

2017

Importance of Auxiliary Theories in Research on University-Community Partnerships: The Example of Psychological Sense of Community

N. Andrew Peterson
Rutgers University, andrew.peterson@ssw.rutgers.edu

Paul W. Speer
Vanderbilt University, paul.w.speer@vanderbilt.edu

Christina Hamme Peterson
Rider University, cpeterson@rider.edu

Kristen Gilmore Powell
Rutgers University, kgilmore@ssw.rutgers.edu

Peter Treitler
Rutgers University, ptreitler@ssw.rutgers.edu

Yuqi Wang
Rutgers University, yuqi.wang@rutgers.edu

Follow this and additional works at: <http://scholarlyrepository.miami.edu/collaborations>



Part of the [Community Psychology Commons](#), [Counseling Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Educational Methods Commons](#), and the [Social Work Commons](#)

Recommended Citation

Peterson, N., Speer, P. W., Peterson, C. H., Powell, K. G., Treitler, P., & Wang, Y. (2017). Importance of Auxiliary Theories in Research on University-Community Partnerships: The Example of Psychological Sense of Community. *Collaborations: A Journal of Community-Based Research and Practice*, 1 (1). Retrieved from <http://scholarlyrepository.miami.edu/collaborations/vol1/iss1/5>

This article is brought to you for free and open access by Scholarly Repository. It has been accepted for inclusion in *Collaborations: A Journal of Community-Based Research and Practice* by an authorized administrator of Scholarly Repository. For more information, please contact repository.library@miami.edu.

Importance of Auxiliary Theories in Research on University-Community Partnerships: The Example of Psychological Sense of Community

Abstract

Psychological sense of community (PSOC) has long been recognized as a key element of successful collaborative initiatives, particularly university-community partnerships. A critical challenge involves the development of auxiliary theories that guide the specification of measurement models in studies of PSOC and other theoretical constructs. Auxiliary theories can be especially useful in clarifying the differences between scales and indexes, and how each is uniquely specified and validated. Scales are based on reflective measurement in which classical test theory can be applied (e.g., reliability estimation, confirmatory factor analysis) to evaluate scores that are hypothesized to be highly correlated and as representing manifestations or reflections of underlying constructs. Indexes are based on formative measurement in which scores are not necessarily correlated but rather are hypothesized as inducing or forming constructs. We reanalyzed original Sense of Community Index (SCI) data from the Block Booster Project to demonstrate the differences between reflective and formative measures and the implications of model misspecification. Results of structural equation modeling showed that although the reflective model was a poor fit to the data, the SCI modeled with formative measures provided a good fit to the data. For more than a decade, researchers have debated the conceptualization and measurement of PSOC. Not addressed in this debate, however, has been the issue of measurement misspecification and how validation practices may lead to inaccuracies in our understanding of PSOC. Key differences between measurement perspectives on PSOC are described and recommendations for future work are discussed.

Keywords

psychological sense of community; theory; measurement; partnerships

Psychological sense of community (PSOC) has been defined as people's feelings of belonging, of significance to one another and to groups, and as shared belief that their needs will be met through their relationships (McMillan & Chavis, 1986). Sarason (1974), whose ideas were central to this definition, observed that a lack of PSOC was too often experienced by people in groups and communities, and that a lack of PSOC could be a destructive force in people's lives. Addressing its consequences and its prevention, he argued, should be a major goal for researchers working to enhance people's health or well-being in the context of the many different environments and social systems in which they live their lives.

Although often studied in neighborhoods or other geographically-based settings (Powell, 2015), PSOC has been examined in a variety of collaborative contexts, particularly university-community partnerships. Examples include case studies of partnerships between university-based urban design studios and municipal planning agencies in Columbia, China, and Germany (Pizarro, 2015), research on the capacity-building efforts of university-community partnerships in rural Pennsylvania and Egypt (Forden & Carrillo, 2015), and an evaluation of a university-community partnership that was focused on redressing health disparities in Missouri (Majee, Goodman, Vetter-Smith & Canfield, 2016). These studies described how PSOC was essential to the success of the partnerships because it served as the foundation on which people sustained their engagement and cooperation over time to address their common problems.

Efforts to develop a theory of PSOC have focused on identifying and explaining different dimensions of the construct (McMillan & Chavis, 1986), as well as describing the expected direction (i.e., positive or negative) and form (i.e., direct or indirect) of relationships between PSOC and other conceptually relevant variables (Mannarini, Rochira, & Talò, 2014; Neal & Neal, 2014; Nowell & Boyd, 2010; 2014). McMillan and Chavis (1986), for example, identified and explained unique dimensions of PSOC. These included: (1) needs fulfillment, a perception that members' needs will be met by the community; (2) group membership, a feeling of

belonging or a sense of interpersonal relatedness; (3) influence, a sense that one matters, or can make a difference, in a community and that the community matters to its members; and (4) emotional connection, feelings of attachment or bonding with others that are rooted in shared history, place or experiences.

Several context-specific, quantitative instruments also have been developed to assess PSOC in studies of university-community partnerships. The Sense of Community Index (SCI) was developed to assess the impact of the Block Booster Project (Florin & Wandersman, 1990; Florin, Chavis, Wandersman & Rich, 1992), and it was published in Perkins, Florin, Rich, Wandersman, and Chavis (1990). This project was funded by the Ford Foundation to create a partnership between the Citizens Committee for New York, a nonprofit organization that provided technical assistance to block associations, and university-based researchers. The SCI was used in a longitudinal evaluation that examined outcomes of block associations. It was administered to members of block associations as well as unaffiliated community residents. Results of the evaluation also were used to inform the development and testing of a technical assistance approach that was intended to enhance the capacity of block associations. Similarly, the Brief Sense of Community Scale (BSCS) (Peterson, Speer, McMillan, 2008) was developed to assess the impact over time of a health promotion initiative conducted in a rural county located in the Midwestern U.S. That project was funded by the Centers for Disease Control and Prevention to support the collaboration between community residents and university staff in community-based participatory research. As part of an outcomes evaluation, PSOC was assessed among community residents prior to and after the intervention in both target and comparison counties. In other studies, PSOC instruments were developed specifically to assess the construct among people who were already actively engaged in university-community partnerships. An example is the Community Organization Sense of Community Scale (COSOC) (Hughey, Speer & Peterson, 1998), which has been used extensively in longitudinal studies of community organizing initiatives (e.g., Speer, Peterson, Armstead & Allen, 2013; Speer, Peterson, Zippay & Christens, 2010) and substance abuse prevention coalitions (Powell & Peterson, 2014) in which university researchers partnered with community residents and professional organizers to change sociopolitical and other environmental conditions that contributed to social and health disparities.

Although the empirical work to date on PSOC has been useful, theory development has received less attention from researchers (Nowell & Boyd, 2010; 2014). Scholars have traditionally considered any theory as consisting of two fundamental and equally important parts (Bagozzi & Phillips, 1982; Blalock, 1968; Costner, 1969; Edwards, 2011). First, consistently with the theoretical work to date on PSOC, a theory typically involves the identification and explanation of relevant constructs, as well as a description of the hypothesized relationships between constructs. Second, a theory should include the hypothesized relationships between constructs and measures (Edwards & Bagozzi, 2000). This latter part of a theory was labeled by Blalock (1968) as the auxiliary theory.

Of critical importance to an auxiliary theory is the hypothesized direction of causality in the relationships between constructs (e.g., schematic models or conceptual terms that are intended to represent real phenomena) and measures (e.g., scores from observations or people's responses to scale items). An example relevant to the literature on university-community partnerships is that of alcohol intoxication. Several researchers have described the development of university-community partnerships to address the potentially harmful consequences of alcohol intoxication (e.g., Davis, Jason, Ferrari, Olson, & Alvarez, 2005; Ozer, & Wright, 2012; Smith, Wise, Rosen, Rosen, Childs & McManus, 2014). As an example of an auxiliary theory, Diamantopoulos (2011) discussed alternatives for how researchers might conceptualize the construct of alcohol intoxication in relation to its measures. One way would be to consider alcohol intoxication as a cause of that which an instrument measures; for example, people's symptoms of feeling intoxicated after they had consumed various amounts of alcohol (i.e., if one is intoxicated, a measure would be conceptualized to capture intoxication as reflected in respondent perceptions or symptoms). Conversely, researchers might conceptualize alcohol intoxication as being formed by, or as a result of, what an instrument is measuring. This could involve observing the amount of actual alcohol consumed, the variety of different types of alcoholic beverages consumed, or the pace with which alcohol is consumed. It would be hypothesized that the level of alcohol intoxication would be indicated by the amount of alcohol consumed.

