Can Nonprofit Capacity Be Measured?

Mathieu R. Despard1

Abstract
Organizational capacity is the set of structures and functions a nonprofit organization needs to effectively serve the community. Although capacity is defined in the nonprofit literature, no standardized measures exist, making it difficult to accurately assess organizational capacity. Data from a survey of nonprofit human service organizations (N = 1,221) that participated in a capacity-building demonstration project are used to assess the fit of two conceptual models of capacity using confirmatory factor analysis. Results indicated that a model that measured capacity with more than 40 performance-related indicators did not fit the data well. However, a model using fewer (19) indicators of organizations’ self-assessed capacity-building progress fit the data well and was invariant by tenure. Implications for measuring nonprofit organizational capacity are discussed.

Keywords
nonprofit capacity, human services, capacity building, confirmatory factor analysis, measurement

Introduction
Nonprofit human service organizations (NPHSOs) are a critical part of the social safety net in the United States, providing a range of critical services such as child care, emergency housing, and child abuse prevention (Boris, de Leon, Roeger, & Nikolova, 2010). NPHSOs are under increasing pressure to demonstrate their impact in communities (Benjamin, 2013; Campbell, 2002; Ebrahim & Rangan, 2010) yet experience a host of capacity deficits in areas such as staff supervision and strategic

1University of Michigan, Ann Arbor, MI, USA

Corresponding Author:
Mathieu R. Despard, School of Social Work, University of Michigan, 1080 South University Avenue, Ann Arbor, MI 48109, USA.
Email: mdespard@umich.edu

Capacity deficits hinder NPHSO’s ability to achieve community impact. Program implementation is affected by factors such as training, supervision, and leadership (Aarons, Sommerfeld, & Walrath-Greene, 2009; Austin & Claassen, 2008; Bond, Drake, McHugo, Rapp, & Whitley, 2009; Fixsen, Naoom, Blase, Friedman, & Wallace, 2005). A lack of sufficient financial resources (Besel, Williams, & Klak, 2011; Calabrese, 2013; Salamon & Geller, 2007; Sontag-Padilla, Staplefoote, & Gonzalez Morganti, 2012; Weerawardena, McDonald, & Mort, 2010) makes it hard for NPHSOs to consistently and reliably implement programs. NPHSOs also struggle with program evaluation (Carman, 2007; Carman & Fredericks, 2008; Hoefer, 2000; Innovation Network, 2012; Leake et al., 2007; Pejsa, 2011), making it difficult to assess client outcomes.

Capacity building is a strategy to help NPHSOs improve their performance (De Vita & Fleming, 2001; Light, 2004). Standardized measures exist for leadership (Avolio, Bass, & Jung, 1999; Posner & Kouzes, 1993), organizational climate and culture (Aarons & Sommerfeld, 2012; N. R. Anderson & West, 1998; Cooke & Rousseau, 1988; Glisson et al., 2008), and readiness for change (Lehman, Greener, & Simpson, 2002). However, no standardized measures exist for NPHSO capacity, which hinders efforts to assess the effectiveness of capacity-building interventions (Wing, 2004).

The purpose of this article was to assess the validity of a set of capacity measures using a large sample (N = 1,221) of small, community-based NPHSOs engaged in a capacity-building demonstration. Validated capacity measures can be used to assess (a) areas where organizational performance might be strengthened, (b) capacity-building intervention outcomes, and (c) factors such as staff performance evaluation and administrative supports that affect program implementation (Fixsen, Blase, Naoom, & Wallace, 2009).

**How Organizational Capacity Is Defined and Measured**

Organizational capacity refers to the resources, skills, and functions a nonprofit organization (NPO) needs to fulfill its mission across multiple domains (Connolly & York, 2003; Doherty, Misener, & Cuskelly, 2013; Eisinger, 2002; Light, 2004; McKinsey & Company, 2001; Millesen, Carman, & Bies, 2010). Strategic management theory posits that organizations make purposive decisions to strengthen their responses to changing environments (Barney & Hesterly, 2012; Hill & Jones, 2008; Oster, 1995), which helps explain the capacity-building motive. Resource dependence theory views organizations as striving for growth and sustainability (Froelich, 1999; Hillman, Withers, & Collins, 2009; Pfeffer & Salancik, 1978), which helps explain the importance of strengthening grant writing, contracting, and fundraising efforts. NPOs’ roles as principals accountable to public and private funders as agents (Eisenhardt, 1989) helps explain the importance of evaluation and financial management capacity.

