Performance management, evaluation capacity, and organizational learning: a study of nonprofit organizations

Shahid Umar

University at Albany, State University of New York, sumar@albany.edu

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Public Administration Commons, and the Public Policy Commons

Recommended Citation

https://scholarsarchive.library.albany.edu/legacy-etd/1528

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
PERFORMANCE MANAGEMENT, EVALUATION CAPACITY, AND ORGANIZATIONAL LEARNING: A STUDY OF NONPROFIT ORGANIZATIONS

By
Shahid Umar

A Dissertation

Submitted to the University at Albany, State University of New York in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Department of Public Administration and Policy
Rockefeller College of Public Affairs and Policy

2015
ABSTRACT

Most evaluation capacity building (ECB) approaches use staff training as the method to build evaluation capacity (EC) in organizations. ECB has rarely been understood and empirically studied at the organizational level. The present study fills this gap by using an organizational learning perspective. Specifically, the study was a response to the following research question: do organizations with high EC differ from organizations with low EC in terms of OLC? OLC was measured using five measures of organizational learning: staff empowerment, staff development, external alignment, experimentation, and clear organizational goals. Data were collected from 26 nonprofit organizations (NPOs) located in the New York State Capital Region: 13 organizations with high EC and 13 organizations with low EC. To avoid common method bias, multiple sources of information were used to measure the dependent variable.

The study found support for an organizational learning approach to ECB: three of the five organizational learning measures (i.e., staff development, experimentation, and clear organizational goals) were found to be significant predictors of EC. The study contributed to both the EC and performance management literatures by 1) bringing these two streams of literature closer together and 2) empirically connecting them with organizational learning literature. NPOs and public organizations attempting to build EC and implement performance management systems should pay attention to staff development, introduce a culture of experimentation, and clearly define their goals.
ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who made this study possible. I am especially grateful to my dissertation committee members, John Rohrbaugh, Jeffrey Straussman, and Ellen Rubin for their support and guidance in completing this dissertation. I would not have accomplished an enormous amount of work in advancing my dissertation to completion without their dedicated support. I owe a very special thanks to John Rohrbaugh whose help, guidance, and encouragement made completion of this dissertation possible.

I am also grateful to my family for their love and support. I wish to thank my parents, Ghulam Fatima and Umar Hayat, my sisters, and brothers for their unconditional love and support. A very special thanks to my wife, Fakiha Shahid, for her sacrifices, patience, encouragement and love, and taking care of our beloved daughter, Hiba Umar. She has been understanding throughout my Ph.D. I also wish to thank my mother-in-law, Rahila Khanum, who has helped to care for Hiba.

Finally, I am grateful to the U.S. government, United Stated Education Foundation in Pakistan (USEFP), the Fulbright Program, and Institute of International Education (IIE) for providing me funding to study in the U.S. Special thanks to the U.S. citizens and taxpayers who paid for my doctoral studies.
This work is dedicated to John W. Rohrbaugh who is a brilliant teacher, a great mentor and friend, and a wonderful human being.
# Table of Contents

CHAPTER 1: INTRODUCTION ........................................................................................................ 1

CHAPTER 2: LITERATURE REVIEW ...................................................................................... 10

  Conceptual Framework ........................................................................................................ 12
  What is Evaluation Capacity? ............................................................................................. 17
  Evaluation Capacity and Organizational Learning Capacity ............................................ 19
    1a. Staff Empowerment and Evaluation Capacity ......................................................... 20
    1b. Staff Empowerment and Performance Management .............................................. 23
  2a. Staff Development and Evaluation Capacity .............................................................. 27
  2b. Staff Development and Performance Management .................................................. 30
  3a. External Alignment and Evaluation Capacity ............................................................. 32
  3b. External Alignment and Performance Management .................................................. 34
  4a. Experimentation and Evaluation Capacity ................................................................. 37
  4b. Experimentation and Performance Management ..................................................... 38
  5a. Clear Organizational Goals and Evaluation Capacity ............................................... 40
  5b. Clear Organizational Goals and Performance Management ..................................... 43

Revised Framework ............................................................................................................. 46

CHAPTER 3: METHOD OF STUDY ...................................................................................... 49

  Sample and Procedures .................................................................................................... 49
  Study Measures ................................................................................................................. 52
    Independent Measures ..................................................................................................... 52
    Dependent Measure of Evaluation Capacity (EC) .......................................................... 56
  Analysis ............................................................................................................................... 58

CHAPTER 4: RESULTS ........................................................................................................ 59

  Demographic Characteristics of Respondents ................................................................. 59
  Psychometric Properties of the Measures ........................................................................ 61
    Factor Analysis of the Independent Measures ............................................................... 61
    Internal Reliability of the Independent Measures ......................................................... 64
    Independent Measures .................................................................................................. 64
    Factor Scores and Organization Level Data ................................................................. 67
    Dependent Measure of Evaluation Capacity ............................................................... 67
    Bivariate Analysis of Study Measures ......................................................................... 69
Multivariate Analysis and Hypothesis Testing ................................................................. 71
  Discriminant Analysis ........................................................................................................ 71
  Logistic Regression Analysis ............................................................................................. 78
  Hypothesis Testing ............................................................................................................... 79
  Ordinary Least Square Regression Analysis ..................................................................... 81

CHAPTER 5: DISCUSSION AND CONCLUSION ................................................................. 84
  Findings Consistent with Previous Studies ........................................................................ 85
  New Findings ...................................................................................................................... 87
  Strengths and Weaknesses of the Study ............................................................................ 97
  Conclusion .......................................................................................................................... 100

REFERENCES ...................................................................................................................... 102
  Appendix 1 ......................................................................................................................... 120
  Appendix 2 ......................................................................................................................... 123
List of Figures

Figure 2.1 Conceptual Framework of Evaluation Capacity Building..........................15

Figure 2.2 Revised Conceptual Framework of Evaluation Capacity Building...............48
List of Tables

Table 1.1 Comparison between Performance Management and Evaluation Capacity …………7
Table 4.1 Demographic Characteristics of Survey Participants …………………………………..60
Table 4.2 Rotated Factor Loadings for Survey (With Parsimax Rotation) …………………..63
Table 4.3 Measure Reliability Coefficients for Independent Measures .................................65
Table 4.4 Measure Reliability Coefficients for Dependent Measure .................................68
Table 4.5 Bivariate Correlations among Study Measures ..................................................70
Table 4.6 Descriptive and t-statistics for Revenues in Million Dollars ..................................72
Table 4.7 Tests of Univariate and Multivariate Normality ...................................................74
Table 4.8 Discriminant Analysis of Evaluation Capacity Level ........................................76
Table 4.9 Discriminant Classification of Evaluation Capacity ...........................................77
Table 4.10 Logistic Regression Results .............................................................................80
Table 4.11 Regression Analysis .........................................................................................83
CHAPTER 1: INTRODUCTION

Nonprofit organizations (NPOs) and public organizations worldwide have been under continuous scrutiny to be more accountable and efficient. The roots of such movements could be traced in the 1980s and 1990s globally when some countries such as New Zealand, Australia and Great Britain introduced new performance measurement systems to make the public sector more cost effective and responsive. This reform movement for accountability has become known as New Public Management (NPM). A significant part of NPM is to evaluate organizational success in terms of achieving outcomes and holding managers responsible for those outcomes (Radin, 2006). The United States (U.S.) federal government responded to this movement for accountability through the introduction of the Government Performance and Result Act (GPRA). Under this act, all U.S. government agencies are required to have strategic plans with clear organizational goals and performance measures. The performance of agencies is judged by the extent to which they achieve these goals (Gualmini, 2008; Lynn, Heinrich, & Hill, 2000).

One of the most commonly used tools to judge organizational performance and to hold NPOs and public organizations accountable is program evaluation. Although accountability through program evaluation has been in use for many decades, it has assumed greater importance as a result of the performance measurement movement and the introduction of legislative measures such as the GPRA (Newcomer, 2004). Program evaluation is “the use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social action in ways that improve social conditions” (Rossi, Lipsey, & Freeman, 2004, p. 29). In principle, organizations and their leadership could be held accountable through program evaluation. Program evaluation also informs policy makers concerning the effectiveness of their
approaches toward the solution of social problems, tests their assumptions and indicates how far they are in the process of achieving policy goals. Consequently, scarce resources are saved and evidence-based practices are promoted (Rossi et al., 2004; Waters, 2011).

Due to the ever increasing demand for accountability, NPOs and public organizations have paid greater attention to building an appropriate capacity to be able to undertake evaluation (Brown & Reed, 2002; Naccarella, Pirkis, Kohn, Morley, Burgess, & Blashki, 2007; Preskill & Boyle, 2008a; Satterlund, Treiber, Kipke, Kwon, & Cassady, 2013). Evaluation capacity building (ECB) is “an intentional process to increase individual motivation, knowledge, and skills, and to enhance a group or organization’s ability to conduct or use evaluation” (Labin, Duffy, Meyers, Wandersman, & Lesesne, 2012, p. 308) and evaluation capacity (EC) is the ability of an organization to do evaluation and use evaluation findings (Cousins, Goh, Elliott, & Bourgeois, 2014b). In addition to accountability demands, organizations also need to develop EC due to reporting mandates, the desire to be competitive in funding, and avoid external evaluations (Carman, 2013; Carman & Fredericks, 2009; Gilliam, Barrington, Davis, Lacson, Uhl, & Phoenix, 2003; Huffman, Thomas, & Lawrenz, 2008; Kapucu, Healy, & Arslan, 2011; Miller, Kobayashi, & Noble, 2006; Milstein, Chapel, Wetterhall, & Cotton, 2002; Moreau & Cousins, 2014; Preskill & Boyle, 2008a; Smits & Champagne, 2008; Stevenson, Florin, Mills, & Andrade, 2002). As a consequence, ECB is not an option for organizations but rather a strategy for survival in a competitive environment. Consequently, ECB has become a major focus of attention for both scholars and practitioners; the literature on ECB has expanded considerably in recent years.

---

1 A clear difference between EC and ECB could be given using the vocabulary of impact evaluation. EC is ability of an organization to do and to use evaluation and should be considered as an outcome measure. ECB should be considered as an intervention process to improve that outcome.
This ever increasing literature on ECB, however, lacks a strong empirical foundation (Bourgeois & Cousins, 2013; Clinton, 2014; Cousins, Goh, Clark, & Lee, 2004). Most of the approaches toward explicating ECB have focused on individual staff skills-building, assuming that organizational factors are not important (Taylor-Ritzler, Suarez-Balcazar, Garcia-Iriarte, Henry, & Balcazar, 2013). However, some scholars (e.g., Cherin & Meezan, 1998; Cousins et al., 2004; Cousins et al., 2014b; Owen, 1998; Preskill & Boyle, 2008b; Suarez-Balcazar, Taylor-Ritzler, Garcia-Iriarte, Keys, Kinney, Ruch-Ross & Curtain, 2010; Taylor-Powell & Boyd, 2008) have pointed out that ECB is not possible without taking into account organizational variables. More specifically, one group of scholars have argued that ECB should be conceptualized as an organizational, not individual, learning system and that an organization with an appropriate level of EC should be thought as a learning organization (Cousins et al., 2004; Cousins et al., 2014b; Preskill & Boyle, 2008a, b; Preskill & Torres, 1999b). Consistent with this school of thought, the present study argued that, until we know how an organization learns and what capacities are required for a learning organization, an appropriate ECB process will not be fully understood.

It is imperative, of course, to distinguish between ECB and EC and to highlight the importance of doing so. Most studies in the ECB literature do not attempt to differentiate between EC and ECB. However, several researchers (Cousins, Bourgeois, & Associates, 2014a; Cousins et al., 2014b; Naccarella et al., 2007; Nielsen, Lemire, & Skov, 2011b) are right to point out that clarifying the dimensions of the EC concept is important for the purpose of designing and developing any ECB initiative. Ultimately, this will make the measurement of ECB possible. In addition, clarifying the dimensions of the EC concept makes it possible to investigate its discriminant validity with respect to other variables such as OLC. OLC can be defined as “an organization’s shared assumptions and mechanisms (in terms of processes or culture) that
contribute to its capabilities to sustain and improve performance unfettered” (Hock-Hai, Xinwei, Kwok-Kee, Choon-Ling, & Lee, 2006, p. 264). Since ECB is considered as a portion of an organizational learning system, researchers first should attempt to answer more basic questions such as how EC is associated with OLC and how organizations with exemplary EC differ in terms of OLC from organizations with lower EC. Following the theoretical framework of organizational learning, the present study was an attempt to fill this gap and was response to the following research question: do agencies with well-developed EC differ from agencies with lesser EC in terms of OLC? The answer to this question also would pertain, at least in part, to a larger question the scholars consistently have raised: what uniquely characterizes an organization with higher EC? To answer these research questions, data from 26 NPOs in the New York State Capital Region were collected: 13 organizations with high EC and 13 organizations with low EC. The details for the selection of NPOs with high and low EC is discussed in the method section of this study. Before moving to the next section, it is essential to explain how EC is related to parallel work undertaken in public administration and termed performance management.

Performance management refers to the implementation of a performance measurement system, that is, the collection of performance data on a regular basis and use\(^2\) of those data to improve performance (Moynihan, 2008). This definition suggest that performance management has two components, 1) performance measurement and 2) use of performance data. Performance measurement is the ongoing monitoring of program through collection of data to improve organizational performance. Performance measurement is implemented in public organizations

\(^2\) In the performance management literature, the success of performance measurement is assessed in terms of use of performance data. In performance data use, purposive performance information (PPI) use is widely discussed in the literature. PPI use means that managers use performance information for specific tasks, such as managing their programs or making resource allocation decisions. The use of performance information for various purposes is beyond the scope of this study.
largely due to external political pressures and institutional isomorphism (Moynihan 2005; Rivenbark & Menter, 2006). For example, the U.S. federal government has mandated government agencies to implement performance measurement systems through legislative measures such as the GPRA of 1993 and the Modernization Act of 2012 (Rainey & Jung, 2015). Performance measurement only make data available; performance management requires this performance information to be used. The real value of performance measurement is achieved when it is used in organizations in a purposeful manner (Moynihan & Landuyt, 2009; Hvidman & Andersen, 2014; Behn, 2003; Boyne & Chen, 2007; Greener, 2012). One of issues raised in the literature regarding performance measurement is how to develop performance measures which yield valid and reliable data that could be used purposefully (Ho, 2006; Streib & Poister, 1999). Many studies have investigated how to ensure the use of performance data and what individual factors (e.g., Ammons & Rivenbark, 2008; Kroll, 2014; Kroll & Vogel, 2014), organizational factors (e.g., Moynihan & Pandey, 2005; Moynihan et al., 2012b; Moynihan & Landuyt, 2009), or external environmental factors are associated with the use of performance data (e.g., Berman & Wang, 2000; Ho, 2006; Kroll, 2015; Newcomer & Caudle, 2011; Yang & Hsiesh, 2007).

The definitions of performance management and EC suggest that both of these concepts are two names that have been attached in different fields of study to the same organizational phenomenon: both of these involve collection of performance information and use of that information for managing and improving programs and for making decisions. Apart from definitions, there are two other commonalities between performance management and EC: the operating environment and the ultimate organizational goal. Both performance measurement and program evaluation operate in political and contested environments and data in both cases could
be used by different actors in alternative ways for various political ends (e.g., Bamberger, Rugh, & Mabry, 2011; Greene, 1988; Johnsen, 2005; Moynihan, 2008; Pritchett, 2002). Additionally, although the performance information coming from program evaluation and performance measurement could be used differently, arguably the ultimate goal of both advancing performance management and building EC in an organization is to improve performance, increase efficiency, and hold organizations accountable (e.g., Brown & Reed, 2002; Naccarella et al., 2007; Newcomer & Caudle, 2011; Poister, Pasha, & Edwards, 2013; Preskill & Boyle, 2008; Satterlund et al., 2013). These parallels are made more explicit in Table 1.1. Therefore, based on the evident similarities, it is argued that both of these concepts could be treated and understood as largely the same. The present study defined EC and performance management as the organizational capability to undertake program evaluation, that is, to collect performance information and to use that information to improve programs and make informed decisions. Both performance management and EC are treated as synonymous in this study.

Answering the research question raised in this study is very important because, until it is known what uniquely characterizes an organization with better performance management and higher EC, any effort toward ECB in NPOs and public organizations would be limited.³ This study added to the understanding of differences in organizational processes associated with higher EC. This study made several contributions to the EC and performance management literatures both in theory and in practice. First, this study allowed for a better

³ Stated in different terms, organizational goals must be specified before organizational objectives can be set.
<table>
<thead>
<tr>
<th>Comparison Aspects</th>
<th>Performance management</th>
<th>Evaluation capacity</th>
<th>Is this aspect common?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td><strong>Performance measurement</strong>: an ongoing data collection of program accomplishment</td>
<td><strong>Program evaluation</strong>: a systematic process of collection of program performance data</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td><strong>Use of performance information</strong> to improve performance and make better decisions</td>
<td><strong>Use of the evaluation information</strong> to improve programs</td>
<td></td>
</tr>
<tr>
<td>Operating Environment</td>
<td>Contested and political</td>
<td>Contested and political</td>
<td>Yes</td>
</tr>
<tr>
<td>Goal</td>
<td>The ultimate claimed goal of performance management is improving performance, being efficient, and holding organizations accountable</td>
<td>The ultimate claimed goal of building evaluation capacity is improving performance, being efficient, and holding organizations accountable</td>
<td>Yes</td>
</tr>
</tbody>
</table>
understanding of the relationship between EC and OLC. Many studies have established a link between ECB and OLC (Andrews, Motes, Floyd, Flerx, & Fede, 2005; Fetterman & Wandersman, 2007; Miller & Campbell, 2006; Preskill & Torres, 1999a; Smith, 2007). However, no study has investigated the relationship between EC and OLC. Although the organizational learning literature defines OLC very precisely and has developed and validated measures of it, few studies in the EC literature and in the performance management literature have referenced the OLC construct or drawn on the organizational learning literature. In contrast, the present study used an organizational learning perspective and five OLC measures—staff empowerment, staff development, external alignment, experimentation, and clear organizational goals—to investigate its relationship with performance management and EC.

Second, as mentioned, the ECB literature lacks a strong empirical knowledge base (Bourgeois & Cousins, 2013; Clinton, 2014; Cousins et al., 2004). Most of the studies might be called “reflective case studies” and may not be even called “empirical studies at all” (Cousins et al., 2004, p. 131). However, the present study makes an important contribution due to its use of a rigorous research method. The present study was conducted using a robust quantitative method with data from 26 NPOs.

Third, the study also is important from the viewpoint of practice. Many scholars have argued that the practice of ECB should be guided by robust scientific knowledge (Wandersman, 2014a, b). Since this study was an effort to make a contribution to social scientific knowledge, therefore, ECB initiatives could draw on this study. Once the nature of the relationship between EC and OLC is known, practitioners could direct their efforts toward organizational learning processes which are associated with higher EC and better performance management. In this way,
unlike the current state of practice which is more driven by anecdotal evidence, the study can play a key role in establishing a science of practice.