An auxiliary theory is particularly useful because it serves as the basis for the specification of measurement models that can be tested empirically to

(dis)confirm researchers' conceptualizations or operationalizations of constructs (Sajtos & Magyar, 2015). Of vital concern is the possibility that a particular instrument, when tested from one theory, could be judged as invalid; whereas the same instrument, with the appropriate measurement model that is based on an alternate theory, could be found valid. For instance, Flaherty, Zwick, and Bouchey (2014) have called for abandonment of the SCI because the hypothesized measurement model did not provide an adequate fit to the data from samples of participants. We assert, however, that the SCI may in fact be found to be valid if tested using the correctly specified measurement model that is based on an explicit auxiliary theory. Misspecification of measurement models can also result in Type I or Type II errors when researchers test relationships between measures of particular constructs and those of other relevant variables. The estimation of the relationships between the SCI and other variables, such as neighboring behavior, may be much stronger when tested using the appropriate measurement model than they are when tested using a misspecified measurement model.

Unfortunately, no prior work has systematically addressed auxiliary theory development for PSOC. This situation is not unique to research on PSOC. Miguel, Ornelas, and Maroco (2015) and Peterson (2014), for example, discussed the lack of auxiliary theory development for the construct of empowerment. Several explanations have been offered for the failure to consider issues that are relevant to this crucial aspect of theory (e.g., Bollen & Davis, 2009). One explanation is that much of our training in measurement has been based on an implicit assumption of a particular direction of causality between constructs and measures, which has led to an emphasis on the use of procedures such factor analysis and Cronbach's alpha even when they are not appropriate. A second, and related, explanation is that researchers are often uncertain about how to deal with the technical challenges involved with achieving statistically identified structural equation models in tests of particular auxiliary theories that assume a different direction of causality between constructs and measures. Although neglected, an auxiliary theory is vital to development of the knowledge base for PSOC because it provides the means through which researchers can carefully and intentionally bridge their abstract ideas with concrete reality and generate a theory that is amenable to more precise observation and scrutiny. There are several issues for

researchers to consider as they begin the work of auxiliary theory development for PSOC.

Considerations in Developing an Auxiliary Theory

The most basic issue to consider in developing an auxiliary theory for PSOC is the direction of causality between the construct and its measures (Willoughby, Holochwost, Blanton, & Blair, 2014). As stated previously, one approach is to conceptualize a construct as a cause of its measures. Reflective measurement is consistent with this approach in which measures are viewed as manifestations or reflections of an underlying construct (Bollen, 1989; Nunnally, 1978). Examples of PSOC studies adopting a reflective measurement perspective include those that applied confirmatory factor analysis to test scores from scales designed to assess different dimensions of PSOC (e.g., Barati, Samah, & Ahmad, 2012; Wombacher, Tagg, Bürgi, & MacBryde, 2010). Instruments based on reflective measurement are often referred to as scales (Diamantopoulos & Winklhofer, 2001), and researchers frequently use techniques that are consistent with classical test theory, such as reliability estimation or confirmatory factor analysis, to evaluate scores that are derived from scales (Streiner, 2003a; Worthington & Whittaker, 2006).

In contrast, another approach is to conceptualize measures as causes or defining characteristics of a construct (Bollen & Lennox, 1991; Streiner, 2003b). Formative measurement is consistent with this approach in which scores are viewed as inducing or forming a construct (Bollen & Davis, 2009). Examples of studies that have explicitly applied a formative measurement perspective are not present in the PSOC literature, but may be found in the literature on socio-economic status (SES) (Hardin, Chang, Fuller, & Torkzadeh, 2011). These studies often consider SES as a construct that is formed by a combination of education, income or occupation. As noted by Nunnally and Bernstein (1994), researchers generally consider people as lower or higher in SES because they have less or more education or income; variations in people's education or income are not generally thought to occur because they have lower or higher SES. Instruments based on formative measurement are often referred to as indexes, such as an SES index, and researchers have recommended criteria involving indicator collinearity and criterion validity to assess the quality of indexes (Diamantopoulos, Riefler, & Roth, 2008).

The differences between reflective (scale) and formative (index) perspectives can be seen more clearly in their distinctive measurement models. Figure 1 illustrates the application of these measurement models to PSOC. The models in this figure incorporate the conventions used in structural equation modeling to depict latent variables or PSOC constructs (represented by circles), observed variables or PSOC measures (represented by rectangles), as well as the hypothesized direction of the relationships between PSOC constructs and PSOC measures (symbolized by arrows). As can be seen in Figure 1, researchers applying a reflective measurement perspective to develop scales that assess PSOC would view the relationship as emanating from the construct and directed toward observed measures, suggesting that variation in a PSOC construct would lead to variation of PSOC measures. This model implies that reflective measures of PSOC would be hypothesized as strongly correlated since they represent the same underlying PSOC construct. Alternatively, as shown in Figure 1, researchers applying a formative measurement perspective to develop indexes to assess PSOC would view the relationship as emanating from the observed PSOC measures and directed toward the PSOC construct, suggesting that variation in PSOC measures would lead to variation in the PSOC construct. The curved, double-headed arrows in the formative model in Figure 1 indicate that there may be some correlation between the formative measures of PSOC due to their relationships with the PSOC construct; however, unlike reflective measures which are expected to be highly correlated, there is no expectation of strong correlations between formative measures (Hardin et al., 2011; MacCallum & Browne, 1993). Moreover, proponents of this measurement perspective (e.g., Diamantopoulos & Siguaw, 2006) consider strong correlations between formative measures as a crucial shortcoming, similar to multicollinearity in regression analysis (Peterson, Gischlar & Peterson, 2017).

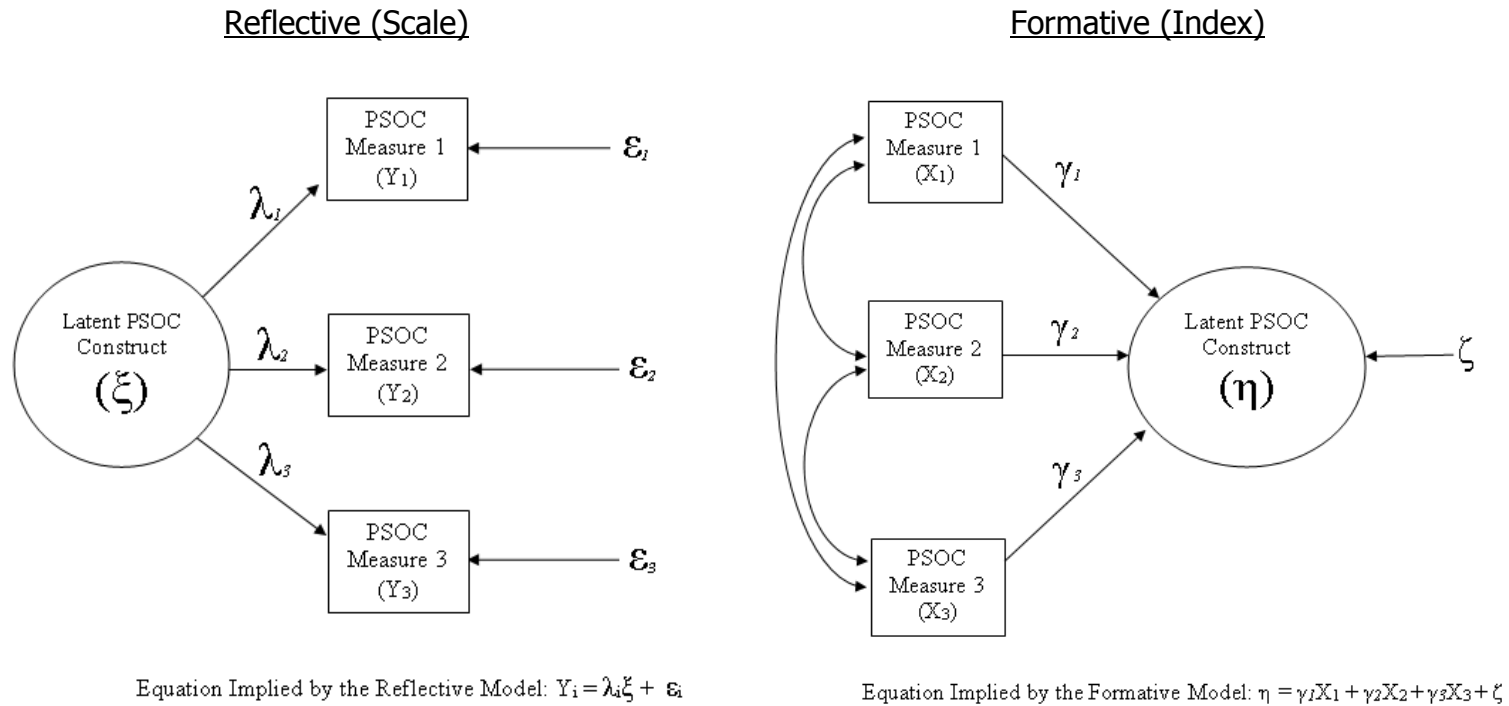


Figure 1. Example reflective (scale) and formative (index) measurement models for psychological sense of community (PSOC). Note: In isolation, the formative model is underidentified.

