8-2018

Measurement Invariance and Construct Validity of the Sense of Community Index

Erin Renee Story
University of Tennessee

To the Graduate Council:

I am submitting herewith a dissertation written by Erin Renee Story entitled "Measurement Invariance and Construct Validity of the Sense of Community Index." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Social Work.

William R. Nugent, Major Professor

We have read this dissertation and recommend its acceptance:

Shandra S. Forrest-Bank, Hillary N. Fouts, Lisa M. Reyes-Mason, John G. Orme

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Measurement Invariance and Construct Validity of the Sense of Community Index

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Erin Renee Story

August 2018

Dedication

This work is dedicated to my Mom, my Dad, my husband, and my son. Mom and Dad, thank you for teaching me the importance of critical thinking, questioning, and learning, and for making sure I always knew that I was capable of all three. Mom, thank you for instilling in me your grit, strength, and stubbornness. Dad, thank you for giving me your flexibility, resilience, and openness. I owe everything I am to the two of you. To my husband, Ali, thank you for the innumerable ways in which you helped me achieve this milestone – for encouraging, comforting, and pushing me. I love you and am so glad you're mine. To my son, Elias, you are my motivation, my inspiration, my heart, and my soul. You are my greatest pride and deepest joy.

Acknowledgements

First, I want to thank my family, nuclear and extended, for their support, encouragement, and humor. I love you all and am so grateful for your influence in my life. I also want to thank Claire Craven, Monica Keele, and Mahasin Habeel for loving me, listening to my venting, and reminding me that I can when I think I can't. I love you all to the moon and back. I would like to thank Becky O'Berry for being a wonderful friend, and for your support while I tried to figure out how to balance practice with pursuing my PhD.

I also want to thank my committee members. Dr. Nugent, thank you for your patience, your encouragement, and your support. Thank you for helping me focus my research and finish despite many setbacks and delays. Dr. Forrest-Bank, thank you for your confidence in me and for all the opportunities you provided me to participate in research, write, present, and publish. Dr. Reyes-Mason, Dr. Orme, and Dr. Fouts, thank you for your incisive, clear, and exceedingly helpful feedback, advice, and comments. I deeply appreciate you contributing your expertise to this project.

Finally, I would like to thank Kellie Cripe, Shannon Trecartin, Matthew Moore, Laurel Proulx, Tara Veerman, Matthew Cuellar, and Autumn Lowry for their support, commiseration, help, and friendship. I owe all of you a great debt for my survival in this process.

Abstract

This multipart dissertation focuses on the psychometric quality of neighborhood social measurement instruments, specifically the Sense of Community Index. The first paper is a systematic literature review of neighborhood social instruments. The findings indicated that while initial evidence of psychometric quality of these instruments was promising, there is further validation work to be done on these instruments. The author recommended further validation of the neighborhood social characteristics measurement tools, specifically the Sense of Community Index. The second and third papers followed the research recommendations of the first paper. The second paper examined the measurement equivalence of the Sense of Community Index among rural, suburban, and urban neighborhoods. The findings of this paper indicated metric invariance of the SCI, and partial scalar invariance. Specifically, results indicated the presence of differential item functioning for three items. Recommendations include the exclusion of those three items in future research comparing sense of community among different neighborhood types and further examination of the measurement invariance of the SCI. The third paper examined the evidence of convergent and divergent validity of the SCI. The findings of this study indicated evidence of convergent and divergent validity for the SCI. Recommendations for future research include replication and further validation of the SCI.

Keywords: sense of community, measurement, Sense of Community Index, neighborhood effects

Table of Contents

List of Tables

List of Figures

Introduction

Sense of community is of great importance to an individual's mental health and well-being. It is particularly important in protecting against depression and loneliness (Fowler, Wareham-Fowler, & Barnes, 2013; Kitchen, Williams, & Chowhan, 2012; Li, Sun, He, & Chan, 2011; O'Brien, Hassinger, & Dershem, 1994). In addition to being important on the individual level, it is also an important neighborhood characteristic. The characteristics of the neighborhood in which one lives have powerful influence on the development and functioning of the residents. While there are many types of neighborhood characteristics (e.g. physical, social, economic, and so on) for the sake of simplicity and because of the nature of their interaction, only neighborhood social and structural characteristics will be discussed.

Neighborhood structural characteristics include such characteristics as crime rates, poverty rates, employment, and educations rates. Neighborhood social characteristics include such characteristics as social cohesion, social control, and sense of community. Both subsets have important effects on residents. For example, neighborhood effects research consistently finds that neighborhood poverty rates have significant impacts on educational outcomes (Galster et al., 2007; Sanbonmatsu, Kling, Duncan, & Brooks-Gunn, 2006) and physical and mental health outcomes (Kling, Liebman, & Katz, 2007; Leventhal & Brooks-Gunn, 2003; Sanbonmatsu, Kling, Duncan, & Brooks-Gunn, 2006). There is evidence that neighborhood social characteristics play a mediating role in the relationship between neighborhood poverty and resident outcomes. For example, Aneshensel and Sucoff (1996) found that neighborhood socioeconomic characteristics increased resident perceptions of the neighborhood as threatening, and negative resident perceptions of the neighborhood had a statistically significant negative association with resident mental health. Given the influence of neighborhood social

characteristics in general, and sense of community specifically, it is important to have instruments that elicit valid scores to measure these constructs. Unfortunately, there is a general lack of strongly validated social instruments (Sampson, 2003). Without sufficient evidence of validity, any inferences drawn from scores must be interpreted with great caution. The purpose of this dissertation was to review and evaluate the current literature regarding the validity evidence of instruments measuring neighborhood social characteristics, and then address gaps in the validation literature.

The first paper was a systematic literature review examining the current state of the neighborhood social measurement validity literature. Twenty validation studies were identified for inclusion in this review. The quantity and quality of validity evidence for these instruments was examined. The three most promising scales were the Sense of Community Index (SCI: Perkins et al., 1990), the Characteristics of the Residential Environment Scale (CRE: Handal, Barling, & Morrissy, 1981), and the Brief Sense of Community Scale (BSCS: Peterson, Speer, & McMillan, 2008). These scales showed good content validity through examination of items. They also had the most evidence of various types of validity. Despite promising findings of the validation studies conducted, there were very few forms of validity evidence collected for each instrument. The SCI had the most types of validity evidence examined, with three types of validity evidence examined. In addition to little validity evidence, there is a scarcity of validity studies for each evidence. The SCI was the most validated instrument in this review. However, the SCI still has some validity issues that need to be addressed. No study examined the measurement invariance or convergent and divergent validity of the SCI. Recommendations included further validation of instruments measuring social characteristics and of the SCI in

particular, as this is the most widely used instrument for measuring sense of community (Chavis & Pretty, 1999).

The second and third papers followed the research recommendations of the first paper. The second paper examined the measurement invariance of the SCI, and the third paper examined the evidence of convergent and divergent validity of the SCI. Data for both studies were collected online using Qualtrics. The sample was collected using snowball and convenience sampling via Facebook posts and advertisements.

The second paper examined the SCI for measurement invariance among rural, suburban, and urban neighborhoods. Metric and scalar invariance were examined using confirmatory factor analysis models. Results indicated metric invariance for the SCI and partial scalar invariance. Specifically, results indicated the presence of differential item functioning for items 4, 10, and 12 of the SCI. Recommendations for future research included dropping these items from the scale when used to compare levels of sense of community among rural, suburban, and urban neighborhoods.

The third paper examined the SCI for evidence of convergent and divergent validity. A CFA was performed to examine convergent validity of the SCI with other measures of sense of community and divergent validity with measures of neighborhood disorder. The results indicated evidence of both convergent and divergent validity of the SCI. Additionally, findings indicated that the Neighborhood Disorder Scale (NDS: Ross & Mirowsky, 1999) loaded both on the neighborhood disorder latent construct and the sense of community latent construct. Recommendations for future research include replication and further validation of the SCI as well as further examination of the NDS.