The present study also contributes to the performance management literature. The study not only bridged the EC and performance management literatures, it also informs how NPOs and public organizations attempting to introduce performance management systems can make these systems successful. The present study also is unique in that it used an organizational learning perspective to understand performance management and brought these two streams of literature closer through empirical investigation. Although there are some studies in the performance management literature which understood performance management in terms of organizational learning (e.g., Dooren, 2011; Moynihan, 2005; Moynihan & Lavertu, 2012), these studies did not use validated organizational learning measures. The present study is first to use organizational learning validated measures and investigate how these measures are associated with performance management.
CHAPTER 2: LITERATURE REVIEW

For nearly 40 years, pronouncements about the manner in which organizations acquire knowledge and modify their actions accordingly have been divided between the scholarly literature on "organizational learning" (e.g., Argyris & Schon, 1978, 1996) and the practice-oriented literature on "learning organizations" (e.g., Senge, 1990, 1994). Both areas of work, however, appear to have fallen short. Empirical studies that would support the emergence of a more unified theory of organizational learning did not make their way into key academic journals. Further, the few concrete prescriptions for developing learning organizations--especially at mid-management rather than executive levels of organization--have proved difficult to implement. Now, however, the study of organizational learning capacity (OLC) appears to be bringing these two literatures together in ways much more useful to both theory and practice.

OLC can be defined as the level of collective capability in acquiring, sharing, interpreting, and recalling knowledge (Huber, 1991). By the mid-1990s, organizational scholars at both Harvard University and the Massachusetts Institute of Technology began to converge--through organizational observations and interviews--on the identification of key organizational factors, sometimes termed "mechanisms," that might increase OLC. Garvin's (1993) article in the *Harvard Business Review* and later Nevis and others' (1995) article in the *Sloan Management Review*.

\[\text{For detailed difference between organizational learning and learning organization, see Easterby-Smith, 1997, 1998; Tsung, 1997. Organizational learning is intentional use of learning processes at various levels (i.e., individual, group and organizational) in a way and to the direction that creates satisfaction for stakeholders (Dixon, 1999). According to Tsung (1997) “organizational learning” is a descriptive approach, that is, it tells how an organization learns. In this approach, most of the studies are undertaken with academic rigor. However, “learning organization” is a prescriptive approach which tells how an organization should learn and how an ideal type of organization with learning processes appears. The prescriptive approach is used by practitioners and consultants and usually lacks a strong empirical foundations (Easterby-Smith, 1997, 1998; Tsang, 1997). The detailed difference between these two concepts entails another study and is beyond the scope of the present study.}\]
Review both provided practical but also theoretically-grounded conceptual frameworks for the subsequent study of OLC.⁵

Goh and Richards (1997) appear to have been the first organizational researchers to report on the development and use of a survey instrument intended to assess OLC. A more extensive validation (Goh, 2001) of their Organizational Learning Survey (OLS) suggested that its five measures were highly correlated and, in fact, might be uni-dimensional.⁶ In about the same period, Watkins and Marsick (1997) introduced their Dimensions of the Learning Organization Questionnaire (DLOQ). The initial validation study was undertaken somewhat later (Yang, Watkins, & Marsick, 2004) which, as in the case of the OLS, suggested low discriminant validity for most if not all of its seven dimensions.⁷ A third questionnaire, the Learning Organization Survey (LOS), has emerged more recently (Garvin, Edmondson, & Gino, 2008). Its development appears not to have been informed by either prior DLOQ or OLS studies, and, in fact, it has been somewhat widely adopted by various private firms and public organizations without requisite validation.⁸ The details of the five measures used in this study are discussed in the following method section.

---

⁵Assertions about organizational learning capacity were based largely on extensive observations and substantive interviews. The organizational-level focus of these articles should be distinguished from parallel work of the same period that recommended specific managerial practices (see, for example, McGill, Slocum, & Lei, 1993; Ulrich, Jick, & Von Glinow, 1992).

⁶Goh (2001) reported bivariate correlations between the OLS measures that ranged from .43 to .76; half of the correlations were .62 or higher. Eventually, Goh, Quon, and Cousins (2007) presented the OLS as a multi-item measure of a single factor.

⁷Yang and others (2004) reported bivariate correlations between the DLOQ measures that ranged from .58 to .77; nearly half of the correlations were .67 or higher.

⁸A weak validation study has appeared that suggested seven instead of the ten original LOS dimensions may show convergent validity, but their discriminant validity is questionable (Singer, Moore, Meterko, & Williams, 2012). In particular, the reported bivariate correlations between these seven measures ranged from .42 to .90; half of the correlations were .65 or higher.
This review of pertinent literature for the present study is rooted in organizational learning theory. While reviewing this literature, careful attention was paid to nonprofit organizations (NPOs) and public organizations. The previous literature reviews conducted on this topic have been insufficient for several reasons. Some review efforts have addressed only small parts of evaluation capacity (EC) such as evaluation use (Johnson, Greenseid, Toal, King, Lawrenz, & Volkov, 2009; King & Thompson, 1983; Shulha & Cousins, 1997), empowerment evaluation (Fetterman & Wandersman, 2007; Miller & Campbell, 2006), participatory evaluation (Smits & Champagne, 2008) or are by now outdated (Cousins et al., 2004; Preskill & Boyle, 2008b; Preskill & Torres, 1999a; Stockdill, Baizerman, & Compton, 2002; Torres & Preskill, 2001). Although Labin and others (2012) conducted an excellent research synthesis of the evaluation capacity building (ECB) literature, they did not include theoretical work, literature reviews, or class room based ECB studies, that is, only 61 empirical studies reporting ECB initiatives. In addition, there has been no review which anchors ECB in organizational learning theory. The present review, in this way, is an important contribution in the ECB literature and, by extension, performance management.

**Conceptual Framework**

The present study used the ECB model developed by Cousins and his colleagues (Cousins et al., 2014b). To increase an understanding of ECB, they developed an ECB model based on a review of both empirical and theoretical studies (Figure 2.1). Unlike most of approaches that have focused on building individual skills to increase evaluation capacity and have assumed that organizational factors are not important (for example, Gibbs, Napp, Jolly, Westover, & Uhl, 2002; Huffman et al., 2008; Satterlund et al. 2013), Cousins and colleagues present ECB as an organizational change process rooted in organizational learning theory. That
is, ECB should be understood as an organizational change process which is not possible unless an organization knows how to learn and be adaptive (Bourgeois & Cousins, 2013; Cousins & Bourgeois, 2014a; Cousins & Earl, 1995; Cousins, Goh, & Clark, 2006; Cousins et al., 2004; Cousins et al., 2014b). In the context of the present study, the model of Cousins and his colleagues proposes that the ECB process and OLC are inextricably linked to each other. That is, an organization with well-developed EC will have higher OLC and vice versa. This present study tested this model.

Complex environments, stiff competition and technological changes necessitate organizations to change. This change entails a capacity of individuals, groups, and organizations to continuously learn from both internal circumstances and external environment. This is the reason why organizational learning is considered as a precondition to the survival of an organization (Garvin, 1993; Garvin, Edmondson, & Gino, 2008; Hock-Hai et al., 2006). An organization which continuously learns from its internal circumstances and external environment is called a learning organization. A learning organization is “a place where employees excel at creating, acquiring and transforming knowledge” (Garvin et al., 2008, p. 110). In order for an organization to build EC, it must undertake a learning process. ECB in an organization can be considered as a key component of an organizational learning system (Cherin & Meezan, 1998; Cousins et al., 2004; Owen, 1998; Preskill & Boyle, 2008b; Taut, 2007b; Taylor-Powell & Boyd, 2008). Additionally, for an organization to become a learning organization, it must undergo both incremental and transformational change (Jenlink, 1994).

Organizations undergo change all the time, of course, but the pace of organizational change varies from one organization to another. Some organizations are rather static and are not able to change at a pace necessitated by environmental factors. Others have better capabilities to
cope with environmental changes and to adapt to those change. The organizations of the latter category are deemed to be learning organizations and have OLC (Bess, Perkins, & McCown, 2011; Cousins et al., 2004; Garvin, 1993; Jenlink, 1994; Marsick & Watkins, 2003; Senge, 1990). Stated differently, organizational change and OLC are directly associated with each other.

Since ECB is a kind of organizational change, it is asserted here that OLC and EC are directly associated.

A brief description of the ECB framework (Cousins et al., 2014b) is shown in Figure 2.1. EC antecedent conditions are divided into two categories: sources of knowledge, skills and abilities (KSA) and organizational support structures. KSA refers to the ways individuals gain competencies to undertake evaluation in an organization. KSA could be built before or after joining an organization, could be the result of formal or informal training inside the organization, and could be enhanced through being part of an evaluation process. KSA is first built at the individual level which naturally leads to an increase in KSA at organizational level. KSA enhances the EC of an organization to do evaluation. Organizational support structures which promote EC include such factors as low job formalization, formal and informal incentives, and a well-functioning communication structure. These organizational support structures lead to the EC of an organization to do evaluation.
Sources of Knowledge

Capacity to DO Evaluation

Evaluative Inquiry

Capacity to USE Evaluation

Organizational Support Structure

Organizational Learning Capacity

Antecedent Conditions Affecting Evaluation
Evaluation Capacity and Processes
Evaluation Consequences

Mediating Conditions

Organizational Consequences

Figure 2.1 Conceptual Framework of Evaluation Capacity Building

Adapted from Cousins and others (2014b, p. 14)
Evaluation capacity and processes include three categories: capacity of an organization to do evaluation, evaluative inquiry, and capacity of an organization to use evaluation. Capacity to do evaluation can be measured from the degree of planning and framing evaluation; development of instruments; data collection, analyses, and reporting; and the presence of ‘soft skills’ such as the ability of individuals to work in teams and resolve conflict. The existing of capacity of an organization leads to evaluative inquiry. Evaluative inquiry refers to “the nature of and extent to which evaluation is actually occurring within the organization” (Cousins et al., 2014, p. 16), which could be undertaken by either internal or external evaluators. This evaluation may be formative or summative, participatory or non-participatory. Evaluative inquiry also could be measured from the number of evaluations undertaken in a given period. Evaluation information may have various uses such as instrumental use (for termination or improvement of a program), educational use (for individual and organizational learning), and persuasive use (that a program has value). The mediating conditions which increase use of evaluation include but are not limited to doing evaluation in a timely manner, demand for evaluation information, accessibility to evaluation information, and involvement of stakeholders in the process of evaluation.

The consequences of ECB are both the collective understanding of the members of an organization concerning the way it operates and also OLC. The construct of OLC has five essential building blocks: clear mission and vision; leadership which empowers employees and promotes culture of learning in an organization; a culture of experimentation and inquiry for the solution of problems; the capacity the organization and its members to transfer knowledge; and a priority on effective teamwork to solve organizational problems. These building blocks span the construct of learning organization, support each other, and are known to be highly correlated with organizational support structures and sources of KSA (Cousins et al., 2014b).
What is Evaluation Capacity?

Much effort in the EC literature during last decade has been focused on definition and conceptualization. However, there is no, as yet, consensus among scholars about nature of EC or how to build it (Bourgeois & Cousins, 2013; Brown & Reed, 2002; Fleming & Easton, 2010; Naccarella et al., 2007; Preskill & Boyle, 2008a; Satterlund et al., 2013; Trevisan, 2002). There may be three at least reasons for what Nielsen and others (2011b, p.325) call “conceptual pluralism.” First, trainers and consultants use various ECB approaches such participatory, empowerment, or collaborative methods. Each approach has its own assumption about learning and organizational change. As a result, for every perspective EC is something different and requires a different effort for ECB. Secondly, evaluation as both a profession and an area of study is very diverse, drawing attention from fields such as education, psychology, public administration, economics, social work, and public health. Each of these fields has its own traditions, methods, and assumptions. Consequently, this diversity does not allow much agreement on one set of EC approaches. Although this diversity of fields of study makes the evaluation profession very rich, it also creates the problem of conceptual pluralism. Third and last, as mentioned, the knowledge base of the field of evaluation lacks a substantial scientific base. This is largely because EC as a field of study is dominated by consultants and practitioners. Very rarely do EC studies use academic rigor which would allow the creation of a common knowledge base (Nielsen et al., 2011b).

Another part of the problem is that most studies only explain ECB processes and strategies (Cousins & Bourgeois, 2014b) and less is written about what constitutes EC. For instance, ECB is viewed as a way of providing skills and knowledge to members of the organization (Naccarella et al., 2007; Taylor-Ritzler et al., 2013), building a culture of evaluation
in an organization (Brown & Reed, 2002), and assigning the role of transferring evaluation skills to other members through one key person (Suarez-Balcazar et al., 2010). Thus, ECB refers to using various strategies such as training and technical assistance (Satterlund et al., 2013); it is considered to be a process aimed at developing the support structure and resources for sustainable evaluation education (Fleming & Easton, 2010).

Similarly, empowerment evaluation, which also is considered as an ECB approach, pays more attention providing various planning and evaluation tools to individuals (Wandersman & Snell-Johns, 2005) with less clarity on what is EC. In the same vein, the most commonly used definition of ECB largely addresses processes and strategies and a little about EC: “a context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites” (Stockdill et al., 2002, p. 8). Many ECB initiatives (e.g., Arnold, 2006; Fleming & Easton, 2010; Gilliam et al., 2003; King & Volkov, 2005) have conceptualized ECB in this way.

According to some scholars, however, such ways of conceptualizing ECB have some problems. These definitions describe how EC is built but do not tell what is being built. When one is not aware of what to build, then it will inevitably be difficult to know how to build it (Cousins & Bourgeois, 2014b; Naccarella et al., 2007; Nielsen et al., 2011b). A clear definition of a concept such as EC will not only encourage the measurement of EC in an organization, but it also will enable tests of empirical relationships with organizational processes such as OLC. As

---

9 Some scholars such as Taut (2007a) have argued that any attempt to measure EC or ECB is not a good idea because a single definition can be dangerous for practical reasons. Although a definition of a complex concept can be developed, its measure may not always be applicable to many contexts or may not accurately or fully depict the construct.
a result, it will allow for an empirical understanding of ECB. One effort to define EC is made by Nielsen and others (2011): EC “refers to an organization’s ability to bring about, align and sustain its objectives, structure, processes, culture, human capital and technology to produce evaluative knowledge that informs on-going practices and decision-making to improve organizational effectiveness” (Nielsen et al., 2011b, p. 338). To some degree Labin (2014) has conceptualized EC in a similar way. The indicators of the EC construct, according to Labin are resources, processes, policies and practices; funding, planning, undertaking and utilizing; and accessibility communication and dissemination of evaluation information. Gibbs and others (2002) have defined EC “as the extent to which an [organization] has the necessary resources and motivation to conduct, analyze, and use evaluations” (Gibbs et al., 2002, p. 261). For Cousins and others (2014b) EC is ability of an organization to do and use evaluation. It allows members of an organization to increase their knowledge of evaluation methods, actually undertaking evaluation and using evaluation information on a sustained basis in a way that EC becomes part of organizational culture, what they term as “organizational readiness for evaluation” (Cousins et al., 2004, p. 107)

Based on the foregoing definitions, one can assert that an organization has EC if it has policies, resources, skills and motivations to undertake evaluation and use evaluation information to the extent that it could be called “organizational readiness for evaluation.” ECB covers those organizational processes that specifically develop these policies, skills, and motivations.

**Evaluation Capacity and Organizational Learning Capacity**

Organizational learning capacity (OLC) as a construct is based in the principles of organizational learning theory and can be understood as the foundation of any learning
organization. Based on the organizational learning literature, the following five dimensions of organizational learning, which measure OLC, were included in this study.

1. Staff empowerment
2. Staff development
3. External alignment
4. Experimentation
5. Clear organizational goals

Drawing on the ECB and organizational learning literature, relationship between EC and each of these OLC dimensions is discussed in the following.

1a. Staff Empowerment and Evaluation Capacity

Staff empowerment, a primary component of OLC, refers to creating an egalitarian culture and encouraging organizational members to own their work, make decisions about their job, and be responsible for the results (Goh, 2000). Without empowering staff, an organization cannot become a learning organization (Bess et al., 2011; Goh & Richards, 1997; Goh, 2000; Goh, Quon, & Cousins, 2007; Hock-Hai et al., 2006; Marsick, 2013; Marsick & Watkins, 2003). In combination with other dimensions of a learning organization such as clear organizational goals and staff development, staff empowerment is considered a resource required for organizational change (Bess et al., 2011). Leaders of learning organizations play important roles to empower staff members toward a collective vision (Goh & Richards, 1997; Goh, 2000; Goh et al., 2007; Hock-Hai et al., 2006; Marsick & Watkins, 2003). Staff empowerment has been positively associated with increase in the performance of employees and the achievement of organizational goals (Bess et al., 2011; Marsick, 2013; Marsick & Watkins, 2003), as well as employee job satisfaction and organizational innovation (Goh et al., 2007).
There is not a single study, however, which has established a direct empirical relationship between EC and staff empowerment. Some studies reporting ECB initiatives have somewhat tangentially mentioned staff empowerment which may allow for an indirect association with EC level. Cherin and Meezan (1998), for example, point out that human service organizations should empower their staff to promote a culture of inquiry which in turn will increase the EC level. According to the authors, the reason for lower EC (i.e., lack of utilization of evaluation knowledge) is an absence of involvement of staff in the process of undertaking evaluation. Since external evaluators are hired for undertaking evaluation, employees may feel powerless and alienated in such situations. Stated differently, staff will feel empowered if they are involved in an evaluation process which in turn could increase EC. Miller and others (2006) and Anderson, Chase, Johnson, Mekiana, McIntyre, Ruerup, and Kerr (2012) also are of the view that internal involvement in evaluation is a way to empower staff and lead to their ownership of the evaluation process.

Most of the ECB initiatives use collaborative approaches such as empowerment and participation. These approaches assume that inclusion and empowerment of employees leads to promotion of an evaluation culture and organizational learning that ultimately increases the EC level. These approaches pay special attention to the opinion of staff in planning and undertaking evaluation. Advocates of participatory evaluation state that this approach is a way both of empowering stakeholders, including staff, at various levels and of promoting learning. As a result of this sense of inclusion and empowerment, staff actively participate in evaluation activities and learn from this process (Anderson et al., 2012; Rotondo, 2012; Wharton & Alexander, 2013); the EC level is increased. Similarly the principle of inclusion of empowerment evaluation states that stakeholders, including staff, also should be involved in the preliminary
processes of planning and designing evaluation and decision making (Fetterman, 2005; Fetterman & Wandersman, 2007; Wandersman et al., 2005). Empowerment evaluation advocates believe that the failure to empower organization members could be counterproductive; it could demotivate employees producing no change or even lowering of the EC level.

In his five-year participatory ECB initiative, Arnold (2006) found an indirect association between staff empowerment and EC for 4-H educators in Oregon. His participatory approach to include educators in the process of evaluation was used as a mean of empowerment. Before this empowerment of the educators, they had viewed evaluation as a tool in the hands of administration to control educators. As a result of this participatory initiative, educators felt empowered to have been included in the evaluation process. This changed their attitude toward evaluation, so that they believed evaluation was an important activity to achieve organizational objectives. This change in the staff’s attitude, in turn, led to an increase in EC.