Figure 1 also includes the equations that are implied by the application of reflective and formative measurement perspectives, respectively. These equations are based on the work of Bollen and Lennox (1991), who described mathematically the differences between these measurement models. The measurement model for PSOC in Figure 1 implies an equation where Y_i is a reflective PSOC measure, the Greek letter xi (ξ) refers to its associated PSOC construct (e.g., the emotional connection dimension of PSOC), and the Greek letter lambda (λ) refers to the effect of ξ on Y_i . The Greek letter epsilon (ϵ) refers to the error term of Y_i , representing measurement error. This equation indicates that PSOC measures from items in a scale, such as people's scores from the BSCS representing the extent to which they agree or disagree with statements like "I feel connected to this partnership¹" (Peterson, Speer, & McMillan, 2008, p. 71), are a function of their true scores on the underlying trait or state, plus measurement error. In this example, the underlying state is the emotional connection dimension of PSOC (i.e., feelings of attachment or bonding with others based on shared history, place or experiences) as defined by McMillan and Chavis (1986). Because models based on this perspective indicate that measures are caused by constructs, reflective measures are also referred to as effect indicators (Bollen & Lennox, 1991; Edwards, 2011).

As opposed to the reflective model in Figure 1, the formative model implies an equation where X_i are formative PSOC measures, the Greek letter eta (η) refers to the associated PSOC construct, and the Greek letter gamma (γ_i) refers to the effect of X_i on η . The Greek letter zeta (ζ) refers to the residual, which is considered that part of η that is not explained by X_i . This equation indicates that individuals' levels on a PSOC construct are a function of their scores from the PSOC items. For example, the SCI (Perkins et al., 1990) and Sense of Community Index – 2 (SCI-2) (Abfalter, Zaglia, & Mueller, 2012; Chavis, Lee, & Acosta, 2008) include items that seem to have been developed from a formative measurement perspective. Items from these instruments also were designed to assess the emotional connection dimension of PSOC as described in the McMillan and Chavis (1986) theory. One such item in the SCI-2 was worded: "Members of this partnership have shared important events together, such as holidays, celebrations, or disasters" (Abfalter et al., 2012, p. 403). As

¹ Items used as examples in this section have been modified to reference "partnership" rather than "neighborhood," "block" or "community".

an index, the scores derived from this and other items of the SCI-2 may be considered not as reflecting an underlying construct but rather as forming it. Because models based on this approach imply that constructs are caused by their measures, formative measures are also referred to as causal indicators (Blalock, 1964; Bollen & Davis, 2009; Hardin et al., 2011).

The measurement perspective that researchers apply to the development of their instruments has important implications for their studies of PSOC, particularly for issues of dimensionality, internal consistency reliability, and validity (Diamantopoulos et al., 2008; MacKenzie, Podsakoff, & Podsakoff, 2011; Streiner, 2003b). Dimensionality refers to whether measures are hypothesized as representing a single dimension of a construct or multiple dimensions of a construct (Bollen & Lennox, 1991; Willoughby et al., 2014). Reflective measures that are derived from items in a scale, or subscale, are assumed to be unidimensional, with each item redundantly capturing the essence of the construct (Peterson et al., 2017). For example, in the BSCS, which also was designed to assess each dimension of PSOC as conceptualized by McMillan and Chavis (1986), two items were included in a subscale to assess needs fulfillment: "I can get what I need in this partnership" and "This partnership helps me fulfill my needs" (p. 71). As stated earlier, needs fulfillment is a dimension of PSOC that was defined by McMillan and Chavis (1986) as a perception that members' needs will be met by the community. The scores from these two BSCS items were hypothesized as representing one underlying construct (i.e., the dimension of needs fulfillment). As can be seen in this example, items that assess a single dimension are designed with the same theme. As a result, reflective measures can be considered as conceptually interchangeable (Jarvis, MacKenzie, & Podsakoff, 2003). This implies that deleting a particular item from this subscale of the BSCS would not change the overall meaning of the construct of needs fulfillment. Instruments developed in this way can have what DeVellis (2016) described as useful redundancy, which refers to multiple items in a scale having the same meaning but are not worded or structured in quite the same way.

Contrary to reflective measures, formative measures are assumed to be multidimensional. Instruments developed from this perspective are comprised of items that are designed to capture different features or

dimensions of a construct (Peterson et al., 2017). The SCI-2 (Chavis et al., 2008), for example, also included several items to assess needs fulfillment. Example items are: "When I have a problem, I can talk about it with members of this partnership" and "Partnership members and I value the same things." These items each appear to involve different features of needs fulfillment. Because formative measures are not hypothesized as representing the same underlying construct, they cannot be considered as conceptually interchangeable. This implies that deleting a particular formative measure would change the overall meaning of the construct (MacKenzie, Podsakoff, & Jarvis, 2005). In addition, redundancy among items that are developed using this perspective is viewed as undesirable (Jarvis et al., 2003; Petter, Straub, & Rai, 2007), and researchers advocating for the use of formative measures generally recommend that redundant items are eliminated during the index development process (Diamantopoulos & Sigua, 2006; Diamantopoulos & Winklhofer, 2001).

Scales and indexes also differ in ways that have important implications for internal consistency reliability. Because reflective measures are viewed as representing the same underlying dimension and as interchangeable, they also should be strongly correlated (Jarvis et al., 2003; Worthington & Whittaker, 2006). Higher correlations among measures will result in estimates of stronger internal consistency reliability, such as coefficient alpha (Cronbach, 1951). Applied specifically to research on PSOC, internal consistency reliability can be regarded as a characteristic of scores from an overall PSOC scale or subscale that was administered to a particular sample (Streiner, 2003a). It refers to "the extent to which a sample's patterns of responses to items or objects are consistent or repeatable across items" (Helms, Henze, Sass, & Mifsud, 2006, p. 632). Likewise, the loadings (λ_i) shown in Figure 1 that involve the strength of the relationships between PSOC measures and the underlying PSOC construct, and which can be determined using confirmatory factor analysis, would be stronger as a result of stronger relationships between reflective measures (Bollen, 1989; Roberts & Thatcher, 2009).

One important caveat is that estimates of internal consistency can be influenced greatly by the length of an instrument (Cortina, 1993; Lord, Novick, & Birnbaum, 1968). More specifically, scores from instruments with greater numbers of items will have higher coefficient alphas. Cortina

(1993), for example, showed convincingly that scores from an instrument with six items and a weak average item correlation of .30 had a coefficient alpha of .72. When Cortina (1993) increased the number of items to 12, which is the same number of items included in the SCI (Perkins et al., 1990), the coefficient alpha increased to .84; and, when the number of items was increased to 18 items, the coefficient alpha increased to .88. This phenomenon explains why scores from indexes such as the one-factor SCI (Perkins et al., 1990) could have coefficient alphas ranging from .69 to .80 (Long & Perkins, 2003) or why scores from the one-factor SCI-2, which has 24 items, could have a coefficient alpha of .94 (Chavis et al., 2008). Scores from instruments with substantial numbers of items that are weakly correlated can produce high coefficient alphas.