These studies will help advance the social work goals of meeting three of the Grand Challenges proposed by the American Association of Social Work and Social Welfare (www.aaswsw.org). These studies will help to address the grand challenges of ensuring healthy development for all youth by providing validity evidence for the SCI, which is used in social science research to examine the relationship between sense of community and adolescent well-being (Pretty, Andrewes, & Collett, 1994; Pretty, Conroy, Dugay, Fowler, & Williams, 1996). Strong validity evidence provides support for the claim that it is sense of community, and not some other construct, that is influencing these important outcomes. This research will aid social work researchers and practitioners toward eradicating social isolation by providing evidence for construct validity of the SCI. With evidence that the construct being measured is sense of community, researchers and practitioners can have confidence that low levels of sense of community identified with the SCI are legitimate, and improvements in SCI scores in response to interventions are, in fact, improvements in sense of community levels. Finally, these three papers will aid social work researchers and practitioners in their work to stop family violence by providing them with information regarding the validity evidence of instruments often used to measure social characteristics associated with child maltreatment, which will help them make decisions regarding the validity of findings from studies using these measures. It also advances work with this challenge by providing evidence of construct validity for scores on the SCI, supporting the use of the SCI in future studies examining the relationship between sense of community and child maltreatment.

References

Aneshensel, C. S. and Sucoff, C. A. (1996). The neighborhood context of adolescent mental health. *Journal of Health and Social Behavior, 37*(4), 293 -310.

Chavis, D. M. & Pretty, G. M. H. (1999). Sense of community: Advances in measurement and application. *Journal of Community Psychology, 27*(6), 635-642.

Fowler, Wareham-Fowler, & Barnes, 2013;

Galster, G., Marcotte, D. E., Mandell, M., Wolman, H., & Augustine, N. (2007). The influence of neighborhood poverty during childhood on fertility, education, and earnings outcomes. *Housing Studies, 22*(5), 723-751.

Handal, P. J., Barling, P. W., & Morrissy, E. (1981). Development of perceived and preferred measures of physical and social characteristics of the residential environment and their relationship to satisfaction. *Journal of Community Psychology, 9,*118-124.

Kitchen, P., Williams, A., & Chowhan, J. (2012). Sense of community belonging and health in Canada: A regional analysis. *Social Indicators Research, 107*(1), 103-126.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica, 75*(1), 83-119.

Leventhal, T. & Brooks-Gunn, J. (2003). Moving to Opportunity: An experimental study of neighborhood effects on mental health. *Research and Practice, 93*(9), 1576-1582.

Li, Y., Sun, F., He, X., & Chan, K. S. (2011). Sense of community and depressive symptoms among older earthquake survivors following the 2008 earthquake in Chengdu China. *Journal of Community Psychology, 39*(7), 776-785.

O'Brien, D. J., Hassinger, E. W., & Dershem, L. (1994). Community attachment and depression among residents in two rural Midwestern communities. *Rural Sociology, 59*(2), 255-265.

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.

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.

Ross, C. E. & Mirowsky, J. (1999). Disorder and decay: The concept and measurement of perceived neighborhood disorder. *Urban Affairs Review, 34*(3), 412-432.

Sampson, R. J. (2003). The neighborhood context of well-being. *Perspectives in Biology and Medicine, 46*(3), S53-S64. Doi: 10.1353/pbm.2003.0059.

Sanbonmatsu, L., Kling, J. R., Duncan, G. J., & Brooks-Gunn, J. (2006). Neighborhoods and academic achievement: Results from the Moving to Opportunity experiment. *The Journal of Human Resources, 41*(4), 649-691.

Chapter 1

A Review of the Neighborhood Social Measurement Validation Literature

Abstract

Neighborhood characteristics have important influence on resident outcomes including education, mental and physical health, and family functioning. While research concerning neighborhood effects consistently finds that these variables have significant impacts on resident outcomes, there remains a paucity of instruments that measure neighborhood social characteristics with acceptable psychometric properties (Sampson, 2003). The purpose of this review was to identify and examine validity studies in the current neighborhood social characteristics measurement literature. The review identified 20 validation studies examining fifteen instruments. The findings of the review indicate that while most of the instruments included in the review show promising psychometric properties, the limited amount of validation and replication studies suggests that caution should be used when making practice or policy decisions based on inferences from the scores of these instruments.

Keywords: neighborhood measurement, neighborhood social characteristics, neighborhood effects

Background

Since the late 1990s, there has been growing recognition of the importance of neighborhood effects research, especially research examining neighborhood social mechanisms (Sampson, Morenoff, & Gannon-Rowley, 2002). Neighborhood effects research consistently indicates that an individual's neighborhood has an important influence on outcomes throughout the lifespan, with impacts on various outcomes such as family functioning, education, and physical and mental health. While there are a host of neighborhood characteristics, this section will focus on structural characteristics and social characteristics because of the nature of their relationship. The following sections are organized by ecological level affected by neighborhood, beginning with the family context and moving to the individual level. This order was chosen to highlight the mediating role that neighborhood social characteristics play in the relationship between neighborhood structural characteristics and resident outcomes.

The Role of Neighborhood Structural Characteristics

Family functioning. Research suggests important connections between the neighborhood context and family functioning. The relationship between poverty and maltreatment is particularly well supported in the neighborhood effects literature (Chaffin et al., 1996; Drake & Pandey, 1996; Merritt, 2009). In a study examining the relationship between neighborhood structural factors and child maltreatment rates in three California counties, impoverishment, immigrant concentration, and access to alcohol were all significantly related to higher neighborhood rates of child maltreatment (Freisthler, 2004). Coulton and colleagues (1995) found that neighborhood impoverishment had the strongest impact on maltreating rates. These findings were supported by Drake and Pandey (1996), who found that neighborhood poverty was

strongly associated with all forms of child maltreatment and had a particularly strong association with neglect.

Education. One of the more recently studied relationships in neighborhood effects research is that between neighborhood characteristics and educational outcomes. There is increasing evidence that neighborhood characteristics, particularly economic factors, play an important role in the educational outcomes of children and adolescents. For example, teens residing in high poverty neighborhoods have been shown to have increased risk of high school drop-out, which affects later opportunities for success (Galster et al., 2007; Harding, 2003). In fact, Harding (2003) found that living in a high poverty neighborhood (greater than 20% of the residents below the federal poverty line) almost doubled the odds of dropping out of high school for non-African American youth and slightly more than doubled the odds for African-American youth as compared to youth living in low poverty (less than 10% below the poverty line) neighborhoods. These results were reinforced by Galster and colleagues (2007), who found that in addition to being associated with greater likelihood of teen parenthood and decreased likelihood of graduating from high school, neighborhood poverty, defined as the percentage of the census tract population living below the federal poverty line, was also associated with a lower rate of college graduation and lower earnings as a young adult.

