

# Toward an Evidence-Based Approach to Building Evaluation Capacity

**Btissam El Hassar, Cheryl Poth, Rebecca Gokiert, and Okan Bulut**  
University of Alberta

**Abstract:** *Organizations are required to evaluate their programs for both learning and accountability purposes, which has increased the need to build their internal evaluation capacity. A remaining challenge is access to tools that lead to valid evidence supporting internal capacity development. The authors share practical insights from the development and use of the Evaluation Capacity Needs Assessment tool and framework and implications for using its data to make concrete decisions within Canadian contexts. The article refers to validity evidence generated from factor analyses and structural equation modelling and describes how applying the framework can be used to identify individual and organizational evaluation capacity strengths and gaps, concluding with practice considerations and future directions for this work.*

**Keywords:** *assessment tool, early childhood development field, evaluation capacity, evaluation capacity assessment, measurement, structural equation modelling*

**Résumé :** *Les organisations doivent évaluer leurs programmes autant à des fins d'apprentissage que d'imputabilité, ce qui exige une capacité d'évaluation interne. Il est cependant toujours difficile d'avoir accès à des outils qui permettent de produire des informations valides sur les capacités internes en évaluation. Les auteurs et auteures de cet article font part de leurs expériences relatives à la création et à l'utilisation d'un outil servant à l'évaluation des besoins relatifs au renforcement des capacités en évaluation ainsi que son utilisation pour la prise de décisions concrètes en contexte canadien. L'article traite des données générées par l'analyse de facteurs et par la modélisation par équations structurelles, et décrit la façon dont le cadre permet de déterminer les points forts et les lacunes des personnes et des organisations en matière d'évaluation. Il se termine avec des considérations pratiques et des suggestions d'orientation future pour ce type de travail.*

**Mots clés :** *outil d'évaluation, domaine du développement de la petite enfance, capacité d'évaluation, évaluation de la capacité d'évaluation, mesure, modélisation par équations structurelles*

Global demand exists for the practice of evaluation (Nielsen, Lemire, & Christie, 2018), as well as targeted evaluation capacity (EC) building initiatives for both organizations and individual evaluators (Buchanan & Kuji-Shikatani, 2014).

**Corresponding author:** Btissam El Hassar, 7720 175 Street NW, Edmonton, AB, T5T 0G8;  
[btissam@brilliantlyu.com](mailto:btissam@brilliantlyu.com)

A key challenge is that there exist limited empirically based assessment tools for monitoring the development of EC at the individual and organizational levels, especially tools that are context-specific (e.g., Bourgeois, Toews, Whynot, & Lamarche, 2013; Cheng & King, 2016; Fierro, 2012). The need for EC measures is pressing, as organizations are required to evaluate their programs for both learning and accountability purposes (Chelimsky, 2006; Patton, 2011; Wilcox & King, 2013). Organizations could benefit from developing their own collective and individual member EC to conduct credible and insightful evaluations. In this research and practice note, we introduce the Evaluation Capacity Needs Assessment (ECNA) tool, which we developed for early childhood development organizations in Alberta. Specifically, our goal is to demonstrate the process of using the tool to inform evaluation capacity development decisions for organizations. We provide a brief description of the tool development process<sup>1</sup> and our approach to generating validity evidence. Finally, we offer an illustrative example to guide practitioners in using the ECNA along with practical implications.