Formative measures from an index are not required to correlate because they are hypothesized as representing different features of a construct (Bollen & Lennox, 1991; Willoughby et al., 2014). Consequently, neither reliability estimation (i.e., internal consistency) nor confirmatory factor analysis are appropriate for testing formative measures (Howell, Breivik, & Wilcox, 2007; Streiner, 2003b). Because many of the items from the SCI (Perkins et al., 1990) and the SCI-2 (Chavis et al., 2008) appear to have been developed from a formative measurement perspective, it is imperative for researchers studying PSOC to recognize that estimates of internal consistency reliability are not appropriate for evaluating scores from these instruments. This issue also can have important implications for item development (Peterson et al., 2017). When researchers apply a reflective measurement perspective, items are viewed as sampled from the same conceptual domain and are constructed to have the same meaning (Bollen & Lennox, 1991). When a formative measurement perspective is applied, each item is intentionally developed to capture a unique aspect of the construct, and researchers have to consider the inclusion of items that fully represent their conceptualization of the construct and which are non-redundant (Petter et al., 2007).

Researchers developing scales or indexes to assess PSOC also must consider the different criteria that are recommended to evaluate the validity of measures. The validity of reflective measures that are derived from PSOC scales can be evaluated, in part, by empirically testing the fit a measurement model similar to that presented in Figure 1 (Bagozzi, Yi, & Phillips, 1991; Roberts & Thatcher, 2009). Validity is determined by the

extent to which the reflective PSOC measures represent the PSOC construct as indicated by a good fit of the model to data from a sample. It also has been suggested that researchers using confirmatory factor analysis to examine the degree of correspondence between a construct and its measures apply specific criteria for loadings (e.g., λ_i greater than .60) (DiStefano & Hess, 2005), with higher loadings indicating greater validity of reflective measures. Although model-to-data fit is similarly important to establishing the validity of formative measures, the criteria proposed also involve issues of collinearity, prediction of a criterion variable(s), and the residual (ζ) in Figure 1 (Diamantopoulos et al., 2008; MacKenzie et al., 2005), with a smaller proportion of variance in the PSOC construct attributed to the residual indicating greater validity.

Empirical Example Demonstrating the Differences between Reflective and Formative Measures

We reanalyzed original SCI data (Perkins et al., 1990) from the Block Booster Project to provide an empirical example that demonstrates the differences between reflective and formative measures and the implications of model misspecification. We used data from Time 1 only and included cases ($n = 546$) with complete data for all of the items used in our analysis. In addition to the original SCI, we analyzed the five-item neighboring behavior scale (NBS) used by Long and Perkins (2003), who also reanalyzed the original SCI data, and the two-item perceived safety scale (PSS) used in the original Block Booster study. These items are shown in the Appendix. Because of the method bias resulting from the mixed use of positively and negatively worded items that has been shown with the SCI (Peterson, Speer & Hughey, 2006) and other measures of PSOC (Stevens, Jason, Ferrari, Olson & Legler, 2012), we excluded negatively worded items from this analysis. All analyses were performed using structural equation modeling procedures of Amos 23.0. Maximum likelihood procedures were used for all models.

Our focus was on testing reflective and formative models of the SCI. Our objective was to determine which of the models provided the best fit to the data. As can be seen in Figure 2, three specific models were tested: Model 1, a reflective model of the one-factor SCI; Model 2, a reflective model of the SCI predicting the NBS and PSS; and, Model 3, a formative model of the SCI predicting the NBS and PSS. We tested a one-factor

model of the SCI in Models 1 and 2 because Perkins et al. (1990) introduced the SCI as a unidimensional scale. In addition, because formative measurement models must be tested within a larger model to achieve identification (Petter et al., 2007), we included the NBS and PSS with reflective measures in the model that included the SCI with formative measures. We also included the NBS and PSS in Model 2, as well as in Model 3, in order to demonstrate how the estimated relationships between the SCI and the NBS and PSS can differ based on model specification.

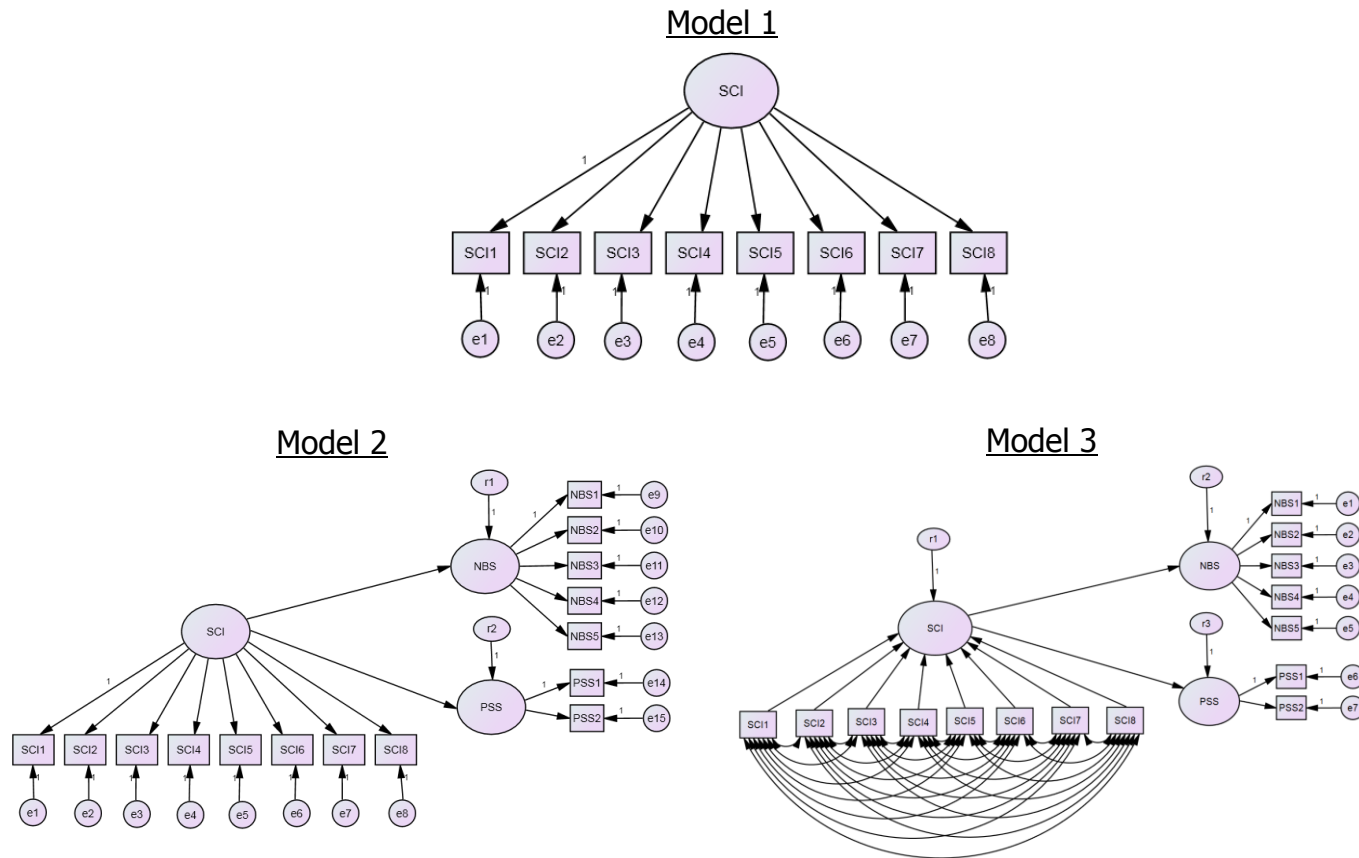


Figure 2. Models tested as an empirical example demonstrating the differences between reflective and formative measures. Items of the Sense of Community Index (SCI), Neighboring Behavior Scale (NBS) and Perceived Safety Scale (PSS) are in the Appendix.

Importantly, we found that Model 3 in Figure 2 was underidentified. The topic of model identification in structural equation modeling is complex and involves “the extent to which a unique set of values can be inferred for the unknown parameters from a given covariance matrix of analyzed variables that is reproduced by the model” (Byrne, 2010, p. 34). A vital issue is whether there are enough data points (i.e., variances and covariances of the observed variables) to estimate the number parameters (e.g., regression coefficients between latent and observed variables) in a particular model. Underidentified models are those in which the number of parameters to be estimated is greater than the number of data points available to the researcher. As stated previously, formative measurement models must be tested within a larger model to achieve identification. However, there was an additional identification issue with Model 3 that involved the scaling of the latent SCI variable. In a structural equation model, each latent variable must be assigned a scale for the model to be identified. Generally, a researcher can scale a latent variable by setting the factor loading for a reflective measure to one. This approach can be seen, for example, in Model 1 in Figure 2, which shows that the path from the latent SCI variable to SCI1 has been set to one. In our Model 3, however, there is not a reflective measure for the latent SCI variable. In this case, “a researcher may scale the latent variable by setting the latent variable’s path to another latent variable to one” (Bollen & Davis, 2009, p. 502). We, therefore, addressed the issue of underidentification in Model 3 by setting the path from the latent SCI variable to the latent NBS variable to one.