McDonald, Rogers, and Kefford (2003) reported a similar five-year initiative to increase EC in the Agriculture Division of Australia. As a result of this initiative and consistent with Arnold (2006), the staff felt positively about evaluation and had a sense of empowerment. Before starting this initiative, the staff had been nervous and fearful for evaluation being undertaken in the Division. This sense of inclusion and empowerment led to positive changes in the attitudes of the staff about evaluation, promoted learning, and, in turn, increased EC. Although these studies were undertaken using case study method and did not use any validated measures for staff empowerment, this research does allow some inferences about its positive relationship with EC, represented in hypothesis 1.
1b. Staff Empowerment and Performance Management

Although there is not a single study in the performance management literature which defined and used staff empowerment measure similar to this study, many alternative concepts and measures such as “decentralization,” “participation” and “employee involvement” have been used in the performance management literature. These studies may allow for an indirect association of performance management with staff empowerment. In her review of literature on performance management, Dooren (2011), for example, argued that decentralization (staff empowerment) should be an essential part of any reform agenda to make performance management successful. Performance management should be decentralized at the level of managers and front-line supervisors who are cognizant of situational needs. Such measures will improve the capacity to use performance information and, consequently, improve performance. In the same vein, decentralization is considered as part and parcel of public management reforms all over the world. Specifically, decentralization and empowerment are essential components of performance management movement advocated by New Public Management (NPM) (Jreisat, 2002; Newcomer & Caudle, 2011). Although decentralization and empowerment are popular slogans in performance management reforms, these concepts have rarely been defined precisely, and their measures have rarely been validated. However, the premise that success of performance management reforms depends on decentralization has found mixed empirical support. Some studies have found that these two are negatively associated and others found support for positive association. Both types of empirical studies are reviewed in the following.

de Maillard and Savage (2012), for example, studied decentralization and performance management in British and French contexts. While comparing performance in Britain and France in policing, de Maillard and Savage (2012) observed that, historically, authority was
decentralized in the British policing system, whereas, the French policing system was more centralized. However, in their quest for better performance management, the policing systems in both countries have moved to more centralization of authority and decision making. An explanation for this could be that performance management is a reason for standardization and therefore could lead to centralization of authority in an organization. Based on this observation, it could be argued that the introduction of a performance management system could lead to lack of staff empowerment. However, the inferences based on this study should be made cautiously because this is a qualitative observation of the authors based on case studies alone.

In his empirical study of performance management in Danish schools, Nielsen (2014) found that some dimensions of decentralization are positively associated with performance management, whereas others are negatively related. Specifically, the author found that decentralization (of goal setting) is negatively associated with performance management. The decentralization of goals could cause huge damage to the success of performance management. For a performance management system to work successfully, goal setting should not be decentralized and left to lower-level managers and supervisors to decide. In the absence of centralization of goal setting, the priorities of managers may change, and they would not be able to work toward achieving centrally set goals. Additionally, task autonomy, defined as “the level of authority over production processes” (Nielsen, 2014, p. 450) was hypothesized to positively moderate the association between performance management and organizational performance. That is, task autonomy could play a significant role in performance improvement through performance management. However, they did not find support for this assertion. The author further argued that this may be because autonomy conflicts with traditional values of political and democratic accountability related to bureaucratic organizations. Moreover, the author also
found that managerial autonomy over human resources was positively associated with the success of performance management. Limited managerial autonomy could hinder managers in implementing performance management. Managerial autonomy becomes even more important if stakeholders and leadership hold managers accountable for performance achievements. This finding is consistent with the argument advanced by some other scholars (e.g., Moynihan & Landuyt, 2009; Moynihan & Pandey, 2010). Although the concept of ‘task autonomy’ and decentralization used by Nielsen is somewhat similar to the staff empowerment concept used in this study, Nielsen (2014) undertook this study at the managerial level, not the staff level. Therefore, these findings should be taken with caution. Consistent with the partial support for positive association between decentralization (empowerment) and success of performance management, two more studies are reviewed in the following.

In their study of Spanish public health care, Casasnovas, McDaid and Costa-Font (2009) found that decentralization (empowerment) of local health care staff could lead to success of performance management. Although decentralization has some costs such as loose central control, absence of coordination, and transitional monetary costs, decentralization and autonomy could go a long way toward improving performance management and holding health care staff accountable. This allows local health care management to tailor strategy according to local needs, get local political support, and be innovative in solving health care issues. The centralization (lack of staff empowerment) of a system does not allow involvement of stakeholders for their input in improving performance. Decentralization and empowerment of local health care staff allows for using performance management tools to effectively monitor performance of local health care staff and improve performance. Stated differently, performance management could be effective only when authority is decentralized.
Consistent with these findings, Nitta, Wrobel, Howard, and Jimmerson-Eddings (2009) found that part of the reason for the failure of performance management and reorganizational reforms in Little Rock School District (LRSD) was inability of leadership to empower staff. Nitta and others (2009) reported findings of their investigation of the failure of the LRSD School Board’s efforts to introduce new performance management and reorganization reforms. At the heart of these reforms was the introduction of a performance measurement system meant to increase transparency and ensure accountability. The authors found that the major reason for the failure of these performance measurement reforms was the inability of school leadership to empower school staff to implement the new performance measurement system. The authors further argued that the success of reforms, including performance measurement, depended on communicating a vision for change to lower level staff and empowering them to own their work and be responsible for that work. These views are also consistent with other authors (e.g., Dooren, 2011; Jreisat, 2002). Along similar lines, performance management can be understood in terms of a learning culture in an organization that could be promoted through staff empowerment and participation (e.g., Moynihan, 2005).

Although a few studies reviewed here found negative or no relationship, most of the studies found or posited a positive association between decentralization (staff empowerment) and the success of performance management (Casasnovas et al., 2009; Dooren, 2011; Jreisat, 2002; Nitta et al., 2009). Moreover, the studies which found negative or no relation between these two variables were either conducted at the managerial level, not the staff level (Nielsen, 2014) or those inferences were merely based on anecdotal evidence (de Maillard & Savage, 2012). Based on these findings, it could be argued that staff empowerment and performance management are
positively associated. This relation can be represented in the following hypothesis and was tested more immediately and directly in the present study.

*Hypothesis 1*: Staff empowerment and EC level in an organization are positively associated.

2a. **Staff Development and Evaluation Capacity**

Staff development is one of the key dimensions of OLC. Learning organizations consistently focus on staff development through training, education, and professional development opportunities (Bess et al., 2011; Cousins et al., 2004; Garvin et al., 2008; Goh, 2000; Marsick & Watkins, 2003). In a learning organization, training and professional development opportunities are provided at both the individual and team levels. These development opportunities are made available for both new and established employees and before initiating a new endeavor. Staff development has been associated repeatedly with better organizational performance (Garvin et al., 2008; Marsick & Watkins, 2003).

The extant EC literature suggests that staff development (i.e., professional development and training) is the most commonly used method for ECB initiatives (Naccarella et al., 2007; Taylor-Ritzler et al., 2013). According to Nielsen and others (2011b) staff development is one of the primary indicator of EC. In addition, Cousins and others (2004) indirectly link staff development and EC level. In their meta-analysis of 36 empirical studies on organizational learning and evaluation, Cousins and others (2004) developed an ECB framework. They argued that staff development activities are part of organizational support structure which promotes a culture of evaluation and organizational learning. This culture of evaluation and organizational learning increases the EC level. Stated differently, staff development and EC are indirectly associated and, in the absence of staff development activities, EC level will not increase and may
Taut (2007b) affirmed this point of view based on her efforts to increase EC level in international development organizations using an action research method. She used training, in addition to other strategies such as self-evaluation, to increase EC level of the organizations. Consistent with Cousins and others (2004), she observed that a culture of learning including staff development initiatives in an organization is important to any increase in EC level.

Although many other formal and informal strategies are used for the purpose of staff development (e.g., coaching and mentoring, team meetings, appreciative inquiry, and communities of practice) (Preskill & Boyle, 2008b), the most commonly used strategies for staff development meant to increase the EC level are training and professional development (Gibbs et al., 2002; Huffman et al., 2008; Preskill & Boyle, 2008a; Satterlund et al., 2013). Most of ECB initiatives using training as a staff development strategy have reported an increase of EC (Adams & Dickinson, 2010; Arnold, 2006; Gilliam et al., 2003; Kuzmin, 2012; Naccarella et al., 2007; Stevenson et al., 2002; Taut, 2007b). Naccarella and others (2007), for example, attempted to increase the EC level of the Division of General Practice, in the Australian government using training and training of trainers as strategies for staff development. The authors reported that these staff development efforts increased the EC level; measured through quality of evaluation reports and staff feedback following this staff development initiative. In their descriptive study Gilliam and others (2003) also reported results of an effort to increase the EC level of several NPOs and public organizations by training project officers and health department personnel as a strategy for staff development. The authors reported that these staff development efforts markedly increased the EC level.

These assertions of linking staff development efforts and increase in EC levels could be problematic for various reasons. These studies did not measure EC using any validated tools.
Most of these studies are descriptive case studies with conclusions based on subjective judgments of the authors. Moreover, staff development efforts were accompanied by other ECB strategies such as technical assistance, manuals, guidance, and technology transfer (Gilliam et al., 2003; Naccarella et al., 2007). Consequently, it is difficult to determine which of these strategies may have increased EC levels. A few studies, however, using staff development to increase EC level have attempted to address some of these issues. Adams and Dickinson (2010), for example, used only staff development as a strategy, that is, an ECB initiative called Easy Evaluation meant to increase the EC level of public health and community organizations in New Zealand. Participants were trained and given opportunity to apply their learning practically.

Adams and Dickinson (2010) reported an increase in EC level, that is, self-reported staff satisfaction with their learning opportunity, as a result of this staff development effort. Although other studies have also used this method (e.g., Satterlund et al., 2013), such self-reported measures of EC level can be challenged with respect to their validity. The study also has other limitations. As a descriptive case study, it does not allow for any strong conclusion about the association between two variables. Another set of problems with such staff development interventions is both dosage and quality (Leviton, 2014; Preskill, 2013; Suarez-Balcazar & Taylor-Ritzler, 2013). A couple of days of modest training may not be enough for an objective increase in EC level on a sustainable basis. No study as yet has been conducted using a robust research method which would establish a valid link between staff development and EC. The present study is an attempt to fill this gap. Based on the weak but consistent findings from these cited studies, however, hypothesis 2 was tested in this study.
2b. Staff Development and Performance Management

Introduction of a performance management system in an organization can be understood as an organizational change process (Moynihan, 2005; Moynihan, Pandey, & Wright, 2012a). Reichers, Wanous, and Austin (1997) argue that a change process always creates uncertainty and skepticism. Specifically, those who are opposed to the change do not see the usefulness of the organizational change. One of the reasons for cynicism is the lack of information about change. Reichers and others (1997) suggested that to lower uncertainty and cynicism, staff should be informed about change and provided training regarding what is the purpose of the change. In case efforts are not made to reduce this cynicism, the change will have a limited chance to be successful. Similarly, Dransfield (2000) argued that, when a performance management system is introduced in an organization, staff teams are expected to change routines, and use the performance management system; this process requires an organizational commitment to development to adopt to new system.

Yang and Hsiesh (2007) also argued that public managers working with performance measurement must be familiar with inputs and processes. However, they often lack an understanding of outputs, outcomes, and evaluative methods. To successfully implement a performance measurement system, managers and staff should be provided with technical training to be able to undertake performance measurement. In their empirical study in the Taiwan context, they found that an organizational investment in a program of technical training was a significant predictor of successful performance management. Based on these findings, the authors suggested that, while implementing a performance measurement system, staff should be provided with evaluation training to be able to perform performance management tasks successfully. Along similar lines, Newcomer (2002) suggested that one of the biggest challenges
of performance management reforms in public sector is lack of staff development to do performance measurement and use that information. Traditionally, public managers and staff are not trained for undertaking performance measurement and program evaluation. Stated differently, there is big gap between what they should be able to do and their present skills. This gap should be filled by the increasing the organizational commitment to staff development.

In the same vein, in his prescriptive work on performance measurement, Hatry (2006) is also of the view that providing staff training is very important to successful implementation of a performance measurement system. Technical performance measurement training ought to provide skills to staff and managers for identifying and measuring performance outcomes and for using performance information in day-to-day business. Various empirical studies conducted in different contexts also confirm this association. For example, Rivenbark and Menter (2006) argued that community-based NPOs lack the capacity to measure performance. In the absence of a strong performance measurement system, local governments providing funding to these community-based NPOs for the delivery of services cannot hold these NPOs accountable. Using a case study method, the authors found that in Aspen, Colorado the local government successfully built the performance measurement capacity of NPOs. As a result of these staff development efforts, NPOs were able to build a performance measurement system and meet reporting mandates of the local government. Drawing on their case study findings, the authors suggested that performance accountability can be ensured by increasing the performance measurement capacity of community-based NPOs through a program of staff training. Although the findings of case studies are not generalizable, it could be inferred that the determination to provide staff development opportunities are positively associated with the success of performance management.
Consistent with Rivenbark and Menter (2006), Nitta and others (2009) reported findings of their investigation of failure of the Little rock School District (LRSD) School Board’s efforts to introduce new management and reorganization reforms. At the heart of these reforms was the introduction of a performance measurement system meant to increase transparency and ensure accountability. They found that one of the reasons for the failure of these performance measurement reforms was inadequate and ineffective training that had been poorly organized for school personnel. As a result of lack of staff development opportunities, the schools’ personnel were not able to understand the new system and communicate well with their colleagues. Consequently, the authors concluded, the new performance measurement system could not advance. Had staff received sufficient training, the authors further argued, the outcomes would have been different.

Based on these studies, it could be argued that, if an organization has provided a sufficient set of opportunities for staff development prior to the process of implementation, the new system of performance management is more likely to be successful. It is, therefore, argued that there is a direct positive link between adequate staff development and the success of a performance management system. This relation can be represented in the following hypothesis and was tested more immediately and directly in the present study.

\textit{Hypothesis 2}: Staff development and EC level in an organization are positively associated.

\textbf{3a. External Alignment and Evaluation Capacity}

External alignment refers to "ongoing practices within the organization that allow members to respond effectively to challenges and opportunities in the organization’s….external environments" (Bess et al., 2011, p. 39). External alignment enables an organization to collect
information about environmental changes and learn from that information. One of the leading factors that contribute to organizational change is external alignment. An organization failing to align itself with external environment has fewer chances to learn and survive (Jenlink, 1994). A learning organization adopts policies and practices pertaining specifically to external alignment. The role of leadership in aligning an organization with its environment is crucial (Bess et al., 2011; Goh & Richards, 1997; Marsick, 2013; Marsick & Watkins, 2003). In fact, most important criterion used to judge performance of an organization may be its long-term survival. To that end, only organizations which align themselves with their environment manage to survive. By the same token, external alignment and success of an organization are closely connected (Bess et al., 2011; Fiol & Lyles, 1985).

Some of the leading reasons for NPOs and public organizations to increase EC level are external forces and pressures such as accountability, legislative mandates, and reporting requirements (Bourgeois & Cousins, 2013; Brown & Reed, 2002; Cousins & Bourgeois, 2014a; Kapucu et al., 2011; King & Volkov, 2005; King, 2007; Preskill & Boyle, 2008b; Stockdill et al., 2002; Volkov, 2008). The organizations that respond to these challenges are, in fact, attempting to align themselves with external environment. Stated differently, ECB initiatives, in most of the cases, are introduced to respond to the environmental circumstances.

Although there is not a single empirical study in the EC literature which exclusively deals with external alignment and indicates a direct relationship between EC and external alignment, some scholars have discussed tangentially this issue. For instance, Bourgeois and Cousins (2013) have argued that external alignment is one of the factors that lead to an increase in the EC level of an organization. That is, they would expect a positive association between external alignment and EC. These observations are consistent with the findings of an empirical study which
indirectly established an association between these two variables (Lawrenz, Thomas, Huffman, & Clarkson, 2008). The study reported results of an effort to increase EC in the education sector where external alignment was a small part the ECB initiative. That is, the school administration wanted to align school curriculum and assessment policies with external assessments methods. The authors reported that overall this initiative was successful and increased the EC level of the school. Admittedly based on these few instances, it is contended that external alignment and EC are positively associated. This relation can be stated in hypothesis 3.

3b. External Alignment\textsuperscript{10} and Performance Management

Public organizations work in a political environment, and stakeholders’ support is critical for their survival and legitimacy. Specifically, success of reforms such as performance management largely depends on support from political stakeholders. Stakeholders’ support strengthens performance management in that poor performance will not lead to budget cuts or blame-avoidance behavior which could be disastrous for performance management (Ho, 2006; Kroll, 2015; Newcomer & Caudle, 2011; Yang & Hsiesh, 2007). Streib and Poister (1999) observed that the real challenge in performance measurement is finding valid and reliable measures in consultation with stakeholders that provide information regarding the achievement of organizational goals. Historically, performance measures were decided from the view point of managers, but stakeholders were involved in the process. This has changed in recent years, however, and there are many studies which provide evidence of involving stakeholder in performance management. A number of studies have empirically investigated the relationship between performance management and external alignment. Most of the studies found a positive

\textsuperscript{10} In the performance management literature, external alignment has been studied under different names such as “environmental factors,” “networking behavior” or “stakeholders’ involvement.”
association between these two variables (e.g., Berman & Wang, 2000; Ho, 2006; Yang & Hsiesh, 2007). In what follows, these studies are critically reviewed.

Ho (2006) studied midwestern U.S. cities with populations between 10,000 and 200,000, to learn about the perception of political stakeholders regarding use of performance tools for various purposes. The author found that support and familiarity of stakeholders with performance measurement was positively associated with the use of performance information. Additionally, familiarity and support of political stakeholders with performance measurement also created enthusiasm among other stakeholders, so that the use of performance information was even greater. He also found that perception that stakeholders do not pay attention to performance measurement was negatively associated with the use of performance information. The involvement of citizens in the development of performance measures also was found to be positively associated with the use of performance information. The author concluded that the involvement of citizens and political stakeholders in performance management could go a long way toward its success. This is because stakeholders’ involvement could support the system and provide legitimacy for performance measurement.

These findings are consistent with those of Yang and Hsiesh (2007) who found in the context of Taiwan that stakeholders’ involvement and their support was positively associated with performance measurement adoption and its success. Based on these findings, the authors suggested that, while implementing a performance measurement system, the participation of stakeholders should be ensured to perform performance management tasks successfully. In his excellent systematic review of performance management literature, Kroll (2015) also found that stakeholders’ involvement is an important driver for improving performance management. Similarly, in their study of performance measurement in the U.S. counties, Berman and Wang
(2000) found that counties with greater use of performance information had greater stakeholder support as compared to low-use counties. They measured stakeholders’ support by the frequency with which managers asked elected officials for their support for their performance measurement system. They found that almost 85% managers in high-use counties requested elected officials for their support, as compared to 33% managers in low-use counties who approached elected officials. Likewise, high-use counties were more likely to seek support for their performance measurement system from advisory boards as compared to low-use counties.

In the same vein, drawing on knowledge of the performance management literature and personal experience with performance management systems in various contexts, Newcomer and Caudle (2011) found that lack of support from stakeholders (including legislators, differ levels of government, or other organizations involved in performance measurement) is one of the reasons for the failure of performance management systems globally. The authors further argue that the greatest benefit would derive when all stakeholders are in agreement regarding the purpose of a performance management system. The involvement of stakeholders plays a significant role in the use of performance data and the success of a performance management system. This views is also supported by Kroll (2015) and Hatry (2006).

Although these studies define and measure external alignment (stakeholders’ involvement) differently, they do inform an expected positive association between performance management and external alignment. This relation can be stated in form of the following hypothesis and was tested in this study:

*Hypothesis 3:* External alignment and EC level in an organization are positively associated.
4a. Experimentation and Evaluation Capacity

Experimentation is the “degree of freedom employees enjoy in the pursuit of new ways of getting the job done and freedom to take risks” (Goh & Richards, 1997, p. 578). Experimentation is another dimension of organizational learning capacity (OLC). Learning processes in an organization include experimentation for the purpose of education and development of employees, as well as innovative thinking toward the solution of problems. Experimentation is an important component of learning processes and practices for an organization to become a learning organization. Many theorists have argued that there ought to be an established mechanism in an organization to experiment with new ideas and evaluate those ideas (Cousins et al., 2014b; Garvin et al., 2008; Goh et al., 2007; Marsick & Watkins, 2003). To develop a culture of experimentation, there should be leadership encouragement, support, and rewards for taking calculated risks in an organization (Cousins et al., 2004; Cousins et al., 2014b; Goh, 2000; Goh et al., 2007) and, of course, this also should be accompanied by psychological safety (Garvin et al., 2008). In the organizational learning literature, experimentation has been associated with learning outcomes, both at individual and organizational levels, and attainment of organizational goals (Fiol & Lyles, 1985; Goh & Richards, 1997; Marsick & Watkins, 2003).