Conceptual models and accompanying assessment tools have been developed to guide nonprofit capacity-building efforts. McKinsey & Company (2001) defined capacity as comprised of 58 indicators in the following seven domains: (a) aspirations
(e.g., mission, vision), (b) strategy (e.g., program development), (c) organizational skills (e.g., budgeting, evaluation, fundraising), (d) human resources (e.g., board composition, senior management team experience), (e) systems and infrastructure (e.g., technology, facilities), (f) organizational structure (e.g., governance, job descriptions), and (g) culture. Each indicator has an ordinal response scale: clear need for increased capacity and basic, moderate, and high levels of capacity already in place.

The Marguerite Casey Foundation (2012) modified the McKinsey tool, defining four capacity domains—originally identified and described by Connolly and York (2003):

1. **Leadership capacity**: ability of organizational leaders to inspire, prioritize, make decisions, provide direction, and innovate;
2. **Adaptive capacity**: ability of an NPO to monitor, assess, and respond to internal and external changes;
3. **Management capacity**: the ability of an NPO to ensure the effective and efficient use of organizational resources; and
4. **Operational/technical capacity**: the ability of an NPO to implement key organizational and programmatic functions (Marguerite Casey Foundation, 2012).

Each indicator is assessed at one of four levels. For example, senior management team is an indicator in the management domain, the highest level that is defined in part as “extensive and varied experience in nonprofit and for-profit management; team drawn from extraordinarily diverse backgrounds and experiences” (Marguerite Casey Foundation, 2012).

The TCC Group’s (2010) Core Capacity Assessment Tool (CCAT) is a 146-item survey that assesses the same four domains as the Marguerite Casey Foundation (MCF) tool. Each capacity domain has a set of subcapacities, such as organizational learning under adaptive capacity. Scores on a 300-point scale are generated from the CCAT for each capacity and subcapacity to indicate strong, satisfactory, and challenging levels of performance.

Based on the Compassion Capital Fund (CCF) Demonstration Program, which focused on small NPHSOs, Minzner et al. (2014) defined capacity as comprised of organizational development (e.g., governance, human resources, finances), program development (e.g., service delivery, evaluation), revenue development, leadership development, and community engagement. A survey instrument to measure changes in capacity among NPHSOs has been used in several studies of the CCF demonstration (Abt Associates & Branch Associates, 2009; Francis, Minzner, Fisman, & Fox, 2011; Minzner et al., 2010; Minzner et al., 2014). Survey items in each domain are measured using binary (yes/no) and ordinal (self-assessed progress) response choices.

The conceptual models and assessment tools outlined above offer NPOs ways to self-assess capacity. However, to the best of the author’s knowledge, none of these tools have been validated. Minzner et al. (2014) acknowledged the lack of a standardized measure as a limitation of the CCF demonstration impact study. Consequently, how capacity is defined and measured is an open question. The lack of validated
capacity measures may be due to the paucity of studies of NPHSO capacity, the difficulty of conceptualizing capacity (Wing, 2004), and the challenge of defining universal performance standards (Herman & Renz, 2008).

Why Measuring Capacity Is Important

Several studies have been conducted to assess outcomes of capacity-building interventions among NPHSOs. Most of these studies found evidence of self-reported gains in various capacities such as board governance and program evaluation (Andrews, Motes, Floyd, Flerx, & Fede, 2005; Bryan & Brown, 2015; Francis et al., 2011; Leake et al., 2007; Minzner et al., 2010; Minzner et al., 2014; Patrizi, Gross, & Freedman, 2006; Sobeck, 2008). However, other studies have found only modest impacts of capacity building on service provision (Markovitz, Magged, Florez, & Klein, 2008; Popescu & Dewan, 2009). It is unclear whether and how capacity building affects the effectiveness of NPHSOs in achieving positive client outcomes.

Although the relationship between NPHSO capacity and service quality and client outcomes has not been well studied, evidence from the field of implementation science suggests such a link. Systematic reviews of hundreds of studies in human services have identified organizational factors (e.g., funding, leadership quality, management support, supervision, training, and organizational culture) that affect program and service quality and client outcomes, (Fixsen et al., 2005; Gearing et al., 2011; Panzano et al., 2005). Although not explicitly labeled as such, these organizational factors are indicators of various capacity domains identified in the nonprofit literature. Thus, because capacities may affect service quality and client outcomes, it is important to accurately measure capacity to better understand these relationships.

