Abstract Title Page

Title:

Individual and School-Level Capacity to Critically Evaluate Research: A Multilevel Organizational Analysis.

Authors:

Henry May, University of Delaware
Elizabeth Farley-Ripple, University of Delaware
Horatio Blackman, University of Delaware
Rui Wang, University of Delaware
Kati Tilley, University of Delaware
Samantha Shewchuk, University of Delaware

Background / Context:

Beginning with the *No Child Left Behind Act* of 2001 (NCLB) and reinforced by the *Education Sciences Reform Act* of 2002 (ESRA) and the *Every Student Succeeds Act* of 2015, the federal government created explicit expectations for the role of research in informing decisions about education programs, policies, and practice, and it also specified new expectations for what constituted research knowledge that was worth using (e.g., randomized experiments). Unfortunately, we don't know much about whether federal legislation has actually led schools and individual educators to make better and more frequent use of research to inform their decisions.

The IES-funded Center for Research Use in Education is developing and validating a survey to measure the role of different forms of evidence (including research) in school-based decisions through collection of multilevel data from district and school staff nationwide.

Purpose / Objective:

To motivate new approaches to increasing research use and evidence-based decision-making in schools, our team is conducting a series of research studies, the first of which involves developing and validating a survey to measure depth of research use by schools and teachers through collection of multilevel data from district staff, school administrators, and classroom teachers. In this paper, we present results from a large-scale survey of more than 3,000 educators from over 200 schools. The results presented in this paper focus on our survey scale of educators' capacity to critique research.

Setting:

Educators from a national sample of schools completed a battery of surveys during the 2016-17, 2017-18, 2018-19, and 2019-20 school years (data collection will conclude in fall 2019). The final sample of schools is approximately 20% rural, 50% suburban, and 30% urban.

Population / Participants:

The sampling frame included instructional staff (e.g., teachers, coaches) or responsible for supervising instruction (e.g., principals). The survey has an overall response rate of just over 50%, with many schools achieving response rates above 70%. The respondent sample includes about 4% school administrators, 2% district staff, 9% instructional coaches/specialists, 14% special education teachers, 59% additional classroom teachers, and 12% other instructional staff (e.g., ESL teachers).

Research Design:

This is a large-scale survey research study with three phases. The first two phases involved piloting the survey in separate samples of 30-schools. A nationally-representative field test was conducted in 2018-19 involving 100 schools and continues during the 2019-20 school year, with a final sample goal of 300 schools. The data for this paper will be augmented with data from the fall 2019 field trial prior to presenting at SREE in spring 2020.

Data Collection and Analysis:

Data were collected through online administration of the survey using the Qualtrics platform. Analyses involve descriptive statistics, t-tests, and two-level HLM analyses.

Findings / Results:

Psychometric analysis of the 11-item Capacity scale demonstrated very high reliability (Cronbach's Alpha = .97), with item-total correlations ranging from .82 to .89 and similar distributions of responses on the Likert scale across all items. The distribution of total scores was markedly skewed right, with the vast majority of respondents scoring in the "not confident" to "somewhat confident" range (see Figure 3). The mean score was 1.02 (i.e., "Somewhat Confident"), the standard deviation of scores was .77 points, and the interquartile range was 0.36 points to 1.54 points, suggesting that fewer than half of respondents were "Mostly" or "Very" confident to critically evaluate research. More specifically, only 26% of respondents had total scores equal to or greater than 1.5 (i.e., approaching or exceeding "Mostly Confident"), and only 4% of respondents had total scores equal to or greater than 2.5 (i.e., approaching "Very Confident").

However, school-level results were quite different from the individual level results. The vast majority of schools included one or more respondents with relatively high total scores on the Capacity scale. More specifically, 90% of schools had a least one respondent with a Capacity score of at least 1.5 (i.e., approaching or exceeding "Mostly Confident"), while 40% of schools had a least one respondent with a Capacity score of at least 2.5 (i.e., approaching "Very Confident").

Interestingly, staff role (teacher, administrator, etc.) was found to be unrelated to Capacity scores (F = 0.73, p = .60). However, prior training and experiences were found to be significant predictors of Capacity scores.

Experiences in degree programs and professional development (PD) that emphasized research were associated with substantially higher Capacity scores, with standardized effects just over one-half a standard deviation, or about one-third of the distance between "Not at All Confident" and "Somewhat Confident." Larger standardized effects were associated with research conference attendance (+.62) and participation in a formal Research-Practice Partnership (+.81). The experience of participating in a professional learning community (PLC) was not significantly associated with Capacity to critique research.

