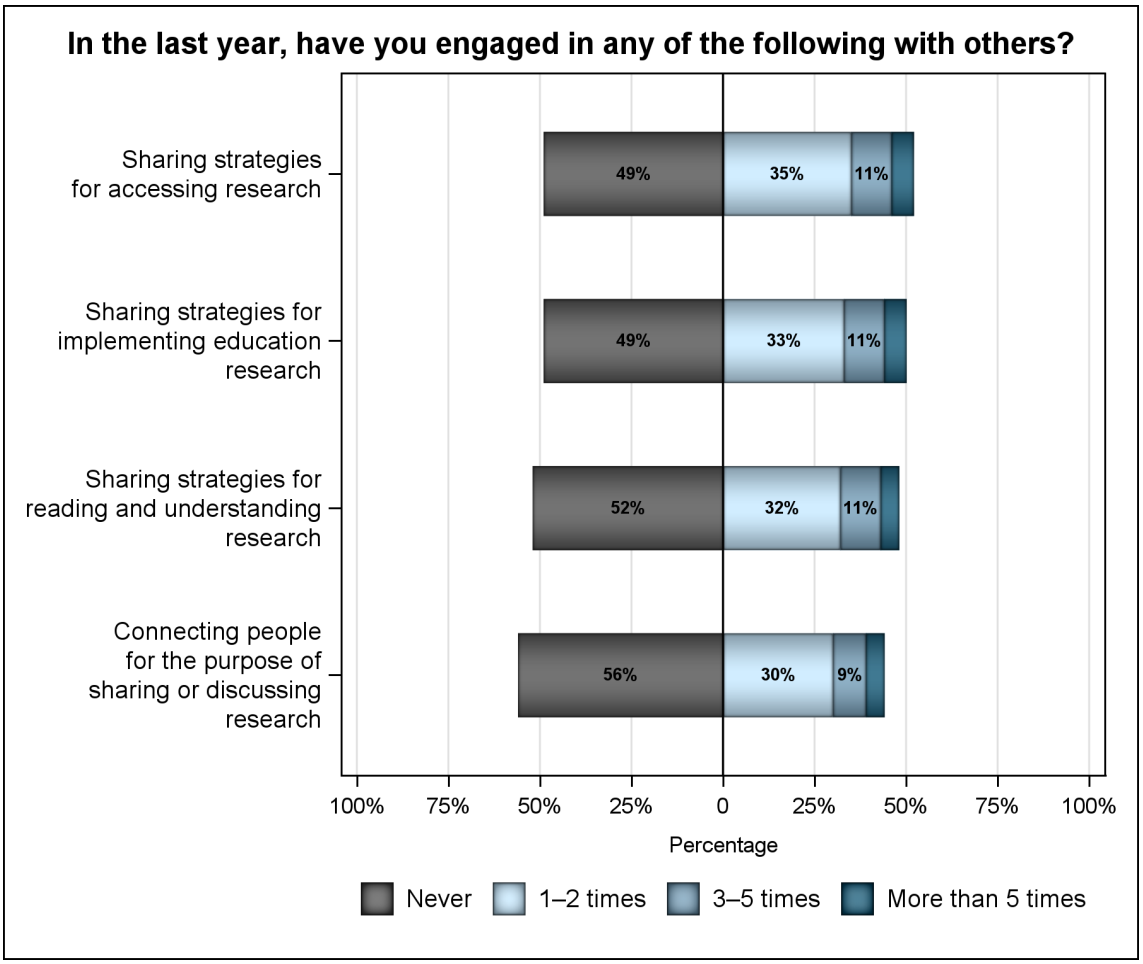


Note. Item response N ranges from 3,810 to 4,072

Capacity Building

Respondents were also asked to report on their brokerage capacity-building efforts in the past year (see Figure 47). Our sample of school-based practitioners reported that they occasionally engaged in capacity building efforts of this type. For example, 51% of respondents indicated that last year, they shared strategies for accessing research at least once. Additionally, 52% reported never sharing strategies for reading and understanding research. Fifty-one percent of respondents reported sharing strategies for implementing research at least once over the past year. Approximately 63% of respondents indicated that they engaged in at least one of these capacity-building activities.

Figure 47. ***Brokerage Activities that Build Capacity as Reported by School-Based Practitioners***