## CONTEXTS SURROUNDING THE DEVELOPMENT OF THE ECNA TOOL

### *Evaluation capacity needs in Alberta's early childhood development organizations*

The ECNA tool was one solution to the long-identified need in the early childhood development field in Alberta for support and resources in evaluation and evaluation capacity building. Dr. Rebecca Gokiert, the associate director of the Community-University Partnership for the Study of Children, Youth and Families based at the University of Alberta, received many requests to support programs in evaluation, connecting to expertise on campus, or linking students for practica and thesis research. Given the high demand, Dr. Gokiert initiated conversations with organizations in the early childhood development field to understand their EC needs and explore potential solutions. The field of early childhood development is interdisciplinary and intersectoral but shares the same goal of improving children's health and well-being. To respond more systematically to these needs, the Evaluation Capacity Network was developed with support from a Social Sciences and Humanities Research Council of Canada (SSHRC) Partnership Development Grant in 2014 (see Gokiert et al., 2017). The diversity (e.g., focus, workplace size, and evaluation resources) of the organizations involved in the network necessitated a way to assess strengths and gaps in EC in order to generate comprehensive understandings and inform an evidence-based approach for addressing EC needs. As part of the SSHRC-funded project's work, development of an assessment instrument and generation of validity evidence became El Hassar's doctoral dissertation study conducted under the supervision of Poth and Gokiert.

## ***Challenges for developing a valid evaluation capacity needs assessment tool***

Defining the construct and describing the relevant components of EC has been an ongoing effort for more than two decades (e.g., [Compton, Baizerman, Preskill, Rieker, & Miner, 2001](#); [Stockdill, Baizerman, & Compton, 2002](#); [King, 2020](#)). The goal of building EC is to provide organizations and their members with the necessary skills, resources, and infrastructure to conduct and use credible and meaningful evaluations ([Cousins, Goh, Elliott, & Bourgeois, 2014](#); [King, 2020](#)). As straightforward as this may sound, defining, conceptualizing, and measuring EC poses practical challenges ([Cousins et al., 2014](#); [Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012](#); [Preskill & Boyle, 2008](#)). Efforts have advanced our understandings of the EC construct. For example, [Gagnon, Aubry, Cousins, Goh, and Elliott \(2018\)](#) empirically investigated a framework and instrument developed by [Cousins et al. \(2014, 2008\)](#) focusing on the federal, provincial, and municipal levels of government and on not-for-profit organizations in Canada. [Gagnon et al. \(2018\)](#) provided important insights and highlighted the need for generating an evidence-based approach to EC in order to address the current ambiguity surrounding its conceptualization and measurement (e.g., [El Hassar, 2019](#); [Gagnon et al., 2018](#); [Labin et al., 2012](#); [Preskill & Boyle, 2008](#)). Specifically, there is a need to provide organizations with practical and meaningful information generated by a tool for assessing EC strengths and gaps that are context-specific and lead to valid conclusions.

A logical first step toward addressing the challenges faced by organizations entails having an assessment tool to inform the empirical definition and conceptualization of EC, as well as to guide the subsequent development of initiatives that address the areas in need of improvement. A review of existing instruments (e.g., [Bourgeois et al., 2013](#); [Cousins et al., 2008](#); [Nielsen, Lemire, & Skov, 2011](#); [Taylor-Ritzler, Suarez-Balcazar, Garcia-Iriarte, Henry, & Balcazar, 2013](#)) demonstrates that there have been significant contributions to the EC literature; however, there remain some unanswered questions regarding these instruments. Among these questions, the most important concern the reliability of the existing EC instruments, the validity of the theoretical frameworks proposed, and their transferability to other contexts beyond those for which they were developed ([Cheng & King, 2016](#); [Labin, 2014](#)). As one way of addressing these questions, researchers have called for more robust approaches to conceptualizing and measuring EC ([Gagnon et al., 2018](#); [King, 2020](#)). To answer this call, El Hassar's dissertation study focused on providing a bridge between theory, measurement, and practice to specify the individual and organizational components that contribute to building EC in the context of the early childhood development field in Alberta, with the central aim of informing EC evidence-based initiatives.

