American Journal of Evaluation
2014, Vol. 35(1) 95-99

DOI: 10.1177/1098214013499440
aje.sagepub.com
SAGE

Moving From Science to Practice in Evaluation Capacity Building

Yolanda Suarez-Balcazar[1] and Tina Taylor-Ritzler[2]

Abstract

A synthesis of the state of the literature is discussed in this section of the Evaluation Capacity Building (ECB) forum organized around four critical questions: (1) What is ECB? (2) How can we make it happen? (3) How do we know it is happening? and (4) What is its impact? The authors argue that to move the field of ECB forward we need to envision the science of ECB, not as the sole activity of creating new knowledge but in a close congruent and reciprocal relationship with practice. By adhering to a science-practice model, we conduct research that directly responds to and contributes to practice thus creating strong synergies between ECB practitioners and researchers. Research on ECB needs to be informed by real issues happening in practice, and practice of ECB needs to be informed by the new knowledge created. We must strengthen the science to refine our practice and strengthen the practice to refine the science.

Keywords

evaluation capacity building, science, practice, capacity building

The current vast body of literature about evaluation capacity building (ECB) has primarily focused on answering four critical questions: (1) What is ECB? (2) How can we make it happen? (3) How do we know it is happening? and (4) What is its impact? The first of these questions addresses conceptualization and definition; the second question addresses issues of process; the third question addresses measurement; and the final question addresses outcomes. We argue that to move science forward we need to make a stronger connection between the science and the practice of ECB.

What Is ECB?

Much has been written about the conceptualization of ECB. It has been widely conceptualized as a dynamic and complex contextual process. One of the most frequently used definitions comes from Stockdill, Baizerman, and Compton (2002) who affirmed that ECB is "intentional work to

[1] University of Illinois at Chicago, Chicago, IL, USA
[2] Dominican University, River Forest, IL, USA

Corresponding Author:
Yolanda Suarez-Balcazar, University of Illinois at Chicago, 1919 West Taylor St., Chicago, IL 60612, USA.
Email: ysuarez@uic.edu

continuously create and sustain overall organizational processes that make quality evaluation and its uses routine.'' During the last decade, evaluators have increasingly used ECB to assist nonprofit organizations in developing evaluation skills in order to document the impact of their programs (Baizerman, Compton, & Stockdill, 2002; Carman & Fredericks, 2008; Compton, Glover-Kudon, Smith, & Avery, 2002; Garcia-Iriarte, Suarez-Balcazar, & Taylor-Ritzler, 2011; Huffman, Thomas, & Lawrenz, 2008; King, 2002; Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012; Milstein, Chapel, Wetterhall, & Cotton, 2002; Stevenson, Florin, Mills, & Andrade, 2002; Stockdill, Baizerman & Compton, 2002). Although a number of complex conceptual models are available to further explicate what ECB is and what factors are involved (e.g., Connolly & York, 2002; Cousins, Goh, Clark & Lee, 2004; Nielsen, Lemire, & Skov, 2011; Preskill & Boyle, 2008), slight attention has been given to the empirical validation of such models. One such effort is the model proposed by Suarez-Balcazar et al. (2010) and validated by Taylor-Ritzler, Suarez-Balcazar, Garcia-Iriarte, Henry, and Balcazar, (2013).

Future research needs to focus on strengthening existing models, validating those that are not validated, and further applying validated models to larger sample sizes and to diverse organizations. Furthermore, current models may be strengthened by establishing stronger relationships between existing models and their corresponding assessment instruments as to integrate and synthesize currently agreed upon components of evaluation capacity (Labin et al., 2012; Taylor-Ritzler, Suarez-Balcazar, Garcia-Iriarte, Henry, & Balcazar, 2013).

How Can We Make It Happen?

ECB has become a prevalent activity among evaluators. Several studies, particularly those describing case studies and examples of successful ECB efforts, are available in the literature (e.g., Garcia-Iriarte et al., 2011; King & Volkov, 2005).

ECB practices generally involve an evaluator providing training, technical assistance, consultation, and other activities to one or more staff within an organization or system (Duffy, Labin, & Wandersman, 2007; Garcia-Iriarte et al., 2011). The goals of ECB practices are typically for staff within the target organization to regularly document the implementation of their programs, and interpret the results, in order to understand and strengthen program implementation, to improve program outcomes, and to meet the accountability requirements of funders and accrediting bodies (Gibbs, Napp, Jolly, Westover, & Uhl, 2002; King, 2002; Mackay, 2002; Milstein et al., 2002; Preskill & Boyle, 2008; Taut, 2007). Although some efforts are available to examine theoretical frameworks to explain the training of evaluation skills, mostly adult learning theories and transformational learning models, future research need to empirically identify the learning theories underpinning the building of evaluation skills and knowledge.

How Do We Know It Is Happening?