The purpose of this study was to assess the validity of a NPHSO capacity survey—to answer Wing’s (2004) question, “How can an abstract concept be concretely measured?” (p. 154). Despite the availability of assessment tools to help NPHSOs make sense of their capacity, none have been validated—at least as reported in the academic literature. Establishing a validated instrument can promote more accurate assessments of NPHSO organizational performance, capacity-building outcomes, and the link between organizational factors and program implementation effectiveness and client outcomes.

Method

Sample

Data used for this study came from a baseline survey administered on a rolling basis from January through November 2006 to a sample of 1,221 NPHSOs. These NPHSOs were sought and determined eligible by 44 intermediary organizations (e.g., local United Ways, universities, large NPOs) to receive capacity-building assistance through the CCF Demonstration Program sponsored by the federal Administration for Children and Families (ACF). The CCF demonstration targeted grassroots
faith- and community-based organizations serving lower-income individuals and families with a wide array of programs including elderly/disabled services, job training, hunger, homelessness, pregnancy prevention, and family services. Participating NPHSOs were required to complete baseline surveys as a condition of receiving one-on-one technical assistance, group training, and financial assistance from intermediaries (Abt Associates & Branch Associates, 2009; Minzner et al., 2014). Thus, the data used for this study were from a nonprobabilistic sample of NPHSOs, characteristics of which are described in Table 2.

Measures

Measures for this study came from the baseline survey of the CCF demonstration study (Abt Associates & Branch Associates, 2009) and included 215 items, 54 of which related to organizational capacity in five domains: organizational, program, revenue, leadership development, and community engagement. Data were accessed through the Inter-university Consortium for Political and Social Research of the University of Michigan. The survey included two types of measures: (a) concrete indicators of capacity and (b) level-of-focus (LOF) indicators. Concrete indicators included items promoting binary (i.e., yes/no) responses to questions such as “Does your organization have a strategic plan?” and “In the past 12 months, has your organization conducted or participated in an assessment of organizational strengths/needs?” LOF is a term study authors (Abt Associates & Branch Associates, 2009) used to label survey items measuring NPHSOs’ self-assessed progress in achieving a capacity, such as “expanding services to include a new group of service recipients or geographic area,” based on a 5-point Likert-type scale (p. 14). Response choices were “not an area of focus at this time,” “know we should work on this but lack time or resources,” “have developed plans or ideas to work on this, but haven’t implemented,” “have implemented steps to address focus area,” or “not a focus because satisfied with achievement in this area.” Items were excluded from the analysis if they were not capacity indicators based on prior research (e.g., Connolly & York, 2003; Eisinger, 2002; Light, 2004; Millesen et al., 2010).

Analysis

Confirmatory factor analysis (CFA) is a technique to assess how well a hypothesized latent factor structure fits observed data by assessing factor loadings, variance, and covariance (Bowen & Guo, 2012; Schreiber, Nora, Stage, Barlow, & King, 2006). As a theory-driven technique (Schreiber et al., 2006), CFA was used instead of exploratory factor analysis (EFA) because latent variables—types of NPHSO capacity—have been previously defined (e.g., Connolly & York, 2003; Millesen et al., 2010; Minzner et al., 2010; Minzner et al., 2014).

CFA was used to test the fit of the hypothesized factor structure of the CCF survey instrument as a measurement model of NPHSO capacity. The CCF survey instrument is a five-factor (organizational development, program development, revenue development, leadership development, and community engagement) model identified from
Table 1. Model Specification.

<table>
<thead>
<tr>
<th>Model</th>
<th>Latent variables</th>
<th>Observed indicators</th>
<th>STV ratio</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCF (full)</td>
<td>Organizational development</td>
<td>27</td>
<td>11:1</td>
<td>1,367</td>
</tr>
<tr>
<td></td>
<td>Program development</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Resource development</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leadership development</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Community engagement</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCF (LOF only)</td>
<td>Program development</td>
<td>6</td>
<td>32:1</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>Management capacity</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Board development</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Resource development</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. STV = subject to variable; CCF = Compassion Capital Fund; LOF = level of focus.