## **METHODS FOR DEVELOPING AND GENERATING VALIDITY EVIDENCE FOR THE ASSESSMENT TOOL**

The ECNA tool was developed using a community-based participatory approach ([Minkler & Wallerstein, 2011](#)) that involved interdisciplinary and intersectoral

partners from the academic, governmental, and not-for-profit sectors. We created the ECNA items by drawing on contributions of existing evaluation capacity tools (e.g., [Bourgeois et al., 2013](#); [Cousins et al., 2008](#); [Taylor-Ritzler et al., 2013](#)) and on the fields of measurement ([Messick, 1989](#)), psychology ([Burns, Kotrba, & Denison, 2013](#)) and organizational change ([Burke, 2017](#)). After developing a draft instrument, we engaged the partnership in expert consultations, and then we administered the instrument with a sample representing early childhood development organizations in Alberta. Drawing partners from diverse academic institutions, government agencies, and community organizations, they contributed academic and practice-based expertise in evaluation, measurement, and early childhood development. These experts contributed to the validation process that was guided by the Standards for Educational and Psychological Testing ([American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014](#); for our validation process and results see [El Hassar, 2019](#)). According to this approach, validation is considered a dynamic process requiring multiple sources of evidence, quantitative and qualitative, that support the intended conclusion ([American Educational Research Association et al., 2014](#)). Our rationale for using this approach is the focus on establishing a transparent process and evidence to make the most sensible justification for using the instrument ([Messick, 1989](#)), and providing potential users guidance to make an informed decision about using the ECNA ([El Hassar, 2019](#)).

The ECNA was administered to organizations in Alberta, resulting in a sample of 329 respondents working in more than 100 early childhood development organizations. The majority of the respondents were female (92%), with diverse professional roles, including 33.8% in management positions, 15% in leadership roles, and 22.5% in front-line positions. The majority (87%) of the respondents reported having some evaluation experience, and 44.8% reporting five or more years of evaluation experience. The data from the tool administration were analyzed with exploratory and confirmatory factor analyses to evaluate the individual and organizational factors, followed by structural equation modelling to assess the framework underpinning the instrument. The resulting validity evidence supported the use of the ECNA tool within early childhood development context to generate meaningful insights for informing the development of targeted EC initiatives (see [El Hassar, 2019](#)).

## **EVALUATION CAPACITY NEEDS ASSESSMENT: FRAMEWORK AND COMPONENTS**

We used confirmatory factor analysis (CFA) and structural equation modelling (SEM) to evaluate the degree to which the instrument produced valid conclusions about EC needs within our study context. We initially used CFA to identify the underlying factor patterns for the individual and the organizational EC items, producing two measurement models that adequately assessed individual EC based on the goodness-of-fit indices (SRMR = 0.050, RMSEA = 0.060 [90% CI

(0.049, 0.071)], TLI = 0.90, CFI = 0.92) and organizational EC (SRMR = 0.057, RMSEA = 0.055 [90% CI (0.040–0.069)], TLI = 0.95, CFI = 0.96). The original ECNA tool consisted of 26 items and 33 items measuring knowledge and skills at the individual and organization levels, respectively. The analyses also provided statistical evidence, which we used to evaluate the reliability of the items and eliminate the less reliable ones (e.g., items measuring multiple factors; for the statistical evidence, see [El Hassar, 2019](#)). The revised ECNA tool contained 17 (rather than 26) items measuring individual factors and 12 (instead of 33) items measuring organizational factors. [Table 1](#) lists the items included in the revised ECNA tool for measuring each factor (for the factor scores, see [El Hassar, 2019](#)).

For the second validation technique, we used structural equation modelling to assess the relationships between the individual and organizational EC factors in terms of the direction and magnitude of influence. Our analysis, illustrated in [Figure 1](#), provided empirical evidence elucidating the nature of the interaction between the individual and organizational EC factors. In the figure, arrows indicate the directionality of each influence (e.g., Organizational Culture positively influences Organizational Leadership), and the numbers indicate the relative strength of the influence. For example, Organizational Culture has an influence on Organizational Leadership ( $r = 0.74$ ), explaining 54% of the variance ( $R^2 = 0.54$ ), which in turn influences Organizational Commitment to Evaluation ( $r = 0.35$ ), explaining 12% of the variance ( $R^2 = 0.12$ ). These findings suggest that the leadership of an organization plays a significant role in predicting other EC components, such as its commitment to conducting and using evaluation. The model also provided two complementary and statistically significant paths through which an organization and its employees can develop their EC, as Organizational Culture influences Individual Motivation indirectly via Individual Attitude (Path A;  $\beta = 0.037$ , 95% CI [0.021, 0.106]) and also indirectly via Individual Skills (Path B;  $\beta = 0.041$ , 95% CI [0.028, 0.108]).