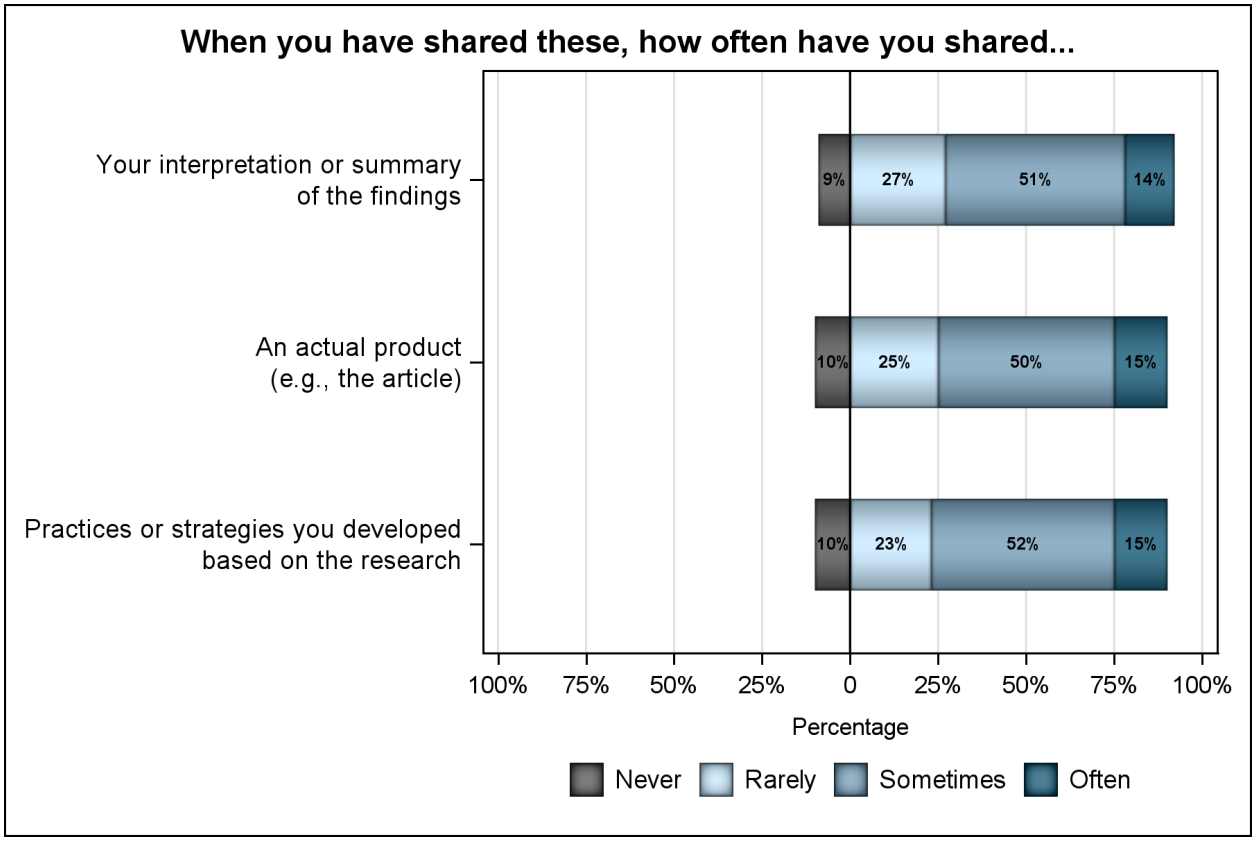
Note. Item response N ranges from 3,915 to 3,932

We conducted exploratory and confirmatory factor analyses of the individual items for *Brokerage - Capacity Building* and the results suggested that was one latent variable. The factor loadings for this scale are included in Appendix D.

Interpreting Information

As shown in Figure 48, it was common that when respondents shared research and engaged in capacity building, they shared the actual research product and also provided their own interpretation or summary of the findings. Specifically, 65% percent of respondents who shared information reported that they "Sometimes" or "Often" shared their own interpretation or summary of the research findings.

Figure 48. ***Brokerage Activities that Interpret Information for Others***



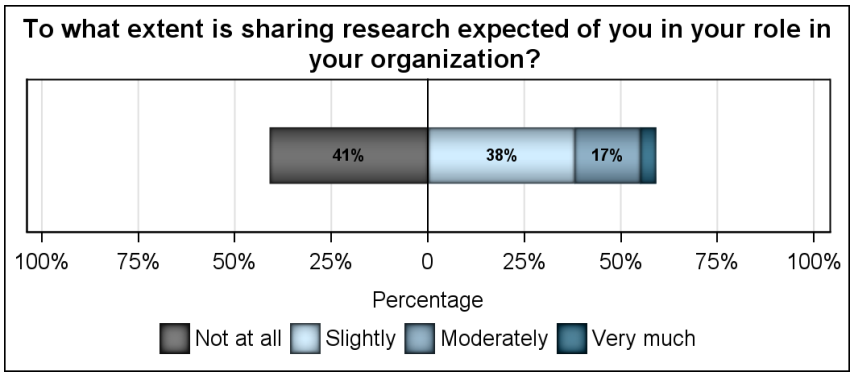
Note. Item response N ranges from 2,109 to 2,118

We conducted exploratory and confirmatory factor analyses of the individual items for *Brokerage - Interpretation* and the results suggested that was one latent variable. The factor loadings for this scale are included in Appendix D.

Expectations for Sharing Research

As shown in Figure 49, only 21% of respondents indicated that sharing research is expected of them "Moderately" or "Very Much." Forty-one percent (41%) of respondents indicated that there is no such expectation.