The EC literature has paid scant attention toward culture of experimentation. There are only a few studies which have directed attention toward this important dimension of OLC and attempted to link EC and culture of experimentation (Cousins et al., 2004; Preskill & Torres, 1999a). In their theoretical work, Cousins and others have argued that experimentation in an organization is associated with EC level and that the EC level of an organization could be increased by the introduction of a culture of experimentation (Cousins et al., 2004; Cousins et al., 2014b). Consistent with these ideas, Bourgeois and Cousins (2013) and Preskill and Torres
(1999b) also have pointed out that experimentation should be part of the organizational culture as an ongoing process to make an ECB initiative successful and increase EC. Experimentation and a culture of inquiry allow members to challenge organizational processes without fear and to learn which, in turn, increases EC. Consistent with the point of view of Cousins and others (2004) and Preskill and Torres (1999b), Taut (2007b), in her action research project meant to increase EC level of international development organizations found that, without a culture of inquiry and learning, efforts to increase EC will have limited success.

Fiol and Lyles (1985) and Garvin and others (2008) have argued both that organizational learning necessitates a culture of experimentation and that organizational culture, supportive of experimentation, makes it possible to unlearn outdated methods to solve organizational problems and find innovative solutions. In the EC literature, scholars such as Bourgeois and Cousins (2013) and Preskill and Torres (1999a) have framed the ECB process as an organizational learning system. Stated differently, increasing EC is not possible without an organization becoming a learning organization; a learning organization is one which has a culture of experimentation. In this way, a culture of experimentation and EC are positively associated, according to these scholars, that is, an organization with higher EC is more likely to have a culture of experimentation. This relation can be stated in hypothesis 4.

4b. Experimentation and Performance Management

Most performance management studies have attempted to establish a relationship of performance management with individual factors (Kroll, 2014) such as motivation and the attitude toward performance management (Ammons & Rivenbark, 2008; Kroll & Vogel, 2014; Moynihan, Pandey, & Wright, 2012b), as well as external factors such as political support (Moynihan et al., 2012b; Yang & Hsieh, 2007). However, Taylor (2014) is of the view that an
explanation of the extent of use of performance data could be sought using organizational
cultural perspective. He argues that the use of performance data depends on organizational
culture: if the organizational culture is supportive of experimentation, there is higher probability
that public managers will use performance data for various purposes. This view is also supported
by other studies (Moynihan & Landuyt, 2009; Moynihan & Pandey, 2010). Consistent with
Taylor (2014), Moynihan & Pandey (2010) posited that an organizational culture which
encourages experimentation and rewards innovation (their term is developmental culture) plays
significant role in the success of performance management. In their empirically study of local
government senior managers, they found that an innovative and risk-taking culture is an
important predictor of performance data use. They further suggested that, if managers have
flexibility and freedom to do experimentation with the performance process, they are more likely
to use performance data in decision making. Theoretically, this may be because performance
data were not used for rewarding or punishing employees but, instead, used formatively for
learning and improving. Conversely, an organizational culture that discourages innovation and
experimentation is less likely to use performance data. Two years later, these findings were
confirmed by the same authors in another study (Moynihan et al., 2012a).

Similarly, based on survey data collected from chief executives in mid-sized U.S. cities,
Folz, Abdelrazek, and Chung (2009) found that the success of performance measurement
depends on culture of experimentation. That is, if an organization has a culture of
experimentation, it is more likely to be successful in implementing performance measurement.
Experimentation and a trial-and-error process allows staff to refine performance measures and
generate performance information which helps them make informed decisions. Stated differently,
if the culture of organization does not allow for experimentation, the organization will not learn
from the process of performance measurement implementation. Consistent with this, in their study of Swedish local governments Johansson and Siverbo (2009) also reported that performance management (measured as higher performance data use) was positively associated with an experimentation culture (measured in terms of change routines). The authors argued that successful adoption of performance management comes with the adoption of experimental processes over a long period of time. Since Swedish local governments encourage this experimentation, they frequently use performance information. As a result of such experimentation with change routines, both staff and organizations learn what works and what does not work. In this process of experimentation, staff use performance information.

Based on this strong empirical evidence, it could be argued that performance management is positively associated with a culture of experimentation. That is, if organizational culture permits experimentation while staff perform various tasks, there will be greater chance for performance data to be used. This relationship is hypothesized in the following:

Hypothesis 4: Experimentation and EC level in an organization are positively associated.

5a. Clear Organizational Goals and Evaluation Capacity

Clear organizational goals or goal specificity refers to the understanding of mission, vision, performance measures and direction of an organization by its employees. The clear understanding of organizational goals leads to an increase in feedback and useful evaluation of an employee’s work. It also enables an organization to adopt a clear strategy for achieving those goals. Clear organizational goals not only increase the performance of employees but also helps an organization to achieve those goals easily (Wright & Davis, 2003). At the opposite, organizational goal ambiguity leads to employee role ambiguity and confusion in an organization (Pandey & Wright, 2006). A clear mission, organizational goal specificity, and understanding of
their importance are building blocks of organizational learning. In the absence of a clear mission and goals, an organization cannot innovate and learn. Top management plays a crucial role in communicating a clear mission and organizational goals to employees (Bess et al., 2011; Cousins et al., 2004; Cousins et al., 2014b; Goh, 2000).

There are some ECB frameworks which include clearly defined program/organizational goals as part of an ECB strategy. Two of these frameworks are catalyst-for-change and getting-to-outcomes. Empirical evidence suggests that both of these frameworks have led successfully to greater EC. Garcia-Iriarte and others (2010) report findings of an ECB initiative using a catalyst-for-change approach toward ECB in which one person (with leadership responsibility) is provided with evaluation skills who then trains others. The study found that success of the ECB initiative depended on clearly defining the goals of the program. A leader, who is assigned the responsibility of ECB, should communicate clear goals of the program and demonstrate how achieving those goals will lead to success of the program.

Another ECB approach which advocates including clear program/organizational goals is known as getting-to-outcomes as described by Wandersman (2014a). This approach is considered a “how-to” for empowerment evaluation. In this approach, before starting an ECB initiative, special attention is paid to setting clear, well-defined, and measureable organizational goals. The following steps include efforts toward achieving those goals. Accountability through measurement of goals is the essence of this approach. According to the authors, this approach has been used successfully for ECB and for achieving organizational goals Wandersman, Snell-Johns, Lentz, Fettermann, Keener, Livet, Imm, and Flasphohler (2005) and other advocates of empowerment evaluation (Diaz-Puente, Yague, & Afonso, 2008) have supported these claims.
A direct empirical relationship between clear organizational goals and EC has been established in very few studies (e.g., Bourgeois & Cousins, 2013; Carman & Fredericks, 2009), although multiple studies have linked EC and clear organizational goals with indirect evidence. The studies which have linked indirectly these two variables have highlighted the importance of clear organizational goals in the process ECB (e.g., Clinton, 2014; Connally, 2002; Fourny, Gregson, Sugerman, & Bellow, 2011; Kapucu et al., 2011; King & Volkov, 2005; Labin, 2014; Labin et al., 2012; Preskill & Caracelli, 1997; Shulha & Cousins, 1997; Taylor-Ritzler et al., 2013). Taylor-Ritzler and others (2013), for example, established indirect link between clear organizational goals and EC in their empirical study of 169 NPOs. The study found that organizational factors, which included clear organizational goals, are predictors of a greater use of evaluation findings. The study, however, had several limitations, including a low response rate of 20% and a mixing of a few items about goal clarity with other leadership issues. Clinton (2014) found evidence of link between organizational goals and ECB process (although not directly with EC) in a study using data from 292 initiatives from two public health programs in New Zealand. The findings of the study suggested that clear program/organizational goals were associated with various successful ECB activities.

Bourgeois and Cousins (2013) undertook one of the few studies which directly linked EC and clear organizational goals. In their study of the Canadian federal government, the authors found that organizations with exemplary EC have very clearly defined goals. The members of those organizations shared a good understanding of the purpose and goals of their organizations. They developed this understanding through formal and informal meetings, retreats, and strategic planning sessions. Conversely, the members of organizations with low EC did not understand the goals of their organizations. In other words, their organizations were suffering from goal
ambiguity. Similarly, in his study of K-12 schools, Trevisan (2002) highlighted the role of a clear organizational mission in determining the EC level of an organization. The author concluded that a clear mission and EC level are positively associated and noted that the schools which did not have a clear mission (i.e., to undertake evaluation of their counselling programs) had lower EC. Surprisingly, the school districts with better resources had lower EC, apparently due to their inability to clearly define their mission. Consistent with Trevisan (2002), Carman and Fredericks (2009) in their empirical study on NPOs found that clear organization goals and EC are positively associated. The authors conducted a survey of evaluation practices in 189 randomly selected NPOs and found that the group of NPOs which had a clear mission and organizational goals were successfully undertaking evaluation and had higher EC. At the opposite, the group of organizations which had no clear mission and organizational goals had less EC.

Although these studies appear to establish a direct or indirect relationship between clear organizational goals and EC, these studies have limitations. None of these studies offer an explicit definition of what are clear organizational goals or EC. In addition, most of these studies were conducted using a case study method which does not allow more than an informed judgment concerning a link between these two variables. Nonetheless, the empirical evidence does justify the positive relation represented in hypothesis 5.

5b. Clear Organizational Goals and Performance Management

All public sector reform movements in recent years such as NPM and Reinventing Government require organizations to clearly define their goals, set strategic plans, and collect data on performance measures to establish whether an organization is achieving those goals (Frederickson & Frederickson, 2006; Moynihan, 2008). Specifically, the U.S. federal government has mandated government agencies to publish strategic plans that include clear
organizational goals and related performance measures through legislative measures such as the GPRA of 1993 and the Modernization Act of 2012 (Rainey & Jung, 2015). Although performance measures, strategic planning and goal setting, and the use of these measures to assess organizational performance have been the subject of controversy (Moynihan, 2008; Radin, 2006), goal clarity remains at the heart of performance management reform (Rainey & Jung, 2015). Notwithstanding the controversy of strategic planning and defining organizational goals (Radin, 2006), many studies have consistently found that clear organizational goals and performance management are positively associated in different contexts.

In their study of local government senior managers, Moynihan and others (2012a) found that performance information use (performance management) was positively associated with clear organizational goals. The study was undertaken to investigate the role leadership plays in clarifying organizational goals and promoting performance management practices. The authors found that transformational leadership played important role in this relationship: the leader promotes performance data use by clearly defining organizational goals. The theoretical explanation for this relationship is that goal clarity gives a clear indication to managers regarding those tasks are critical to achieving those goals and how to accomplish those tasks. Consequently, the managers are more likely to use task data for accomplishing those tasks.

This positive association between performance management and goal clarity also was confirmed in other studies involving same authors (Moynihan & Pandey, 2005; Moynihan et al., 2012b; Moynihan & Landuyt, 2009). In their study of public and nonprofit organizations, Moynihan and others (2012b) found that goal clarity and performance management are positively associated. However, it is pertinent to mention that in this study goal clarity was used as a control variable. In their study of state government health and human services officials
engaged in performance management activities, Moynihan and Pandey (2005) found that employees’ perception of organizational effectiveness (measured as the degree to which employee think their organization is accomplishing its core mission) was positively associated with clear organizational goals. In their study of Texas state agencies, Moynihan and Landuyt (2009) also reported the same findings. The authors found that mission orientation (clear organizational goals) is positively associated with organizational learning (measured in terms of use of performance information for learning).

In their theoretical work, O’Toole and Meier (2015) reviewed the performance management literature to clarify conflicting findings on performance management studies coming from various contexts (O’Toole & Meier, 2015). They theorized that performance is affected by two types of contextual factors: external (environmental and political) and internal (organizational) factors. They argued that goal clarity is one of the most important organizational factors affecting performance management. Theoretically, O’Toole and Meier (2015) recognized that organizations come in existence to achieve certain goals. Practically, however, it is very challenging, especially in the public sector, for an organization to clearly define goals. Such goal ambiguity may be the result of conflict between competing organizational goals. Goal ambiguity also could arise in public organizations from the need to serve multiple stakeholders and public interest groups. Irrespective of its causes, goal ambiguity affects the performance of organizations negatively. The authors theorized that goal clarity has a positive effect on performance management. Latham and others (2008) argued that clear goals are positively associated with performance management because clear goals affect choice and persistence. In other words, specific goals guide individuals and teams to prioritize tasks and accomplish what is more important as compared to less important.
Based on this empirical evidence, performance management and goal clarity would appear to be positively associated. This relationship is hypothesized in the following and analytically tested in this study:

_Hypothesis 5:_ Organizational goal clarity is positively associated with EC level.

**Revised Framework**

Based on the literature reviewed for this study, the following conclusions can be drawn. As shown in Figure 2.1, Cousins and colleagues (2014b) do not directly see relationship between EC and OLC. However, the literature presented here clearly suggests that EC and OLC are directly associated, that is, there is positive association between EC and at least five dimensions of OLC as indicated in Hypotheses 1-5. This relationship is depicted in Figure 2.2. In addition, Cousins and colleagues’ (2014b) framework indicates that OLC is a consequence of EC. That is, there is a causal link between EC and OLC (i.e., ECB naturally increases OLC). However, the authors directly did not link EC and OLC. What they did not propose is that this relationship could be two-way. The literature presented here clearly shows that there may be a two-way relationship between EC and OLC indicated by a positive association between EC and OLC; the direction of this relationship, however, is unknown. The direction of this relationship is not possible to determine without experimental or longitudinal studies.

The process to increase EC in an organization is considered as an organizational learning process (Cousins et al., 2004; Cousins et al., 2014b; Preskill & Torres, 1999a) and, for an organization to learn, it must have some OLC. Stated differently, first an organization should have the OLC to learn—to increase EC level. Once an organization starts learning, OLC could also be increased. In this way, there is a two-way relation between EC and OLC as depicted in Figure 2.2. Moreover, based on the literature cited here, it is argued that the organizations with
higher EC are different than those with the low EC. That is, the organizations with a higher EC level have higher OLC, and organizations with the lower EC level have lower OLC. The present study was an attempt to test this conjecture empirically.
Figure 2.2 Revised Conceptual Framework of Evaluation Capacity Building Based on Cousins and others
In this chapter the research design that was used for data collection and data analysis for the study is presented. This chapter is divided into three parts. First, the sample and procedures for the study are described. Second, the measures of the organizational learning capacity (OLC) and evaluation capacity (EC) construct are explained. Third, the data analysis strategy is presented.

Sample and Procedures

This study was undertaken as part of a research project conducted in collaboration between the Rockefeller College of Public Affairs and Policy and the United Way of the Greater Capital Region (UWGCR).11 UWGCR identified a list of 72 nonprofit organizations (NPOs)12 that it had funded historically or that recently had expressed interest in future UWGCR funding.

A systematic selection of 36 organizations was made based on initial assessment of EC levels of the 72 NPOs. To judge whether a NPO had high or low EC, two criteria were used.  

---

11 The United Way of the Capital Region is a part of the United Way of America, a charitable organization which raises funds in partnership with other organizations and distributes those funds among local NPOs with the help of its regional branches.
12 NPOs and public organization face similar pressures to be accountable, efficient, and effective (Carman & Fredericks, 2009; Kapucu, Healy, & Arslan, 2011; Seiden & Sowa, 2011). As a result of these external pressures from stakeholders, funders, or taxpayers, both NPOs and public organizations must pay special attention toward building EC and performance management systems to be able to demonstrate future improvements. The process of ECB and implementation of performance measurement systems do not appear to be different in NPOs and public organizations. Additionally, the performance management and EC literatures do not distinguish between these two types of organizations. The literature reviewed in the previous section is a testament to this statement. Based on these similarities, it is argued that NPOs and public organizations could be understood and studied from the same lens of analysis. The present study, therefore, also does not distinguish between NPOs and public organizations. As discussed in this section, the data were collected from NPOs and the implications of findings are discussed for both NPOs and public organizations.
First, in the spring of 2014, four staff members of UWGCR who were familiar with the NPOs assessed these organizations using a checklist of four EC indicators: organization collects data to document program outputs and outcomes, organization prepares useful reports on program performance and effectiveness, organization is one of the best agencies for its program monitoring and evaluation, and organization does not seem to be engaged in very much program improvement (see Appendix 1, Table 1). Staff members were asked to check any relevant indicators of EC. In case they did not know about an organization’s EC, they left all four indicators unchecked. Based on these four UWGCR staff members’ work, the frequency of checkmarks for each EC indicator for each organization was computed. These frequencies were used as the first criterion to assess each organization’s EC.

Second, in the summer of 2014 UWGCR required interested NPOs to submit a letter of intent to request funding. The letter of intent included a question concerning EC. Based on the replies given by the NPOs on this question, three to six UWGCR volunteers assessed each organization’s EC through rating each organization on this question, “Organization demonstrates the willingness to collect performance data and make adjustments over the two-year funding cycle.” UWGCR volunteers rated each organization based on a four-point rating scale: strongly agree, agree, disagree, and strongly disagree. UWGCR volunteers ratings were coded from “1” for strongly disagree to “4” for strongly agree. The arithmetic mean of rating scores made by the UWGCR volunteers for each organization was calculated and used as the second criterion to assess each organization’s EC.

Before finalizing the list of selected NPOs, inter-rater reliability and convergent validity of these criteria were tested. Organizations where these two sources of information did not agree with each other were excluded. Based on both UWGCR staff assessments and
UWGCR volunteer ratings a systematic selection of 36 organizations was made so that a rectangular distribution of EC levels might be approximated. The aim in this selection process was to generate approximately the same proportion of organizations at high and low EC while achieving a representative and balanced sample with respect to organizational size, geographic location, and service mission. UWGCR subsequently determined that four of these 36 organizations were inappropriate for inclusion, providing various reasons for declination but primarily recent leadership transitions. The UWGCR sent email invitations to leaders of the remaining 32 organizations encouraging participation in the collaborative project. Altogether 26 of these 32 nonprofit organizations agreed to join in the project, a participation rate of over 0.80.

At each organization, seven-page printed questionnaires were distributed to staff members systematically selected across job titles. The number of questionnaires varied between five and nine depending on the size of the organization. A total of 170 questionnaires were distributed and 154 questionnaires were returned, a response rate of over 0.90.\(^{13}\) Upon completion of a questionnaire, it was sealed by the respondent in an unmarked envelope to preserve the anonymity of each participant. All individual envelopes then were enclosed and sealed in a larger unmarked envelope to preserve the anonymity of each organization. Questionnaires were coded independently by two project staff members; every discrepancy in the separate coding was reconciled before the data were added to the project database. For each organization, a three-digit numerical code was attached to its own records.

Overall less than 3% data of the data for the OLC measures were missing. Approximately 14% of the data specifically for the EC measure were missing, since employees at lower

\(^{13}\) The arithmetic mean for the number of questionnaires distributed in each of the 26 organizations was 6.5; the mode was 7.0.
organizational levels appeared unfamiliar with the full extent of ECB efforts. About 10% of the respondents, excluding gender and age-related questions, did not answer questions such as their current position, years in current position, and years in current organization.