The fit indices that we interpreted are considered to be robust measures of model-to-data fit. These included the discrepancy chi-square (χ^2), the χ^2 / degrees of freedom (df) ratio, the Goodness of Fit Index (GFI), the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI) and the Root Mean Square of Error Approximation (RMSEA). Although the χ^2 statistic is reported, we acknowledge that it is often considered too stringent and an unrealistic standard. To interpret the fit indices, non-significant χ^2 values, χ^2 / df ratios less than 2, and higher values (i.e., greater than .95) on the GFI, CFI and TLI indicate acceptable fit, while smaller RMSEA values are desirable. According to Browne and Cudeck (1992), guidelines for interpreting the RMSEA include: <.05 = good fit; .05 to .08 = acceptable fit; .08 to .10 = marginal fit; > .10 = poor fit.

Prior to testing the reflective and formative models of the SCI, we conducted two analyses. First, we tested the fit of the hypothesized one-factor reflective models for the NBS and PSS to the data from the sample, and we computed estimates of their internal consistency reliability. We found that these measurement models for the NBS and PSS provided a good fit to the data, $\chi^2(13) = 17.440$, $p = .180$; χ^2/df ratio = 1.342; GFI = .991; CFI = .994; TLI = .990; RMSEA = .025 (90% CI = .000, .052). Cronbach's alphas for the NBS and PSS were .75 and .64, respectively. Second, we estimated the relationship between the observed, composite scores of the SCI, NBS and PSS. Composite scores were computed as the mean of items in each instrument. Our interest was in comparing the proportion of variance explained in the NBS and PSS by the SCI in this analysis with the results of our structural equation modeling. In this analysis, the observed variable for SCI was found to be a significant predictor of both observed variables, with a standardized regression coefficient of .22 ($p < .001$) and a squared multiple correlation (R^2) of .05 for the observed NBS variable, and a standardized regression coefficient of .20 ($p < .001$) and a squared multiple correlation (R^2) of .04 for the observed PSS variable.

Results of our structural equation modeling are shown in Table 1 and Figures 3 and 4. As can be seen in Table 1, the one-factor reflective model for the SCI provided a poor fit to the data. The χ^2 value was statistically significant, the χ^2/df ratio was greater than the threshold of 2, and the TLI and CFI values were below the .95 cutoff. The GFI was at the .95 threshold. In addition, the RMSEA for Model 1 was .090, indicating a marginal model-to-data fit. Similarly, Model 2, which included a reflective model of the SCI predicting the NBS and PSS with reflective measures, also provided a poor fit to the data. For Model 2, the χ^2 value was again statistically significant, χ^2/df ratio was greater than the threshold of 2, and the GFI, TLI, and CFI values were below the .95 cutoff. The RMSEA value was .054, however, indicating a good model-to-data fit for Model 2. Overall, however, five of the six fit indices indicated that Model 2 failed to provide a good fit to the data. Although this model did not fit the data, we present more detailed results of the analysis in Figure 3. In that model, the R^2 was .07 for the latent NBS variable and the R^2 was .11 for the latent PSS variable.

Table 1. Overall Fit Statistics

Measures of Fit	Models		
	Model 1: Reflective Model of the SCI	Model 2: Reflective Model of the SCI Predicting the NBS and PSS	Model 3: Formative Model of the SCI Predicting the NBS and PSS
χ^2	108.428	225.905	90.340
<i>df</i>	20	88	60
<i>p</i> -value	<.001	<.001	.007
χ^2/df Ratio	5.42	2.567	1.506
GFI	.950	.945	.979
CFI	.833	.897	.977
TLI	.766	.877	.960
RMSEA	.090	.054	.030
(90% CI)	.074, .107	.045, .062	.016, .043

Note: Items of the Sense of Community Index (SCI), Neighboring Behavior Scale (NBS), and Perceived Safety Scale (PSS) are shown in the Appendix.

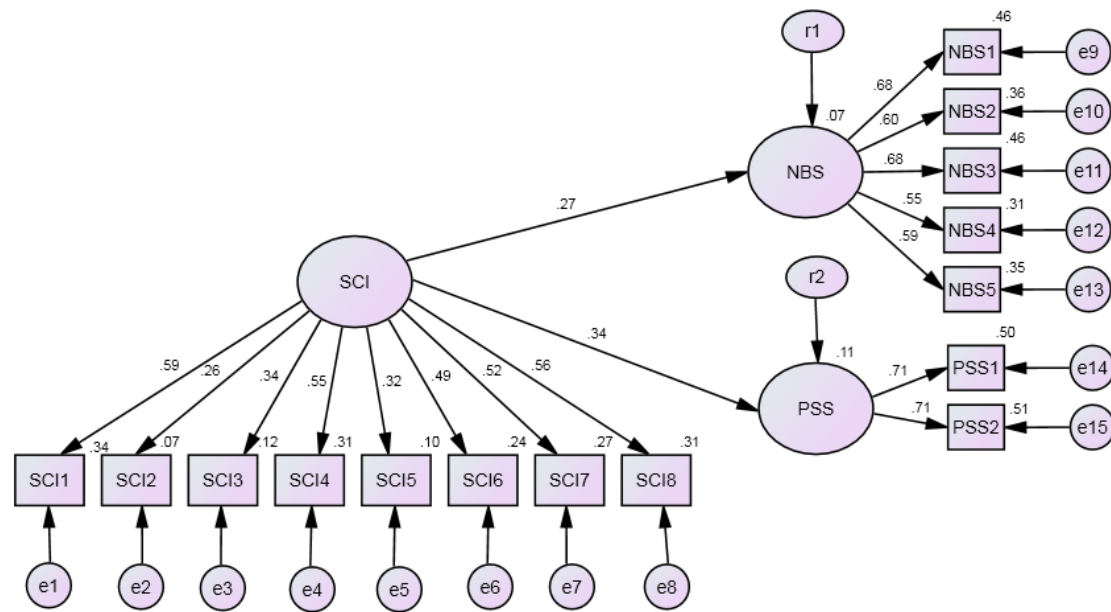


Figure 3. Results of structural equation modeling of the reflective measurement model (Model 2). Values corresponding to paths between latent variables and observed variables, as well as paths between latent variables, represent standardized regression coefficients. All other values are squared multiple correlation coefficients. All of the standardized regression coefficients were statistically significant, $p < .001$.

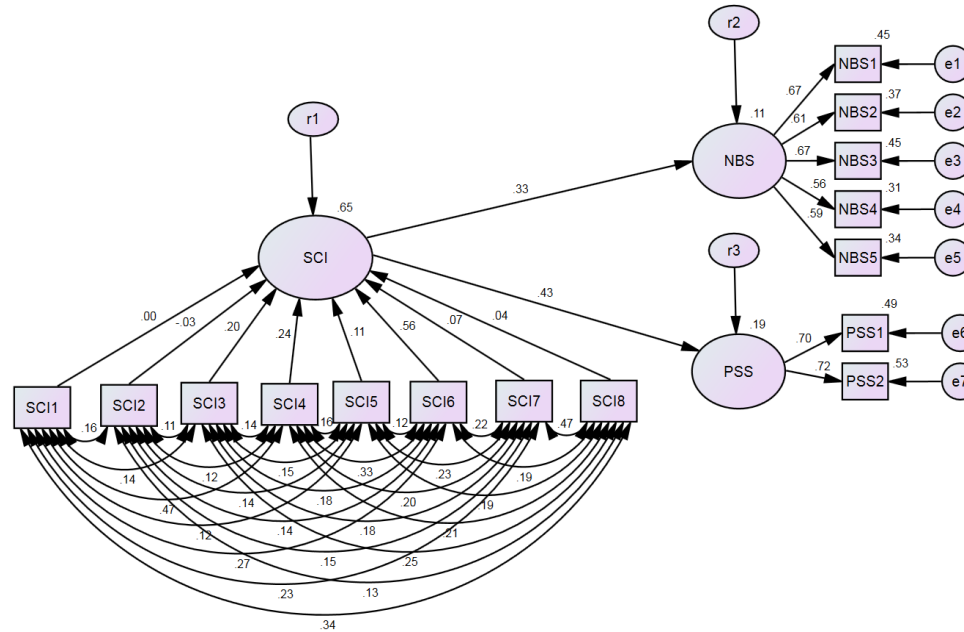


Figure 4. Results of structural equation modeling of the formative measurement model (Model 3). Values corresponding to paths between latent variables and observed variables, as well as paths between latent variables, represent standardized regression coefficients. Values to double-headed arrows are correlation coefficients. All other values are squared multiple correlation coefficients. All of the standardized regression coefficients for paths between the NBS and PSS items and latent variables, as well as those values for paths between latent variables, were statistically significant ($p < .001$). Two of the standardized regression coefficients corresponding to SCI items were statistically significant, SCI4 ($p < .05$), and SCI6 ($p < .001$), while the value for SCI3 approached significance ($p = .052$)