Figure 49. ***Expectations of Sharing Research in the Organization***

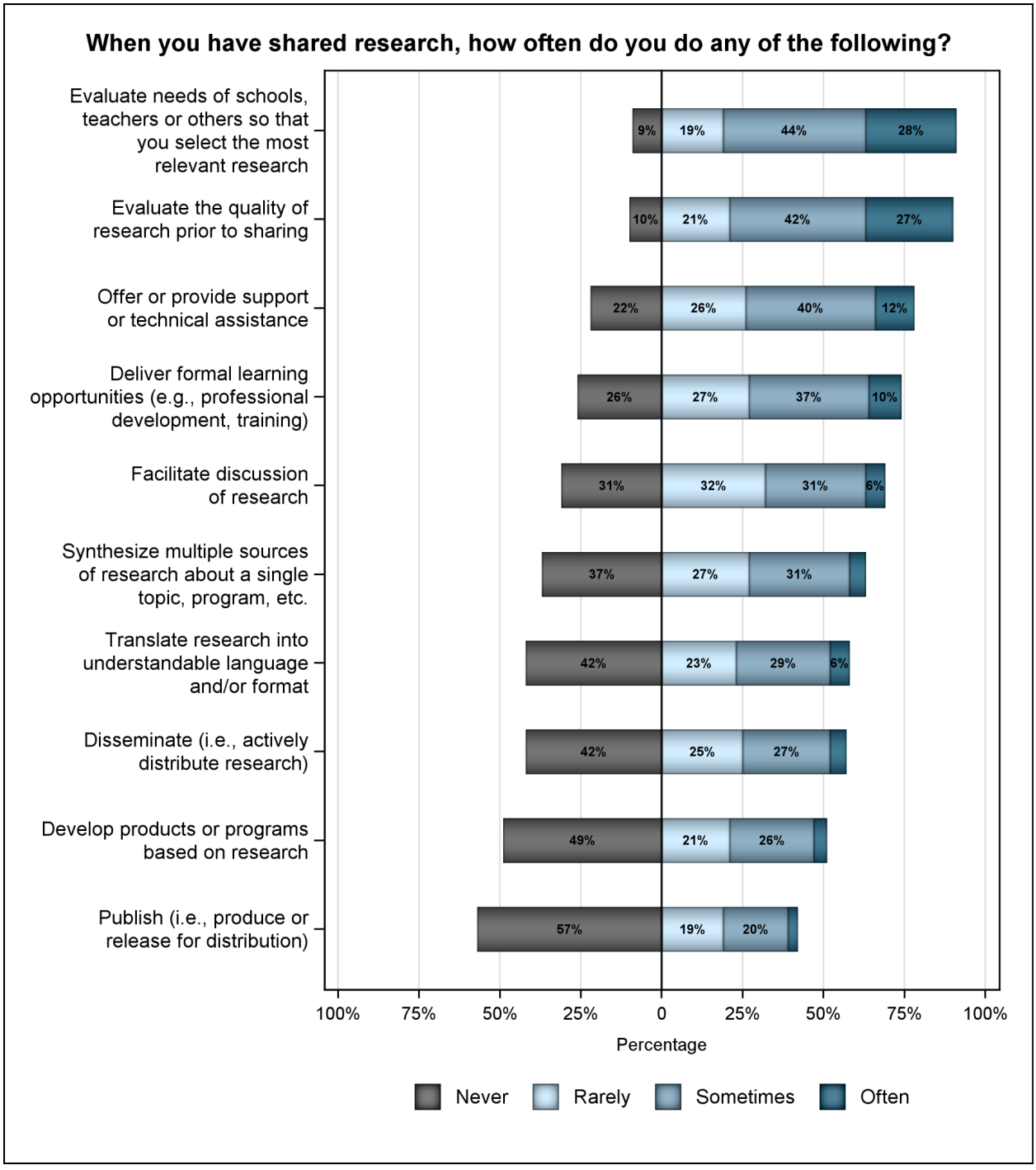


Note. Item response N = 3,958

Brokering Activities

If respondents indicated that they shared a research artifact and engaged in at least one form of capacity building, they were asked what activities they engaged in as part of the sharing process (see Figure 50). These items focused on evaluation of the evidence or context, dissemination, and assistance with moving the evidence into practice. When sharing research, 27% of respondents 'often' evaluate the quality of research prior to sharing. Approximately 10% of respondents reported never engaging in this activity. More than a third (37%) reported 'sometimes' or 'often' translating research into understandable language and/or format. Few respondents (12%) reported 'often' to the item 'offer or provide support or technical assistance' when sharing research.

Figure 50. ***Research Brokering Activities Reported by School-Based Practitioners***



Note. Item response N ranges from 2,065 to 2,121

We conducted exploratory and confirmatory factor analyses of the individual items for *Brokerage - Activities* and the results suggested that was one latent variable. The factor loadings for this scale are included in Appendix D.

Survey Scales Based on Brokerage Items

As described in the previous section entitled, "*Methods for Establishing Validity and Reliability of the SEE-S*," we conducted exploratory and confirmatory factor analyses (EFA and CFA) on sets of items from the *Brokerage* section of the survey. Detailed results from these factor models (i.e., final factor loadings) are included in Appendix D of this report. Once the final set of items for a scale were identified, internal consistency reliability was estimated using Cronbach's alpha, and scale scores were calculated as a simple average across items using raw item responses (i.e., coded 1, 2, 3, 4). Descriptive statistics and reliability estimates for the survey scales for the *Brokerage* portion of the survey are included in Table 9.

Table 9. ***Descriptive Statistics and Reliability of Survey Scales***

Construct	N	# of Items	α	Range	Mean	Standard Deviation
Brokerage-Information Sharing						
Research Products	3824	5	0.85	1-4	1.61	0.62
Local Knowledge	3743	4	0.83	1-4	2.03	0.78
Experts/State and Federal Sources	3830	4	0.86	1-4	1.66	0.68
Brokerage - Capacity Building	3891	4	0.92	1-4	1.70	0.79
Brokerage - Interpretation	2104	3	0.84	1-4	2.70	0.72
Brokerage - Activities	2003	10	0.91	1-4	2.22	0.69

Conclusions

Results from our work to develop and validate the *Survey of Evidence in Education for Schools* (*SEE-S*) suggest three primary conclusions and implications. The first conclusion is that the *SEE-S* measures have demonstrated validity and reliability, as evidenced by the content and face validity of the items, the construct validity demonstrated through exploratory and confirmatory factor analyses, and the moderate to high reliability for each of the final scales and subscales. Second, the graphical displays and descriptive statistics for the scales and items presented in this technical report suggest that direct engagement with external research is relatively low in general, and that there are some critical gaps separating the research and practice communities. For example, the *Depth of Use* scores show that external research was used less often than local research and local data analysis for organizational decisions. Also, school-based practitioners more often report using professional development materials than they do actual research products. However, there are aspects of connections between and brokerage of evidence between the research and practice communities that hold promise for improving the use of research. Implications of these results are discussed in detail in a separate descriptive research report. The third, which is more of an implication, is that the observed variation in *Depth of Use* scores and subscale scores at the school level and individual level begs for additional analyses in order to explore in what contexts and under what conditions schools and individual educators are more likely to engage in deeper use of research. The fourth implication is that the *SEE-S* survey and scales hold promise as a tool to be used by researchers, district leaders, and state agencies for measuring schools' engagement with research, and for monitoring improvement in use of research as efforts to build this capacity are undertaken.

Limitations

The field trial sample of the *Survey of Evidence in Education for Schools* consisted of 154 schools across the United States. Though the sample is not a random sample of the nation's schools, the field trial included a sample of schools from eighteen different states that mirror the national proportions of elementary, middle, and high schools, and represent districts from urban, suburban, and rural settings. Still, the large number of schools who were initially sampled but

declined to participate in the survey suggests that the final sample should not be considered as truly representative of the nations' schools.