Prior CCF studies (Minzner et al., 2010; Minzner et al., 2014). Survey items were hypothesized to load on these five factors based on how CCF study results were reported. For example, the outcome for the survey item “Does your organization conduct formal measurement/assessment of the results and benefits of the services provided to individuals or families?” is reported by Minzner et al. (2010) and Minzner et al. (2014) in the program development capacity domain.

In addition, the CCF model was assessed only using the LOF survey items, representing NPHSOs’ self-assessed progress in building capacity. Like the full CCF model, the LOF-only model incorporates hypothesized domains from prior CCF studies. In measuring self-assessed capacity-building progress, the LOF-only model is well aligned with the life stages perspective of NPHSO capacity (Bess, 1998; Hasenfeld & Schmid, 1989; Light, 2004; Simon, 2001; Stevens, 2002), which views capacity as a developmental process.

Because factor analysis results should be cross-validated to assess factor stability (Bowen & Guo, 2012; de Vet, Adèr, Terwee, & Pouwer, 2005), two random samples without replacement were taken from the CCF sample: a calibration sample (n = 611) and a sample (n = 610) to validate the selected measurement model. There were no statistically significant differences between the calibration and validation samples for NPHSO size, tenure, type, and paid executive director.

Model specification steps recommended by Bowen and Guo (2012) were followed. Observed indicators (survey items) with measurement error were specified as loading onto latent variables. To set scales, the path of each latent variable to the first observed indicator was fixed to one. In Table 1, latent variables, observed indicators, subject-to-variable (STV) ratios, and degrees of freedom were specified for both models. STV ratios represent the number of observations per variable used in factor analyses (de Vet et al., 2005); ratios of at least 3:1 are recommended to produce reliable results (P. Kline, 2000). Both models exceed recommended STV ratios and are overidentified (Hoyle, 1995; R. B. Kline, 2011).
Models were estimated using mean- and variance-adjusted weighted least squares (WLSMV) in Mplus version 7.3 (L. K. Muthén & Muthén, 1998-2012). WLSMV produces a polychoric correlation matrix, which is recommended for nominal and ordinal data (Brown, 2006; Flora & Curran, 2004; Garrido, Abad, & Ponsoda, 2013; Holgado-Tello, Chacón-Moscoso, Barbero-García, & Vila-Abad, 2008; B. Muthén, du Toit, & Spisic, 1997; Norris & Lecavalier, 2010). Root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker–Lewis index (TLI) were used as fit indices, which are recommended for use with the WLSMV estimator (Bowen & Guo, 2012; Schreiber et al., 2006). RMSEA is an absolute fit index, which reflects the extent to which sample data are reproduced by the specified model, whereas CFI and TLI are incremental fit indices, which compare fit improvement of the specified model with a nested baseline model (Hu & Bentler, 1998).

The following recommended cutoff values were used: RMSEA < .06 (90% upper bound confidence interval [CI] < .08), CFI > .95, and TLI > .95 (Bowen & Guo, 2012; Hu & Bentler, 1999; Schreiber et al., 2006; West, Taylor, & Wu, 2012). The chi-square goodness-of-fit indicator was excluded because it is typically statistically significant with larger samples ($n \geq 400$; Dimitrov, 2010; Hoe, 2008; Hoyle, 1995; Kenny, 2014).

### Table 2. Study Sample Description.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>% or M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tenure—years of providing services ($n = 1,175$)</strong></td>
<td></td>
</tr>
<tr>
<td>Younger (less than 6 years)</td>
<td>43%</td>
</tr>
<tr>
<td>Older (6 or more years)</td>
<td>57%</td>
</tr>
<tr>
<td><strong>Type ($n = 1,205$)</strong></td>
<td></td>
</tr>
<tr>
<td>Faith-based organization</td>
<td>52%</td>
</tr>
<tr>
<td>Community-based organization</td>
<td>48%</td>
</tr>
<tr>
<td><strong>Size ($n = 773$)</strong></td>
<td></td>
</tr>
<tr>
<td>Smaller (less than US$100,000 in revenue)</td>
<td>58%</td>
</tr>
<tr>
<td>Larger (US$100,000 or more in revenue)</td>
<td>42%</td>
</tr>
<tr>
<td><strong>Paid executive director ($n = 827$)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>62%</td>
</tr>
<tr>
<td>No</td>
<td>38%</td>
</tr>
<tr>
<td><strong>Communities served ($n = 1,214$)</strong></td>
<td></td>
</tr>
<tr>
<td>Urban only</td>
<td>55%</td>
</tr>
<tr>
<td>Large town only</td>
<td>12%</td>
</tr>
<tr>
<td>Rural area only</td>
<td>11%</td>
</tr>
<tr>
<td>Suburban only</td>
<td>4%</td>
</tr>
<tr>
<td>Two or more types</td>
<td>18%</td>
</tr>
</tbody>
</table>

Note. CCF = Compassion Capital Fund.