Due to the difficulty in constructing true experimental designs in neighborhood effects research, most of the evidence supporting the relationship between neighborhood characteristics and educational outcomes has been correlational. However, with the Moving To Opportunity (MTO) program, researchers have begun to identify the causal effects of moving out of high poverty areas on children's educational outcomes. The MTO demonstration was a program instituted by the U. S. Department of Housing and Urban Development that investigated the

impact of neighborhood poverty on children's academic performance and behavioral problems (Sanbonmatsu, Kling, Duncan, & Brooks-Gunn, 2006). Participants in the program were divided into three groups: a control group, a Section 8 housing voucher group, and an experimental group. Participants in the control group received no housing vouchers and resided in public housing, participants in the Section 8 group received traditional Section 8 housing vouchers with no restrictions on where they could live, and participants in the experimental group received section 8 housing vouchers but were restricted to moving to neighborhoods with a poverty rate of less than 10% (Leventhal & Brooks-Gunn, 2003). In their analysis of the MTO data, Sanbonmatsu and colleagues (2006) found statistically significant neighborhood effects for African American children in the group living in neighborhoods with a poverty rate of less than 10%. The authors calculated z-scores based on the student's standardized reading and math test scores. The z-scores had a mean of 0 and a standard deviation of 1. The African American children in the experimental group showed a statistically significant improvement of eight tenths of one standard deviation in achievement test scores on reading. The mean reading z-score for the African American children in the experimental group was .084, versus -.058 for the control group and .069 for the children in the section 8 group. One explanation for this gap in achievement offered by Ainsworth (2002) is that differences in neighborhood norms regarding academic behavior, such as homework completion, affect academic achievement. This exemplifies the interactive nature of neighborhood social and structural characteristics and is discussed in greater detail later.

Physical and mental health. There is increasing evidence that the neighborhood environment has significant impacts on resident physical and mental health. As has been the case in other areas of neighborhood effects research, most of the research on the relationship between

neighborhood environment and physical and mental health focuses on the impact of neighborhoods' poverty rates. In their analysis of the MTO data, Kling, Liebman, and Katz (2007) examined the impact of neighborhood characteristics on adult economic self-sufficiency; adult mental and physical health; and adolescent mental health, education, risky behavior, and physical health. While there were no significant neighborhood effects on adult economic self-sufficiency, there was a significant negative relationship between neighborhood poverty rates and adult mental health outcomes. That is, as neighborhood poverty rates decreased, there were improvements in adult mental health outcomes. The results also indicated significant negative relationships between neighborhood poverty and physical health outcome benefits for female adolescents in the section 8 housing and experimental groups. Among female adolescents in the experimental group, there was a significant improvement in overall physical health, mental health, decreases in risky behavior, and improvements in educational outcomes. The effects were largest for mental health, followed by improved education outcomes and decreases in risky behavior (Kling et al., 2007).

In their analysis of the MTO data, Leventhal and Brooks-Gunn (2003) found that boys who moved to lower poverty neighborhoods had significantly fewer symptoms of anxiety and depression. This study also found that boys who moved to less poor neighborhoods had fewer substance addiction problems than the boys in the control group (Leventhal & Brooks-Gunn, 2003). However, Kling and colleagues (2007) found adverse effects for male youths in the section 8 and experimental (neighborhoods with less than 10% poverty rate) groups, the largest of which were for substance abuse and injuries. According to the researchers, these results were most likely due in part to the male youth continuing to have contact with adult male figures and with their original neighborhood (Kling et al., 2007). These findings supported those reported by

Sanbonmatsu and colleagues (2006), who found that behavior problems mean scores were higher for boys of all ages and ethnicities in the experimental group after moving compared to control group boys' mean scores.

The Role of Neighborhood Social Characteristics

Neighborhood social characteristics. Neighborhood social characteristics are those qualities that are associated with how the residents of a particular neighborhood interact with each other, are connected to each other, and how they psychologically experience their neighborhood. Examples of neighborhood social characteristics are cohesion, social control, social disorganization, and sense of community (*for a review see* Sampson, Morenoff, & Gannon-Rowley, 2002). These traits are contrasted with neighborhood structural characteristics such as poverty rates, crime rates, employment, and education rates (*see* Coulton et al., 2007); or physical characteristics like walkability, architectural dilapidation, and empty lots. Neighborhood structural characteristics and neighborhood social characteristics are related, however there is no overlap in their conceptual dimensions. While neighborhood social characteristics have not been as extensively studied as structural and physical characteristics, there is growing evidence that these characteristics play an important role in the relationship between neighborhood structural characteristics and resident outcomes such as education, mental health, physical health, and child maltreatment (Ainsworth, 2002; Coulton, Crampton, Irwin, Spilsbury, & Korbin, 2007; Curry, Latkin, & Davey-Rothwell, 2008; Echeverria et al., 2008).

Family functioning. The relationship between neighborhood structural characteristics and family functioning, particularly child maltreatment, has been investigated to a greater degree, and these studies have yielded stronger relationships, than research investigating the relationships between neighborhood social process and maltreatment (Coulton et al., 2007).

Findings from research examining the impact of neighborhood social characteristics are mixed, with some findings supporting and others refuting the role of social factors in predicting maltreatment rates. For instance, Garbarino and Sherman (1980) found that, for families of similar socioeconomic status, neighborhoods with higher than expected child maltreatment risk were much more characterized by social disorganization and a lack of neighborhood cohesion – findings that were later supported by Garbarino and Kostelny (1992). Their findings also indicated that neighborhoods with high risk profiles for child maltreatment were more likely to provide less adequate childcare options, less social exchange between neighbors, and were characterized by more isolation and depression (Garbarino & Sherman, 1980; Garbarino & Kostelny, 1992). In a study of households in North and South Carolina, Zolotor and Runyan (2006) found that a one point increase in household social capital (which included neighborhood cohesion and sense of community) was associated with lower odds of neglectful parenting (odds ratio:.72 ; 95% CI:.57 - .91). In a study testing a structural model for neighborhood effects on maltreatment, Kohen and colleagues (2008) found that neighborhood structural risk factors were significantly related to lower neighborhood cohesion, which was significantly and negatively associated with family functioning and maternal depression. Positive family functioning was in turn predictive of fewer punitive parenting behaviors and more consistent parenting (Kohen et al.,2008).

Education. Ainsworth (2002) examined the influence of neighborhood structural and social characteristics on the educational behaviors of a nationally representative sample of high school students. The neighborhood structural characteristics included high-status residents, residential stability, economic deprivation, and racial heterogeneity. Ainsworth (2002) proposed that the effects of these variables on education behavior were mediated by neighborhood social

characteristics (i.e. collective socialization, peer-group culture, social capital, occupational opportunity, and institutional factors). The findings indicated a greater prevalence of high status residents in the neighborhood significantly predicted more time spent doing homework and higher math and reading scores for children and adolescents, and that several neighborhood social characteristics mediated this relationship, with collective socialization being the strongest mediator of neighborhood effect on educational outcomes (Ainsworth, 2002). That is, the presence of higher status residents impacts educational outcomes through socialization of norms that encourage behavior that positively impacts homework practices and math and reading scores.

Mental health. In a study examining how neighborhood structural and racial characteristics impact adolescent mental health outcomes, Aneshensel and Sucoff (1996) found that negative resident perceptions of the neighborhood had a significant negative impact on mental health and mediated the relationship between poverty and mental health. Specifically, the findings indicated that the more threatening the neighborhood was perceived to be, the more common the symptoms of anxiety, depression, and behavioral disorders were among residents. In addition, neighborhood cohesion was significantly and negatively related to depression so that as cohesion increased, depression symptoms decreased. Neighborhood socioeconomic characteristics did not affect neighborhood cohesion, although they were associated with increased perception of the neighborhood as threatening (Aneshensel & Sucoff, 1996).

Physical health. In their analysis of cross-sectional data from the Multi-Ethnic Study of Atherosclerosis, Echeverria and colleagues (2008) examined the relationship between neighborhood problems (e.g. crime, heavy traffic, and excessive noise), neighborhood social cohesion, and various health outcomes such as smoking, drinking, and low levels of exercise. As

expected, residents of neighborhoods with fewer problems showed significantly less depression, less smoking behavior, and less heavy drinking than those in high problem neighborhoods (Echeverria et al., 2008). Residents of less cohesive neighborhoods also had higher levels of depression, were more likely to smoke, and less likely to walk for exercise.