The measurement of ECB has been somewhat complex. Several ECB and related assessment instruments have been suggested in the literature which mostly focus on Congressional Budget Office staff perceived knowledge of their capacity to conduct an evaluation and their knowledge of the evaluation processes (e.g., Botcheva, White, & Huffman, 2002; Cousins, Goh, & Elliott, 2007; Danseco, Halsall, & Kasprzak, 2009; Preskill & Torres, 2000; Taut, 2007; Texas Christian University, 2005; Volkov & King, 2007). Taylor-Ritzler et al., (2013) conducted one of the few empirical validations of an ECB assessment instrument that measured organizational factors, individual factors, and utilization and sustainability factors. Despite the availability of this validated instrument, existing evidence suggests that we do not know enough about how to measure contextual and cultural factors that impact ECB.

What Is the Impact of ECB?

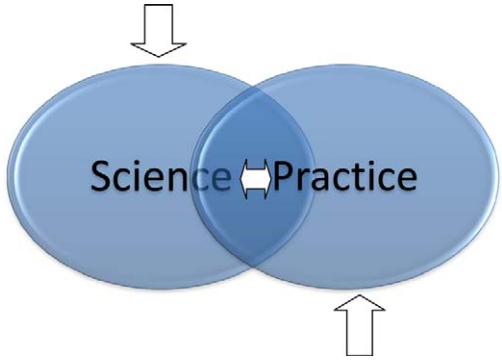
Knowing how to measure ECB does not necessarily tell us if programming is successfully strengthened at the organizational level, nor does it necessarily reveal if ECB is sustained over time. Ultimately, our efforts must address a different question—the fundamental question that ECB was designed to address—does building organizational evaluation capacity lead to better program outcomes? Therefore, we must scientifically evaluate whether organizations use evaluation results and evaluative thinking to improve what they do in a way that benefits consumers?

The evidence base for the utility of ECB in improving organizational program outcomes is, at best, nascent. There are now several theoretical models of ECB and even a validated instrument. However, as Labin, Duffy, Meyers, Wandersman, and Lesesne (2012) point out, most of the empirical research involves descriptive case study designs and little use of reliable and valid instruments. As a result, although ECB holds much promise in improving program outcomes for program participants and accountability for organizations, there is, as yet, little evidence that ECB practices actually help organizations systematically improve services that lead to positive outcomes for their program participants. Here we make a call to the field to advance our knowledge of ECB. We argue that without a strong science, practice will continue to advance in uneven and unsubstantiated ways that cannot be built upon and further developed and refined.

Moving the Field From Science to Practice to Practice to Science

We posit that to move the field forward we need to envision the science of ECB, not as a sole activity in itself of creating new knowledge but in close congruent relationship with practice. The model below illustrates the essence of our proposal.

By adhering to a science-practice model, we are committed to conducting research that directly responds to and contributes to practice and creates strong synergies between ECB practitioners and researchers. This means that the development of theory, both in ECB conceptual models and measurement, as well as the how to at the organizational level must be intertwined and must inform one another in a dynamic, fluid, contextual, and continuous process. Research on ECB needs to be informed by real issues happening in practice, and practice of ECB needs to be informed by the new knowledge created. In order to strengthen ECB, we need to create partnerships with organizations conducting ECB and offer measurement instruments they can use to enhance their practice of evaluation. Methodologically sound measurement instruments assessing critical components and domains of ECB can serve as diagnostic tools for organizations and produce practical results for evaluation practitioners informing the areas in need of intervention. We must strengthen the science to refine our practice and strengthen the practice and write about it to refine the science.



To build a strong ECB science, we must address the shortcomings of our current state of knowledge. The next generation of ECB studies should be ambitious. We now need to explicate the adult learning theories that underlie our work to build the capacity of individuals and organizations and better identify the intervention components of our ECB efforts. We must find the needed dosage of training and support to make ECB effective and sustainable. Strengthening our study designs to make them more reliable can help us identify ECB effects and more reliably attribute them to our ECB interventions.

Additionally, to strengthen our science, we need to adopt study designs that build on what we currently know. The future of ECB efforts need to include quasi and true experimental designs that are longitudinal in nature and that include multiple respondents from numerous programs and organizations. Perhaps even randomize control trials in which an organization follows an ECB process and another organization follows a different model—maybe consultant-based evaluation or traditional training. In summary, we should use stronger research designs to investigate subsequent critical research questions: Can organizational evaluation capacity be sustained over time? Under what conditions? What is the impact of ECB on program delivery? What are the outcomes? And why is it better than other forms of evaluation? What are critical mediators and moderators of these effects? And what contextual factors impact ECB? In sum, the existing state of the science and practice of ECB offers great opportunities to continue to grow the body of knowledge.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Baizerman, M., Compton, D. W., & Stockdill, S. H. (2002). New directions for ECB. *New Directions for Evaluation*, *93*, 109–119.

Botcheva, L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcome measurement practices in community agencies. *American Journal of Evaluation*, *23*, 421–434.