Only one of the Training items was associated with higher Capacity to critique research—taking a Research Design course as an undergrad or grad student was associated with a standardized effect of +.62 standard deviations, or about 40% of the distance between "Not at All Confident" and "Somewhat Confident." Having taken an Introductory Statistics Course or another course on understanding and interpreting research, despite being quite common, was not significantly associated with Capacity to critique research.

Conclusions:

These results from of our large-scale survey align well with recent results from previously published research and theory related to educators' capacity to critique research (e.g.,

Supovitz and Klein, 2003; West and Rhoton, 1994; Reichardt, 2000; Coburn & Talbert, 2006). Simply put, most educators report relatively low confidence in critiquing research. This occurs despite widespread participation in pre-service and in-service training and experiences intended (explicitly or implicitly) to improve engagement with and capacity to critically evaluate research. On the other hand, these results suggest that most, if not the majority of schools have at least one staff member who feels confident in their ability to critically evaluate research.

Appendices
Not included in page/word count.

Bibliography

Asen, R., Gurke, D., Connors, P., Solomon, R., & Gumm, E. (2012). Research evidence and schoolboard deliberations: Lessons from three Wisconsin school districts. *Educational Policy, 26* , 33–63.

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

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.

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

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.

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

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

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

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

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.

Hood, P. (2003). *Scientific research and evidence-based practice*. San Francisco, CA: WestEd.

Ingram, D., Seashore Louis, K., & Schroeder, R. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record 106* (6), 1258–87.

Kennedy, M. M. (1982a). Working knowledge. In M. M. Kennedy (Ed.), *Working knowledge and other essays* (pp. 1-28). Cambridge, MA: The Huron Institute.

Kennedy, M.M. (1982b). Evidence and decision. In M. M. Kennedy (Ed.), *Working knowledge and other essays* (pp. 59-103). Cambridge, MA: The Huron Institute.

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

Nelson, C. E., Roberts, J., Maederer, C., Wertheimer, B., & Johnson, B. (1987). The utilization of social science information by policy makers. *American Behavioral Scientist, 30*, 569-577.

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.

Rich, R. F. (1977). Use of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding. In C.H. Weiss, (ed.) *Using social research in public policymaking* (pp. 199-211). Lexington: Lexington Books.

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.

Wayman, J. C. & Stringfield, S. (2006). Technology-supported involvement of entire faculties in examination of student data for instructional improvement. *American Journal of Education 112*(4), 549-571.

Weiss, C. H. (1980). Knowledge creep and decision accretion. *Knowlege: Creation, diffusion, utilization, 1*(3), 381-404.