Figure 4 shows additional results for Model 3, including correlation coefficients (r), standardized regression coefficients, and squared multiple correlation coefficients. Figure 4 also shows the relationship between the latent SCI variable, which was formed by scores from the index, and the latent NBS and PSS variables. As can be seen in Figure 4, 11% of the variability in the latent NBS variable was explained. This proportion of variability in the NBS that was explained in Model 3 ($R^2 = .11$) was larger than that which was explained in Model 2 ($R^2 = .07$), as well as that which was explained when we examined only the observed variables representing the SCI and NBS ($R^2 = .05$). Likewise, the proportion of variability in the PSS that was explained in Model 3 ($R^2 = .19$) was larger than that which was explained in Model 2 ($R^2 = .11$), as well as that which was explained when we examined only the observed variables representing the SCI and PSS ($R^2 = .04$). Furthermore, the correlations between SCI items that are shown in Figure 4 indicate that they were, in general, weakly correlated. These correlations ranged from the lowest ($r = .11$) between SCI2 and SCI3 to the highest ($r = .47$) between SCI1 and SCI5 and ($r = .47$) between SCI7 and SCI8. The majority of these correlations (i.e., 17 of the 28 correlations between the SCI items) that are shown in Figure 4 were below .20. These findings are consistent with our expectations and support the validity of formative measures. However, only a few of the SCI items were significant predictors of the latent SCI variable. In addition, SCI items explained only 65% of variability in the latent SCI variable ($R^2 = .65$). Those paths between items and the latent SCI variables that had relatively strong standardized regression coefficients included SCI3 ("I can recognize most of the people who live on my block"), SCI4 ("I feel at home on this block"), and SCI6 ("If there is a problem on this block people who live here can get it solved"). These findings suggest that the validity of the SCI could be improved by the addition of items which may more fully form the construct.

Conclusions

The decision to apply a particular measurement perspective to PSOC should be driven by theory (i.e., the auxiliary theory), which clearly specifies the direction of the relationship between a construct and its measures (Edwards & Bagozzi, 2000). Heretofore, auxiliary theories have not been addressed systematically in research on PSOC nor university-

community partnerships. Perhaps because of this oversight, researchers of PSOC have frequently used indexes in their studies of the construct while relying on procedures consistent with classical test theory (e.g., estimation of internal consistency reliability, confirmatory factor analysis) to analyze scores from these instruments (e.g., Abfalter et al., 2012; Carrillo, Welsh, & Zaki, 2015; Chavis et al., 2008; Chipuer & Pretty, 1999; Flaherty, Zwick, & Bouchey, 2014; Lindblad, Manturuk, & Quercia, 2013; Long & Perkins, 2003; Obst & White, 2004; Perkins et al., 1990; Peterson, Speer, & Hughey, 2006; Stevens, Jason, & Ferrari, 2011; Townley et al., 2013). When providing direct tests of the validity of the SCI, many of these researchers found, as we did with Model 1 and Model 2 in our results, that the hypothesized reflective model of the SCI did not adequately fit their data, and some have called for abandonment of the SCI based on these findings.

Our results also showed that when a formative model of the SCI was tested, it provided a good fit to the data from the sample of participants. Moreover, our reanalysis of original SCI data demonstrated that several problems can result from the misspecification of measures. One critical problem is that formative measures from indexes that are judged as invalid may actually be found to be valid when tested using appropriate methods. Another problem can involve the bias that is introduced when researchers attempt to estimate the relationships between PSOC and other theoretically relevant variables. Several studies (i.e., Jarvis et al., 2003; MacKenzie et al., 2005; Petter et al., 2007; Taing, Johnson, & Jackson, 2010) have shown how large Type I or Type II errors can occur when constructs that are modeled appropriately using a formative measurement perspective are misspecified and modeled using a reflective measurement perspective. Using Monte Carlo simulation, MacKenzie et al. (2005) showed that the misspecification of measurement models can inflate parameter estimates by as much as 400% or deflate parameter estimates by as much as 80%. Our results were consistent with these findings. These issues can have crucial implications for research on university-community partnerships and other interventions that aim to promote PSOC, as well as research that tests associations between PSOC and other conceptually relevant variables.

We assert that knowledge of university-community partnerships in general, and PSOC in particular, can be improved by the development of

auxiliary theories and, consequently, greater clarity about the differences between scales and indexes and how each is uniquely specified and validated. This paper provides an introduction to readers about these issues and offers a tangible way to improve research in these areas. We contend that a theory of PSOC (or any other construct of interest to researchers studying university-community partnerships) is incomplete without an auxiliary theory. In particular, a construct assessed using a reflective measurement perspective may not necessarily be the same if it is assessed using a formative measurement perspective. Diamantopoulos (2011) demonstrated this point with the example of the construct of alcohol intoxication, which was discussed earlier in this paper. He described alcohol intoxication as a construct that could be measured reflectively, such as assessing people's symptoms of feeling intoxicated after they had consumed various amounts of alcohol, or formatively, such as observing people's actual consumption of different types of alcoholic beverages. Although either measurement perspective could be used in the assessment of alcohol intoxication, Diamantopoulos (2011) argued that the construct assessed using reflective measures could be more accurately labeled as self-perceived alcohol intoxication. The construct assessed using formative measures could be more accurately labeled as behavioral consumption of intoxicating beverages. This example also demonstrates that constructs themselves are not inherently reflective or formative (Baxter, 2009; MacKenzie et al., 2011), but rather can be defined differently based on the auxiliary theory. This issue is especially relevant to PSOC. The construct of PSOC that is assessed using an instrument such as the BSCS may be quite different than that which is assessed using the SCI or SCI-II, and theory should be developed to explain these differences.

There are several limitations to our discussion of scales, indexes, and auxiliary theory development for PSOC, which should be acknowledged. First, our discussion was limited to a focus on the relationships between measures and first-order constructs. It should be recognized, however, that the principles we described can be generalized to higher-order multidimensional models for PSOC (Peterson, 2014). In these types of models, dimensions also can be hypothesized as relating to more general (i.e., higher-order) constructs either reflectively or formatively (Law & Wong, 1999; MacKenzie et al., 2005). Second, our discussion was limited to introductory issues involving reflective and formative measurement

perspectives on PSOC. There has been considerable debate about the limitations of the application of formative measurement to the assessment of first-order constructs (e.g., Bagozzi, 2007; Edwards, 2011; Iacobucci, 2010). Our purpose here is not to advocate for a particular measurement perspective; rather, our goal is to clarify for readers the distinctions between scales and indexes and describe implications of researchers' measurement choices. Third, our analysis of the SCI was limited by the particular variables we used to achieve statistical identification of our formative measurement model. As stated earlier, a formative measurement model must be tested within a larger structural equation model to achieve statistical identification. We included the NBS and PSS as criterion variables in our analysis to address this issue. It should be recognized, however, that the strength of relationships between formative measures and a latent construct can be influenced by the particular criterion variables used to achieve statistical identification (Howell et al., 2007). Finally, our analysis was also limited by its focus on a version of the SCI that included a dichotomous response-option format. Researchers should attempt to replicate our results with other versions of the SCI that include types of response-option formats which can increase variability of measures.