In summary, neighborhood research consistently indicates that the neighborhood in which an individual resides has important influence on the development, functioning, and wellbeing of the residents. Structural characteristics are the most commonly examined in neighborhood effects research. Poverty in particular has been shown to have strong links to such problems as child maltreatment, increased high school dropout risk, teen parenthood, depression, and poor physical health. The findings discussed above indicate that social characteristics play important roles as mechanisms through which the effects of structural characteristics on resident outcomes are translated. In order to understand how these mechanisms or characteristics work, it is vital that neighborhood effects researchers have access to instruments measuring social characteristics that have satisfactory psychometric qualities.

Purpose

Confidence in the findings of neighborhood effects research relies heavily on the quality of the instruments used. Unfortunately, there are few instruments used to measure neighborhood social characteristics with known and accepted psychometric properties (Sampson, 2003). Without sufficient evidence of the psychometric quality of the instruments used, any results and inferences drawn from the scores on these instruments must be interpreted with caution. Furthermore, there seems to be no existing systematic review of validation studies of instruments measuring neighborhood social mechanisms. The purpose of this paper is to review and evaluate the current literature regarding the validity evidence of instruments measuring neighborhood

social characteristics. The research questions guiding this review are a) How many validation studies examining neighborhood social instruments are there? b) What is the quality of these validation studies? and c) What is the quality of the validity evidence presented for the instruments? This paper will contribute to the literature by offering a systematic review of validation studies of neighborhood social instruments. The assessment of validity evidence will be guided by contemporary unitary validity theory as articulated by Messick (1989).

Review Methodology

Inclusion and Exclusion Criteria

The studies included in this review were selected based on three specific criteria: 1) the studies must have been psychometric evaluations of measurement tools, 2) the measurement tools must have been designed to measure neighborhood social characteristics (e.g. neighborhood social cohesion, collective efficacy, and so on), and 3) the studies must have been conducted in the United States. The final criterion mitigates potential validity concerns related to cultural differences in interpretations or understanding of the constructs measured. Only studies examining the validity evidence of the measures were included. Studies in which measurement tools were utilized to investigate relationships between variables or constructs were not included, as the use of the measurement tools was a means to examining the relationships between variables and not to examine the validity characteristics of the instruments themselves. This was done with the primary purpose of identifying how many studies focused solely on validation exist for each instrument. Measurement tools designed to measure only structural neighborhood characteristics (e.g. poverty rates, crime rates, architectural dilapidation) or physical health behavior (such as walking, smoking, and so on) were not included. While these characteristics can impact or be influenced by neighborhood social characteristics, they are not in themselves

social characteristics as defined in this paper and thus were excluded from the review. International validation studies were excluded as were studies using instruments translated into different languages to limit validity issues related to translation and cultural differences in how the constructs are conceptualized. Subtle differences in word meanings or difficulty translating idioms between languages could result in major differences in how items are understood by respondents, resulting in differential item functioning or differential test functioning.

Literature Search

A systematic review was executed to identify studies of the psychometric quality of instruments measuring neighborhood social characteristics. The search was completed using electronic research data bases (*Social Work Abstracts, Sociological Abstracts, PsychInfo, Medline/PubMed, and Criminal Justice Abstracts).* The search included studies published in English between the years 1980 and 2014 using two search term strings. The first search string used was "neighbor*OR commun*, AND social". The second search string used was "(neighborhood OR community) AND (measure*OR instrument*)". This specific date range was chosen because of 1980's proximity to Bronfenbrenner's development and dissemination of his social ecological model. The end date of 2014 was chosen because this was the year in which the literature search was completed. In the advanced search settings of the databases, there is an option to include tests and measures in your study search. The search for this study was supplemented by selecting this option. Further studies were identified through scanning the references of articles identified in the electronic search. Grey literature was identified through grey literature collections and databases (*greylit.org, ifp.nyu.edu/category/grey-literature, greynet.org)* using the same search terms as above.

Selection of Studies

The author read the abstracts of studies identified by the methodology reviewed above. After reviewing the abstracts, studies that met the inclusion criteria were saved. The author then read the retained studies in full. Studies assessing the validity evidence of instruments designed to measure the social characteristics of neighborhoods were then retained and included in the review.

Results

The original search yielded 44 studies for examination. The abstracts of these studies were read, and 21 were discarded due to failure to meet the requirements for inclusion. The remaining studies were then read in full, and three more were excluded from the review for failing to meet the inclusion criteria. In total, 20 studies were identified for inclusion in the review.

Instruments

The studies are organized by instrument and discussed individually to allow for ease of evaluation of methodology and evidence for validity.

Neighborhood Cohesion Instrument. Buckner (1988) developed and tested the Neighborhood Cohesion Instrument (NCI). The NCI is a 37-item instrument used to measure the construct of neighborhood cohesion, which is a combination of psychological sense of community, resident attraction to neighborhood, and neighboring behavior. The instrument was examined for evidence of known group validity and reliability. Three neighborhoods with differing reputations for cohesion were selected for this study. These neighborhoods were mostly white and middle class, and were pre-defined with history, landmarks, names, and census neighborhood

boundaries. From these neighborhoods, 100 households were randomly selected. The instrument was delivered in person, then mailed back by the respondents. To examine for evidence of known-groups validity, Buckner (1988) ranked the mean cohesion values for each neighborhood. The rank ordering of mean values across neighborhoods was in the predicted order for all items, meaning the neighborhoods that were expected to score as more cohesive, did. A one-way ANOVA computed for the final instrument showed significant differences among the cohesion index values for all neighborhoods. The internal consistency and stability coefficients were .95.

Chicago Neighborhood and Disability Study/Baltimore Memory Study. Cagney, Glass, Skarupski, Barnes, Schwartz, and Mendes de Leon (2009), examined the evidence for validity and reliability for a measure of neighborhood cohesion and disorder that was developed using questions from the Chicago Neighborhood and Disability Study (CNDS) and the Baltimore Memory Study (BMS). The combined instrument was a 14-item survey that measured both the social and physical characteristics of the neighborhood. The social items included questions such as “Do you see neighbors and friends talking outside in the yard or on the street?”. The physical items included questions such as “Do you see poorly maintained sidewalks or broken curbs?”.

The instrument was examined for evidence of convergent validity, factorial validity, and reliability. The CNDS sample consisted of 6158 adults 65 or older who were residents of three adjacent neighborhoods in South Side Chicago. These neighborhoods were chosen for ethnic diversity. The definition of neighborhood for the CNDS was not addressed. The study states that the data came from a "geographically defined community area of three adjacent neighborhoods"(pg. 417). The data were collected through in-home interviews. The BMS sample consisted of 1140 adults age 50 to 70 from 65 adjoining Baltimore neighborhoods. The neighborhoods were chosen to ensure variability of characteristics of interest. From these

neighborhoods, six households with telephone numbers were randomly selected. The BMS used neighborhoods that were pre-defined as "neighborhood statistical areas" by the city department of planning (Schwartz et al., 2004). The data were collected through structured interviews at a study clinic.

Convergent validity was examined through correlations between neighborhood cohesion, disorder, and related variables such as resident income level and number of vacant houses. The results showed moderate negative correlations between cohesion and disorder, use of public assistance, resident low income, and high levels of vacancy. There were also moderate negative correlations between disorder and education level, housing, and high socioeconomic status. Cohesion was positively correlated to higher education levels, housing, and higher socioeconomic status, while disorder was positively correlated with use of public assistance, low income, and high levels of housing vacancy. Factorial validity was examined using confirmatory factor analysis. The results indicated that the items loaded on two latent factors, as expected. Factor one was labeled as cohesion/exchange, and Factor two was labeled social and physical disorder (CFI = .953, RFI = .968, RMSEA = .046). Reliability was examined by calculating Cronbach's α and Intra-neighborhood Correlation Coefficients (ICC) for each factor on each original scale. The CNDS cohesion scale had an estimated α of .71 and ICC of .08, while α was .91 and the ICC was .11 for the disorder scale. The BMS cohesion scale had an α of .76 and ICC of .05, while the disorder scale had an α of .74 and ICC of .33.