The two paths have a cumulative influence on Individual Motivation. These findings highlight the importance of approaching EC conceptualization, measurement, and practice from a systems perspective in which the relationships among EC factors are critical to conceptualizing, measuring, and implementing EC. Taken together, the findings contribute to practice by showing organizations how to assess their specific EC needs and then refocus their efforts to address those needs. [Table 2](#) represents a summary of the EC needs within our sample. Once we established the items measuring each factor, we used a simple mean calculation of the estimated factor scores to generate [Table 2](#).

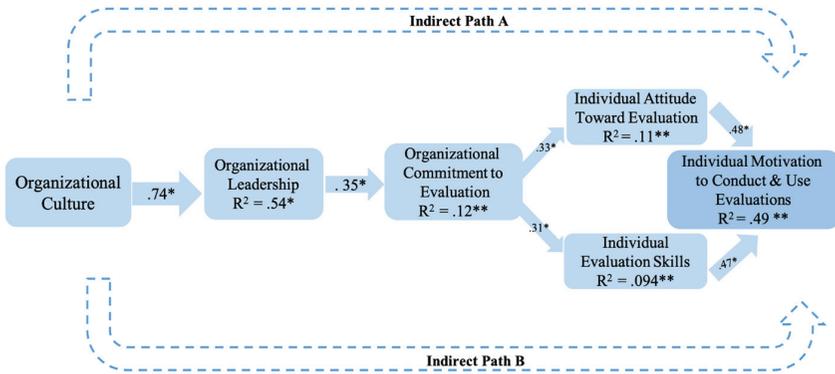
## IMPLICATIONS FOR DEVELOPING EVALUATION CAPACITY PRACTICES

### *Practical use of the ECNA tool*

The ECNA tool has important implications for practice. Considering the early childhood development practitioners who were assessed in this study, we can see

**Table 1.** Individual and organizational evaluation capacity factors and items measuring them

Factors	Items measuring individual EC	Factors	Items measuring organizational EC
Individual skills	<p>I know what skills to look for in an external evaluator.</p> <p>I have the skills to oversee an external evaluator.</p> <p>I know how to use evaluation findings in decision making.</p> <p>I know how to make organizational level changes based on evaluation findings.</p>	<p>Organizational culture</p> <p><i>My organization</i></p> <p>...</p>	<p>... encourages staff to express their opinions.</p> <p>... involves staff when making long-term plans.</p> <p>... gives staff the opportunity to reflect on organizational goals.</p> <p>... reviews its mission, vision, and values with staff.</p> <p>... builds ideas in collaboration with staff members.</p>
Individual attitude	<p>I think evaluation can be useful in determining the effectiveness of my organization.</p> <p>I think evaluation findings can be beneficial to my sponsor(s)/funder(s).</p> <p>I think evaluation can improve transparency in an organization.</p> <p>I would be concerned if an organization doesn't evaluate its activities.</p> <p>I have an ethical responsibility to participate in evaluation as needed.</p> <p>I think it is important that staff members get involved in evaluation.</p> <p>I think internal allocation of funding for an evaluation is justified.</p> <p>I think credentialed training (e.g., course, certificate, diploma) in evaluation is important.</p> <p>I think noncredentialed training (e.g., webinar, coaching) in evaluation is important.</p> <p>I think external funding for an evaluation is justified.</p> <p>I am open to staff being provided with the opportunities to learn the skills necessary to conduct evaluations.</p> <p>I am open to adopting new ideas in my day-to-day activities based on evaluation findings.</p>	<p>Organizational leadership</p> <p><i>The leadership of my organization</i></p> <p>...</p> <p>Organizational commitment</p> <p><i>My organization</i></p> <p>...</p>	<p>... resolves interpersonal conflicts in a positive manner.</p> <p>... celebrates staff members' achievements.</p> <p>... promotes evaluative thinking.</p> <p>... dedicates funds to conduct an evaluation.</p> <p>... dedicates funds to ensure ongoing evaluation.</p> <p>... has resources in place to undertake evaluation on an ongoing basis.</p> <p>... has the commitment from external stakeholders to ensure evaluation sustainability.</p>