Another limitation of this study is the reliance on self-reporting, which can lead to socially desirable responses. While many of our items include follow-up items that were used as validity checks and filters, we are still reliant on individuals' reports of what is happening in their schools. The case-studies conducted as part of this larger project allow us to explore the role of research in decision-making in schools in more detail and depth, albeit again through the perspective of those participating in interviews. Results from the case studies are presented in another research report from our center (E.g., one case study is How is evidence enacted in schools? A mixed methods multiple case study of "deep use" schools (Farley-Ripple et al., 2022).)

Some additional limitations include variation in individuals' conceptualizations of research (and our survey's inability to capture that well), our focus on instrumental use of research (as distinct from other ways in which research can be used), and the potential tensions and distinctions between individual versus organizational use of research.

Ultimately, we believe this study was successful in developing and validating a large-scale survey of research use in schools, despite the limitations inherent in survey research. We hope that other researchers and practitioners find the survey and the data it yields to be informative and helpful.

Information for Others Seeking to Use the *Survey of Evidence in Education for Schools*

If you are interested in using the *Survey of Evidence in Education for Schools,* please contact the Center for Research Use in Education at the University of Delaware at crue-info@udel.edu.

References

Allen, C. D., & Penuel, W. R. (2015). Studying teachers' sensemaking to investigate teachers' responses to professional development focused on new standards. *Journal of Teacher Education, 66*(2), 136–149.

Backer, T. E. (1993). Information alchemy: Transforming information through knowledge utilization. *Journal of the American Society for Information Science, 44*(4), 217-221.

Bass, B. M. (1983). *Organizational decision making.* Homewood, IL: Richard D. Irwin, Inc.

Behrstock-Sherratt, E., Drill, K., & Miller, S. (2011). Is the supply in demand? Exploring How, When, and Why Teachers Use Research. Washington, DC: American Institutes for Research.

Bertrand, M., & Marsh, J. A. (2015). Teachers' sensemaking of data and implications for equity. *American Educational Research Journal*, *52*(5), 861–893.

Broekkamp, H. & Hout-Walters, B. V. (2007). The gap between educational research and practice: A literature review, symposium and questionnaire. *Educational Research and Evaluation, 13*(3), 203-220

Campanelli, P. (1997). Testing survey questions: New directions in cognitive interviewing. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *55*(1), 5–17. https://doi.org/10.1177/075910639705500103

Caplan, N. (1979). The two-communities theory and knowledge utilization. *American Behavioral Scientist, 22*(3), 459-470.

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145-170.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3-12.

Coburn, C. E., Touré, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making at the district central office. *Teachers College Record, 111*(4), 1115–1161.

Coburn, S. E., & Stein, M. K. (2010). *Research and practice in education. Building alliances, bridging the divide.* Lanham, MA: Rowman & Littlefield Publishers, Inc.

Coburn, C. E. & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education, 112*(4), 469-495.

Conrad, F. & Blair, J. (2004). Data quality in cognitive interviews: The case for verbal reports. In S. Presser, J. Rothgeb, M. Couper, J. Lessler, E. Martin, J. Martin, E. Singer (Eds.), *Questionnaire Development Evaluation and Testing Methods.* 67–87. Hoboken, NJ: John Wiley and Sons.

Corcoran, T. B., Fuhrman, S. H., & Belcher, C. L. (2001). The district role in instructional improvement. *Phi Delta Kappan, 83*(1), 78–84.

Corcoran, T., McVay, S., & Riordan, K. (2003). *Getting it right: The MISE approach to professional development.* Philadelphia, PA: Consortium for Policy Research in Education.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika 16***,** 297–334. https://doi.org/10.1007/BF02310555

David, J. L. (1981). Local uses of Title I evaluations. *Educational Evaluation and Policy Analysis, 3*(1), 27-39.

Davies, H. T. O., & Nutley, S. (2008). *Learning more about how research-based knowledge gets used.* William T. Grant Foundation: New York, NY

Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies, 47*(2), 108-121.

Dedoose Version *9.0.17*, web application for managing, analyzing, and presenting qualitative and mixed method research data *(2021)*. Los Angeles, CA: SocioCultural Research Consultants, LLC www.dedoose.com.

Dunn, W.N. (1980). The two-communities metaphor and models of knowledge use. *Knowledge, 1*, 515-536.

Farley-Ripple, E.N. (2008a). *Reconsidering use: Exploring the depth of evidence use practices in district-level decision-making.* [Paper presentation]. Annual Meeting of the American Educational Research Association 2008, New York City, New York.

Farley-Ripple, E.N. (2008b). *Accountability, Evidence and School District Decision-Making.* University of Pennsylvania, unpublished doctoral dissertation.

Farley-Ripple, E.N. (2012). Research use in central office decision-making: A case study. *Education Management, Administration and Leadership*, *40*(6), 784-804.

Farley-Ripple, E.N., & Cho, V. (2014). Depth of use: How district decision-makers did and did not engage with evidence. In A. R. Shoho, B. Barnett, & A. Bowers (Eds.), *Using Data in Schools to Inform Leadership and Decision Making. International Research on School Leadership Series, Vol. 5.* Charlotte, NC: Information Age Publishing, Inc.

Farley-Ripple, E.N., May, H., Karpyn, A., Tilley, K., & McDonough, K. (2018) Rethinking connections between research and practice in education: A conceptual framework. *Educational Researcher, 47*(4): 235-245. https://doi.org/10.3102%2F0013189X18761042

Farley-Ripple, E. N., Tilley, K., Mead, H., Van Horne, S., & Agboh, D. (2022) How is evidence use enacted in schools? A mixed methods multiple case study of “deep-use” schools, A Report from the Center for Research Use in Education. University of Delaware. Supported by the Institute of Education Sciences, U.S. Department of Education, through

Grant R305C15001. Retrieved on August 10, 2022 from http://www.research4schools.org/wp-content/uploads/2022/06/Case-Studies-of-Deep-Use-Schools-Final-Report-2022.pdf

Finnigan, K. S., Daly, A., & Che, J. (2012). *The acquisition and use of evidence district-wide.* [Paper presentation]. Annual Meeting of the American Educational Research Association 2012, Vancouver, Canada.