Characteristics of the Residential Environment Scale. Handal, Barling, & Morrissy (1981) developed and tested the Characteristics of the Residential Environment Scale (CRE), a 24-item instrument intended to measure perceived and preferred social and physical neighborhood characteristics. The data were collected from a random sample of housing units within a chosen

planned community. The authors state this community was chosen because it had a racial balance and was a cross section of income levels. From the original sample of 132 residents, 120 surveys were completed. The neighborhood sample consisted of one government subsidized housing community, thus there was no diversity of neighborhood type. A battery of survey instruments was delivered in person, completed by residents, and picked up the next day.

The analyses included intercorrelations and a stepwise multiple regression with satisfaction as the dependent variable to test for evidence of convergent validity. Preferred Physical Characteristics (PPC), Preferred Social Characteristics (PSC), & Ideal Social Characteristics (ISC) correlated moderately and significantly (p <.01) with satisfaction. Ideal Physical Characteristics (IPC) were weakly correlated to satisfaction, and the Social Discrepancy Scale (SDS) and the Physical Discrepancy Scale (PDS) were moderately negatively correlated with satisfaction. The discrepancy scores indicate the difference between the perceived scores and the ideal scores.

Morrissy and Handal (1981) also examined evidence for criterion validity and reliability of the CRE. For the reliability study, a convenience sample of 50 individuals was collected, including staff, residents, and trainees at the Chicago teaching hospital's outpatient department of psychiatry. The CRE questionnaire was submitted twice to the respondents with one-week interval between instances. The neighborhood study sample was collected using random selection of 40 individuals from two neighborhoods in Chicago. The neighborhoods were majority white neighborhoods and were defined as established, named neighborhoods. Data were collected via phone interviews.

To examine for evidence of criterion validity, a stepwise regression analysis was computed using neighborhood satisfaction as the criterion, and the six CRE subscale scores as

predictors: PSC, PPC, ISC, IPC, PDS, and the SDS. In one neighborhood, satisfaction was only significantly predicted by SDS ($R = .67$, $df = 38$, $p < .001$). In the other neighborhood, PPC, PSC, SDS, and IPC significantly predicted satisfaction ($R = .86$, $df = 38$, $p < .001$). The one-week test retest reliability of the subscales was as follows: PSC = .86, PPC = .80, ISC = .67, IPC =.78, SDS = .84, and PDS = .79.

Communities Advancing Resilience Toolkit Assessment Survey. Pfefferbaum, Pfefferbaum, Nitiema, Houston, and Van Horn (2015) examined the factorial validity and reliability of the Communities Advancing Resilience Toolkit Assessment Survey (CART), a 24-item instrument used to measure community resilience. The CART consists of five domains: Connection and Caring, Resources, Transformative Potential, Disaster Management, and Information and Communication. The sample was a purposive sample of 64 members of affiliated volunteer responder groups. The definition of neighborhood was not addressed in this study. Data were collected through a web-based survey. A Bayesian confirmatory factor analysis (BCFA) was used to assess if the factor structure represented by the data denoted the five domains of the instrument. The BCFA confirmed the five-factor structure with a posterior predictive probability (PPP) of .5, which indicates a good fit of the data to the model. The PPP is a goodness of fit statistic that examines the discrepancy between the proposed model and the data (Gelman, Menge, & Stern, 1996). All intercorrelations of the five domains were positive and statistically 219significant. The Cronbach's α statistics for all the subscales were .75 or higher.

Sense of Community Index. Chavis, Hogge, McMillan, and Wandersman (1986) developed and tested the Sense of Community Index (SCI) for evidence of content validity and convergent validity. The version of the SCI examined in this study was a 44-item instrument intended to measure four domains of sense of community: membership, integration and fulfillment of needs,

influence, and shared emotional connection (McMillan & Chavis, 1986). The data were taken from the Neighborhood Participation Project of 1979. Participants were randomly selected from interviews from this dataset. The original sample consisted of 39 blocks, yielding 1213 adult residents. One hundred residents were selected from this pool, along with 21 judges from three cities. Only one neighborhood was selected from the original data. Neighborhood was operationally defined as blocks. First, expert judges were asked to rate sense of community profiles which consisted of the 44 items from the Neighborhood Participation Project. The profiles were rated 1-5 with 5 being the highest level of sense of community. These scores were then averaged and those averages were used as predictors. All predictors correlated significantly with SCI scores. To test for convergent validity, scores from the SCI were correlated with a one item sense of community score. The correlation between the SCI scores and the sense of community score was .52.

Long and Perkins (2003) examined the factorial validity of the 12-item short form SCI (Perkins et al., 1990), developed a Brief Sense of Community Index (BSCI), and examined the new BSCI for factorial validity, construct validity, and reliability. The BSCI was developed using items from the short-form SCI by discarding poorly performing items, resulting in an 8-item survey. The original Block Booster Project telephone survey data was analyzed for this paper. The five neighborhoods were chosen based on high crime rates. Residents on each block were then randomly selected. The first wave sample included 1081 residents, and the second wave sample included 638 residents. Neighborhood was operationally defined as residential blocks. Data were collected in two waves over two years via telephone surveys of randomly selected residents of each block, a visual assessment of the neighborhood physical environment, examination of police records, and a survey of block association members.

To assess the factorial validity of the SCI, Long and Perkins (2003) conducted a confirmatory factor analysis on the original data and found that it did not support the theoretical four factor conceptualization of sense of community. The BSCI initially fared no better with poor or mixed fit on a one-dimension model and no improvement using a four-dimension structure. After replacing the poorly performing items, an exploratory factor analysis showed improved fit for a three-factor solution. The BSCI successfully discriminated between the communities. The analyses revealed positive, significant multilevel correlations between the BSCI overall score and the subscales, as well as large effect sizes at the block level. In addition, the BSCI scores were highly correlated with measures of place attachment, neighboring, informal social control, and participation in block activities. The instrument's ICCs ranged from .04 to .07.

Obst and White (2004) re-examined the evidence of factorial validity and reliability of a 10-item version of the SCI. Participants in this study were self-selected volunteers. The sample consisted of 219 university freshmen. This being the case, the authors made changes to item wording to make the items more applicable to the sample (i.e. replacing the word "neighborhood" with "university"). Two items were also removed from the original 12-item scale because they could not be adapted to fit the sample context. Data were collected through paper surveys. To examine the factorial validity evidence, a confirmatory factor analysis was conducted. The results indicated that the original four factor structure did not fit the data adequately. Obst and White (2004) then revised the four-factor structure by changing the factors on which six items loaded and conducted the CFA again. The results to this second analysis displayed adequate fit to the data with factorial invariance across community types (neighborhood, university campus, and student group). The new subscales had moderate alphas

ranging from .71 to .80, and the internal consistency coefficients for the 10 item SCI ranged from .80 to .84.

Brief Sense of Community Index – Disability (BSCI-D). Townley and Kloos (2009) examined the evidence for convergent validity, factorial validity, and reliability for a sense of community measure developed for the severely mentally ill. The instrument was developed using the BSCI (Long & Perkins, 2003), and adapted for use with individuals with severe mental illness by adding three items assessing community acceptance of disability taken from a subscale of the Neighborhood Social Climate scale (Wright & Kloos, 2007). This adaptation resulted in an 11-item scale. The BSCI-D was expected to reflect the factor structure of the original BSCI: social connections, mutual concerns, and community values. The sample used in this study was a convenience sample. Participants were recruited by case managers at the supportive housing mental health centers in which the participants lived. In total, 424 adults from 17 mental health centers participated in the study. Neighborhood was not explicitly defined by the researchers, however the respondents answered the items in reference to the supportive housing community in which they lived. The data were collected in two waves through interviews conducted by 24 graduate students and research assistants. Only the second wave data was included in this study.