*Covariates defined by prior CCF research (Minzner et al., 2010; Minzner, Klerman, Markovitz, & Fink, 2014).

*Retrospective observations at follow-up. Lower $n$ due to attrition.*
To improve model fit, two strategies were used. First, modification indices were reviewed to identify sources for improved fit by correlating the measurement errors of within-factor observed indicators. This is an acceptable procedure if conceptually substantiated, such as when observed indicators share the same question stems or constructs (J. C. Anderson & Gerbing, 1988; Bowen & Guo, 2012; Schreiber et al., 2006) and correspond to the same latent variable (Hooper, Coughlan, & Mullen, 2008). For example, in the full CCF model, the measurement errors of the following two items were correlated to improve fit: “[O]rganization keeps records on the number of individuals or families enrolled in/served through programs” and “organization keeps records on individual service recipients’ outcomes.” Second, observed indicators were dropped if they met two or more of the following criteria: low factor loading (<.32), cross-loading (i.e., indicators that load on more than one latent variable according to modification indices), low $R^2$ (<.20), and/or source of multiple high-residual correlations (Bowen & Guo, 2012).

A final analytical step was to assess group invariance by tenure—the length of time the NPHSO had been in operation—using the DIFFTEST command in Mplus 7.3 with Satorra–Bentler scaled chi-square. This step was taken to assess whether the best fitting model showed consistent fit for both younger (i.e., less than 6 years in operation) and older (i.e., 6 or more years in operation) NPHSOs. A model with unconstrained factor loadings was compared with a nested model with factors loadings constrained to be equal by tenure. A statistically significant result of the DIFFTEST would suggest retaining the unconstrained model, which means that there was systematic variance by tenure. In contrast, a nonsignificant DIFFTEST result would indicate that the model was invariant (R. B. Kline, 2011) by tenure, that is, applicable to both younger and older NPHSOs.

Missing Values
Up to 20% of missing data is considered acceptable for CFA (Peng, Harwell, Liou, & Ehman, 2007). Mean covariance coverage across 77 analysis variables was 88%, and the average rate of missing values was 3.8%. All but three items had a missing value rate of less than 20%. These three items—annual performance reviews for paid staff, written job descriptions for paid staff, and obtained funding from a new source in the last 12 months—had a missing values rate of 31%, 30%, and 29%. To examine missing data patterns, dummy variables were created for missing values on these three items and regressed on covariates. Having no paid staff was a significant predictor ($p < .001$) of all three items. Similarly, missing data for several other items with non-negligible but less than 20%, missing values were also predicted by other observed variables. All “don’t know,” “refused,” and “not applicable” responses were coded as missing values in the CCF outcome study (Campbell, n.d.). Thus, it is highly likely that most missing values represented “not applicable” responses as predicted by other variables, suggesting data were missing at random (MAR; Graham, 2009).

Missing data were handled using full information maximum likelihood (FIML) in Mplus version 7.3 (L. K. Muthén & Muthén, 1998-2012), so all available information was used. FIML as a means of handling missing data in structural equation modeling has been found to produce unbiased and efficient estimates (Asparouhov & Muthén, 2010; Enders & Bandalos, 2001).
Sample Characteristics

Table 2 displays characteristics of the study sample. Reflecting the purpose of the CCF demonstration to reach small NPHSOs, most (57%) had annual expenditures of below US$100,000, compared with 30% of all public charities in the United States (McKeever, 2015). Type of NPHSO was roughly evenly divided between faith- and community-based organizations. Most NPHSOs had a paid executive director—though the average number of paid staff was very low—and most had been providing services for 10 years or less. The most common primary programmatic areas were services for at-risk children and youth (70%) and education (51%).

Model Fit

Initial model fit based on recommended cutoff values (Hu & Bentler, 1999) was good based on RMSEA but not for CFI and TLI (see Table 3), yet model respecifications improved fit for both the full CCF and LOF-only models (see Table 4).