Theoretical work is required to improve research on PSOC and university-community partnerships. Auxiliary theories that advance our understanding of the directional relationship between PSOC constructs and associated measures must be improved. Without increasing attention to the alignment of evaluation methods with the conceptualizations anchoring PSOC instrument development, our science will be substandard. Researchers, editors, and reviewers all share in the responsibility for strengthening our conceptual and empirical efforts. Given the absence of explicit consideration of auxiliary theory in relation to PSOC in the past, it is incumbent for scholars to carefully and straightforwardly articulate their measurement perspective, and to provide direct justification for their particular conceptualizations in relation to PSOC. Contributions that enhance PSOC and the university-community partnerships we seek to promote require this kind of commitment.

References

- Abfalter, D., Zaglia, M. E., & Mueller, J. (2012). Sense of virtual community: A follow up on its measurement. *Computers in Human Behavior, 28*(2), 400-404. doi: 10.1016/j.chb.2011.10.010
- Bagozzi, R. P. (2007). On the meaning of formative measurement and how it differs from reflective measurement: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods, 12*(2), 229-237. doi: 10.1037/1082-989X.12.2.229
- Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly, 36*(3), 421-458. doi: 10.2307/2393203
- Bagozzi, R. P., & Phillips, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly, 27*(3), 459-489. Retrieved from <http://www.jstor.org/stable/2392322>
- Barati, Z., Samah, B. A., & Ahmad, N. (2012). Sense of community and citizen participation in neighborhood council in Iran. *Journal of American Science, 8*(1), 655-661. Retrieved from http://free-journal.umm.ac.id/files/file/090_7974am0801_655_661.pdf
- Baxter, R. (2009). Reflective and formative metrics of relationship value: A commentary essay. *Journal of Business Research, 62*(12), 1370-1377. doi: 10.1016/j.jbusres.2008.12.004
- Blalock, H. M. (1964). *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina.
- Blalock, H. M. (1968). Theory building and causal inferences. *Methodology in Social Research, 155*-198.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley-Interscience.
- Bollen, K. A., & Davis, W. R. (2009). Causal indicator models: Identification, estimation, and testing. *Structural Equation Modeling: A Multidisciplinary Journal, 16*, 498-522. doi: 10.1080/10705510903008253
- Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*(2), 305.
- Browne, M. W. & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230-258.

- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming (2nd edition)*. New York: Routledge/ Taylor & Francis Group.
- Carrillo, A. M., Welsh, E., & Zaki, N. (2015). The revolution, New Egypt, and psychological sense of community. *Journal of Community Psychology, 43*(1), 107-118. doi: 10.1002/jcop.21713
- Chavis, D. M., Lee, K. S., & Acosta, J. D. (2008). *The sense of community (SCI) revised: The reliability and validity of the SCI-2*. Paper presented at the Second International Community Psychology Conference, Lisboa, Portugal.
- Chipuer, H. M., & Pretty, G. M. (1999). A review of the sense of community index: Current uses, factor structure, reliability, and further development. *Journal of Community Psychology, 27*(6), 643-658. doi: 10.1002/(SICI) 1520-6629(199911)27:6<643::AID-JCOP2>3.0.CO;2-B
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.
- Costner, H. L. (1969). Theory, deduction and rules of correspondence. *American Journal of Sociology, 75*(2), 245-263. Retrieved from <http://www.jstor.org/stable/2776106>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.
- Davis, M. I., Jason, L. A., Ferrari, J. R., Olson, B. D., & Alvarez, J. (2005). A collaborative action approach to researching substance abuse recovery. *The American Journal of Drug and Alcohol Abuse, 31*, 537-553. doi: 10.1081/ADA-200068110
- DeVellis, R. F. (2016). *Scale Development: Theory and Applications*. Newbury Park, CA: Sage Publications.
- Diamantopoulos, A. (2011). Incorporating formative measures into covariance-based structural equation models. *MIS Quarterly, 35*(2), 335-358.
- Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research, 61*, 1203-1218. doi: 10.1016/j.jbusres.2008.01.009
- Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management, 17*, 263-282. doi: 10.1111/j.1467-8551.2006.00500.x

- Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, 38(2), 269-277. doi: <http://dx.doi.org/10.1509/jmkr.38.2.269.18845>
- DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment*, 23(3), 225-241. doi: [10.1177/073428290502300303](http://dx.doi.org/10.1177/073428290502300303)
- Edwards, J. R. (2011). The fallacy of formative measurement. *Organizational Research Methods*, 14, 370-388. doi: [10.1177/1094428110378369](http://dx.doi.org/10.1177/1094428110378369)
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5(2), 155-174. doi: <http://dx.doi.org/10.1037/1082-989X.5.2.155>
- Flaherty, J., Zwick, R. R., & Bouchey, H. A. (2014). Revisiting the sense of community index: A confirmatory factor analysis and invariance test. *Journal of Community Psychology*, 42(8), 947-963. doi: [10.1002/jcop.21664](http://dx.doi.org/10.1002/jcop.21664)
- Florin, P., Chavis, D., Wandersman, A. & Rich, R. (1992). The block booster project: A systems approach to understanding and enhancing grassroots organizations. In R. Levine & H. Fitzgerald (Eds.), *Analysis of dynamic psychological systems* (pp 215-243). New York: Plenum.
- Florin, P. & Wandersman, A. (1990). An introduction to citizen participation, voluntary organizations, and community development: Insights for empowerment through research. *American Journal of Community Psychology*, 18, 41-54.
- Forden, C.L. & Carrillo, A.M. (2015). A cross-cultural comparison of campus-community partnerships to build evaluation capacity in rural Pennsylvania and urban Cairo. *Community Psychology in Global Perspective*, 1 (1), 86-104. doi: [10.1285/i24212113v1i1p86](http://dx.doi.org/10.1285/i24212113v1i1p86)
- Hardin, A. M., Chang, J. C., Fuller, M. A., & Torkzadeh, G. (2011). Formative measurement and academic research: In search of measurement theory. *Educational and Psychological Measurement*, 71(2), 281-305. doi: [10.1177/0013164410370208](http://dx.doi.org/10.1177/0013164410370208)
- Helms, J. E., Henze, K. T., Sass, T. L., & Mifsud, V. A. (2006). Treating Cronbach's alpha reliability coefficients as data in counseling

- research. *The Counseling Psychologist*, 34(5), 630-660. doi: 10.1177/0011000006288308
- Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. *Psychological Methods*, 12(2), 205. doi: 10.1037/1082-989X.12.2.205
- Hughey, J., Speer, P. W. & Peterson, N. A. (1999). Sense of community in community organizations: Structure and evidence of validity. *Journal of Community Psychology*, 27, 97-113. doi: 10.1002
- Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20(1), 90-98. doi: 10.1016/j.jcps.2009.09.003
- Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30(2), 199-218. doi: 10.1086/376806
- Law, K. S., & Wong, C. S. (1999). Multidimensional constructs in structural equation analysis: An illustration using the job perception and job satisfaction constructs. *Journal of Management*, 25(2), 143-160. doi: 10.1177/014920639902500202
- Lindblad, M. R., Manturuk, K. R., & Quercia, R. G. (2013). Sense of community and informal social control among lower income households: The role of homeownership and collective efficacy in reducing subjective neighborhood crime and disorder. *American Journal of Community Psychology*, 51(1-2), 123-139. doi: 10.1007/s10464-012-9507-9
- Long, D. A., & Perkins, D. D. (2003). Confirmatory factor analysis of the sense of community index and development of a brief SCI. *Journal of Community Psychology*, 31(3), 279-296. doi: 10.1002/jcop.10046
- Lord, F. M., Novick, M. R., & Birnbaum, A. (1968). *Statistical Theories of Mental Test Scores*. Oxford, England: Addison-Wesley.
- MacCallum, R. C., & Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin*, 114(3), 533-541. doi: <http://dx.doi.org/10.1037/0033-2909.114.3.533>
- MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *Journal*