West, R. F., & Rhoton, C. (1994). School district administrators' perceptions of educational research and barriers to research utilization. *ERS Spectrum, 12*(1), 23-30.
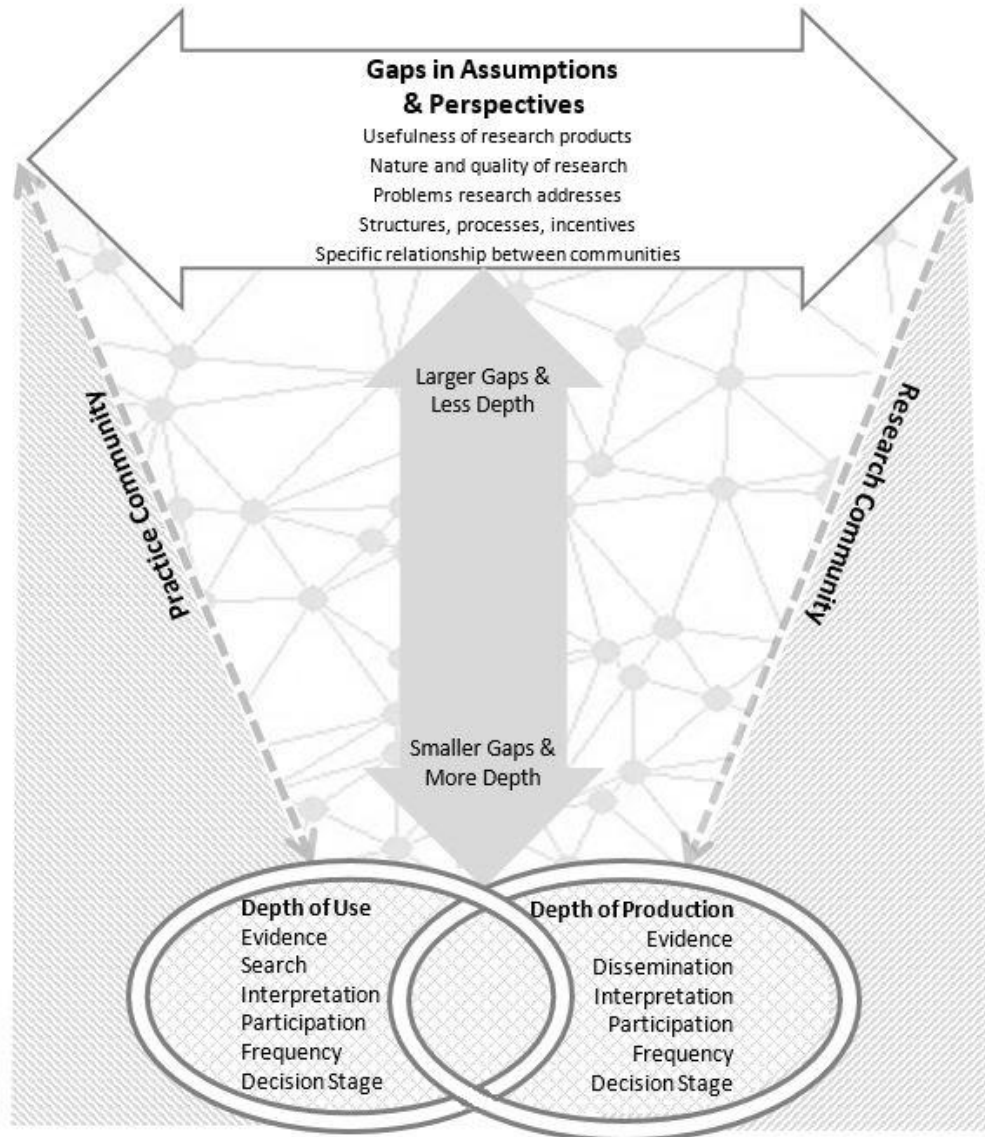
Gaps in Assumptions & Perspectives
Usefulness of research products
Nature and quality of research
Problems research addresses
Structures, processes, incentives
Specific relationship between communities

Practice Community

Research Community

Larger Gaps & Less Depth

Smaller Gaps & More Depth

Depth of Use
Evidence
Search
Interpretation
Participation
Frequency
Decision Stage

Depth of Production
Evidence
Dissemination
Interpretation
Participation
Frequency
Decision Stage

Figure 1. Conceptual Framework for Research Use in Schools

Please rate how confident you feel to determine whether…

	Not at all	Somewhat	Mostly	Very confident
…a research study conducted appropriate statistical analyses.	O	O	O	O
...a research design was appropriate for the research questions posed.	O	O	O	O
...a research study had an adequate sample size.	O	O	O	O
...results from a research study might be dismissed because they are actually attributable to something that the study missed.	O	O	O	O
...a program evaluation demonstrated real impacts versus improvement that would have happened even without the program.	O	O	O	O
… a comparison group is a good match to the treatment group.	O	O	O	O
...research supported (or not) inferences about the causal effects of a new program.	O	O	O	O
...the surveys and assessments used in a research study were reliable and valid.	O	O	O	O
…results from a research study are generalizable to different schools, districts, etc.	O	O	O	O
...results from a research synthesis (i.e., combining results across multiple research studies) are trustworthy.	O	O	O	O
...research evidence provided by a vendor is trustworthy, versus slanted to support their products.	O	O	O	O