The full CCF model still failed to achieve fit thresholds, but the LOF-only model achieved good fit. This finding was corroborated by running the model with a separate validation sample, which resulted in good fit. All 19 observed indicators had statistically significant loadings on the four latent variables (i.e., resource, board, and program development, and management capacity; see Appendix A). All four latent variables had good reliability (α = .72, .80, .72, and .74 for resource, board, program development, and management capacity, respectively). The final LOF-only model
included six sets of correlated measurement errors, three sets each within the board
and program development latent variables. All of these correlated errors were for
items that had a common or conceptually similar question stem (see Appendix B for a
complete list of LOF survey items).

**Group Invariance Test for LOF-Only Model**

The final analytical step was to determine whether the LOF-only model was invariant
measure based on tenure, that is, applicable to both younger (in operation for less than
6 years) and older (in operation for 6 or more years) NPHSOs. Testing invariance by
tenure was informed by the life stages perspective of NPOs (Light, 2004; TCC Group,
2010), that is, whether the LOF-only model represented a valid way to measure capac-
ity irrespective of an NPHSO’s life stage. The group invariance test using the
DIFFTEST procedure in Mplus by tenure was nonsignificant (p = .67), indicating the
LOF-only model had equally good fit for younger and older NPHSOs.

**Poor-Performing CCF Items**

Table 5 lists survey items that were dropped from the full CCF model because these
items failed to sufficiently load on a latent factor (≥ .32), had low $R^2$ values, and/or
cross-loaded on more than one latent factor. The nine poor-performing items reflect
multiple capacity domains, including four of five CCF survey items concerning tech-
ology, which were in the CCF organizational development domain.

**Discussion**

The goal of this study is to assess the validity of a model for measuring NPHSO capac-
ity using baseline survey data from a federal capacity-building demonstration project.
Although the full CCF model does not fit the data well, the LOF-only model fits the
data well for both younger and older NPHSOs. The 19-item LOF-only model could be

---

**Table 5. Poor-Performing Full CCF Survey Items.**

<table>
<thead>
<tr>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. NPO engages in partnerships with other organizations</td>
</tr>
<tr>
<td>2. NPO has a fundraising plan</td>
</tr>
<tr>
<td>3. Primary activity of the board is outreach</td>
</tr>
<tr>
<td>4. Primary activity of the board is reviewing executive director performance</td>
</tr>
<tr>
<td>5. NPO has access to the Internet</td>
</tr>
<tr>
<td>6. The number of functioning computers is sufficient</td>
</tr>
<tr>
<td>7. The Internet is used in support of organizational activities</td>
</tr>
<tr>
<td>8. The Internet is used to support the organization’s website</td>
</tr>
<tr>
<td>9. LOF: Organization incorporates new approaches to providing services</td>
</tr>
</tbody>
</table>

*Note. CCF = Compassion Capital Fund; NPO = nonprofit organization; LOF = level of focus.*
used as a scale to measure an NPHSO’s self-assessed progress in strengthening board, program, and resource development and management capacity. This scale could be useful to help an NPHSO, including its board of directors, monitor progress in implementing its strategic plan or gauge the results of having received technical assistance.

However, the LOF-only scale has limitations as a subjective measure of capacity-building progress. Unlike the full CCF model and other capacity assessment tools, the LOF-only scale does not include more objective measures of capacity, such as whether the organization has a strategic plan. Thus, the LOF-only scale can be regarded as a tool for organizational self-assessment regarding capacity building but not as a way for NPHSOs to benchmark performance against peer organizations and best practices.

Why might it be difficult to measure capacity as a multidomain construct? It is possible that the full CCF model did not fit the data well because there may be other unnamed capacity domains that more accurately reflect sets of organizational functions. Other conceptual models such as the TCC Group’s (2010) CCAT or the Marguerite Casey Foundation’s Organizational Capacity Assessment Tool might be validated as more accurate models of NPO capacity. Another possibility is that there are simply too many indicators in the full CCF model. As observed indicators in a model increase, RMSEA declines, yet CFI and TLI tend to indicate worse model fit (Breivek & Olsson, 2001, and Kenny & McCoach, 2003, as cited in Fan & Sivo, 2007), which is what occurred in this study. Thus, the LOF-only model may have fit the data better than the full CCF model in part simply because there were fewer indicators.