Finnigan, K. S., Daly, A. J., & Che, J. (2013). Systemwide reform in districts under pressure: The role of social networks in defining, acquiring, using, and diffusing research evidence. *Journal of Educational Administration, 51*(4), 476–497. https://doi.org/10.1108/09578231311325668

Finnigan, K. S., & Daly, A. J. (2014). Conclusion: Using research evidence from the schoolhouse door to Capitol Hill. In K.S. Finnigan & A.J. Daly (Eds.), *Using research evidence in education* (pp. 177-187). New York: Springer International Publishing.

Gross, B., Kirst, M., Holland, D., & Luschei, T. (March, 2005). Got you under my spell? How accountability policy is changing and not changing decision making in high schools. *Holding high hopes: How high schools respond to state accountability policies* (pp.43-80). Philadelphia: Consortium for Policy Research in Education.

Hemsley-Brown, J. (2009). Using evidence to support administrative decisions. In Kowalski, T. J., & Lasley, T. J. II (Eds.), *Handbook of data-based decision making in education* (pp. 272-285). New York: Taylor & Francis.

Hill, H., & Briggs, D. (2020). Education leaders? Knowledge of causal research design: A measurement challenge. *Annenberg EdWorking Papers.*

Honig, M. I. (2003). Building policy from practice: District central office administrators' roles and capacity for implementing collaborative education policy. *Educational Administration Quarterly, 39*(3), 292-338

Honig, M. I., & Coburn, C. (2008). Evidence-based decision making in school district central offices: Toward a policy and research agenda. *Educational Policy*, *22*(4), 578-608.

Honig, M. I., & Venkateswaran, N. (2012). School-central office relationships in evidence use: Understanding evidence use as a systems problem. *American Journal of Education, 118*(2), 199-222.

Honig, M. I., Venkateswaran, N., & McNeil, P. (2017). Research use as learning: The case of fundamental change in school district central offices. *American Educational Research Journal*, *54*(5), 938–971.

Huberman, M. (1990). Linkage between researchers and practitioners: A qualitative study. *American Educational Research Journal, 27*(2), 363-391.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Kennedy, M. M. (1982). *Working knowledge and other essays*. Cambridge: The Huron Institute.

Landry, R., N. Amara, & M. Lamari. (2001). Utilization of social science research knowledge in Canada. *Research Policy, 30*, 333-349.

Lavis, J. N., Robertson, D., Woodside, J. M., McLEOD, C. B., & Abelson, J. (2003, June 6). *How can research organizations more effectively transfer research knowledge to decision makers?* Wiley Online Library. https://doi.org/10.1111/1468-0009.t01-1-00052

Leithwood, K., & Steinbach, R. (1995). *Expert problem solving: Evidence from school and district leaders.* SUNY Press.

Malin, J. R., Brown, C., & Trubceac, A. S. (2018). Going for broke: A multiple-case study of brokerage in education. *AERA Open, 4* (2), 1-14. https://doi.org/10.1177/2332858418769297

Massell, D., Goertz, M. E., & Barnes, C. A. (2012). State education agencies' acquisition and use of research knowledge for school improvement. *Peabody Journal of Education, 87*(5), 609-626.

Maynard, L. M. (2006). The role of repetition in the practice sessions of artist teachers and their students. *Bulletin of the Council for Research in Music Education, 167,* 61-72.

National Research Council (2012). *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*. In S.R. Singer, N.R. Nielsen, and H.A. Schweingruber (Eds.), Washington, DC: The National Academies Press. https://doi.org/10.17226/13362.

Neal, J. W., Neal, Z. P., Kornbluh, M., Mills, K., & Lawler, J. (2015). Brokering the research-practice gap: A typology. *American Journal of Psychology, 56*(3–4), 422–435.

Neeleman, A. (2019). The scope of school autonomy in practice: An empirically based classification of school interventions. *Journal of Educational Change, 20*(1), 31-55. https://doi.org/10.1007/s10833-018-9332-5

Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., & Allen, A. R. (2016). *Findings from a National Study on Research Use among School and District Leaders.* (Technical Report No. 1). National Center for Research in Policy and Practice.

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). https://doi.org/10.1177/2332858417705370

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*. Thousand Oaks, CA: Sage Publications, Inc.

Reichardt, R. (2000). *The state's role in supporting data-driven decision-making: A view of Wyoming.* Aurora, CO: Mid-Continent Research for Education and Learning.

Scott, J., Lubienski, C., DeBray, E., Jabbar, H. (2014). The intermediary function in evidence production, promotion, and utilization: The case of educational incentives. In K.S. Finnigan & A. J. Daly (Eds.), *Using research evidence in education* (pp. 69–89). New York, NY: Springer International Publishing.

Spillane, J. P. (1998). State policy and the non-monolithic nature of the local school district: Organizational and professional considerations. *American Educational Research Journal 35*(1), 33-63

Supovitz, J. A., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement.* University of Pennsylvania: Consortium for Policy Research in Education.

U.S. Department of Education (2019). *Digest of Education Statistics,* 2018. National Center for Education Statistics (NCES 2020-009), Table 105.50.

Weick, K. E. (1995). *Sensemaking in organizations.* Thousand Oaks: Sage Publications.

Weiss, C.H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 65-92). Washington, DC: The Aspen Institute.

West, R. F., & Rhoton, C. (1994). School district administrators' perceptions of educational research and barriers to research utilization. *ERS Spectrum, 12*(1), 23-30.