**Figure 1.** Relationships among the individual and organizational evaluation capacity factors based on the structural equation modelling

\* $p < .05$ . \*\* $p < .001$ .

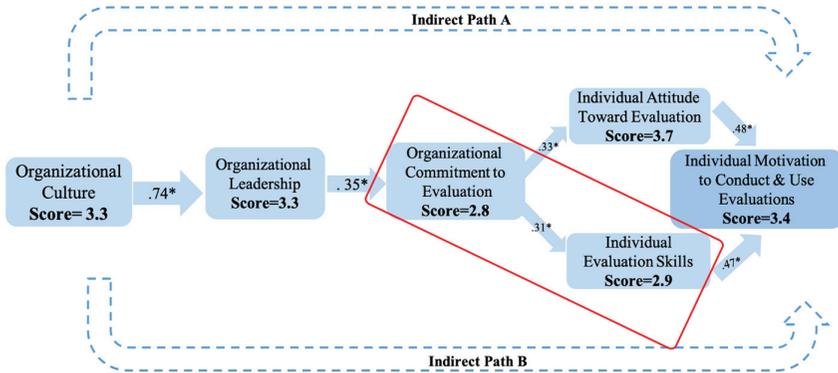
(El Hassar, 2019)<sup>1</sup>

**Table 2.** Factor scores for individual and organizational evaluation capacity

Evaluation capacity factors	Factor score	Level	Average evaluation capacity
Individual Attitude	3.7	High	3.1
Individual Evaluation Skills	2.9	Moderate	
Individual Motivation	3.4	Moderate	
Organizational Culture	3.3	Moderate	3.3
Organizational Commitment to Evaluation	2.8	Moderate	
Organizational Leadership Evaluation	3.3	Moderate	

Note. First, calculate the mean of respondents for each factor. Second, calculate the average of that mean to obtain the total factor score. Finally, calculate the means of the three individual and organizational EC factors separately to find the average EC. Interpretation of the Level of Evaluation Capacity: Limited: Score  $\leq 2$ ; Moderate:  $2 < \text{Score} \leq 3.4$ ; High:  $3.5 \leq \text{Score}$

that overall individual and organizational EC in Alberta’s early childhood development organizations appear to be at a moderate level. The two factors that are at the lower end of the interpretative scale are Organizational Commitment to Evaluation and Individual Skills (see Figure 2). Based on our findings, the former factor appears to influence the latter. If organizations are not able to provide the funding and other necessary resources for evaluation, it is not surprising that individual skills are much lower as well. One way for organizations to use these findings in practice would be to make a continuous effort to provide the resources, which could



**Figure 2.** Two factors requiring attention within the context of our study

\* $p < .05$ , \*\* $p < .001$ .

include time for professional development, needed to build individuals' skills. To understand which specific evaluation skills need development, we included a section in the ECNA tool that asks targeted questions about evaluation skills and their importance to the respondent's job. This section allows organizations to dig deeper into the skills that are needed and to provide specific solutions to develop them.