**Study Measures**

The seven-page printed questionnaire contained 17 OLC measures (88 items altogether), 23 EC questions, as well as 18 demographic codes.

**Independent Measures**

Out of 17 OLC measures embedded in the questionnaire, five measures were relevant to the present study: staff empowerment, staff development, external alignment, experimentation and clear organizational goals. The present study examined how organizations with high EC differ from organizations with low EC with respect to these five selected OLC measures. Each of these measures already had been validated. For some items, six-point Likert-type response scales soliciting the degree of a respondent’s agreement with the statement were used: strongly agree, generally agree, agree a little, disagree a little, generally disagree, and strongly disagree. For the remaining items, five-point frequency scales of occurrence were used: almost always, often, sometimes, rarely, never or almost never.

This study used the Dimensions of the Learning Organization Questionnaire (DLOQ) to measure staff empowerment, staff development, and external alignment. The DLOQ was developed in the 1990s to assess learning culture in for-profit organizations by Watkin and colleagues (Marsick, 2013). The DLOQ measures seven aspects of a learning organization: inquiry and dialogue, team learning, embedded system, staff empowerment, continuous learning (staff development), system connection (external alignment), and strategic leadership (Song, Joo, & Chermack, 2009). Of these seven, only three measures from the DLOQ (i.e., staff
empowerment, staff development, and external alignment) were relevant and used for the study. The original DLOQ version, comprised of 43 items, is claimed to have been validated by a number of studies (Dahanayake & Gamlath, 2013; Marsick & Watkins, 2003). However, a shorter version of DLOQ which consists of 21 items was derived from the original version and partially validated by Yang, Watkins, and Marsick (2004). The shorter version of DLOQ is considered as reliable as the original DLOQ version and was used to measure staff empowerment, staff development, and external alignment for this study. Using a sample of 836, Yang and others (2004) have at least partially validated staff empowerment, staff development, and external alignment with good evidence of reliability (Cronbach’s alpha =.81, .84, and .80, respectively). The DLOQ also has been partially validated in international contexts such as Sri Lanka (Dahanayake & Gamlath, 2013), Korea (Song et al., 2009), Iran (Nielsen, Rasmussen, & Sadegh-Sharifirad, 2011), and Latin America (Hernandez & Watkins, 2003).

**Staff empowerment** is one of the key measures of OLC. In an organization which empowers staff, “people are involved in setting, owning, and implementing a joint vision; responsibility is distributed close to decision making so that people are motivated to learn toward what they are held accountable to do” (Marsick & Watkins, 2003, p. 139). This study used four items to measure staff empowerment. Three of these were Yang and colleagues’ (2004) items partially validated in the shorter version of the DLOQ to measure staff empowerment. These three items were “This organization recognizes people for taking initiative,” “This organization gives people control over the resources they need to accomplish their work,” and “This organization supports employees who take calculated risks.” A fourth item from the DLOQ that measured staff empowerment also was included in the questionnaire: “This organization gives us
choices in our work assignment” (see Appendix 1, Table 2). Some modifications in the wording of these items were made to achieve consistency in the questionnaire design.14

**Staff development** is another core measure of OLC. Staff development is the form of learning designed “into work so that people can learn on the job; opportunities are provided for ongoing education and growth” (Marsick & Watkins, 2003, p. 139). This study used Yang and colleagues’ (2004) items partially validated in the shorter version of the DLOQ to measure staff development. The three items were “We help each other learn,” “People are rewarded for learning,” and “People are given time to support learning” (see Appendix 1, Table 2). Some modifications in the wording of these items were made to achieve consistency in the questionnaire design.15

**External alignment** is yet another measure central to the construct of OLC. In an organization which is externally aligned, “people are helped to see the effect of their work on the entire enterprise; people scan the environment and use information to adjust work practices; and the organization is linked to its communities” (Marsick & Watkins, 2003, p. 139). This study used Yang and colleagues’ (2004) items partially validated in the shorter version of the DLOQ to measure external alignment. These three items are “This organization encourages people to think from a regional perspective,” “This organization encourages people to get answers from across

---

14 The original wording of four items measuring staff empowerment was as follows.
   1. My organization recognizes people for taking initiative.
   2. My organization gives people control over the resources they need to accomplish their work.
   3. My organization supports employees who take calculated risks.
   4. My organization gives people choices in their work assignments.

15 The original wording of three items measuring staff development was as follows.
   1. In my organization, people help each other learn.
   2. In my organization, people are rewarded for learning.
   3. In my organization, people are given time to support learning.
the organization when solving problems,” and “This organization works together with the outside community to meet mutual needs” (see Appendix 1, Table 2). Some modifications in the wording of these items were made to achieve consistency in the questionnaire design.\textsuperscript{16}

\textbf{Experimentation} is the “degree of freedom employees enjoy in the pursuit of new ways of getting the job done and freedom to take risks” (Goh & Richards, 1997, p. 578). The Learning Organization Survey (LOS) was developed in the 1990s to assess OLC in for-profit organizations (Garvin et al., 2008) and is the only OLC instrument that includes a measure of experimentation. The original version of the LOS consisted of 55 items to measure three building blocks (ten measures) of organizational learning: supportive learning environment (including measures of psychological safety, appreciation for differences and openness to new ideas, and time for reflection); concrete learning processes and practices (including measures of experimentation, information collection, analysis, education and training, and information transfer); and leadership that reinforces learning. Of these ten, only the measure of experimentation is relevant for this study. Singer and others (2012) developed a shorter form of the LOS that consists of 27 items, and they partially validated it using a sample of 6996, reporting an acceptable level of reliability for experimentation (Cronbach’s alpha =.86). This study used the same four items as previously to measure experimentation. These items were “This program experiments with new ways of working,” “This organization experiments with new service or program offerings,” “This organization has a formal process for conducting and

\textsuperscript{16} The original wording of three items measuring external alignment was as follows.
1. My organization encourages people to think from global perspective.
2. My organization encourages people to get answers from across the organization when solving problems.
3. My organization works together with the outside community to meet mutual needs.
evaluating experiments or new ideas,” and “This program employs pilot projects or simulations when trying out new ideas” (see Appendix 1, Table 2). Some modifications in the wording of these items were made to achieve consistency in the questionnaire design.17

Clear organizational goals or goal specificity “represents the degree to which employees believe that they understand or can explain the direction, purpose, and performance measures of the organization” (Wright & Davis, 2003, p. 76). No validated measure pertaining to organizational goals appeared in the DLOQ or LOS. A measure of clear organizational goals was developed by Wright and Davis (2003) in their study of public sector organizations using a sample of 385 of New York State employees at an acceptable level of reliability (Cronbach’s alpha =0.73). Using a sample of 570 managers, Pandey and Wright (2006) also used a parallel measure (termed goal ambiguity) with good evidence of reliability (Cronbach’s alpha of 0.73). This study used the same four items to measure clear organizational goals: “This organization’s mission is clear to almost everyone who works here,” “It is easy to explain the goals of this organization to outsiders,” “This organization has explicitly stated goals,” and “This organization has objectives that are specific and well-defined” (see Appendix 1, Table 2).

Dependent Measure of Evaluation Capacity (EC)

EC was assessed with both dichotomous and continuous measures. As mentioned above, NPOs with high and low EC were identified using UWGCR staff assessments and UWGCR volunteer ratings. In addition to these two separate sources of information, employee perceptions

17 The original wording of four items measuring experimentation was as follows.
   1. This unit experiments frequently with new ways of working.
   2. This unit experiments frequently with new product or service offerings.
   3. This unit has a formal process for conducting and evaluating experiments or new ideas.
   4. This unit frequently employs prototypes or simulations when trying out new ideas.
concerning the EC of their organizations also were sought in the questionnaire. In a section of the questionnaire dedicated to EC, four questions pertained to evaluation activities including the frequency of program performance data collection, the frequency of preparing useful evaluation reports, the frequency of staff meetings to discuss performance data collected, and the frequency of meetings with other organizations and external stakeholders. Additionally, ten other items were included to seek employee perceptions regarding EC of their organizations (see Appendix 1, Table 3). These items generally regarded leadership support, incentives, and resources available for evaluation. In these items, four-point response scales inquired about the degree of accuracy of each statement: very accurate, moderately accurate, slightly accurate and not accurate. These ten items were “Our leaders are committed to the evaluation of our projects and programs,” “Financial resources are available to implement program evaluation processes,” “There is a plan and a timeline for completing regular and routine evaluations,” “Stakeholders / funders expect our organization to report on program outcomes,” “We have sufficient technical skills to support evaluations and use the findings,” “There are positive incentives and rewards for planning and managing evaluations,” “Meetings are convened for us to plan and prepare for anticipated evaluations,” “Computer hardware and software are available to well manage evaluation data,” “Our program staff have a genuine interest in supporting evaluation processes,” and “We participate in meetings to review and discuss the findings of past evaluations.” Based on these three sources of information, EC as dichotomous dependent measure was created, coding NPOs as high EC and as low EC. Additionally, EC as a continuous dependent measure was created using all three sources of information, namely, UWGCR staff assessments, UWGCR volunteer ratings, and questions included in the questionnaire. The validation of the both the dichotomous and continuous dependent measures of EC is detailed in the next chapter.
Analysis

Three following analytical techniques were used for data analysis. In order to see how five measures of OLC differ in terms of high or low EC, discriminant analysis were used. Discriminant analysis is a multivariate method used when the purpose is to examine how two or more groups are different with respect to multiple measures. Models using discriminant analysis are good for both empirical explanation and prediction (Klecka, 1980; Vlachonikolis, 1980). Since the research question was how organizations with high EC are different than those with low EC with respect to OLC, discriminant analysis was the most suitable technique to examine this difference. This technique not only allowed for explaining the differences between organizations with high EC and low EC, but also is helpful in predicting how far an organization is on its way to building EC. Similarly, using EC as a dichotomous dependent measure (high EC coded as 1 and low EC as 0), logistic regression analysis were also performed. Logistic regression analysis allows to predict the probability of nominal measure at a given independent measure. Finally, using EC as continuous dependent measure, ordinary least square (OLS) regression analysis were performed to see how EC is associated to OLC measures. All these analyses were conducted at 95% power (β=.05), an α=.05 (one-tailed), and 1.2 of standard deviation units effect size (Dµ). To test the difference at the stated power and α levels between two populations, the minimum required sample size is approximately 13 NPOs for each group (Bailey, 1971). In this way, the present study had sufficient sample size to perform these analyses.
CHAPTER 4: RESULTS

In this chapter the results of the study are presented. This chapter is divided into three parts. First, the demographic characteristics of the survey respondents are reviewed. Second, the psychometric properties of the measures are presented. In the third part, the results of discriminant analysis, logistic regression, and ordinary least square (OLS) model are presented.

Demographic Characteristics of Respondents

The demographic characteristics of survey respondents are summarized in Table 4.1. In terms of ethnicity, nearly 80% of the respondents identified themselves as White; only about one respondent in five replied with Black, Hispanic, Asian, or some other self-description. Approximately three-quarters of the respondents were female. Ages of study participants ranged from 20 to 74 years with an arithmetic mean of 43.7 and a standard deviation of 15.4. Nearly half of the respondents were under 40 years of age, and one respondent in five was 60 years of age or older.

Of the 154 study participants, the full hierarchy of organizational responsibility was represented; 14% of the respondents were clerical or support staff, and 21% were senior managers or executives. The length of employment in the organization ranged from one month to 44 years with an arithmetic mean of 7.1 years and a standard deviation of 8.0 years. Nearly one-quarter of the respondents had been employed for less than two years, and another quarter had been employed for 10 years or more. With respect to length of employment in their current positions, this time ranged from one month to 30 years with an arithmetic mean of 4.7 years and a standard deviation of 6.1 years. Nearly two-fifths of the respondents had been placed in their positions for under two years.
Table 4.1
Demographic Characteristics of Survey Participants  
(n = 154)

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>104</td>
<td>79%</td>
</tr>
<tr>
<td>Black</td>
<td>16</td>
<td>12%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>7</td>
<td>5%</td>
</tr>
<tr>
<td>Asian</td>
<td>3</td>
<td>2%</td>
</tr>
<tr>
<td>Others</td>
<td>1</td>
<td>1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>94</td>
<td>73%</td>
</tr>
<tr>
<td>Male</td>
<td>35</td>
<td>27%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-29 years</td>
<td>29</td>
<td>23%</td>
</tr>
<tr>
<td>30-39 years</td>
<td>31</td>
<td>25%</td>
</tr>
<tr>
<td>40-49 years</td>
<td>17</td>
<td>14%</td>
</tr>
<tr>
<td>50-59 years</td>
<td>22</td>
<td>18%</td>
</tr>
<tr>
<td>60-69 years</td>
<td>18</td>
<td>14%</td>
</tr>
<tr>
<td>70 or above</td>
<td>8</td>
<td>6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Position</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clerical/Support</td>
<td>19</td>
<td>14%</td>
</tr>
<tr>
<td>Professional/Technical</td>
<td>38</td>
<td>28%</td>
</tr>
<tr>
<td>Manager</td>
<td>51</td>
<td>38%</td>
</tr>
<tr>
<td>Senior Manager/Executive</td>
<td>28</td>
<td>21%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Years in Current Organization</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1.9 years</td>
<td>32</td>
<td>23%</td>
</tr>
<tr>
<td>2-4.9 years</td>
<td>42</td>
<td>30%</td>
</tr>
<tr>
<td>5-9.9 years</td>
<td>30</td>
<td>22%</td>
</tr>
<tr>
<td>10-14.9 years</td>
<td>12</td>
<td>8%</td>
</tr>
<tr>
<td>15 years or more</td>
<td>22</td>
<td>16%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Years in Current Position</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1.9 years</td>
<td>49</td>
<td>38%</td>
</tr>
<tr>
<td>2-4.9 years</td>
<td>43</td>
<td>33%</td>
</tr>
<tr>
<td>5-9.9 years</td>
<td>20</td>
<td>15%</td>
</tr>
<tr>
<td>10-14.9 years</td>
<td>8</td>
<td>6%</td>
</tr>
<tr>
<td>15 years or more</td>
<td>10</td>
<td>8%</td>
</tr>
</tbody>
</table>
Psychometric Properties of the Measures

Factor Analysis of the Independent Measures

Exploratory Factor Analysis (EFA) was performed using Stata 12.1 version to explore factor structure of the study measures and their discriminant validity. EFA was conducted using Principal Component Factor method. To make the pattern of loadings and factor structure of study measures more independent, orthogonal (parsimax) rotation was used. The factor structure of four measures staff empowerment, staff development, experimentation, and clear organizational goals was explored jointly, and the factor structure of external alignment was explored jointly with three other measures of learning organization (not part of this study); job feedback, demand and job clarity. Rotated factor loadings pattern matrices are summarized in Appendix 2 Table 1 and Table 2.

Of the four items used in the survey for the staff empowerment measure, three items loaded heavily and uniquely on the staff empowerment factor between 0.56 and 0.76 (Table 4.2). One item (‘‘This organization supports employees who take calculated risks’’) which did not load uniquely on the staff empowerment factor was dropped from the analysis. To strengthen the staff empowerment measure, another item was added (‘‘Teams/groups have the freedom to adapt their goals as needed’’) which loaded heavily and uniquely on the staff empowerment factor. Thus, these four items were used to create factor scores for the staff empowerment factor.

Of the three items used in the survey for the staff development measure, two items loaded heavily and uniquely on the staff development factor between 0.58 and 0.75 (Table 4.2). One item (‘‘We help each other learn’’) which did not load uniquely on the staff development factor was dropped from the analysis. To strengthen the staff development measure, another item was added (‘‘Our leaders support requests for learning opportunities and training’’) which loaded
heavily and uniquely on the staff development factor. Thus, these three items were used to create factor scores.

Of the three items used in the survey for the external alignment measure, two items loaded heavily and uniquely on the external alignment factor (Table 4.2). One item (“This organization encourages people to get answers from across the organization when solving problems”) which did not load uniquely on external alignment factor was dropped from the analysis. To strengthen the external alignment measure, two more items were added (“This organization builds alignment of visions across different levels and work groups,” and “This organization actively engages with interest groups that share our mission and goals”) which loaded heavily and uniquely on the external alignment factor. Thus, these four items were retained to create factor scores.

Of the four items used in the survey for the experimentation measure, only two items (“This organization has a formal process for conducting and evaluating new ideas or experiments,” and “This program employs pilot projects or simulations when trying out new ideas”) loaded uniquely on the experimentation factor. These loadings on the experimentation factor were 0.78 and 0.81 (Table 4.2). Two items which did not load uniquely (“This program experiments with new ways of working,” and “This organization experiments with new service or program offerings”) were dropped from the analysis. Thus, only two items were retained to create factor scores for experimentation.

Of the four corresponding items used in the survey for the clear organizational goals measure, three items loaded heavily and uniquely on the clear organizational goals factor from 0.68 to 0.72 (Table 4.2). One item (“This organization has objectives that are specific and well-
### Table 4.2
Rotated Factor Loadings for Survey (With Parsimax Rotation)

<table>
<thead>
<tr>
<th>Survey Item</th>
<th>Factor Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Staff Empowerment</strong></td>
<td></td>
</tr>
<tr>
<td>1 This organization recognizes people for taking initiative.</td>
<td>0.76</td>
</tr>
<tr>
<td>2 This organization gives people control over the resources they need to</td>
<td>0.66</td>
</tr>
<tr>
<td>accomplish their work.</td>
<td></td>
</tr>
<tr>
<td>3 Teams/groups have the freedom to adapt their goals as needed.</td>
<td>0.75</td>
</tr>
<tr>
<td>4 This organization gives us choices in our work assignments.</td>
<td>0.56</td>
</tr>
<tr>
<td><strong>Staff Development</strong></td>
<td></td>
</tr>
<tr>
<td>1 People are rewarded for learning.</td>
<td>0.58</td>
</tr>
<tr>
<td>2 Program staff are given time to support learning.</td>
<td>0.75</td>
</tr>
<tr>
<td>3 Our leaders support requests for learning opportunities and training.</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>External Alignment</strong></td>
<td></td>
</tr>
<tr>
<td>1 This organization works together with the outside community to meet mutual</td>
<td>0.69</td>
</tr>
<tr>
<td>needs.</td>
<td></td>
</tr>
<tr>
<td>2 This organization encourages us to think from a regional perspective.</td>
<td>0.81</td>
</tr>
<tr>
<td>3 This organization builds alignment of visions across different levels and</td>
<td>0.59</td>
</tr>
<tr>
<td>work groups.</td>
<td></td>
</tr>
<tr>
<td>4 This organization actively engages with interest groups that share our</td>
<td>0.72</td>
</tr>
<tr>
<td>mission and goals.</td>
<td></td>
</tr>
<tr>
<td><strong>Experimentation</strong></td>
<td></td>
</tr>
<tr>
<td>1 This organization has a formal process for conducting and evaluating new</td>
<td>0.82</td>
</tr>
<tr>
<td>ideas or experiments.</td>
<td></td>
</tr>
<tr>
<td>2 This program employs pilot projects or simulations when trying out new</td>
<td>0.78</td>
</tr>
<tr>
<td>ideas.</td>
<td></td>
</tr>
<tr>
<td><strong>Clear Organizational Goals</strong></td>
<td></td>
</tr>
<tr>
<td>1 This organization’s mission is clear to almost everyone who works here.</td>
<td>0.68</td>
</tr>
<tr>
<td>2 It is easy to explain the goals of this organization to outsiders.</td>
<td>0.72</td>
</tr>
<tr>
<td>3 This organization has explicitly stated goals.</td>
<td>0.69</td>
</tr>
<tr>
<td>4 This organization seems to be without central purpose or apparent</td>
<td>0.70</td>
</tr>
<tr>
<td>direction.</td>
<td></td>
</tr>
</tbody>
</table>
defined”) which did not load uniquely on the clear organizational goals factor was dropped from the analysis. To strengthen the measure, another item was added (“This organization seems to be without central purpose or apparent direction”) which loaded heavily and uniquely on the clear organizational goals factor. Thus, these four items were used to create factor scores.