Williams, D., & Cole, L. (2007). Teachers' approaches to finding and using research evidence: An information literacy perspective. *Educational Research, 49*(2), 185-206.

Willis, G. (2013), Cognitive aspects of survey methodology. In P. Lavrakas (Ed.), *Encyclopedia of survey research methods*.

Willis, G. B., Artino, A. R. (2013). What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *Journal of Graduate Medical Education, 5*, 353-356.

Appendix A: Summary Statistics of Survey Scales

Table A1. Descriptive Statistics and Reliability of Survey Scales

Construct	N	# of Items	Cronbach's Alpha (α)	Range	Mean	Standard Deviation
Problems of Practice	4112	5	0.88	1-4	2.68	0.61
Processes and Incentives	4042	6	0.89	1-4	2.51	0.61
Processes	4081	3	0.86	1-4	2.62	0.66
Incentives	4068	3	0.82	1-4	2.41	0.67
Products						
Media	4006	5	0.83	1-4	1.96	0.75
Research	4046	4	0.91	1-4	1.60	0.69
Professional Resources	4129	3	0.71	1-4	2.23	0.62
Products - Characteristics						
Ease of Access	4114	6	0.90	1-4	3.42	0.65
Helpful Features	4128	4	0.84	1-4	2.76	0.79
Trustworthiness and Credibility						
Reputation-Based Trustworthiness	4153	4	0.80	1-4	2.47	0.80
Methods-Based Credibility	4112	5	0.89	1-4	3.19	0.72
Confidence to Critique Research	4082	7	0.97	1-4	2.17	0.81
Brokerage-Information Sharing						
Research Products	3824	5	0.85	1-4	1.61	0.62
Local Knowledge	3743	4	0.83	1-4	2.03	0.78
Experts/State and Federal Sources	3830	4	0.86	1-4	1.66	0.68
Brokerage - Capacity Building	3891	4	0.92	1-4	1.70	0.79
Brokerage - Interpretation	2104	3	0.83	1-4	2.70	0.72
Brokerage - Activities	2003	10	0.91	1-4	2.22	0.69

Appendix B: Difference in Problem & Decision Coding Frameworks for Paths A and B

A pattern emerged in the reporting of problems at the individual-level, which was not seen at the organizational-level, of respondents stating that there was no problem and the decision made was proactive in nature. For example, comments like “It wasn’t necessary, just trying something new,” or “Not necessary but our school and district is heading in the direction of offering on-line courses and blended learning options so I was ok with trying it out.” Upon recognizing this significant pattern, the new problem code “Proactive Improvement” was added to the framework to capture these examples. The decision was made to nest the code “Federal/state/local mandate” as a child code of “System,” rather than leaving it as a parent code as in the Path-A data. This decision was made due to the infrequent reporting of mandates, or systemic problems in general, motivating decisions at the individual-level. In other words, unlike the Path-A data, in the Path-B data the code “Federal/state/local mandate” was not prevalent enough to justify parent code status. Similarly, the decision code “Human Resources” was eliminated from the coding framework for Path-B as these types of decisions would be made at the organizational-level and not the individual-classroom level.

Appendix C: Research Citation Coding Rubric

SEE-S ITEM: You indicated that articles, reports, books, or summaries based on independent research or program evaluation influenced the decision:

Please tell us about a research or evaluation study that influenced the decision, including as much information as possible about the study author, title, or web URL.

PRIMARY CODE: Confirmed_Research

CRITERIA	**CODE**
The text-field response includes sufficient information to identify: (1) a specific publication (e.g., book, article, report, etc.) which has been cited in Google Scholar at least 10 times, **OR** (2) an author who has been cited in Google Scholar at least 10 times, **OR** (3) an organization that cites on their website at least one publication (e.g., book, article, report, etc.) which has been cited in Google Scholar at least 10 times, **OR** (4) a specific publication (e.g., dissertation, press article, blog) with embedded references including at least one publication that has been cited in Google Scholar at least 10 times.	YES
Otherwise	NO

SECONDARY CODE: Direct_or_Indirect_Citation

Coded only when the PRIMARY CODE (Confirmed_Research) is **YES**.

The purpose of this code is to record the category of the entity cited in the text-field response.

CRITERIA	**CODE**
Text-field response includes: (1) the name of an author (or authoring organization) that has been cited in Google Scholar at least 10 times, **OR** (2) the title of a publication that has been cited in Google Scholar at least 10 times, **OR** (3) the title of a publication with at least one author (or authoring organization) who has been cited in Google Scholar at least 10 times.	DIRECT
Text-field response includes the name of an individual or entity that cites research produced by others.	INDIRECT

Note that the DIRECT code will apply, almost exclusively, to categories 1 & 2 under the YES code for the Confirmed_Research PRIMARY CODE *above.*

TERTIARY CODE: **Research_Citation_Category**

Coded only when the PRIMARY CODE (Confirmed_Research) is **YES**.

The purpose of this code is to record the category of the entities cited in the text-field response. Multiple codes can be recorded in this field (e.g. "Book" and "Author").