There are a host of measurement challenges that may also account for the poor fit observed for the full CCF model and illustrate how challenging it is to measure capacity. It may be too difficult to accurately measure complex organizational behaviors using binary response options. For example, an indicator of program development in the full CCF model is “Does your organization conduct formal measurement/assessment of the results and benefits of the services provided to individuals or families?” The real answer to this question may not be revealed through a yes-or-no response option. Perhaps an NPHSO conducts outcome measurement for a well-established program but not yet for a new program. In this case, a yes-or-no response would not accurately reflect this NPHSO’s capacity-related behavior.

There were other measurement limitations of the CCF survey to note. Nine items performed poorly and were dropped. Four of these items are related to technology. How NPHSOs use technology may be difficult to measure, particularly given how quickly information technology evolves. Other items may be poorly constructed. For example, another dropped item was “NPO engages in partnerships with other organizations.” “Partnerships” could mean many different things depending on the particular NPHSO. For a youth-serving NPHSO, partnerships could be regarded in relation to receiving referrals from schools, whereas for NPHSOs addressing hunger, partnerships could be regarded in relation to receiving in-kind donations from food retailers and wholesalers. Other dropped items used “outreach” and “new approaches” to service delivery, which may also mean very different things depending on the NPHSO’s mission or are simply too ambiguous. The LOF-only model may have better fit because it is easier for staff to self-assess progress than to attempt to accurately depict organizational performance.
Although the full CCF model did not fit the data well, it is important to note that in the context of CFA, this means the data did not fit the hypothesized relationships between individual survey items and capacity domains. Thus, it may not be accurate to use the full CCF model as a scale with subscales representing the five capacity domains, yet using individual survey items as performance indicators may be sufficient. This raises the question of whether it is helpful or necessary for NPHSOs to use measures that assess domain scores representing latent constructs or if assessing stand-alone indicators helps guide capacity-building efforts.

This study has limitations to note. The CCF demonstration and measures used to assess capacity-building outcomes were designed for small and young NPHSOs—a limited proportion of the 35% of all public charities classified as human services (McKeever, 2015). Thus, findings are not generalizable to large NPHSOs nor to NPOs in other fields, such as arts, culture, and humanities and health care. Also, survey responses were provided by a single representative of each NPHSO, usually the executive director. Respondents may not have offered accurate assessments of organizational capacity because of inconsistent awareness of different organizational functions. For example, an executive director who writes grant proposals and solicits donations may offer an accurate assessment of resource development capacity but not of program development capacity if a different staff member has primary managerial responsibility in this area. The respondent might have also not been with the NPHSO long enough to offer accurate observations of organizational capacity. The length of the survey (more than 200 items) also introduces the possibility of respondent fatigue, particularly considering the very busy work days of executive directors of smaller NPHSOs.

Research on NPHSO capacity could be improved in three ways. First, rather than depend solely on self-report survey responses from a single respondent, responses could be aggregated across multiple staff members, which is a practice with other instruments such as the Organizational Social Climate tool (Aarons & Sommerfeld, 2012). Second, capacity observations could be triangulated by adding observational checklists and/or rating forms based on researchers’ interviews with NPHSO staff members and on-site reviews of organizational records. Third, NPHSO leaders themselves could be more involved in developing capacity instruments using expert review, cognitive interviewing, and pretesting methods.

Given time and resource limitations, it may be unrealistic for most small NPHSOs to engage in the research improvements outlined above, unless this research is conducted as part of a capacity-building effort like the CCF, which offered NPHSOs capacity-building grants and technical assistance. Foundations and government agencies are already engaged in a range of capacity-building efforts with NPHSOs; the research improvements outlined above could strengthen the evaluation of these efforts.

Conclusion

NPHSOs are a growing and important part of the social safety net in the United States. Many NPHSOs are small and in need of strengthened capacity to reliably and
effectively implement programs to meet community needs. Although organizational capacity is difficult to measure accurately, this study finds a 19-item scale NPHSOs and their capacity-building partners can use to measure self-assessed progress in building capacity. Capacity-building efforts of foundations, government agencies, and other intermediaries can be enhanced with measures that have been validated to more accurately measure capacity outcomes. Using validated measures may give intermediaries and NPHSOs increased confidence in baseline assessments that indicate the need to improve performance in certain areas like governance. Use of validated measures also enhance confidence in the results of capacity-building efforts to help NPHSO leaders—amid so many other demands on time and resources—understand which efforts best enhance organizational performance.