**Internal Reliability of the Independent Measures**

Internal reliability of the independent measures was obtained using Cronbach’s coefficient alpha. The alpha level of measures ranged from 0.62 to 0.82. According to Nunnally and Bernstein (1994) 0.70 is an acceptable alpha level. Some other authors also have confirmed that 0.70 is an acceptable alpha level (Acock, 2012; Muijs, 2010). Four of five independent measures achieved the acceptable alpha level. The detail of each of the measure is given in the following.

**Independent Measures**

*Staff empowerment* was measured using the four items. The coefficient alpha for the four-item staff empowerment measure was 0.79, a strong evidence of internal validity. Inter-item correlations of these four items ranged from 0.44 and 0.55, and the item-test correlation (or item-total correlation) of the measure ranged from 0.72 and 0.83. Additional analysis showed that the reduction of further items would decrease the coefficient alpha. The coefficient alpha for the staff empowerment measure with a description of the items is summarized in Table 4.3.

*Staff development* was measured using the three items. The coefficient alpha for the three-item staff development measure was 0.81. Although, the addition of a fourth item (“We help each other learn”), could increase the coefficient alpha to 0.84, it negatively affected the discriminant validity of the measure and, therefore, was not included in the measure. According to Nunnally and Bernstein (1994) coefficient alpha more than 0.80 is strong evidence of internal
Table 4.3
*Measure Reliability Coefficients for Independent Measures (n=154)*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Item-test correlation</th>
<th>Alpha if item deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Staff Empowerment (Cronbach’s Alpha Level 0.79)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This organization recognizes people for taking initiative.</td>
<td>0.78</td>
<td>0.74</td>
</tr>
<tr>
<td>This organization gives people control over the resources they need to accomplish their work.</td>
<td>0.81</td>
<td>0.72</td>
</tr>
<tr>
<td>Teams/groups have the freedom to adapt their goals as needed.</td>
<td>0.83</td>
<td>0.71</td>
</tr>
<tr>
<td>This organization gives us choices in our work assignments.</td>
<td>0.72</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>Staff Development (Cronbach’s Alpha Level 0.81)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>People are rewarded for learning.</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>Program staff are given time to support learning.</td>
<td>0.85</td>
<td>0.77</td>
</tr>
<tr>
<td>Our leaders support requests for learning opportunities and training.</td>
<td>0.83</td>
<td>0.79</td>
</tr>
<tr>
<td><strong>External Alignment (Cronbach’s Alpha Level 0.78)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This organization works together with the outside community to meet mutual needs.</td>
<td>0.78</td>
<td>0.72</td>
</tr>
<tr>
<td>This organization encourages us to think from a regional perspective.</td>
<td>0.69</td>
<td>0.79</td>
</tr>
<tr>
<td>This organization builds alignment of visions across different levels and work groups.</td>
<td>0.81</td>
<td>0.70</td>
</tr>
<tr>
<td>This organization actively engages with interest groups that share our mission and goals.</td>
<td>0.82</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>Experimentation (Cronbach’s Alpha Level 0.62)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This organization has a formal process for conducting and evaluating new ideas or experiments.</td>
<td>n/a</td>
<td>0.35</td>
</tr>
<tr>
<td>This program employs pilot projects or simulations when trying out new ideas.</td>
<td>n/a</td>
<td>0.40</td>
</tr>
<tr>
<td><strong>Clear Organizational Goals (Cronbach’s Alpha Level 0.73)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This organization’s mission is clear to almost everyone who works here.</td>
<td>0.73</td>
<td>0.69</td>
</tr>
<tr>
<td>It is easy to explain the goals of this organization to outsiders.</td>
<td>0.72</td>
<td>0.70</td>
</tr>
<tr>
<td>This organization has explicitly stated goals.</td>
<td>0.77</td>
<td>0.65</td>
</tr>
<tr>
<td>This organization seems to be without central purpose or apparent direction.</td>
<td>0.76</td>
<td>0.66</td>
</tr>
</tbody>
</table>
reliability. The inter-item correlations of these three items ranged from 0.53 and 0.58, and the item-test correlation of the measure ranged from 0.80 and 0.85. The coefficient alpha for the staff development measure with a description of the items is summarized in Table 4.3.

**External alignment** was measured using the four items. The coefficient alpha for the four-item external alignment measure was 0.78. Although dropping one item (“This organization encourages people to think from a global perspective”) could slightly increase the coefficient alpha to 0.79, it negatively affected the construct validity of the measure and, therefore, was not dropped. Nonetheless, 0.78 is far higher than an acceptable alpha level of 0.70. The inter-item correlation of these four items ranged from 0.42 and 0.56, and the item-test correlation of the measure ranged from 0.69 and 0.82. The coefficient alpha for external alignment measure with a description of the items is summarized in Table 4.3.

**Experimentation** was initially measured using the three items. The coefficient alpha for the three-item experimentation measure was 0.56, less than an acceptable alpha level. Further analysis, however, suggested that dropping one item (“This program experiments with new ways of working”) improved the alpha level to 0.62. Since this two-item measure had higher internal reliability than that of the three-item measure, the two-item measure was used in this study. This is the only measure in this study where alpha level was less than 0.70. The coefficient alpha for experimentation measure with a description of the items is summarized in Table 4.3.

**Clear organizational goals** or goal specificity was measured using the four items. The coefficient alpha for the four-item clear organizational goals measure was 0.73. Although the addition of a fifth item (“This organization has objectives that are specific and well-defined”) could increase the coefficient alpha to 0.74, it negatively affected the discriminant validity of the measure and, therefore, was dropped. The inter-item correlation of these four items ranged from
0.38 and 0.44, and the item-test correlation of the measure ranged from 0.72 and 0.77. The coefficient alpha for the clear organizational goals measures is summarized in Table 4.3.

**Factor Scores and Organization Level Data**

Factor scores for each of the five organizational learning capacity (OLC) predictors (i.e., staff empowerment, staff development, external alignment, experimentation, and clear organizational goals) were created using Stata version 12.1. The advantage of using factor scores, instead of mean or total scores, is that factor scores use a regression coefficient weight for each item while creating factor scores. That is, the items which load high on a particular factor are given more weight than those with low loadings when creating factor scores (Acock, 2012). Further, with orthogonal rotation, the factor scores for the predictors are uncorrelated, achieving the greatest discriminant validity in measurement. Since the organization was the unit of analysis in the present study (n=26), individual-level factor scores for each of the five independent measures of OLC were aggregated for participants using their arithmetic means within organizations as the descriptive statistics for each of the 26 NPOs.

**Dependent Measure of Evaluation Capacity**

EC, the dependent measure in this study, was measured using seven items. Of these seven items, as mentioned in the method section, one item was generated by UWGCR volunteers, two items originated with the UWGCR staff, and four items came from the questionnaire.18

---

18 In total, 14 items, as mentioned in method of study, were included in the questionnaire to seek employee perceptions concerning the EC of their organizations. Of those 14 items only four items which converged with UWGCR staff assessments and UWGCR volunteer ratings were used for further analysis. These four items for each organization (to create organizational level data) were created by counting number of respondents who checked certain options about program evaluation activities in an organization (see Table 4.4).
Table 4.4
*Measure Reliability Coefficients for Dependent Measure (n=26 NPOs)*

<table>
<thead>
<tr>
<th>Items</th>
<th>Item-test correlation</th>
<th>Alpha if item deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Average volunteer ratings of EC from initial letter of intent.</td>
<td>0.82</td>
<td>0.74</td>
</tr>
<tr>
<td>2 Number of the United Way staff who checked organization ‘not engaged’ in evaluation.</td>
<td>0.71</td>
<td>0.77</td>
</tr>
<tr>
<td>3 Number of the United Way staff who checked organization ‘prepares useful’ reports.</td>
<td>0.70</td>
<td>0.77</td>
</tr>
<tr>
<td>4 Proportion of respondents who checked 2, 3, or 4^a</td>
<td>0.65</td>
<td>0.78</td>
</tr>
<tr>
<td>5 Proportion of respondents who checked 2, 3, or 4^b</td>
<td>0.73</td>
<td>0.76</td>
</tr>
<tr>
<td>6 Proportion of respondents who checked 1, 2, or 3^c</td>
<td>0.54</td>
<td>0.80</td>
</tr>
<tr>
<td>7 Proportion of respondents who checked 1, 2, or 3^d</td>
<td>0.57</td>
<td>0.80</td>
</tr>
</tbody>
</table>

---

a. Is information about the accomplishment of your organization’s projects and programs ever collected?
   
   If:
   
   no=0
   
   Once a year=1
   Twice a year=2
   About once a month=3
   About once a week=4

b. Are program reports prepared that include some of the performance information collected?
   
   If:
   
   no=0
   
   Once a year=1
   Twice a year=2
   About once a month=3
   About once a week=4

c. There is a plan and a timeline for completing regular and routine evaluations.
   
   If:
   
   Not Accurate (NA)=0
   Slightly Accurate (SA)=1
   Moderately Accurate (MA)=2
   Very Accurate (VA)=3

d. Computer hardware and software are available to well manage evaluation data.
   
   If:
   
   Not Accurate (NA)=0
   Slightly Accurate (SA)=1
   Moderately Accurate (MA)=2
   Very Accurate (VA)=3
Since the organization was the unit of analysis in the present study (n=26), individual-level responses for four items measuring EC, coming from the questionnaire, were aggregated using proportion of respondents checking certain options about program evaluation activities within organization as the descriptive statistics for each of the 26 NPOs. The coefficient alpha for the seven-item EC measure was 0.80 which is a strong evidence of internal validity, as suggested by Nunnally and Bernstein (1994). The inter-item correlation of these seven items ranged from 0.32 and 0.41, and the item-test correlation of the measure ranged from 0.54 and 0.82. Additional analysis showed that the reduction of further items would not improve the coefficient alpha. For each of these seven items, organizations’ scores were standardized. These statistical results for the dependent measure are summarized in Table 4.4.

Then the arithmetic mean of the seven items z scores was computed to create a single measure used as dependent continuous measure for analysis. Dichotomous dependent measure for high EC organization was created by separating the 13 NPOs with positive EC measurements (high EC levels) from the 13 NPOs with negative EC measurements (low EC levels).

**Bivariate Analysis of Study Measures**

Bivariate correlations among study measures at the organizational level are given in Table 4.5. There were only four instances where correlation between two measures was significant at $p < .05$. The remaining 11 correlations between study measures were not significant. This could be because of orthogonal rotation was used for factor analysis which maintains zero correlation among factors. Bivariate correlation among study measures ranged between 0.0 and 0.52. No pair of independent measures shared as much as 25% of the variance in common; the largest correlation ($r=.47$) was between staff development and external alignment.
Table 4.5
Bivariate Correlations among Study Measures (n=26 NPOs)

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clear Organizational Goals</td>
<td>(0.73)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Staff Development</td>
<td>-0.10</td>
<td>(0.81)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Staff Empowerment</td>
<td>-0.15</td>
<td>0.22</td>
<td>(0.79)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. External Alignment</td>
<td>0.39*</td>
<td>0.47*</td>
<td>-0.12</td>
<td>(0.78)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Experimentation</td>
<td>0.04</td>
<td>0.19</td>
<td>-0.13</td>
<td>0.10</td>
<td>(0.62)</td>
<td></td>
</tr>
<tr>
<td>6. Evaluation Capacity</td>
<td>0.52*</td>
<td>0.30</td>
<td>0.00</td>
<td>0.18</td>
<td>0.43*</td>
<td>(0.80)</td>
</tr>
</tbody>
</table>

*p < .05

Cronbach alphas are along the diagonal in parentheses.
The largest bivariate correlation was 0.52 between EC and clear organizational goals. Most of the study measures produced correlations below 0.30 indicating considerable discriminant validity among measures used for the study.

**Multivariate Analysis and Hypothesis Testing**

To investigate how EC is associated with the five independent measures of OLC, three types of analysis were conducted: 1) discriminant analysis to check whether high and low EC organizations differed; 2) logistic regression analysis using EC as a dichotomous dependent measure coded as 1 for the 13 higher EC organizations and 0 for the 13 lower EC organizations;\(^{19}\) and 3) ordinary least squares (OLS) regression analysis where EC was used as continuous measure.

**Discriminant Analysis**

Before testing hypotheses, it is important to mention that difference between the sizes of high and low EC organizations (measured as annual revenues) was not significant. Median annual revenue of low EC organizations in 2013 was 1.40 million dollars and median revenue of high EC organizations was 1.45 million dollars. Mean annual revenue of low EC organizations was 2.54 million dollars with a standard deviation of 3.83 million dollars. Mean annual revenue of high EC organizations was 1.74 with a standard deviation of 0.98 million dollars. To determine if size of low and high capacity organizations was different (which could be a factor associated with EC) the difference between annual revenue was tested as shown in Table 4.6.

---

\(^{19}\) All 13 higher EC organizations produced positive EC scores on the dependent measure. All lower EC organizations produced negative EC scores.
Table 4. 6
Descriptive and t-statistics for Revenues in Million Dollars (Size) in 2013 of Nonprofit Organizations Included in the Study

<table>
<thead>
<tr>
<th></th>
<th>Organizations with Low Evaluation Capacity (n=13)</th>
<th>Organizations with High Evaluation Capacity (n=13)</th>
<th>Difference in Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>1.40</td>
<td>1.45</td>
<td>t-statistic</td>
</tr>
<tr>
<td>Arithmetic Mean</td>
<td>2.54</td>
<td>1.74</td>
<td>0.72</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3.83</td>
<td>.98</td>
<td>p&gt;.05</td>
</tr>
</tbody>
</table>
The difference between size of organizations of both types was not found significant ($p > .05$). Since the difference between sizes of high and low EC organizations was not significant, size could be eliminated as a factor immediately associated with level of EC in an organization.

Discriminant analysis was used to assess how well a straightforward classification of high and low EC organizations was possible using the five OLC factors selected for the present study: staff empowerment, staff development, external alignment, experimentation, and clear organizational goals. The most important assumption made about discriminant analysis is that data follow multivariate normality (Afifi, May, & Clark, 2011; Tabachnick & Fidell, 2012). Therefore, prior to undertaking discriminant analysis, multivariate normality of the data was checked. The test of multivariate normality checks the null hypothesis that all measures together are normally distributed. Univariate and multivariate normality tests are reported in Table 4.7. Multivariate test statistics failed to reject all null hypotheses (Mardia mSkewness test $p < 0.82$, Mardia mKurtosis test $p < 0.31$, Henze-Zirkler test $p < 0.90$, and Doornik-Hansen test $p < 0.84$). Based on these tests, it was inferred that all five factors included in this discriminant analysis had multivariate normality. If a set of measures shows multivariate normality, then individual variables also are normally distributed. Univariate normality tests also failed to reject all null hypotheses (at the .05 level) (Table 4.7). This is confirmed in Kernal Density Graph 4.1 that all five factors included in the study followed univariate normal distributions.

Summary results of the discriminant analysis are reported in Table 4.8. The Wilk’s lambda of 0.49 indicated that high and low EC organizations group differed significantly;
<table>
<thead>
<tr>
<th>Factors</th>
<th>Adj Chi²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff Empowerment</td>
<td>1.08</td>
<td>0.58</td>
</tr>
<tr>
<td>Staff Development</td>
<td>2.38</td>
<td>0.30</td>
</tr>
<tr>
<td>External Alignment</td>
<td>0.47</td>
<td>0.79</td>
</tr>
<tr>
<td>Experimentation</td>
<td>0.82</td>
<td>0.66</td>
</tr>
<tr>
<td>Clear Organizational Goals</td>
<td>1.84</td>
<td>0.40</td>
</tr>
</tbody>
</table>

**Tests for Multivariate Normality**

<table>
<thead>
<tr>
<th>Tests</th>
<th>Adj Chi²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mardia mSkewness</td>
<td>27.38</td>
<td>0.82</td>
</tr>
<tr>
<td>Mardia mKurtosis</td>
<td>1.01</td>
<td>0.31</td>
</tr>
<tr>
<td>Henze-Zirkler</td>
<td>0.01</td>
<td>0.90</td>
</tr>
<tr>
<td>Doornik-Hansen</td>
<td>5.66</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Graph 4.1

Normality Testing: Kernal Density Graphs of the Study Measures

Learning Organization Measures

- **Clear Organizational Goals**
  - Kernel = epanechnikov, bandwidth = 0.2170

- **Staff Development**
  - Kernel = epanechnikov, bandwidth = 0.2679

- **Staff Empowerment**
  - Kernel = epanechnikov, bandwidth = 0.2458

- **Experimentation**
  - Kernel = epanechnikov, bandwidth = 0.2355

- **External Alignment**
  - Kernel = epanechnikov, bandwidth = 0.2452
Table 4.8
*Discriminant Analysis of Evaluation Capacity Level*

<table>
<thead>
<tr>
<th>Eigen Value</th>
<th>Canonical Correlation</th>
<th>Wilks' Lambda</th>
<th>F</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.06</td>
<td>0.71</td>
<td>0.49</td>
<td>4.22</td>
<td>5</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Table 4.9
*Discriminant Classification of Evaluation Capacity*

<table>
<thead>
<tr>
<th>Actual Group</th>
<th>N</th>
<th>Predicted Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High Evaluation Capacity</td>
</tr>
<tr>
<td>High Evaluation Capacity</td>
<td>13</td>
<td>11 (85%)</td>
</tr>
<tr>
<td>Low Evaluation Capacity</td>
<td>13</td>
<td>2 (15%)</td>
</tr>
</tbody>
</table>
consequently the discriminant function was significant. The given F statistic (F=4.22) tests the null hypothesis that the canonical correlation of the discriminant function is zero in the population. Given the \( p \)-value of 0.01, the null hypothesis was rejected; the canonical correlation (0.71) of the discriminant function was shown to be statistically significant. When this discriminant function was used to classify the 26 NPOs into high and low EC groups, 85% of NPOs with high EC and 77% of NPOs with low EC were correctly classified. Stated differently, only 3 of the 13 higher EC organizations may have been mistakenly labelled as such in this study and only 2 of the 13 lower EC organizations may have been initially identified incorrectly (Table 4.9).

**Logistic Regression Analysis**

A stepwise logistic regression model was computed using Stata version 12.1 to determine which of the five factors of OLC were significantly associated with the probability of high EC in an organization. The results of this logistic regression model are reported in Table 4.10. Three of five OLC factors were found to be significant predictors of dichotomous EC classification being used (\( p<0.05 \)): staff development (coefficient of 2.18), experimentation (coefficient of 2.33), and clear organizational goals (coefficient of 2.31). Goodness of fit, likelihood ratio chi-square test suggested that the overall model was statistically significant. Likelihood ratio chi-square tests the null hypothesis that all of the logistic coefficients in the model are equal to zero in the population. The chi-square test value 12.50 (\( p\)-value .01) indicated that the null hypothesis was rejected, evidence of goodness of fit. The pseudo R\(^2\) was 0.35. This also indicated that the logistic coefficients in the population are not equal to zero and that these three factors could be used to predict EC.