Figure 2. Survey Items Measuring Educator Capacity to Critically Evaluate Research.
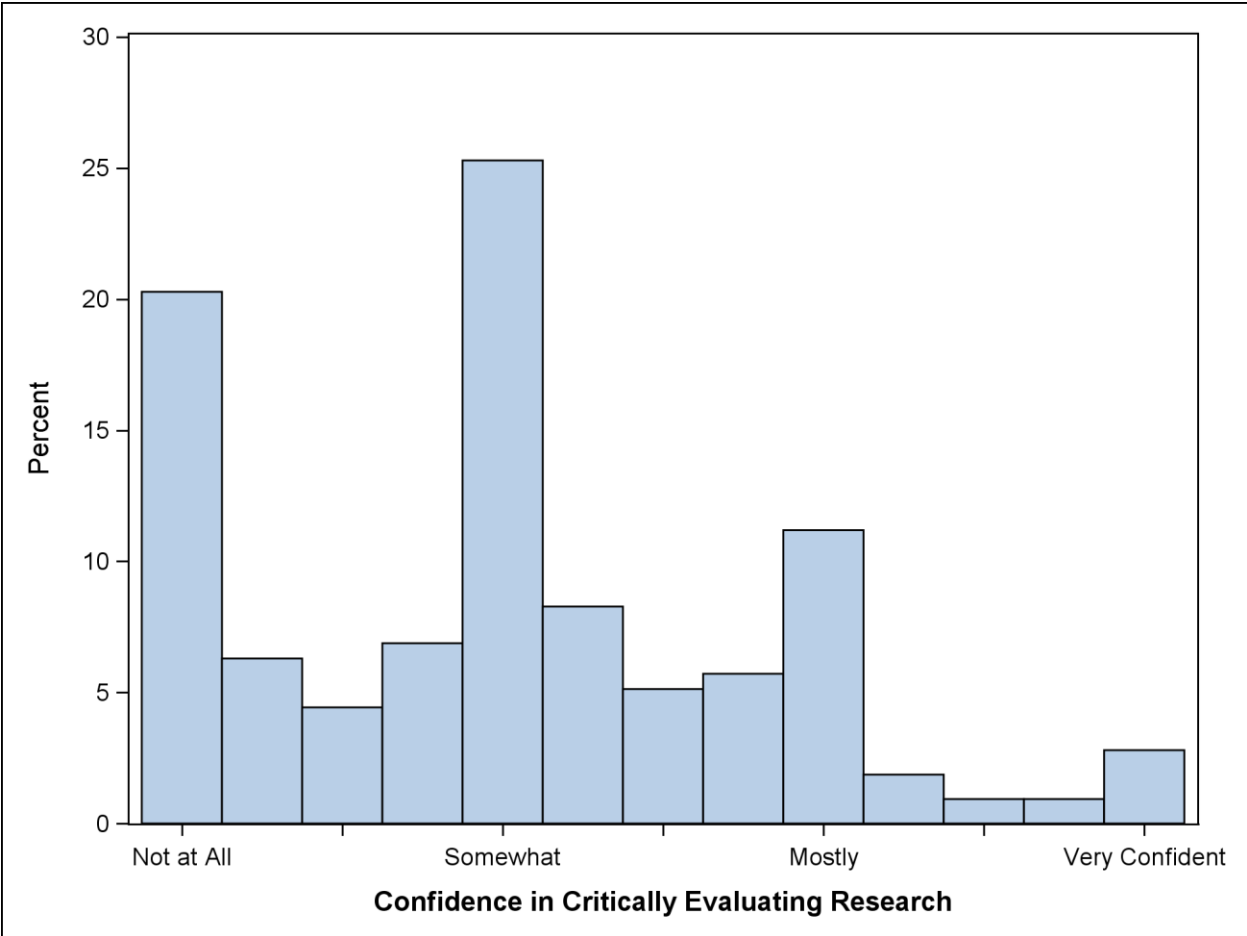
Figure 3. Histogram of Educator's Overall Scale Scores on Confidence in Critiquing Research.

Table 1. Educators' Overall Scale Scores on Confidence in Critiquing Research by 11 Indicators of Prior Experience and Training.

Prior Experience / Training	Percent Yes	Mean Total Capacity Score		Raw Difference	Standardized Difference
		Yes	No		
Experience (Yes/No)					
I was in a graduate program where I conducted research.	52%	1.35	0.88	+0.46***	+0.59
I was in an undergraduate/graduate program that heavily emphasized research use.	51%	1.34	0.87	+0.46***	+0.59
I review research and apply it in my own work.	46%	1.28	0.95	+0.33***	+0.43
I have engaged with research through a Professional Learning Community.	41%	1.16	1.03	+0.13	+0.17
I have conducted action research.	36%	1.41	0.95	+0.46***	+0.59
I have participated in other professional development around critically consuming research.	28%	1.41	0.98	+0.43***	+0.55
I attend research conferences.	21%	1.48	1.00	+0.48***	+0.62
I have been involved in a formal research-practice partnership.	12%	1.64	1.01	+0.63***	+0.81
Training (Yes/No)					
I have taken an Introduction to Statistics course.	75%	1.12	0.95	+0.17	+0.22
I have taken a Research Design course.	49%	1.39	0.91	+0.48***	+0.62
I have taken another course on understanding and interpreting research.	45%	1.20	1.07	+0.12	+0.15

Note. N = 858. Statistical significance adjusted for multiple comparisons using Bonferroni Correction; *p<.05, **p<.01, ***p<.001