- of Applied Psychology*, 90(4), 710-730. doi:
<http://dx.doi.org/10.1037/0021-9010.90.4.710>
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, 35(2), 293-334.
- Majee, W. Goodman, L., Vetter-Smith, M., & Canfield, S. (2016). Healthy communities initiative: A preliminary assessment of the University of Missouri-Sedalia health promotion partnership. *Community Development*, 47(1), 91-105. doi:
 10.1080/15575330.2015.1100643
- Mannarini, T., Rochira, A. & Talò, C. (2014). Negative psychological sense of community: Development of a measure and theoretical implications. *Journal of Community Psychology*, 42(6). 673–688. doi: 10.1002/jcop. 21645
- McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology*, 14(1), 6-23.
- Miguel, M. C., Ornelas, J. H., & Maroco, J. P. (2015). Defining psychological empowerment construct: Analysis of three empowerment scales. *Journal of Community Psychology*, 43, 900-919. doi: 10.1002/jcop.21721
- Neal, Z. P., & Neal, J. W. (2014). The (In) compatibility of diversity and sense of community. *American Journal of Community Psychology*, 53(1-2), 1-12. doi: 10.1007/s10464-013-9608-0
- Nowell, B., & Boyd, N. (2010). Viewing community as responsibility as well as resource: Deconstructing the theoretical roots of psychological sense of community. *Journal of Community Psychology*, 38(7), 828-841. doi: 10.1002/jcop.20398.
- Nowell, B., & Boyd, N. (2014). Sense of community responsibility in community collaboratives: Advancing a theory of community as resource and responsibility. *American Journal of Community Psychology*, 54(3-4), 229-242. doi: 10.1007/s10464-014-9667-x
- Nunnally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychological theory*. New York: McGraw-Hill.
- Obst, P. L., & White, K. M. (2004). Revisiting the sense of community index: A confirmatory factor analysis. *Journal of Community Psychology*, 32(6), 691-705. doi: 10.1002/jcop.20027

- Ozer, E. J., & Wright, D. (2012). Beyond school spirit: The effects of youth-led participatory action research in two urban high schools. *Journal of Research on Adolescence, 22*, 267-283. doi: 10.1111/j.1532-7795.2012.00780.x
- Perkins, D. D., Florin, P., Rich, R. C., Wandersman, A., & Chavis, D. M. (1990). Participation and the social and physical environment of residential blocks: Crime and community context. *American Journal of Community Psychology, 18*(1), 83-115. doi: org/10.1007/BF00922690
- Peterson, C. H., Gischlar, K. L. & Peterson, N. A. (2017). Item construction using reflective, formative, or Rasch measurement models. *Journal for Specialists in Group Work, 42*(1), 17-32. doi: 10.1080/01933922.2016.1264523
- Peterson, N. A. (2014). Empowerment theory: Clarifying the nature of higher-order multidimensional constructs. *American Journal of Community Psychology, 53*(1-2), 96-108. doi: 10.1007/s10464-013-9624-0
- Peterson, N. A., Speer, P. W., & Hughey, J. (2006). Measuring sense of community: A methodological interpretation of the factor structure debate. *Journal of Community Psychology, 34*(4), 453-469. doi: 10.1002/jcop.20109
- Peterson, N. A., Speer, P. W., & McMillan, D. W. (2008). Validation of a brief sense of community scale: Confirmation of the principal theory of sense of community. *Journal of Community Psychology, 36*(1), 61-73. doi: 10.1002/jcop.20217
- Petter, S., Straub, D., & Rai, A. (2007). Specifying formative constructs in information systems research. *MIS Quarterly, 31*(4), 623-656. Retrieved from <http://www.jstor.org/stable/25148814>
- Pizarro, R. E. (2015). Challenges of implementing sustainable urban design plans through community–university partnerships: lessons from Colombia, China, and Germany. *Current Opinion in Environmental Sustainability, 15*, 48-56. doi: 10.1016/j.cosust.2015.11.005
- Powell, K. H. (2015). A theoretical exploration of PSOC in an ecological context: The town-gown example. *Community Psychology in Global Perspective, 1*(2), 1-22. doi: 10.1285/i24212113v1i2p1.
- Powell, K.G. & Peterson, N.A. (2014). Pathways to effectiveness in substance abuse prevention: Empowering organizational characteristics of community-based coalitions. *Human Service*

- Organizations: Management, Leadership, & Governance*, 38, 471-486. doi: 10.1080/23303131.2014.935839
- Roberts, N., & Thatcher, J. (2009). Conceptualizing and testing formative constructs: Tutorial and annotated example. *ACM SIGMIS Database*, 40(3), 9-39. doi: 10.1145/1592401.1592405
- Sajtos, L. & Magyar, B. (2015). Auxiliary theories as translation mechanisms for measurement model specification. *Journal of Business Research*. doi: 10.1016/j.jbusres.2015.12.007.
- Sarason, S. B. (1974). *The psychological sense of community: Prospects for a community psychology*. Jossey-Bass.
- Smith, E. P., Wise, E., Rosen, H., Rosen, A., Childs, S., & McManus, M. (2014). Top-down, bottom-up, and around the jungle gym: A social exchange and networks approach to engaging afterschool programs in implementing evidence-based practices. *American Journal of Community Psychology*, 53, 491-502. doi: 10.1007/s10464-014-9656-0
- Speer, P. W., Peterson, N. A., Armstead, T. L. & Allen, C. T. (2013). The influence of participation, gender, and organizational sense of community on psychological empowerment: Moderating effects of income. *American Journal of Community Psychology*, 51, 103-113. doi: 10.1007/s10464-012-9547-1
- Speer, P. W., Peterson, N. A., Zippay, A. & Christens, B. (2010). Participation in congregation-based organizing: A mixed-method study of civic engagement. In M.R. DeGennaro and S. Fogel (Eds.), *Using Evidence to Inform Practice for Community and Organizational Change* (pp. 200-217). Chicago: Lyceum Books.
- Stevens, E. B., Jason, L. A., & Ferrari, J. R. (2011). Measurement performance of the Sense of Community Index in substance abuse recovery communal housing. *Australian Community Psychologist*, 23(2), 135-147. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4114391/>
- Stevens, E. B., Jason, L. A., Ferrari, J. R., Olson, B., & Legler, R. (2012). Sense of community among individuals in substance abuse recovery. *Journal of Groups in Addiction & Recovery*, 7, 15-28. doi: 10.1080/1556035X.2012.632319
- Streiner, D. L. (2003a). Starting at the beginning: An introduction to coefficient alpha and internal Consistency. *Journal of Personality Assessment*, 80(1), 99-103, doi: 10.1207/S15327752JPA8001_18

- Streiner, D. L. (2003b). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *Journal of Personality Assessment, 80*(3), 217-222. doi: http://dx.doi.org/10.1207/S15327752JPA8003_01
- Taing, M. U., Johnson, R. E., & Jackson, E. M. (2010). *On the nature of core self-evaluation: A formative or reflective construct*. Paper presented at the 70th annual meeting of the Academy of Management, Montreal, Quebec.
- Townley, G., Katz, J., Wandersman, A., Skiles, B., Schillaci, M. J., Timmerman, B. E., & Mousseau, T. A. (2013). Exploring the role of sense of community in the undergraduate transfer student experience. *Journal of Community Psychology, 41*(3), 277-290. doi: 10.1002/jcop.21529
- Willoughby, M., Holochwost, S. J., Blanton, Z. E., & Blair, C. B. (2014). Executive functions: Formative versus reflective measurement. *Measurement: Interdisciplinary Research and Perspectives, 12*(3), 69-95. doi: 10.1080/15366367.2014.929453
- Wombacher, J., Tagg, S. K., Bürgi, T., & MacBryde, J. (2010). Measuring sense of community in the military: Cross-cultural evidence for the validity of the brief sense of community scale and its underlying theory. *Journal of Community Psychology, 38*(6), 671-687. doi: 10.1002/jcop.20388
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806-838. doi: 10.1177/0011000006288127

Appendix

Sense of Community Index (SCI)

Item	Statement (True/False Response Option Format)
SCI1	I think my block is a good place for me to live.
SCI2	My neighbors and I want the same things from this block.
SCI3	I can recognize most of the people who live on my block.
SCI4	I feel at home on this block.
SCI5	I care about what my neighbors think of my actions.
SCI6	If there is a problem on this block people who live here can get it solved.
SCI7	It is very important to me to live on this particular block.
SCI8	I expect to live on this block for a long time.

Neighboring Behavior Scale (NBS)

The following is a short list of things neighbors might do for each other. Please indicate how many times in the past year, you have been asked to do each one for a neighbor on this block.

Item	Statement (Responses Ranged from Zero to Eight or More)
NBS1	Watch a neighbor's home while they were away.
NBS2	Loan a neighbor some food or a tool.
NBS3	Help a neighbor in an emergency.
NBS4	Offer a neighbor advice on a personal problem.
NBS5	Discuss a problem on the block with a neighbor.

Perceived Safety Scale (PSS)

Item	Statement (Responses: Unsafe; Fairly Safe; Very Safe)
PS1	How safe do you feel during the day?
PS2	How safe do you feel during the night?