Logistic regression coefficients do not have a substantive interpretation, therefore, results of discrete changes from the minimum to the maximum in factors also are presented in Table
4.10. This is the change in the probability of EC if all other predictor factors are at their means, and one predictor measure is changed from its minimum to maximum. These results suggest, for example, that varying observed staff development factor scores from its minimum -1.54 (no effort to develop staff) to maximum 1.22 (substantial staff development opportunities) increased the predicted probability of EC from 0.03 to 0.92, an increase of 0.89, if all other predictor factors are held at their means.

Similarly, change in observed experimentation factor scores from its minimum -1.05 (few new ways of doing work) to maximum 0.80 (staff are encouraged to test new ways of working) increased the predicted probability of EC from 0.07 to 0.85, an increase of 0.77, if all other predictor factors are held at their means. Change in clear organizational goals factor scores from its minimum -0.85 (organizational goals are not very clear) to maximum 0.81 (organizational goals are very clear to everyone in organization) increased the predicted probability from 0.11 to 0.85, an increase of 0.74, if all other predictor factors are held at their means. Since the stepwise logistic regression model was computed, two factors, staff empowerment and external alignment, which were not statistically significant, are not reported.

**Hypothesis Testing**

Hypothesis 1 predicted that staff empowerment and EC level in an organization are positively associated. Contrary to hypothesized relation, stepwise logistic regression model suggested that staff empowerment did not significantly affect an organization’s probability of EC. Stated differently, the null hypothesis that staff empowerment has no effect on EC cannot be rejected.
Table 4.10
Logistic Regression Results for Factors Associated with Evaluation Capacity as Dependent Measures (n=26)

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Logistic Coefficient</th>
<th>Change in Predicted probability from Minimum to Maximum(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>From: factor=min</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To: factor=max</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Difference</td>
</tr>
<tr>
<td>Staff Development</td>
<td>2.18** (1.18)</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.89</td>
</tr>
<tr>
<td>Experimentation</td>
<td>2.33** (1.31)</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.77</td>
</tr>
<tr>
<td>Clear organizational</td>
<td>2.31** (1.26)</td>
<td>0.11</td>
</tr>
<tr>
<td>Goals</td>
<td></td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.74</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.016 (0.52)</td>
<td></td>
</tr>
<tr>
<td>Pseudo R Square</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>LR Chi Square</td>
<td>12.50**</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses
** p<0.05 (one tailed tests)

a. This is the change in probability of evaluation capacity if all other predictor variables are at their means, and one predictor variable is changes from its minimum to maximum.
Hypothesis 2 posited that staff development and EC level in an organization are positively associated. Staff development significantly and positively affected an organization’s probability of EC \( (z=1.84, p=.03 \text{ for a one-tailed test}) \) (Table 4.10). Stated differently, staff development opportunities had a positive effect on EC.

Hypothesis 3 posited that external alignment and EC level in an organization are positively associated. Contrary to the hypothesized relation, stepwise logistic regression model suggested that external alignment did not significantly affect an organization’s EC.

Hypothesis 4 posited that experimentation and EC level in an organization are positively associated. As predicted, experimentation significantly and positively affected an organization’s probability of EC \( (z=1.78, p=.04 \text{ for a one-tailed test}) \) (Table 4.10). Stated differently, a culture of experimentation in an organization had a positive effect on EC.

Hypothesis 5 predicted that organizational goal clarity is positively associated with EC level. As predicted, clear organizational goals significantly and positively affected an organization’s probability of EC \( (z=1.83, p=.04 \text{ for a one-tailed test}) \) (Table 4.10). Stated differently, clear organizational goals and EC were positively associated.

**Ordinary Least Square Regression Analysis**

Finally a stepwise OLS regression model using EC as a continuous dependent measure also was computed to assess which of the five factors of OLC were associated with EC levels in NPOs organizations. This approach had the advantage of not relying on an initial, dichotomous classification scheme for higher EC and lower EC organizations but, rather, using a more exact continuous measure of EC level for each of the 26 NPOs. The results of OLS regression model are reported in Table 4.11. Consistent with the findings of logistic regression analysis, the same three OLC factors, namely, staff development \((\beta=0.28)\), experimentation \((\beta=0.36)\), and clear
organizational goals ($\beta=0.53$) were found to be statistically significant predictors ($p<0.05$) of EC levels. The multiple coefficient of determination was statistically significant ($R^2 = .51$, $p<0.05$); over half of the variability in the observed levels of EC was explained by this three-predictor model. Contrary to two of the hypotheses for the present study, however, neither staff empowerment nor external alignment entered the regression equation to explain additional variability in EC.
<table>
<thead>
<tr>
<th>VARIABLE</th>
<th>(1) Coefficients</th>
<th>(2) Standardized Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff Development</td>
<td>0.34** (0.18)</td>
<td>0.28** (0.18)</td>
</tr>
<tr>
<td>Experimentation</td>
<td>0.49** (0.21)</td>
<td>0.36** (0.21)</td>
</tr>
<tr>
<td>Clear Organizational Goals</td>
<td>0.78** (0.22)</td>
<td>0.53** (0.22)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.02 (0.10)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.51</td>
<td>0.51</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
** p<0.05
CHAPTER 5: DISCUSSION AND CONCLUSION

Due to the ever increasing demand for accountability, NPOs and public organizations have paid greater attention to building an appropriate capacity for evaluation and performance management systems (Brown & Reed, 2002; Naccarella et al. 2007; Preskill & Boyle, 2008a; Satterlund et al., 2013). Consequently, evaluation capacity (EC), evaluation capacity building (ECB) and performance management have received the attention of both scholars and practitioners in recent years. EC and performance management could play important role not only in holding NPOs and public organizations accountable but also in providing information about the success of programs and learning lessons to improve those programs. Drawing on relevant literature, this study claimed that both EC and performance management are identical concepts and treated them as synonyms. This is the first study to bring these two concepts closer and to bridge two streams of literature.

EC as a topic of study, however, has received scant attention in previous public management research. Most of the studies which have investigated EC also have assumed ECB to be a process of individual learning and attempted to establish a relationship between these two (Taylor-Ritzler et al., 2013). Little attention has been paid to understanding EC with respect to organization-level variables. For example, empirical research concerning the relationship between EC and organizational learning capacity (OLC) are few. The present study is unique in that the connection between EC and OLC was investigated using Cousins and colleagues’ (2014b) framework. Three convergent analytical techniques (i.e., discriminant analysis, logistic regression analysis, and ordinary least squares regression analysis) were engaged to establish whether any of five dimensions of OLC (i.e., staff empowerment, staff development, external alignment, experimentation, and clear organizational goals) could link to the level of EC in 26
NPOs. All three multivariate methods supported the findings that EC and OLC, in fact, are positively associated. Specifically, three out of five research hypotheses were empirically supported in the predicted positive direction: staff development, experimentation and clear organizational goals.

**Findings Consistent with Previous Studies**

The extant EC literature suggests that staff development training is the most commonly used method for ECB initiatives (Adams & Dickinson, 2010; Arnold, 2006; Gilliam et al., 2003; Kuzmin, 2012; Naccarella et al., 2007; Stevenson et al., 2002; Taut, 2007b; Taylor-Ritzler et al., 2013). Consistent with the assertions of a number of EC scholars (Nielsen et al., 2011b; Cousins et al., 2004; Preskill & Boyle, 2008a; Preskill & Boyle, 2008b; Gibbs et al., 2002; Huffman et al., 2008; Satterlund et al., 2013; Treiber et al., 2013), this study found that EC and staff development are positively associated. In this study, it was observed that NPOs with higher EC provided more staff development opportunities than those with lower EC. That is, staff development opportunities can increase EC level of an organization. The present study and extant literature offer compelling evidence of a strong positive association between EC and staff development. The practical implication of this finding is that organizations attempting to increase EC should make a strong commitment to expanding staff development opportunities. In short, the role and importance of staff development in building EC cannot be overemphasized.

Consistent with two ECB frameworks-- catalyst-for-change (Garcia-Iriarte et al., 2010) and getting-to-outcomes (Wandersman, 2014a)-- this study found that EC levels in organizations were positively associated with clear organizational goals. Although goal specificity has been posited in some studies to be associated with EC, the present study is the first to empirically test this relationship. It was observed that organizations with higher EC had clearly defined goals,
whereas organizations with low EC exhibited goal ambiguity. This may be due to the fact that clear organizational goals not only increase the performance of employees but also help an organization to achieve performance objectives more easily (Wright & Davis, 2003). Conversely, organizational goal ambiguity leads to employee role ambiguity and confusion in an organization (Pandey & Wright, 2006). Organizational leaders attempting to increase EC should assure that the mission and goals of their organizations are explicit and well understood. In fact, this facet of OLC appears to have the strongest connection to the level of EC observed in the 26 NPOs.

Another important finding of the study was that the level of EC did not depend on the size of the NPO. This finding is consistent with Carman and Fredericks (2010) who found that the type of services provided by NPOs, the size of the organizations, and their various sources of funding do not affect evaluation practices. Stated differently, a large organization could have lower EC, while a small organization could have higher EC. This finding should be very encouraging for smaller organizations with fewer resources. Staff development opportunities do not need to be expensive, experiments with pilot projects can be designed efficiently, and establishing clear organizational goals can be one essential item on a regular meeting agenda.

Consistent with many performance management studies (e.g., Folz et al. 2009; Moynihan & Pandey, 2009; Moynihan et al., 2012a; Siverbo, 2009) the present study also confirmed that experimentation and performance management are positively associated. The theoretical explanation of this relationship could be that an experimentation culture allows staff to learn what works and what does not work. In this trial and error process, staff use performance data (a measure of success of performance management), which in turn, promotes organizational learning. Based on these findings, NPOs and public organizations attempting to introduce performance measurement system should promote experimentation culture in organization. That
is, staff and managers should be given freedom to do their job in new ways. This will not only allow organizational leadership know what works and what does not work during performance management implementation processes, but an experimentation culture also will entail greater performance data use and make the performance management system successful.

Consistent with many studies (e.g., Latham et al. 2008; Moynihan et al. 2012a, 2012b; Moynihan & Pandey, 2005; Moynihan & Landuyt, 2009) it was also found that the success of performance management is positively associated with clear organizational goals. The theoretical explanation for this association is that lack of centrally set goals would not allow political and organizational leadership to hold managers accountable for anything. Similarly, lack of goal clarity leads to goal ambiguity which negatively affects performance management (Moynihan, 2008; Moynihan & Pandey, 2005). Latham and others (2008) argue that clear goals are positively associated with performance management because clear goals affect choice and persistence. In other words, specific goals guide individuals and teams to priorities and accomplish what is more important as compared to less important. The association between clear organizational goals and performance management has been consistently found in various contexts and the present study confirms this in the context of NPOs. Moynihan and colleagues (2012a) found that transformational leadership plays an important role in clearly defining organizational goals.

Based on the findings of the present study and empirical evidence cited in this study, the leaders of organizations attempting to build EC and performance management systems should pay attention to clearly defining and communicating organizational goals.

**New Findings**

ECB generally has been considered to be a set of training activities directed at individual staff members rather than organizational-level interventions associated with institutional
development and change. Some scholars alternatively have conceptualized EC as an organizational process embedded in organizational practices and organizational culture (Cousins et al., 2004; Cousins et al., 2014b; Preskill & Boyle, 2008a, b; Preskill & Torres, 1999b). Unfortunately, there have been few attempts to gather empirical evidence that would support this organizational-level perspective. The present study provided empirical evidence that EC can be considered to be an organizational-level learning process. Specifically, EC was found to be positively associated with organizational processes such as commitment to staff development opportunities, a culture of experimentation, and establishment of clear organizational goals. The ordinary least squares regression model reported here indicated that these three OLC factors can predict over half of the variation in the level of EC observed in 26 NPOs. An appropriate level of EC apparently cannot be attained without simultaneously enhancing OLC, that is, becoming more effective as a learning organization.

Although some theoretical articles have mentioned that experimentation and EC are associated (Cousins et al., 2004; Cousins et al., 2014b), no prior investigation has tested this relationship empirically. The present study documented that a culture of experimentation in an organization was positively associated with EC level in an organization. Experimentation in an organization makes it possible to unlearn outdated methods to solve organizational problems and find innovative solutions (Fiol & Lyles, 1985; Garvin et al., 2008). It appears that organizations with low EC lacked a culture of experimentation and were significantly more resistant to trying out new ways of working. Organizational leaders who are determined to increase EC should not only encourage pilot projects and demonstration programs that could lead to genuine innovation but also establish protocols and procedures to formally assess the problems and advantages that they entail.
The present study was the first empirical investigation of its kind which attempted to differentiate between organizations with low and high EC. It answered the question consistently raised in the EC literature regarding what uniquely characterizes an organization with higher EC. That is, the present study identified key aspects of an organization with optimum level of EC based on data from 26 NPOs. Discriminant analysis suggested, as mentioned earlier, that high EC organizations differ in terms of staff development, experimentation, and clear organizational goals. This finding is significant from both theoretical and practical perspectives. Theoretically, scholars can investigate which of these variables are antecedents or consequences. The present study also is important from the viewpoint of practice. Many scholars have argued that the practice of ECB should be guided by robust scientific knowledge (Wandersman, 2014a, b). Drawing on the present study, practitioners can direct their efforts toward organizational learning processes associated with higher EC. In this way, unlike the current state of practice which is more driven by anecdotal evidence, the study has played a role to establish a science of practice.

One of the most important contributions of the present study in the EC literature is to bring the EC and organizational learning literature closer and to integrate these two streams of literatures empirically. Although some theoretical articles have mentioned OLC in the EC literature, few studies have clearly defined the OLC construct or drawn on the organizational learning literature. Drawing on the organizational learning literature, the present study precisely defined the OLC construct and used previously validated measures to establish its relationship with EC. Future studies can use the measures of OLC validated in the present study to investigate the relationship between organizational learning and EC in different contexts. Almost all measures demonstrated an acceptable level of internal validity and these measures could readily provide a base for future research.
This study would be one of only a few to provide a strong empirical knowledge base to the EC literature which lacks a strong empirical foundation (Bourgeois & Cousins, 2013; Clinton, 2014; Cousins et al., 2004). Most of the studies might be called “reflective case studies” and may not be even called “empirical studies at all” (Cousins et al., 2004, p. 131). However, the present study has made an important contribution due to its use of a rigorous research method. The study was conducted using a robust quantitative method with data from 26 NPOs. Future studies could be conducted using this study as a basis for strengthening the empirical foundations of the EC literature.

The study also found that EC and external alignment were not associated. That is, EC level in an organization and linkages of an organization with outside communities and involvement of stakeholders had no relationship. Some studies in the EC literature have suggested that involvement of outside communities and stakeholders increase EC in an organization (Bourgeois & Cousins 2013; Lawrenz et al., 2008). However, contrary to that line of argument, the present study documented that the level of EC in an organization was not associated with an organization’s linkages to external communities. It could also be argued that an organization could have higher EC without being linked to external communities and stakeholders. Stated differently, an organizational alignment with its environment does not necessarily lead to a greater level of EC.

This lack of support for an association between the EC level of an organization and external alignment, however, should not be interpreted as lack of empirical support for the need to involve outside communities and groups in organizations to build EC. It also should not be considered as evidence against collaborative approaches such as empowerment and participatory
evaluations. This is largely because external alignment as a dimension of OLC comes from for-profit organizational studies and is defined in terms of the link between and organization and its environment and outside community to meet mutual needs. That is, the members of organization are aware of trends and developments in the area of an organization’s services and products. On the other hand, in empowerment and participatory evaluation perspectives, however, local communities and stakeholders are involved in the process of evaluation. In this case, the goal is to empower local communities and stakeholders to monitor and evaluate their own progress.

Since external alignment measure used in this study, and empowerment or participatory evaluations are two different concepts and should not be mixed. The lack of association between EC and external alignment, therefore, should not be considered as lack of empirical support for empowerment or participatory evaluation.

Contrary to the argument put forward in some studies (Anderson et al., 2012; Rotondo, 2012; Wharton & Alexander, 2013), the present study found that there is no relationship between staff empowerment and EC in the 26 NPOs. Staff empowerment was conceptualized in this study as decentralization or transferring authority to lower level staff in an organization. Based on these findings, it could be inferred that transferring authority and decision making at lower level of an organization does not necessarily increase EC in an organization. That is, an organization with centralized authority could have higher EC and an organization with decentralized authority may have lower EC. This also indicates that an organization could attain an optimum EC level without transferring authority to lower level staff of an organization. This finding should be very encouraging for organizations where authority is centralized. They could build an optimum level

---

20 Although empowerment and participatory evaluations differ from each other in many respects and proponents of empowerment evaluation such as Fetterman and colleagues do not want empowerment evaluation to be bracketed with participatory evaluation.
of EC without changing structure of organization and by providing staff development opportunities, experimenting with pilot projects, and establishing clear organizational goals.

This study partially supported collaborative approaches of ECB such as empowerment and participatory evaluations. Although indirectly, the empowerment evaluation approach states that, for effectively materializing results of empowerment evaluation, organizations need to clearly define their program/organizational goals. Consistent with that, the present study found that clear organizational goals and EC concurrently exist in an organization. Organizations with low EC exhibited goal ambiguity. Contrary to the assumption of collaborative approaches, the present study found that ECB is possible without staff empowerment. Collaborative approaches assume that inclusion and empowerment of employees leads to promotion of an evaluation culture and organizational learning that ultimately increases the EC level. These approaches pay special attention to the opinion of staff in planning and undertaking evaluation. Advocates of participatory evaluation state that this approach is a way both of empowering stakeholders, including staff, at various levels and of promoting learning. As a result of this sense of inclusion and empowerment, staff actively participate in evaluation activities and learn from this process (Anderson et al., 2012; Rotondo, 2012; Wharton & Alexander, 2013; Fetterman, 2005; Fetterman & Wandersman, 2007; Wandersman et al., 2005); the EC level is increased.

The present study, however, observed a high level of EC in organizations where staff lacked authority and a low level of EC in organization where authority was transferred to frontline staff. Stated differently, an organization with centralized decision making could have higher EC level without taking into account the opinion of staff in planning and undertaking evaluation activities. These findings are in stark contradiction to the principal of inclusion of empowerment evaluation. This finding, however, should not be considered conclusive and
judgment should not be made about the effectiveness of collaborative approaches of ECB. This is because none of these 26 organizations included in this study appeared to have introduced empowerment evaluation as an approach to ECB. Had any of these organizations introduced empowerment evaluation perspective for ECB, the present study would not have been able to measure the impact of that based on cross-sectional data collected in this study. A full judgement of the effectiveness of an intervention and testing of its assumptions, such as empowerment evaluation as an ECB initiative, is only possible through an experimental research. Consequently, there is need for further investigation from NPOs and public organizations with a large sample to make such conclusive judgements. The best way to judge if a particular ECB approach works and to test its assumptions would be to undertake an experimental study.

The study partially supported Cousins and colleagues’ (2014b) understanding of the construct of EC. Consistent with Cousins and colleagues, the present study found that EC should be understood as the capacity of an organization to do and use evaluation. However, the present study found that EC is more than that. Overall, the study found that EC could be measured using the following indicators: the frequency of program performance data collection, the frequency of preparing useful evaluation reports, the frequency of staff meetings to discuss performance data collected, and the frequency of meetings with other organizations and external stakeholders to collection of performance data, leadership support, incentives to undertake evaluation, and availability of resources (including computer software, money, skills and time) for evaluation. The items that built on these indicators to measure EC had acceptable internal reliability, an indication of measurement validity.