To examine for evidence of convergent validity, correlations were run between subscales and measures of psychiatric distress, loneliness, neighbor relations, and adaptive functioning. The neighborhood attitudes toward mental illness subscale was significantly negatively correlated with the psychological distress and loneliness scales, and significantly positively correlated with the neighbor relations and adaptive functioning scales. The sense of community subscale was significantly correlated with all the outcome measures except for adaptive functioning. All the dimensions were moderately and significantly correlated. A confirmatory

factor analysis (CFA) was conducted to test if the instrument structure would fit theoretical three factor structure of the original BSCI. The results of the CFA indicated the four-factor model showed better fit to the data, indicating the need for a fourth factor measuring community acceptance of disability. Subscale reliabilities are as follows: social connections $\alpha = .75$, mutual concerns $\alpha = .73$, community values $\alpha = .61$, community acceptance of disability $\alpha = .83$.

Brief Sense of Community Scale. Peterson, Speer, and McMillan (2008) developed and tested a Brief Sense of Community Scale (BSCS) and examined the evidence for convergent validity, factorial validity, and reliability. The BSCS is an 8-item scale that measures the original dimensions of sense of community: membership, influence, needs fulfillment, and shared emotional connection (McMillan & Chavis, 1986). The sample was obtained through the random selection of 308 cases from larger study. The larger study used a simple random sample from all residences with phone numbers. The final sample size was 293. The individuals were diverse in terms of age, income, and education. Respondents' ethnicities only included white or Hispanic. The operational definition of neighborhood was not discussed. Data were collected via survey in face to face interviews. Partial correlations were computed to examine for convergent validity.

The BSCS correlated statistically significantly and in the expected direction with community participation, psychological empowerment, and positive mental health. It also correlated significantly and negatively with depression. A confirmatory factor analysis was then conducted to see if the scale fit the hypothesized four factor (fulfillment of needs, group membership, influence, and shared emotional connection) structure. The CFA results supported the scale's four factor first order (χ^2 difference = 25.946, GFI = .979, AGFI=.945, RMSEA = .054) and one factor second order (χ^2 difference = 28.032, GFI = .977, AGFI=.949,

RMSEA=.051) factor structure. The total scale had an estimated coefficient α of .92. The subscale estimated reliability coefficients were all high, ranging from .77 to .94.

Social Capital Scale – PHDCN. Brisson and Usher (2007) examined five bonding social capital items from the survey developed for the Project on Human Development in Chicago Neighborhoods (PHDCN). The survey measured bonding social capital by including four items measuring social cohesion and one item measuring trust of neighbors. The researchers collected probability samples of about 800 residents in each of 10 cities. For families with children, the focus child was randomly selected, but the family respondent was not randomly selected. For families with no children, the respondent was randomly chosen. The final sample consisted of 7437 residents of targeted low-income neighborhoods. Data was collected through a county wide telephone survey. Neighborhood was not explicitly operationally defined.

The items were examined for evidence of convergent validity and measurement invariance. To examine for convergent validity, the authors constructed a correlation matrix. All correlations among the positively worded items were moderate to strong and in the expected direction. The correlations between negatively worded items and positively worded items, however, were weaker. Cross tabulations revealed contradictory responses among some participants on the negatively worded items. Specifically, participants who agreed that the residents of their neighborhood "generally don't get along" also agreed that their neighborhood was "close-knit". This indicated possible validity issues, and the authors suggesting altering these items to improve validity. The authors examined two models with confirmatory factor analysis – one with uncorrelated errors between the negatively worded items, and one with correlated errors, which were used to represent erroneous responses to the negatively worded items. The results indicated that the correlated error model was a significantly better fit. Tau

equivalence indicates whether each item loads on the latent construct with equal weight (Graham, 2006). The analysis showed that tau equivalence did not exist between the effect indicators. The authors then tested for measurement invariance between the ten cities using multi-group analysis and a chi-square difference test. While measurement invariance was found between the reference city and three others, it was not found between the reference city and the remaining six cities. However, the model fit was good for the remaining cities when an unconstrained CFA of the correlated error solution was run. The coefficient alpha for the measure was .7.

General Neighborhood Social Characteristics. Henry, Gorman-Smith, Schoeny, and Tolan (2014) developed and tested the reliability and evidence for convergent and criterion validity of a system of measurement tools that evaluated various social characteristics of neighborhoods including 45 items measuring neighborhood norms, 61 items measuring informal social control, 37 items measuring social connection, 14 items measuring neighborhood resources, 6 items measuring neighborhood change, and 13 items measuring neighborhood problems. A stratified random sample was drawn using groups of eligible Chicago census tracts that had a high density of minority residents. The sample was stratified into five strata by poverty level ranging from 20 - 45% by 5% increments. A random sample of 20 individuals for each tract was collected, stratified by gender and age to ensure equal numbers of participants in each age category and gender. Age was divided into two strata, one for ages 18 – 24 and on for 30 and older. Gender was also divided into two strata, male and female. The final sample consisted of 606 individuals from 30 census tracts. Neighborhood was operationalized as census tracts. Data were collected via in-person surveys and from archival crime records.

To examine for convergent validity, correlations among the scales and subscales were computed. Most aggregate level correlations for the Norms scale were moderate and statistically significant. The correlation between Norms for Neighborhood Management and Norms regarding Crime was strong and statistically significant. However, the correlations between Child Welfare and Crime Norms, Child Welfare and Neighborhood Management, Child Management and Crime Norms, and Child Management and Neighborhood Management were not significant. All individual level correlations were statistically significant. For the Informal Social Control scale, all but four aggregate correlations were statistically significant. The aggregate level correlations between Child Welfare and Neighborhood Management, Child Management and Neighborhood Management, Child Welfare and Neighborhood Organization, and between Neighborhood Organization and Neighborhood Management were all nonsignificant. All individual level correlations for the Informal Social Control Scale were significant. Social Cohesion and Social Resources had a non-significant correlation of almost zero. Correlations were not computed for the Perceived Neighborhood Changes and Neighborhood Resources scales because they did not have any subscales. Correlations were also not computed for the Neighborhood Problems scale because it was a checklist.

To examine for criterion validity, archival crime records were regressed on the subscales. Subscales measuring Norms regarding Adolescent Behavior significantly predicted police reports of property crime and neighborhood violent crime, though only marginally with the latter. All the subscales measuring Informal Social Control significantly predicted violent crime and property crime, but only half of them significantly predicted drug crime. For the subscales measuring Social Connections, the coefficients were significant and negative between Social Cohesion and violent and drug crime. The Neighborhood Resources scale only significantly

predicted violent crime ($p < .01$). The Neighborhood Change scale showed a strong and significant association with all measures of crime; however, the Neighborhood Problems scale showed no significant associations with crime.

Estimated reliability coefficients for each of the subscales were examined by calculating Interclass Correlations (ICC) to assess individual internal consistency and using a generalizability theory (G theory) method of neighborhood variance attribution. No ICC was calculated for the social resources subscale of the neighborhood social connection scale, the neighborhood resources scale, or the neighborhood problems scale. All subscales with calculated ICCs showed high levels of individual level reliability, with ICCs above .7. The results of the G-theory analysis indicated that almost all the subscales had acceptable levels of statistically significant shared variance. The exceptions were the child welfare and adolescent behavior subscales from the Norms scale, for which the shared variance was not significantly different from zero, and the citizen responsibility and neighborhood organization subscales from the Informal Social Control scale.