### *Practical considerations for using the ECNA tool*

If using the ECNA, organizations should first consider whether their context is similar to the early childhood development context in which the tool was built (e.g., interdisciplinary and intersectoral focus, low resources for evaluation, primarily not-for-profit). This question is important because context is key to validity and the validation process is foundational and relevant not only to academic discussions but also to evaluation practice. Making decisions based on weak or nonexistent validity evidence not only may be unethical but also may negatively affect people, organizations, and policies (Abma, 2006; Davies, Newcomer, & Soydan, 2006; Moss, 1998). Validity is not an inherent characteristic of an instrument; it is a fluid concept that is context-dependent. Hence, a survey that measures EC in one context may or may not be appropriate in another. The way we define the factors and the items designed to measure those factors might be influenced by the context (Rivers, Meade, & Fuller, 2009). For example, the language we used to describe items made sense for respondents working in early childhood development organizations but may not be as clear or may mean something different for respondents working in a different organizational context. Context is a complex topic that merits further study. In this paper, our emphasis is on bringing awareness to the importance of context when developing and using assessment tools. Establishing the validity of instruments is especially relevant for evaluators who work in interdisciplinary sectors such as early childhood, where the necessary instruments are not always readily available (Weitzman & Silver, 2013).

Organizations that operate in a context similar to that of our study may administer the short form of the ECNA and then calculate the simple means of individual factors and of the individual and organizational EC scales. The findings from the data can be combined with qualitative evidence (e.g., interviews, group discussions, staff meetings, board meetings) for corroboration and strengthening the validity of the conclusion. For contexts that are very different from ours, it is best for organizations to use the original ECNA and then replicate our validation process to evaluate whether the measurement of individual and organizational EC factors and their relationships in fact remain stable in the different contexts. If not, it will be important to understand what the relationships in the different contexts are and under which conditions they exist. Given that this is a new instrument, we also recommend that others evaluate our validation process and findings to enhance the measurement and practice of EC.

Measuring individual and organizational EC factors and the relationships among them may provide critical insights that can advance the evaluation field toward evidence-based EC practice. However, there are risks if the context is not considered or the validity of conclusions that are reached is not examined. The ECNA we developed is appropriate for assessing evaluation needs within our context to guide development efforts toward building relevant EC. We believe that the ECNA provides the potential for evaluators and researchers in the early childhood development field in Canada and fields with similar characteristics to enhance their context-based understandings of EC. We invite them to build on our work by replicating it in other areas in order to strengthen evidence-based approaches to building EC.

## ACKNOWLEDGEMENTS

This research was partially funded by the generosity of the Social Sciences and Humanities Research Council of Canada (SSHRC) and of the Stollery Children's Hospital Foundation and supporters of the Lois Hole Hospital for Women, both through the Women and Children's Health Research Institute.

The authors gratefully acknowledge the contributions of the Evaluation Capacity Network's Core Research Team, our evaluation stakeholders, and participants.

## NOTE

1 See [El Hassar \(2019\)](#) for a detailed description of the development and validation process and results.

## REFERENCES

[Abma, T. A. \(2006\)](#). The social relations of evaluation. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The Sage handbook of evaluation* (pp. 184–199). Thousand Oaks, CA: SAGE.