Appendix A

A measurement model of self-assessed NPHSO capacity-building progress.

Note. All factor loadings are statistically significant at $p < .001$. NPHSO = nonprofit human service organization.
## Appendix B

### CCF Level of Focus Survey Items.

<table>
<thead>
<tr>
<th>Item abbreviation</th>
<th>Latent variable</th>
<th>Item description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Fundraising plan</td>
<td>Resource development</td>
<td>Developing a fund-development plan (including setting fundraising goals).</td>
</tr>
<tr>
<td>2. New sources government funding</td>
<td></td>
<td>Identifying and pursuing new sources of government funding.</td>
</tr>
<tr>
<td>3. New sources nongovernment funding</td>
<td></td>
<td>Identifying and pursuing new sources of in-kind donations.</td>
</tr>
<tr>
<td>4. New sources in-kind funding</td>
<td></td>
<td>Identifying and pursuing new sources of nongovernment funding.</td>
</tr>
<tr>
<td>5. Client data system</td>
<td>Program development</td>
<td>Developing a way to collect more information about our clients, including number and characteristics of clients as well as how they are helped.</td>
</tr>
<tr>
<td>6. Evaluation system</td>
<td></td>
<td>Strengthening the organization's ability to evaluate its overall effectiveness.</td>
</tr>
<tr>
<td>7. Increase number of clients</td>
<td></td>
<td>Increasing the number of clients served by the organization.</td>
</tr>
<tr>
<td>8. Service new clients or communities</td>
<td></td>
<td>Expanding services to include new group of service recipients or geographic area.</td>
</tr>
<tr>
<td>9. Increase scope of services</td>
<td></td>
<td>Increasing the number or scope of services offered to clients.</td>
</tr>
<tr>
<td>10. Incorporate new services approach</td>
<td></td>
<td>Incorporating a new approach to services to improve quality/effectiveness.</td>
</tr>
<tr>
<td>11. Budgeting process</td>
<td>Management capacity</td>
<td>Putting in place a budgeting process that ensures effective allocation of resources</td>
</tr>
<tr>
<td>12. Financial management system</td>
<td></td>
<td>Developing systems that will help manage the organization’s finances more effectively</td>
</tr>
<tr>
<td>13. Volunteer management</td>
<td></td>
<td>Recruiting, developing, and managing volunteers more effectively.</td>
</tr>
<tr>
<td>14. Staff leadership development</td>
<td></td>
<td>Creating a plan or locating resources to help our executive director and other staff improve their leadership abilities.</td>
</tr>
<tr>
<td>15. Staff professional development</td>
<td></td>
<td>Providing staff with professional development and training to enhance skills in service delivery or skills in administration and management.</td>
</tr>
<tr>
<td>16. Cross-section of community</td>
<td>Board development</td>
<td>Developing a board that represents a cross-section of our community.</td>
</tr>
<tr>
<td>17. Diverse expertise</td>
<td></td>
<td>Recruiting board members with diverse expertise.</td>
</tr>
<tr>
<td>18. Ties to different constituents</td>
<td></td>
<td>Developing a board with ties to different constituencies.</td>
</tr>
<tr>
<td>19. Better understand governance</td>
<td></td>
<td>Providing information to the board so that they can better understand their responsibilities and create plans for improving their performance.</td>
</tr>
</tbody>
</table>

Note. CCF = Compassion Capital Fund.
Acknowledgments
The author thanks anonymous reviewers and the following members of his dissertation committee for their feedback on earlier versions of this article: Gina Chowa, Matthew Howard, Thomas Kelley, Gary Nelson, and Kathleen Rounds.

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

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

References


Glisson, C., Landsverk, J., Schoenwald, S., Kelleher, K., Hoagwood, K., Mayberg, S., & Green, P. (2008). Assessing the organizational social context (OSC) of mental health services:


**Author Biography**

**Mathieu R. Despard** conducts research on the financial security of low-income households and the capacity and effectiveness of nonprofit organizations (NPOs) assisting these households. Previously, he worked in and with NPOs in management, governance, funding, and consulting roles, and coordinated the Nonprofit Leadership Certificate Program at The University of North Carolina (UNC) at Chapel Hill.