Research Citation Category: Format

CRITERIA	CODE
Text-field response includes the title of, or a link that leads to, an article. This includes articles published in peer-reviewed journals, magazines, newspapers, practitioner journals, etc.	Article
Text-field response includes the name of, or a link to, a blog.	Blog
Text-field response includes the title of, or link that leads to, a book.	Book
Text-field response includes the title of an evaluation report or research report.	Evaluation/Research Report
Text-field response includes the name of, or a link to, a audio/visual material (e.g., video, podcast, etc.)	Multimedia
Text field response includes the name of, or a link to, material from a professional development workshop/training (e.g., a PDF, PowerPoint, Word doc).	Professional Development Material
Text-field response includes the name of, or a link to, a website (without citing/pointing to a particular piece of information on that website, e.g., response simply provides ies.ed.gov not a link to a particular article or resource on the site).	Website
Text-field response includes the name of, or a link to, some other format of information not represented by the above codes. These will be reviewed by the team to make sure the other code is warranted.	Other

Research Citation Category: Source

CRITERIA	CODE
Text-field response includes the name of, or a link to the website of/information published by, an advocacy group. These groups have a special interest and use advocacy to influence public opinion & policy (e.g., Learning Disabilities Assoc of America, Stand for Children).	Advocacy Group
Text-field response includes the name of an author.	Author
Text-field response includes the name of, or a link that leads to the website of/info published by, the local education agency's central office or its staff.	Central Office/School District
Text-field response cites a consultant.	Consultant
Text-field response cites a Federal Agency, a Federal Technical Assistance Center, a Federal Comprehensive Center, a Regional Educational Laboratory, or a similarly commissioned organization or its staff.	Federal Agency
Text-field response includes the name of, or a link that leads to the website of/info published by, a newspaper, news website, news magazine (e.g., EdWeek, NY Times, US News, etc.)	Popular Press
Text-field response includes the title of, or a link that leads to the website of/an article published by, a practitioner-focused journal (e.g., Ed Leadership, PDK, etc.)	Practitioner Journal, Magazine, or Similar Publication
Text-field response includes the name of, or a link that leads to the website of/info published by, a professional association (e.g., NEA, AFT, NASSP)	Professional Association
Text-field response includes the name of, or a link that leads to the website of/info published by, a professional development provider, program developer, or publisher (e.g., Pearson, Scholastic, LSI)	Professional Development Provider/Program Developer/Publisher

Research Citation Category: Source (continued)

CRITERIA	CODE
Text-field response includes the name of, or a link that leads to, the website of/info published by, an independent research firm (e.g., AIR, Mathematica, WestEd)	Research Firm
Text-field response includes name of, or a link that leads to, a peer-reviewed academic journal or article.	Peer-reviewed Journal
Text-field response cites an LEA or SEA research\evaluation department or office.	Research\Evaluation Office
Text field response includes the name of, or a link that leads to the website of/info published by a state education agency or its staff, or by an intermediary (regional within a state) service agency or its staff (e.g., regional service center, intermediate unit)	State Agency or Intermediary Unit
Text-field response includes the name of, or a link that leads to the website of/info published by, a research center based at a university (e.g., CRESP, CPRE, NCRPP)	University Research Center
Text-field response includes the name of, or a link to, some other source of information not represented by the above codes. These will be reviewed by the team to make sure the other code is warranted.	Other
List multiple categories as follows, separated by semicolons.	MULTIPLE: entry1; entry2

Appendix D: Factor Loadings for Survey Scales

To examine the overall fit of the factor model, we conducted a confirmatory factor analysis (CFA) of all factors with the exception of the individual sub-scales for *Processes* and *Incentives*. There was evidence that the overall model was a good fit to the data: the RMSEA was 0.04, the Bentler CFI was 0.91, and the Bentler-Bonett NNI was 0.90. The factor loadings in the CFA of the overall model were similar to the factor loadings produced in separate analyses of each scale. We present the factor loadings from individual analyses of factors because we predict that researchers will be interested in adopting a sub-set of scales and will be more interested in the factor loadings of individual scales rather than those from a model that includes all factors.

Factor Loadings for *Problems of Practice* Scale

Scale	Variable	Factor Loading
Problems of Practice	Most education research suggests actionable steps to take in practice	0.62
	Researchers have a solid grasp on evolving problems in schools/districts.	0.86
	Research addresses the most important issues schools/districts face.	0.86
	Research takes into consideration the varying levels of resources available to schools/districts to implement research findings.	0.81
	Research is produced quickly enough for me to make use of it.	0.69

Note: Item correlations range from 0.44–0.76 with a median of 0.58.

Factor Loadings for *Processes and Incentives* Scale

Scale	Variable	Factor Loading
Processes and Incentives	Our school/district has a documented process (e.g. guidelines) for using research to inform decisions.	0.78
	Our school/district provides time to discuss research.	0.81
	Our school/district prioritizes research in decision making.	0.79
	There are school/district incentives for me to use research in my practice.	0.79
	Using research is part of my evaluation as a practitioner.	0.71
	We use research because a supervisor or administrator requires it.	0.67

Note. Item correlations range from 0.48–0.68 with a median of 0.58.

Factor Loadings for *Processes and Incentives* Sub-Scales

Item	Processes	Incentives
Our school/district has a documented process (e.g., guidelines) for using research to inform decisions.	**0.78**	0.03
Our school/district provides time to discuss research.	**0.89**	-0.05
Our school/district prioritizes research in decision making.	**0.74**	0.07
There are school/district incentives for me to use research in my practice.	0.29	**0.54**
Using research is part of my evaluation as a practitioner.	-0.05	**0.85**
We use research because a supervisor or administrator requires it.	0.02	**0.71**

Note. Item correlations for *Processes* sub-scale range from 0.64–0.68 with a median of 0.68, and 0.57–0.63 with a median of 0.59 for the *Incentives* sub-scale.

Factor Loadings for *Products* Scales

Item	Research	Media	Professional Resources
Professional development materials	-0.02	0.01	**0.76**
Conferences/presentations	0.02	-0.02	**0.66**
Books	0.21	0.16	**0.39**
News (e.g., in print or online)	0.03	**0.48**	0.30
Magazine articles	0.19	**0.41**	0.21
Blogs	0.07	**0.77**	-0.12
Multimedia (e.g., podcast, videos)	0.07	**0.58**	0.19
Posts from social media (e.g., pins, tweets)	-0.08	**0.77**	-0.01
Peer-reviewed academic journals	**0.68**	0.12	0.02
Research summaries/briefs	**0.87**	-0.01	0.02
Research/program evaluation reports	**0.91**	-0.07	0.00
Reviews of multiple research studies	**0.92**	-0.01	-0.03

Note: The item correlations for the *Research*, *Media,* and *Professional Resources* sub-scale range from 0.62–0.80 with a median of 0.74, 0.38–0.65 with a median of 0.50, and 0.39–0.54 with a median of 0.43, respectively.