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bourgeois, I., Toews, E., Whynot, J., & Lamarche, M. K. (2013). Measuring organizational evaluation capacity in the Canadian federal government. *Canadian Journal of Program Evaluation, 28*(2), 1–19. Retrieved from <https://journalhosting.ucalgary.ca/index.php/cjpe/article/view/30828>
- Buchanan, H., & Kuji-Shikatani, K. (2014). Evaluator competencies: The Canadian experience. *Canadian Journal of Program Evaluation, 28*(3), 29–47. Retrieved from <https://journalhosting.ucalgary.ca/index.php/cjpe/article/view/30950>
- Burke, W. W. (2017). *Organization change: Theory and practice*. Thousand Oaks, CA: SAGE.
- Burns, G. N., Kotrba, L. M., & Denison, D. R. (2013). Leader-culture fit: Aligning leadership and corporate culture. In J. Passmore (Ed.), *The Wiley-Blackwell handbook of the psychology of leadership, change, and organizational development* (pp. 113–128). Chichester, England: John Wiley & Sons.
- Chelimsky, E. (2006). The purposes of evaluation in a democratic society. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The SAGE handbook of evaluation* (pp. 33–55). Thousand Oaks, CA: SAGE.
- Cheng, S.-H., & King, J. A. (2016). Exploring organizational evaluation capacity and evaluation capacity building: A Delphi study of Taiwanese elementary and junior high schools. *American Journal of Evaluation, 38*(4), 521–539. <https://doi.org/10.1177/1098214016672344>
- Compton, D., Baizerman, M., Preskill, H., Rieker, P., & Miner, K. (2001). Developing evaluation capacity while improving evaluation training in public health: The American Cancer Society's Collaborative Evaluation Fellows Project. *Evaluation and Program Planning, 24*(1), 33–40. [https://doi.org/10.1016/S0149-7189\(00\)00045-8](https://doi.org/10.1016/S0149-7189(00)00045-8)
- Cousins, J. B., Elliott, C., Amo, C., Bourgeois, I., Chouinard, J., Goh, S. C., & Lahey, R. (2008). Organizational capacity to do and use evaluation: Results of a pan-Canadian survey of evaluators. *Canadian Journal of Program Evaluation, 23*(3), 1–35. Retrieved from <https://evaluationcanada.ca/system/files/cjpe-entries/23-3-001.pdf>
- Cousins, J. B., Goh, S. C., Elliott, C. J., & Bourgeois, I. (2014). Framing the capacity to do and use evaluation. *New Directions for Evaluation, 2014*(141), 7–23. <https://doi.org/10.1002/ev.20076>
- Davies, P., Newcomer, K., & Soydan, H. (2006). Government as structural context for evaluation. In I. Shaw, J. Greene, & M. Mark (Eds.), *The Sage handbook of evaluation* (pp. 163–183). Thousand Oaks, CA: Sage.
- El Hassar, B. (2019). *Development and Validation of an Evaluation Capacity Instrument for the Early Childhood Field* (Unpublished doctoral dissertation). University of Alberta, Edmonton, Alberta.
- Fierro, L. A. (2012). *Clarifying the connections: Evaluation capacity and intended outcomes* (Unpublished doctoral dissertation). Claremont Graduate University, Claremont, CA.
- Gagnon, F., Aubry, T., Cousins, J. B., Goh, S. C., & Elliott, C. (2018). Validation of the evaluation capacity in organizations questionnaire. *Evaluation and Program Planning, 68*, 166–175. <https://doi.org/10.1016/j.evalprogplan.2018.01.002>