Carman, J. C., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. In J. G. Carman & K. A. Fredericks (Eds.), *Nonprofits and evaluation. New directions for evaluation* (Vol. 119, pp. 51–71). San Francisco, CA: Jossey Bass.

Compton, D., Glover-Kudon, R., Smith, I. E., & Avery, M. E. (2002). Ongoing capacity building in the American Cancer Society (ACS) 1995–2001. The art, craft, and science of evaluation capacity building. *New Directions for Evaluation*, *93*, 47–61.

Connolly, P., & York, P. (2002). Evaluating capacity-building efforts for nonprofit organizations. *Organizational Development Practitioner*, *34*, 33–39.

Cousins, B., Goh, S., Clark, S., & Lee, L. (2004). Integrating evaluation inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, *19*, 99–141.

Cousins, J. B., Goh, S., & Elliott, C. (2007, November). *Integrating consequences of evaluation into evaluation capacity building inquiry*. Paper presented at the annual meeting of the American Evaluation Association, Baltimore, MD.

Danseco, E., Halsall, T., & Kasprzak, S. (2009). *Readiness assessment tool for evaluation capacity building*. Ottawa, ON: The Provincial Centre of Excellence for Child and Youth Mental Health at CHEO.

Duffy, J., Labin, S., & Wandersman, A. (2007, November). A review of research on evaluation capacity building strategies. In T. Taylor-Ritzler (Ed.), *Building and assessing capacity for evaluation: Creating communities of*

learners among service providers. Symposium presented at the annual meeting of the American Evaluation Association Conference, Baltimore, MA.

Garcia-Iriarte, E., Suarez-Balcazar, Y., & Taylor-Ritzler, T. (2011). A catalyst-for-change approach to evaluation capacity building. *American Journal of Evaluation*, *32*, 168–182.

Gibbs, D., Napp, D., Jolly, D., Westover, B., & Uhl, G. (2002). Increasing evaluation capacity within community-based HIV prevention programs. *Evaluation and Program Planning*, *25*, 261–269.

Huffman, D., Thomas, K., & Lawrenz, F. (2008). A collaborative immersion approach to evaluation capacity building. *American Journal of Evaluation*, *29*, 358–368. doi:10.1177/1098214008318798

King, J. A. (2002). Building the evaluation capacity of a school district. *New Directions for Evaluation*, *93*, 63–79. Retrieved from http://www3.interscience.wiley.com/cgi-bin/fulltext/92013254/PDFSTART

King, J. A., & Volkov, B. (2005). A framework for building evaluation capacity based on the experiences of three organizations. *CURA Reporter*, *35*, 10–16.

Labin, S., Duffy, J., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, *33*, 307–338.

Mackay, K. (2002). The World Bank's ECB experience. *New Directions for Evaluation*, *93*, 81–100. Retrieved from http://www3.interscience.wiley.com/cgi-bin/fulltext/92013252/PDFSTART

Milstein, B., Chapel, T. J., Wetterhall, S. F., & Cotton, D. A. (2002). Building capacity for program evaluation at the Centers for Disease Control and Prevention. *New Directions for Evaluation*, *93*, 27–46.

Nielsen, S. B., Lemire, S., & Skov, M. (2011). Measuring evaluation capacity: Results and implications of a Danish study. *American Journal of Evaluation*. Retrieved from http://aje.sagepub.com/content/early/2011/01/14/1098214010396075

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29*, 443–459. doi:10.1177/1098214008324182

Preskill, H., & Torres, R. T. (2000). The learning dimension of evaluation use. *New Directions for Evaluation*, *88*, 25–38.

Stevenson, J. F., Florin, P., Mills, D. S., & Andrade, M. (2002). Building evaluation capacity in human service organizations: A case study. *Evaluation and Program Planning*, *25*, 233–243.

Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, *93*, 7–25.

Suarez-Balcazar, Y., Taylor-Ritzler, T., Garcia-Iriarte, E., Keys, C. B., Kinney, L., & Ruch-Ross, H., . . . Curtin, G. (2010). Evaluation capacity building: A culturally- and contextually-grounded interactive framework and exemplar. In F. Balcazar, Y. Suarez-Balcazar, & T. Taylor-Ritzler, & Keys (Eds.), *Race, culture and disability: Rehabilitation science and practice* (pp. 307–324). Sudbury, MA: Jones and Bartlett.

Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation*, *28*, 45–59. doi:10.1177/1098214006296430

Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D., & Balcazar, F. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation*, *34*, 190–206.

Texas Christian University. (2005). *Organizational readiness for change* (TCU ORC). Fort Worth, TX: Institute of Behavioral Research. Retrieved March 9, 2011, from http://www.ibr.tcu.edu/evidence/evi-orc.html

Volkov, B., & King, J. (2007). *A checklist for building organizational evaluation capacity*. Retrieved October 27, 2008, from http://www.wmich.edu/evalctr/checklists/ecb.pdf