Measure of Neighborhood Social Context. Coulton, Korbin, and Su (1996) developed and tested a measure of neighborhood context designed to represent both structural and social characteristics. The dimensions represented by the 116 items included resource and service availability, participation in neighborhood activities, interactions with neighbors, intervention behaviors with neighborhood children, neighborhood quality, neighborhood stability, neighborhood change, neighborhood disorder, and neighborhood identity. The researchers collected a purposive sample of 16 block groups, which were chosen based on extremity of risk, eight high risk and eight low risk, of child maltreatment. Risk was established through maltreatment rates. Ten households were then randomly selected from each block group. A total

sample of 156 caregivers was then obtained. Neighborhood was operationalized as block groups. Data were collected via in person questionnaires.

The measure was tested for reliability and evidence of known group validity. The instruments were tested for reliability at the individual level and at the aggregate level. Individual level reliability was estimated by calculating coefficient alpha, and aggregate level reliability was calculated using a G-theory model. Individual reliability estimates ranged from $\alpha = .46$ (intervening with neighborhood children) to $\alpha = .95$ (neighborhood disorder). The measures with the greatest aggregate level reliability were disorder ($\alpha = .94$), mobility ($\alpha = .83$), identity (.74), neighborhood quality (.76), and retaliate (.71). Items measuring neighborhood interaction ($\alpha = -0.46$), stop misbehavior ($\alpha = -1.33$), assist ($\alpha = .02$), and intervening with children ($\alpha = .17$) had little consensus and very low reliability coefficients. Coulton, Korbin, and Su (1996) used two-way ANOVA using neighborhood risk level and predominant racial group as factors to examine the evidence for known groups validity on the eleven scales that had aggregate reliability coefficients greater than .4. Most of the included scales successfully discriminated between high and low risk neighborhoods. The two scales that did not discriminate neighborhood based on risk level (positive neighborhood change and facility quality) discriminated between European American and African American neighborhoods.

Community Safety Scale. Shoffner and Vacc (2002) developed and tested the factor structure and reliability of the 15-item Community Safety Scale (CSS). The CSS was intended to measure community safety through identifying social and physical incivilities. The sample was taken from one urban high school in the southeast and included 1067 high school students. Neighborhood was not explicitly operationalized, but rather left to the students to define for themselves. The students completed the survey during English class. Experts evaluated the items

for content validity. There was 100% agreement among the experts on the inclusion of all items except for item 2. An exploratory factor analysis identified two factors (social incivility, lack of physical incivility) that accounted for 41% of the variance. The overall reliability coefficient was acceptable at $\alpha = .84$. The social incivility subscale was acceptable at $\alpha = .84$, but the physical incivility subscale was lower at $\alpha = .58$.

Neighborhood Disorder. Ross and Mirowsky (1999) developed and assessed a scale meant to measure neighborhood disorder for factorial validity and reliability. The authors used secondary data from the 1995 Survey of Community, Crime, and Health. The neighborhood disorder scale consists of 15 items intended to measure four dimensions: physical disorder, physical order, social disorder, and social order. The sample was a probability sample of Illinois households obtained using random digit dialing. The final sample consisted of 2482 adults. The sample was collected from a variety of neighborhood types including neighborhoods from big cities, suburbs, small cities, small towns, and rural areas. The operational definition of neighborhood, however, was not discussed. Data was collected via surveys. An exploratory factor analysis showed two distinct factors, order and disorder. The structural model, however, yielded disorder and decay as distinct, though highly related, factors. The items representing social characteristics (informal social control, etc.) loaded on the disorder factor, however there was a good deal of overlap between social and physical disorder. The total scale reliability coefficient was $\alpha = .92$.

Neighborhood Youth Support. Hohl (2013) developed and tested the psychometric properties of a measure of neighborhood youth support. The instrument consisted of 25-items organized into five subscales: leadership, respect, active engagement in positive behavior, adult presence, and intervening in negative behavior. The author obtained a convenience sample of 65 adults from one neighborhood in west Philadelphia. Neighborhood was defined as established

neighborhoods with names. Data was collected via surveys distributed at block parties. Exploratory factor analysis using principal axis factor was conducted for each of the five subscales to see if they each represented one factor. Results on the initial scale identified one factor for leadership, respect, and active engagement, but two factors each for intervening in negative behavior and adult presence. One item was removed from the scale, the analysis was performed again, and results identified only one factor for each of the scales.

Evidence for convergent validity was examined by calculating Spearman's rank order correlations between previously validated scores from the same scale and new factor scale scores. Leadership, respect, adult presence, and active engagement were expected to be positively correlated with social cohesion and trust. Informal social control was expected to positively correlate with intervening. Contrary to expectations, leadership was not significantly correlated with social cohesion. All other correlations were as expected. The total scale estimated reliability coefficient was $\alpha = .85$. Subscale reliability coefficients were acceptable and ranged from .63 to .78.

City Stress Inventory. Ewart and Suchday (2002) developed and tested an 18-item measure of neighborhood disorder and exposure to violence. The City Stress Inventory (CSI) was examined for reliability and evidence of convergent validity. The authors collected a convenience sample of 212 9th grade students sampled from two adjacent magnet high schools. The authors argue that since the schools were magnet schools, they attracted students from all areas and social strata of the city. However, the sample was African American and white and included no other ethnicities. Neighborhood was not explicitly operationalized, instead it was left up to student to define. Later, the students wrote down the address where they spent most nights per week and those addresses were mapped onto census block groups. This approach could potentially be

problematic. Since neighborhood wasn't explicitly defined for the participants, the area about which they responded to the items may not coincide with the delineations of the census block into which they were categorized. Data were collected via questionnaires administered at school, and observations of the debates were conducted at school. This process was repeated after 12 months. Neighborhood socioeconomic status and other structural characteristics were gathered from census bureau data.

Cronbach's α coefficients were calculated to examine the internal consistency of the subscales. The neighborhood disorder subscale had an α = .88, and the exposure to violence subscale had an α = .85. Test-retest reliability was examined using Pearson's *r*. The neighborhood disorder subscale had a one-year test-retest coefficient of *r* = .82, and the exposure to violence subscale had a coefficient of *r* = .75. The evidence for convergent validity consisted of correlations between the two subscales and four major structural indices of economic and educational disadvantage (per capita income, unemployment rate, percent of population born to unmarried mothers, and education level). The subscales were moderately and significantly correlated to the structural indices in the expected directions. The subscales were also significantly correlated to mood hostility and dejection while anticipating a peer encounter.

Community Mobilization. Cheadle, Wagner, Anderman, Walls, McBride, Bell, Catalano, & Pettigrew (1998) examined evidence for convergent validity and reliability on a 14-item measure of community mobilization used by the Seattle Minority Youth Health Project. The sample for this study was a purposive sample of 13 neighborhoods chosen based on the size of their youth of color population. The authors then used snowball sampling of individual neighborhood informants. The final survey consisted of 1202 youth, 668 parents, and 108 Community Mobilization leaders. The operational definition of neighborhood was not addressed. Data were

collected through school-based surveys of youth, a neighborhood survey of parents, and a telephone community mobilization survey of the neighborhood leaders and activists. The authors then conducted semi-structured interviews with four Seattle Department of Neighborhoods employees.

Evidence for convergent validity was examined through calculating correlations among several items measuring community mobilization from surveys of youth, parents, and neighborhood leaders (Cheadle et al., 1998). The correlations between the community mobilization survey and youth perception survey were low to moderate (.11, .66), as were the correlations between the community mobilization survey and the parent survey for neighborhood cooperation (-.09,.55). Only around 25% of these correlations were statistically significant. The correlations between the community mobilization survey and the youth survey for neighborhood pride were quite low (-.04, -.39), as were the correlations for the community mobilization survey and parent survey for pride (-.11, -.47). Only one correlation between the community mobilization survey and the parent neighborhood pride survey was statistically significant. Reliability evidence was collected by calculating interclass correlations (ICC). Most of the ICCs calculated fell at .05 or less, with only one > .1.