- Gokiert, R., Kingsley, B., Poth, C., Edwards, K., El Hassar, B., Tink, L., ... Hopkins, S. (2017). Developing an evaluation capacity building network in the field of early childhood development. *Engaged Scholar Journal: Community-Engaged Research, Teaching, and Learning*, 3(2), 59–80.
- King, J. A. (2020). Putting evaluation capacity building in context: Reflections on the Ontario Brain Institute's Evaluation Support Program. *Evaluation and Program Planning*, 80. <https://doi.org/10.1016/j.evalprogplan.2017.05.013>. Medline:28641901
- Labin, S. N. (2014). Developing common measures in evaluation capacity building: An iterative science and practice process. *American Journal of Evaluation*, 35(1), 107–115. <https://doi.org/10.1177/1098214013499965>
- Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, 33(3), 307–338. <https://doi.org/10.1177/1098214011434608>
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–104). New York, NY: Macmillan.
- Minkler, M., & Wallerstein, N. (2011). *Community-based participatory research for health: From process to outcomes* (2nd ed.). San Francisco, CA: John Wiley & Sons.
- Moss, P. A. (1998). The role of consequences in validity theory. *Educational Measurement: Issues and Practice*, 17(2), 6–12.
- Nielsen, S. B., Lemire, S., & Christie, C. A. (2018). The evaluation marketplace and its industry. *New Directions for Evaluation*, 2018(160), 13–28. <https://doi.org/10.1002/ev.20344>
- Nielsen, S. B., Lemire, S., & Skov, M. (2011). Measuring evaluation capacity: Results and implications of a Danish study. *American Journal of Evaluation*, 32(3), 324–344. <https://doi.org/10.1177/1098214010396075>
- Patton, M. Q. (2011). *Essentials of utilization-focused evaluation*. Thousand Oaks, CA: SAGE.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29(4), 443–459. <https://doi.org/10.1177/1098214008324182>
- Rivers, D. C., Meade, A. W., & Fuller, W. L. (2009). Examining question and context effects in organization survey data using item response theory. *Organizational Research Methods*, 12(3), 529–553. <https://doi.org/10.1177/1094428108315864>
- Stockdill, H. S., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, 2002(93), 7–26. <https://doi.org/10.1002/ev.39>
- Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D. B., & Balcazar, F. E. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation*, 34(2), 190–206. <https://doi.org/10.1177/1098214012471421>
- Weitzman, B. C., & Silver, D. (2013). Good evaluation measures: More than their psychometric properties. *American Journal of Evaluation*, 34(1), 115–119. <https://doi.org/10.1177/1098214012461628>
- Wilcox, Y., & King, J. A. (2013). A professional grounding and history of the development and formal use of evaluator competencies. *Canadian Journal of Program Evaluation*,

28(3), 1–28. Retrieved from <https://journalhosting.ualgary.ca/index.php/cjpe/article/view/30949>

## AUTHOR INFORMATION

**Btissam El Hassar** is a founder of B Brilliantly U Coaching and Consulting, focusing on evaluation and change management. She recently earned her doctorate in Measurement, Evaluation, and Cognition from the University of Alberta. Her dissertation work involved developing and validating an evaluation capacity needs assessment (ECNA) tool that is contextually relevant to the early childhood development field in Alberta within the Evaluation Capacity Network. She is an alum of the Alberta Clinical and Community-based Evaluation Research Team (ACCERT).

**Cheryl Poth** is a professor in the Centre for Research in Applied Measurement and Evaluation (CRAME) in the Department of Educational Psychology, in the Faculty of Education, at the University of Alberta and a co-founder of the Alberta Clinical and Community-based Evaluation Research Team (ACCERT). She teaches and conducts research on program evaluation, working with stakeholders from school boards, post-secondary institutions, and federal organizations in the areas of educational programs and health services. She is a regular contributor to the Canadian and American Evaluation Association conferences and publications. Her research interests include evaluation use with particular emphasis on developmental evaluation and evaluation use.

**Rebecca Gokiert** is an associate professor and Associate Dean Research in the Faculty of Extension at the University of Alberta. Her teaching and research focus on early childhood development, psychometrics, and evaluation. Within this portfolio, she explores intercultural realities at the intersection of measurement and development. In 2014, Rebecca received a SSHRC Partnership Development Grant to build the Evaluation Capacity Network (ECN), a provincial community-university collaborative that has since expanded to include partners across Canada and the United States. As director of the ECN, Rebecca leads continued efforts to engage community in broadening understandings of community-driven and culturally relevant approaches to evaluation, and building evaluation capacity within the early childhood sector.

**Okan Bulut** is an associate professor of educational measurement and psychometrics as well as a member of the Centre for Research in Applied Measurement and Evaluation (CRAME) at the University of Alberta. Okan's primary research areas are centred on educational and psychological measurement, evaluation, psychometrics, and technology integration in education. Specifically, his current research interests include computerized adaptive and computer-based testing, automated feedback and score reporting, survey design, psychological testing, and statistical programming using the R programming language.