Factor Loadings for *Characteristics of Products* Sub-Scales

Item ("How important to you are each of the following product characteristics from a research study? The research product…"	Ease of Access	Helpfulness of Features
is actionable (i.e., provides instructions).	**0.65**	0.06
is concise.	**0.79**	0.01
is easy to understand.	**0.96**	-0.10
is easy to access.	**0.94**	-0.07
is free to access.	**0.74**	0.04
can be found online.	**0.55**	0.21
is transmitted verbally.	-0.08	**0.70**
uses graphics to illustrate findings.	0.04	**0.78**
provides a demonstration of findings or models strategies.	0.20	**0.67**
is a summary of multiple studies of the same policy or practice.	0.03	**0.77**

Note. The item correlations for the *Ease of Access* and *Helpfulness of Features* sub-scales range from 0.42–0.82 with a median of 0.61 and 0.44–0.66 with a median of 0.59, respectively.

Factor Loadings for *Reputation-Based Trustworthiness* Scale

Item	Factor Loading
The study was conducted by a researcher or research firm you and/or your colleagues know.	0.58
The study is published in a peer-reviewed journal.	0.56
The study has been highlighted in a major national newspaper, on a national television network, or on a major radio program.	0.84
The number of times the study has been cited (e.g., as reported by Google Scholar).	0.82

Note. The item correlations range from 0.44–0.71 with a median of 0.45.

Factor Loadings for *Methods-Based Credibility* Scale

Item	Factor Loading
The research study produced statistically significant results. Statistically significant means that results are unlikely to occur by chance.	0.76
The sample size was large.	0.78
The study randomly assigned students/teachers/schools to a treatment or control condition.	0.73
The study reported findings for all outcomes, positive or negative.	0.82
The research described the methods used in detail.	0.81

Note. The item correlations range from 0.54–0.76 with a median of 0.59.

Factor Loadings for *Confidence to Critique Research* Scale

Item (How confident do you feel in determining whether…)	Factor Loading
a research study used appropriate statistical analyses?	0.87
a research study had an adequate sample size?	0.89
a program evaluation demonstrated real impacts versus improvement that would have happened even without the program?	0.91
the surveys and assessments used in a research study were reliable and valid?	0.92
results from a research study are generalizable to different schools, districts, etc?	0.92
results from a research synthesis (i.e., combined results across multiple research studies) are trustworthy?	0.92
research evidence provided by a vendor is trustworthy, versus slanted to support their products?	0.89

Note. The item correlations range from 0.76–0.86 with a median of 0.81.

Factor Loadings for *Brokerage - Information Sharing* Sub-scales

Item("In the last year, have you shared any of the following types of evidence with others?)	Research	Local Knowledge	Experts/State and Federal Sources
Articles, reports, books, or summaries based on independent research or program evaluation (paper or web-based)	**0.31**	0.30	0.15
Research or program evaluation conducted by central office staff	**0.80**	-0.09	0.10
Research or program evaluation conducted by teacher(s) or principal(s)	**0.86**	0.13	-0.13
Research or program evaluation led by students or local youth	**0.75**	-0.15	0.10
Other Formal analysis of a school-wide or district-wide data	**0.62**	0.16	0.06
Informal data collected by school/district staff	0.22	**0.44**	0.14
Materials from a program developer or publisher	0.14	0.34	**0.38**
Opinion of national experts	0.08	0.06	**0.66**
Guidance from federal or state departments of education	-0.04	-0.04	**0.93**
Advice from local education leaders (e.g., district superintendent)	0.12	0.16	**0.59**
Other practitioners' experiences/advice	-0.05	**0.75**	0.13
Opinions of parents or other community members	0.10	**0.52**	0.20
My own professional experience	-0.02	**0.88**	-0.08

The item correlations for *Research*, *Local Knowledge*, and *Experts/State and Federal Sources* range from 0.37–0.68 with a median of 0.57, 0.46–0.66 with a median of 0.53, and 0.57–0.68 with a median of 0.58, respectively.

Factor Loadings for *Interpretation* Scale

Item ("When you've shared these, how often have you shared…")	Factor Loading
An actual product (e.g., the article, a link to the article, etc.)	0.76
Your interpretation or summary of the findings.	0.82
Practices or strategies you developed based on the research.	0.81

Item correlations range from 0.59–0.72 with a median of 0.61.

Factor Loadings for *Capacity Building* Scale

Item ("When you've shared these, how often have you shared…")	Factor Loading
Sharing strategies for accessing research	0.86
Sharing strategies for reading and understanding research	0.88
Sharing strategies for implementing education research	0.88
Connecting people for the purpose of sharing or discussing research	0.85

Note. Item correlations range from 0.73–0.78 with a median of 0.75.

Factor Loadings for *Brokerage-Activities* Scale

Item ("When you have shared research, how often do you do any of the following?")	Factor Loading
Evaluate the quality of research prior to sharing	0.40
Evaluate needs of schools, teachers or others so that you select the most relevant research	0.43
Deliver formal learning opportunities (e.g., professional development, training)	0.67
Offer or provide support or technical assistance	0.66
Publish (i.e., produce or release for distribution)	0.76
Develop products or programs based on research	0.78
Disseminate (i.e., actively distribute research)	0.83
Synthesize multiple sources of research about a single topic, program, etc.	0.82
Translate research into understandable language and/or format	0.80
Facilitate discussion of research	0.80

Note. Item correlations range from 0.15–0.74 with a median of 0.51.

Please cite this report as follows:

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.* The Center for Research Use in Education (CRUE) & the Center for Research in Education and Social Policy (CRESP), University of Delaware. Available from http://crue.cehd.udel.edu/