The EC measures typically used in the EC literature are very weak and rarely have been validated. The present study is an important contribution to the EC literature in that a new
measure of EC was developed and validated using three different sources of information. Although various studies have used a variety of measures for EC, none of these has been validated. Future studies can use the EC measure validated in the present study to test in different organizational and cultural contexts. This will be a step forward to building one substantial measure of EC. This will be important from the standpoint of both theory and practice. Future studies could use validated measure of EC to further investigate its relationship with other organizational process and ultimately strengthening the body of EC knowledge. From the standpoint of practice, the EC measure could be used to diagnose and assess EC level in an organization. Assessment of EC probably should be first step to building EC in an organization. Once the level of EC is known in an organization, practitioners and organizational development experts could divert their attention to factors associated with high EC.

The present study is unique in that it used organizational learning perspective to understand performance management and bring these two streams of literature close through empirically investigation. In the performance management literature there are a couple of studies (e.g., Dooren, 2011; Moynihan, 2005; Moynihan & Lavertu, 2012) which understood performance management in terms of organizational learning. In all these studies, organizational learning was defined and understood very narrowly; most of such studies used single-item or two-item measures for organizational learning. Performance management studies have very rarely drawn on organizational learning literature. Kroll (2015) also found relationship between some measures of organizational learning and performance management. However, he did not understood these association in terms of organizational learning. Drawing on organizational literature, the present study used validated measures and established its relationship with performance management. This study found that organizational learning and performance
management go hand in hand: there is a positive association between these two. NPOs and public organizations attempting to implement performance measurement, should pay attention toward staff training, introduce a culture of experimentation, and leadership should pay attention toward defining organizational goals clearly. With the introduction of these three practice, NPOs and public organizations will learn from performance management and use performance information effectively.

The present study provides partial support for the reform movement, known as NPM, advocating for performance management in the public sector globally. At the heart of the NPM reform agenda is decentralization (empowerment) and clearly defining organizational goals. It is assumed that performance management reforms will be successful with decentralization and clear goals. One of these assumptions was supported in this empirical study. That is, performance management and clear organizational goals were found to be positively associated. Public management reforms in recent years have paid greater attention toward goal clarity and strategic planning. For example, the U.S. Federal government has mandated government agencies to publish strategic plans and clearly define their goals through legislative measures such as the GPRA of 1993 and the Modernization Act of 2012 (Rainey & Jung, 2015). However, no association was found between performance management and empowerment. This may be because empowerment and decentralization are not necessary for the success of performance management. Stated differently, a centralized organization could have an effective performance management system. In addition, these findings may be due to definitional and measurement differences. Although decentralization is a highly valued practice in the NPM literature, it is not clear what precisely is meant by decentralization or how to measure it. The present study took empowerment (decentralization) measure from the organizational learning literature which
understands empowerment as creating an egalitarian culture and encouraging organizational members to own their work, make decisions about their job, and be responsible for the results. However, it is possible that NPM literature understands decentralization in terms of transfer of authority from central government to local government. Nonetheless, finding that out was beyond the scope of this study.

Since this study was part of the Evaluation Capacity Assessment Project undertaken in collaboration with the United Way (UW), it is pertinent to discuss the implications of the findings of the study for the UW. The ultimate goal of the UW in establishing performance management systems and building the EC of NPOs is to make them efficient, effective, and accountable. That is, the UW would want to know if the money they give to NPOs is spent appropriately and desired goals are achieved. The findings of the present study suggest that the UW could build EC and introduce performance management successfully in these NPOs through organizational development efforts. That is, the UW should attempt to strengthen performance monitoring and evaluation in NPOs through staff development, introduction of a culture of experimentation, and helping these NPOs to clearly define their goals. None of these organizational development aspects require any monetary resources except staff development. The UW can provide technical training to NPOs’ staff to collect data on performance indicators and maintain performance management systems. It is advisable to maintain standardized and centralized performance measurement system. This will allow NPOs and the UW to compare performance across programs and organizations. Additionally, NPOs’ staff should also be trained to use these performance data for other purposes such as decision making and allocation of resources. The other two organizational development aspects, introduction of a culture of experimentation and goal clarity, do not require any monetary resources. All that is required for
the UW is to help these NPOs direct their attention toward introduction of a culture of experimentation where staff are allowed to experiment with new ways of doing their job and learn from this process of trial and error. In addition, the UW should help these NPOs define their goals clearly and communicate these goals at all levels of organization. Once these organization define their goals clearly, they will produce and use performance information more frequently (Moynihan & Pandey, 2005; Moynihan, Pandey, & Wright, 2012) and promote organizational learning.

**Strengths and Weaknesses of the Study**

One of the limitations of public management research is that employees’ self-reported data are used to undertake these studies. Usually, employees respond positively or quite often similarly to the questions due to various reasons including social desirability. Consequently, the validity of findings of the studies is questionable. The validity of the findings becomes more questionable in studies where data on both independent and dependent measures are gathered from the similar source, which most of the cases are self-reported data (Kelman, 2015). This so called ‘common-method bias’ could be eliminated by collecting information on either independent or dependent measure from different or multiple sources of information (Jakobsen & Jensen, 2015). Unlike many of public management and organizational studies, the present study avoided common-method bias by collecting data on independent and dependent measures from different sources of information and by using multiple criteria to measure the dependent measure (i.e., EC) of the study. For the present study, the data on independent measures were collected using questionnaire. Whereas the dependent measure, EC, was measured, as discussed in study method section, using three sources of information at various time points: the United Way of the Greater Capital Region (UWGCR) staff assessments, UWGCR volunteer ratings, and
employee perceptions concerning the EC of their organizations sought in the questionnaire. Two of these sources of information come from external observers. This is especially important because the method with UWGCR staff was a simple checklist as external observers of the NPOs, the UWGCR volunteers rated EC based on their assessment of submitted text by the organizations, and the employees as internal observers answered both frequency of event questions and also accuracy of statements about the EC level in their organizations. Additionally, the three sources of information had high convergent validity and internal reliability. This made the study avoid common-method bias, and consequently, the findings of the study valid.

Another strength of this study is higher participation and response rate. Organizations participation rate was over 0.80 and survey participants response rate was over 0.90. One of the weaknesses of studies using survey data is lower response rate, which seriously hinders the ability to generalize the findings of the study in population. This problem has been more serious in the EC literature. For instance, Taylor-Ritzler and others (2013) is one of very few studies using surveys to seek response of employees regarding EC levels in their organizations, had a response rate as low as 20%. Such low response rate calls the validity of findings in question. Conversely, the present study had very high response rate and findings are more valid and generalizable in the population.

Although the role of leadership was not directly investigated in this study, some of the measures of the study included items on leadership. Leadership items loaded heavily on the respective factors. Based on this, one might infer that leadership support and willingness might play important role in ECB. However, this premise should further be investigated in future studies. Like all the measures used in this study, the EC literature does not have valid leadership
measure. The future studies should draw on organizational leadership studies for valid measure of leadership and investigate its relationship with EC.

One of the limitations of this study is that only simple association between EC and five dimensions of OLC was tested, which in no way could be considered as causal link between these two. EC was found to be positively associated with three measures of OLC, staff development, experimentation and clear organizational goals. It is not clear whether staff development, experimentation and clear organizational goals are antecedents or consequences of EC. Stated differently, this study does not inform if EC is cause or effect of OLC. In future studies one may want to explore if these dimensions of OLC are antecedents or consequences of EC using structural equation causal models. This will allow practitioners to direct their attention toward processes which cause high EC.

In the present study the data were collected only from 26 NPOs, 13 with low EC and 13 with high EC. This was the minimum sample required to see difference between these two types of organizations. Low sample size could be one the limitation of this study. In future studies, to see difference between organizations with low and high EC, high sample size should be used.

One of the limitations of the EC literature is absence of use to experimental research designs to evaluate effectiveness of ECB initiatives. A number of ECB frameworks have been reported in the EC literature, such as capacity to do and use evaluation, empowerment evaluation, and participatory evaluation. The effectiveness of some of these frameworks has been tested using reflective case studies. However, experimental research design rarely has been used. In future ECB initiatives, the studies should use experimental methods to see if a particular approach to ECB works or not. This will not only answer important theoretical questions in the
EC literature, it will also let practitioners know what works and what does not works in the ECB process.

**Conclusion**

The present study answered an important research question regarding what uniquely characterizes organizations with high EC (i.e., success of performance management), and how these two types of organizations are different from each other, with respect to OLC, using data from 26 NPOs—13 with high EC and 13 with low EC. The OLC was measured using five dimension: staff empowerment, staff development, external alignment, experimentation, and clear organizational goals. Three out of five measures (staff development, experimentation, and clear organizational goals) were found significant predictors of the EC level (success of performance management) in NPOs. The present study makes important contribution to the EC and performance management literatures in a number of ways. This study found that ECB should be considered to be an organizational-level learning process, instead of a set of training activities directed at individual staff members generally understood in the EC literature. This study also found that organizational learning perspective could explain success of performance management in an organization. This study also brought three streams of literature together—the EC, performance management, and the organizational learning literature. This is the first study of its kind in both the EC and performance management literatures which used OLC measures drawing on the organizational learning literature to empirically test their relationship. This also is the only study in EC literature which attempted to differentiate between organizations with high and low EC and investigated how these two types of organizations are different. This study is unique because it validated the EC measure which could be used for further research and to measure the EC level in organizations. The study is also important for avoiding common-method
bias by using multiple method and various sources of information to measure dependent measure. The implications of this study are that ECB and performance management initiatives should focus on staff development, introduce a culture of experimentation, and define organizational goals clearly.
REFERENCES


## Appendix 1

### Table 1

**Evaluation Capacity Assessment Checklist Tool Used by Staff Members of the United Way**

<table>
<thead>
<tr>
<th>ORGANIZATION</th>
<th>Collects data to document program outputs and outcomes</th>
<th>Prepares useful reports on program performance and effectiveness</th>
<th>One of the best agencies for its program monitoring and evaluation</th>
<th>Does <strong>not</strong> seem to be engaged in very much program improvement</th>
<th>I am not familiar with this organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization 1</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Organization 2</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Organization 3</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Organization 4</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Organization n</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Sr#</td>
<td>Measure</td>
<td>Items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----------------------</td>
<td>----------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><strong>Staff Empowerment</strong></td>
<td>1. This organization recognizes people for taking initiative.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. This organization gives people control over the resources they</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>need to accomplish their work.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. This organization supports employees who take calculated risks.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. This organization gives us choices in our work assignment.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td><strong>Staff Development</strong></td>
<td>1. We help each other learn.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. People are rewarded for learning.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. People are given time to support learning.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><strong>External Alignment</strong></td>
<td>1. This organization encourages people to think from a regional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>perspective.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. This organization encourages people to get answers from across</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>the organization when solving problems.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. This organization works together with the outside community to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>meet mutual needs.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td><strong>Experimentation</strong></td>
<td>1. This program experiments with new ways of working.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. This organization experiments with new service or program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>offerings.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. This organization has formal process for conducting and</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>evaluating new ideas or experiments.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. This program employs pilot projects or simulations when trying</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>out new ideas.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td><strong>Clear organizational goals</strong></td>
<td>1. This organization’s mission is clear to almost everyone who works here.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. It is easy to explain the goals of this organization to outsiders.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. This organization has explicitly stated goals.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. This organization has objectives that are specific and well-defined.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3
*Items Included in Questionnaire to Measure Evaluation Capacity*

<table>
<thead>
<tr>
<th>Sr#</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Is information about accomplishments of your organization’s projects and programs ever collected?</td>
</tr>
<tr>
<td>2</td>
<td>Are program reports prepared that include some of the performance information collected?</td>
</tr>
<tr>
<td>3</td>
<td>Are program staff meetings convened to discuss some of performance information collected?</td>
</tr>
<tr>
<td></td>
<td>Are meeting convened with staff from other programs or organizations, external stakeholders, and / or program constituents during which the accomplishment of your organization’s projects and programs are discussed?</td>
</tr>
<tr>
<td>5</td>
<td>Our leaders are committed to the evaluation of our projects and programs.</td>
</tr>
<tr>
<td>6</td>
<td>Financial resources are available to implement program evaluation processes.</td>
</tr>
<tr>
<td>7</td>
<td>There is a plan and a timeline for completing regular and routine evaluations.</td>
</tr>
<tr>
<td>8</td>
<td>Stakeholders / funders expect our organization to report on program outcomes.</td>
</tr>
<tr>
<td>9</td>
<td>We have sufficient technical skills to support evaluations and use the findings.</td>
</tr>
<tr>
<td>10</td>
<td>There are positive incentives and rewards for planning and managing evaluations.</td>
</tr>
<tr>
<td>11</td>
<td>Meetings are convened for us to plan and prepare for anticipated evaluations.</td>
</tr>
<tr>
<td>12</td>
<td>Computer hardware and software are available to well manage evaluation data.</td>
</tr>
<tr>
<td>13</td>
<td>Our program staff have a genuine interest in supporting evaluation processes.</td>
</tr>
<tr>
<td>14</td>
<td>We participate in meetings to review and discuss the findings of past evaluations.</td>
</tr>
</tbody>
</table>
### Appendix 2

#### Table 1

*Rotated Factor Loadings Pattern Matrix (with Parsimax Rotation)*

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor1</th>
<th>Factor2</th>
<th>Factor3</th>
<th>Factor4</th>
</tr>
</thead>
<tbody>
<tr>
<td>COG1 This organization’s mission is clear to almost everyone who works here.</td>
<td>0.34</td>
<td><strong>0.68</strong></td>
<td>0.10</td>
<td>-0.08</td>
</tr>
<tr>
<td>COG2 It is easy to explain the goals of this organization to outsiders.</td>
<td>0.09</td>
<td><strong>0.71</strong></td>
<td>-0.06</td>
<td><strong>0.36</strong></td>
</tr>
<tr>
<td>COG3 This organization has explicitly stated goals.</td>
<td>0.02</td>
<td><strong>0.69</strong></td>
<td>0.43</td>
<td>0.04</td>
</tr>
<tr>
<td>COG4 This organization seems to be without central purpose or apparent direction.</td>
<td>0.13</td>
<td><strong>0.70</strong></td>
<td>0.26</td>
<td>0.06</td>
</tr>
<tr>
<td>SD1 People are rewarded for learning.</td>
<td>0.45</td>
<td>0.20</td>
<td><strong>0.58</strong></td>
<td>0.20</td>
</tr>
<tr>
<td>SD2 Program staff are given time to support learning.</td>
<td>0.26</td>
<td>0.06</td>
<td><strong>0.75</strong></td>
<td>0.36</td>
</tr>
<tr>
<td>SD3 Our leaders support requests for learning opportunities and training.</td>
<td>0.17</td>
<td>0.28</td>
<td><strong>0.75</strong></td>
<td>0.26</td>
</tr>
<tr>
<td>SE1 This organization recognizes people for taking initiative.</td>
<td>0.76</td>
<td>0.10</td>
<td>0.23</td>
<td>0.14</td>
</tr>
<tr>
<td>SE2 This organization gives people control over the resources they need to accomplish their work.</td>
<td>0.66</td>
<td>0.25</td>
<td>0.28</td>
<td>0.24</td>
</tr>
<tr>
<td>SE3 Teams/groups have the freedom to adapt their goals as needed.</td>
<td>0.75</td>
<td>0.23</td>
<td>0.13</td>
<td><strong>0.33</strong></td>
</tr>
<tr>
<td>SE4 This organization gives us choices in our work assignments.</td>
<td>0.56</td>
<td>0.07</td>
<td>0.44</td>
<td>0.00</td>
</tr>
<tr>
<td>EXP1 This organization has a formal process for conducting and evaluating new ideas or experiments.</td>
<td>0.18</td>
<td>0.16</td>
<td>0.09</td>
<td><strong>0.81</strong></td>
</tr>
<tr>
<td>EXP2 This program employs pilot projects or simulations when trying out new ideas.</td>
<td>0.04</td>
<td>-0.08</td>
<td>0.25</td>
<td><strong>0.78</strong></td>
</tr>
</tbody>
</table>
### Table 2
*Rotated Factor Loadings Pattern Matrix (with Parsimax Rotation)*

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor1</th>
<th>Factor2</th>
<th>Factor3</th>
<th>Factor4</th>
</tr>
</thead>
<tbody>
<tr>
<td>EA1 This organization works together with the outside community to meet mutual needs.</td>
<td>0.25</td>
<td>-0.23</td>
<td><strong>0.69</strong></td>
<td>0.09</td>
</tr>
<tr>
<td>EA2 This organization encourages us to think from a regional perspective.</td>
<td>-0.08</td>
<td>0.01</td>
<td><strong>0.81</strong></td>
<td>0.08</td>
</tr>
<tr>
<td>EA3 This organization builds alignment of visions across different levels and work groups.</td>
<td><strong>0.49</strong></td>
<td>-0.16</td>
<td><strong>0.59</strong></td>
<td>0.29</td>
</tr>
<tr>
<td>EA4 This organization actively engages with interest groups that share our mission and goals.</td>
<td>0.34</td>
<td>-0.09</td>
<td><strong>0.72</strong></td>
<td>0.14</td>
</tr>
<tr>
<td>JFB1 Despite the workload, program staff find the time to review how the work is going.</td>
<td><strong>0.56</strong></td>
<td>-0.42</td>
<td>0.27</td>
<td>0.31</td>
</tr>
<tr>
<td>JFB2 Our leaders mentor and coach those they lead.</td>
<td><strong>0.74</strong></td>
<td>-0.14</td>
<td><strong>0.32</strong></td>
<td>0.22</td>
</tr>
<tr>
<td>JFB3 I receive useful evaluations of my strengths and weaknesses at work.</td>
<td><strong>0.80</strong></td>
<td>-0.18</td>
<td>0.08</td>
<td>0.13</td>
</tr>
<tr>
<td>JFB4 Our leaders provide time, resources, and venues for identifying organizational problems and challenges.</td>
<td><strong>0.63</strong></td>
<td>-0.24</td>
<td><strong>0.39</strong></td>
<td>0.31</td>
</tr>
<tr>
<td>DEM1 In this program, we are too busy to invest time in improvement.</td>
<td>-0.33</td>
<td><strong>0.73</strong></td>
<td>-0.15</td>
<td>0.04</td>
</tr>
<tr>
<td>DEM2 Schedule pressure here gets in the way of doing a good job.</td>
<td>-0.15</td>
<td><strong>0.79</strong></td>
<td>-0.07</td>
<td>-0.31</td>
</tr>
<tr>
<td>DEM3 There simply is no time for reflection in this organization.</td>
<td>-0.39</td>
<td><strong>0.68</strong></td>
<td>-0.12</td>
<td>-0.13</td>
</tr>
<tr>
<td>DEM4 Employees in this organization are overly stressed.</td>
<td>0.04</td>
<td><strong>0.78</strong></td>
<td>-0.10</td>
<td>-0.20</td>
</tr>
<tr>
<td>JCL1 My responsibilities at work are very clear and specific.</td>
<td>0.36</td>
<td>-0.24</td>
<td>0.08</td>
<td><strong>0.71</strong></td>
</tr>
<tr>
<td>JCL2 I fully understand which of my job duties are more important.</td>
<td>0.15</td>
<td>0.00</td>
<td>0.15</td>
<td><strong>0.84</strong></td>
</tr>
<tr>
<td>JCL3 I know exactly what I am supposed to do on my job.</td>
<td>0.00</td>
<td>-0.18</td>
<td>0.10</td>
<td><strong>0.84</strong></td>
</tr>
</tbody>
</table>