Discussion

The goal of this review was to identify instruments that have been used to measure neighborhood social characteristics and examine the extent of their evidence for validity. In a literature search, no systematic review of the neighborhood social measurement literature was identified. This paper contributes to the literature by examining and synthesizing the findings of the neighborhood social measurement validation literature.

Twenty validation studies were identified for inclusion in this review. The methodologies and validity evidence for each of the studies were discussed individually. This section will focus on the general state of the validity evidence for neighborhood social measurement and suggest recommendations for strengthening the neighborhood social characteristics measurement literature.

The primary problem identified by the results of this review is the overall lack of validity evidence in the neighborhood social characteristics measurement literature. According to *The Standards for Educational and Psychological Measurement* (AERA et al., 2014), not all forms of validity evidence are necessary for every instrument. However, depending on the intended use of the instrument, there are facets of validity evidence that are necessary to examine before inferences made based on instrument scores can be said to be valid.

No study included in this review examined the predictive validity, concurrent validity, or divergent validity of the instruments. Convergent and factorial validity were the most common types of evidence collected with 12 (Brisson & Usher, 2007; Cagney et al., 2009; Cheadle et al., 1998; Ewart & Suchday, 2002; Handal et al., 1981; Henry et al., 2014; Hohl, 2013; Long & Perkins, 2003; Peterson et al., 2008; Townley & Kloos, 2009) and 10 (Cagney et al., 2009; Chavis et al., 1986; Hohl, 2013; Long & Perkins, 2003; Obst & White, 2004; Peterson et al., 2008; Pfefferbaum et al., 2015; Ross & Mirowski, 1999; Shoffner & Vacc, 1999; Townley & Kloos, 2009) studies respectively examining these types of evidence. Only two studies examined instruments for known-groups validity (Buckner, 1988; Coulton et al., 1996). Ten instruments included in this review only had one type of validity evidence besides reliability coefficients (Buckner, 1988; Cheadle et al., 1998; Coulton et al., 1996; Ewart & Suchday, 2002; Handal et al., 1981; Morrissy & Handal, 1981; Obst & White, 2004; Pfefferbaum et al., 2015; Ross &

Mirowski, 1999; Shoffner & Vacc, 2002), and only one instrument was examined in more than one study that examined additional types of validity evidence (Handal et al., 1981; Handal & Morrissy, 1981). With so few types of validity examined for each instrument, researchers run the risk of using the scores of an instrument to make inferences for which it has not been validated. For example, the Neighborhood Cohesion Instrument (Buckner, 1988) has only been examined for known-groups validity. Notwithstanding, it has been used to examine the association between neighborhood psychosocial characteristics and smoking (Rachele, Wood, Nathan, Giskes, & Turrell, 2016), the role of neighborhood characteristics on aggressive behavior among low-income African American youth (Romero, Richards, Harrison, Garbarino, & Mozley, 2015), and to examine how social support and sense of community impact rural Australian men's self-reported wellbeing (Kutek, Turnbull, & Fairweather-Schmidt, 2011).

Another problem that was identified was the scarcity of validity studies for each instrument. In general, only one validity study was identified for most instruments in this review. The exceptions were the SCI (Chavis et al., 1986), which was examined in three studies (Chavis et al., 1986; Long & Perkins, 2003; Obst & White, 2004), and the CRE (Handal et al., 1981), which was examined in two studies (Handal et al., 1981; Handal & Morrissy, 1981). Validation is an ongoing process requiring replication and extension of evidence. More than one study is necessary in the validation of inferences made from scores on instruments, both to support findings of previous studies and to expand on them. In the case of neighborhood social characteristics measurement, it is especially important to replicate validity findings using different types of neighborhoods and populations because these measures are so often used to understand the impact of different types of neighborhoods on resident outcomes.

There is some difficulty in discussing quality of the instruments included in this review as there have been so few validity studies and so little replication of what validity studies have been conducted. However, what evidence there is suggests that the neighborhood social instruments examined in this review have promising psychometric characteristics. In general, reliability coefficients for these instruments tend to be high to moderate. There are two exceptions, however. The Community Mobilization scale (Cheadle et al., 1998) had a very low interrater reliability with most of the ICCs falling below .05. The physical incivilities subscale of the Community Safety Scale (Shoffner & Vacc, 2002) also had a low reliability level, with an alpha of .58.

In addition to generally acceptable reliability levels, the instruments tend to show promising evidence for the types of validity for which they are examined. A possible exception to this is the Neighborhood Disorder scale (Ross & Mirowsky, 1999), which exhibited low to moderate correlations between the items measuring neighborhood pride and feelings about the neighborhood. Only one of those correlations was statistically significant.

As discussed above, the SCI was the most studied instrument included in this review. The SCI is also the most widely used instrument for measuring sense of community (Chavis & Pretty, 1999; Chipuer & Pretty, 1999). However, as demonstrated in this review, there are still validity and measurement issues to be addressed. For example, the SCI failed to exhibit satisfactory evidence for factorial validity. Both studies examining the factorial validity evidence of the SCI found that the data from the original instrument did not fit the four-factor structure proposed by McMillan and Chavis (1986) (Long & Perkins, 2003; Obst & White, 2004). Each study addressed this differently. While Long and Perkins (2003) revised the instrument to create the BSCI, Obst and White (2004) changed the factors on which certain items loaded and achieved

adequate fit to the four-factor model. Since the SCI is the most widely used sense of community instrument, it is vital that it have a strong foundation of validity evidence to support its use.

Limitations

The first limitation of this paper is that a single author conducted this literature review. While efforts were made to avoid bias, the lack of multiple authors to ensure interrater reliability may have resulted in the presence of bias. A second limitation is the exclusion of international validation studies and studies using non-English versions of the instruments. While this was done to mitigate complicating validity issues of language and culture, it may have resulted in the exclusion of replication studies, additional validation studies of instruments presented in this review, or validation studies of instruments otherwise not included in this review. This may have suppressed the quantity of validity evidence presented for some instruments.

Conclusions and Recommendations for Future Research

To address the issue of insufficient or fragmented validity evidence, researchers should view instruments and validation holistically when conducting validation studies. This requires considering not only the theoretical underpinnings of the instrument, but also the intended use of the instrument's scores. For example, a multidimensional measure of neighborhood cohesion used to examine the effect of neighborhood on resident outcomes should not only be examined for evidence of factorial validity, but also for evidence of content, predictive, convergent, divergent, and known groups validity. This holistic approach to validation ensures that the validity evidence of instruments is not piecemeal and adequately supports the instrument's use. This does not necessarily mean that all forms of validity evidence must be collected in one study, as this could present significant challenges in cost and feasibility. However, it does require the

examination of more than one type of validity evidence for an instrument so as to ensure that the instrument has been validated for all of its intended uses.

It is also imperative that validation of instruments measuring neighborhood social characteristics be a continuous process that includes replication. All but two of the instruments identified in this review were evaluated in only one validation study. To be able to support claims that valid inferences can be drawn from the scores of an instrument requires more than a single study. Furthermore, claims made based on results of research using poorly validated instruments can easily be called into question due to the lack of support for using these instruments. The instruments identified in this review should be used with caution and with the understanding that there is a paucity of validity evidence for many common uses of the instruments.

Confidence in the accuracy of neighborhood effects research requires a solid foundation of measures with robust evidence of validity. The present state of evidence suggests the need for the continued validation of instruments measuring neighborhood social characteristics. The SCI (Chavis et al., 1986) emerged in this review as both the most examined instrument, and the instrument with the most facets of validity evidence presented. Notwithstanding, the SCI still was only examined for three facets of validity evidence (i.e., content, convergent, and factorial) and reliability estimates, and showed problematic results regarding factorial validity evidence. In addition, the evidence for convergent validity was weak, as it consisted of total scores correlated with a one item sense of community score. Given the instrument's widespread use, it is recommended that the SCI (Chavis et al., 1986) be examined further for various types of validity evidence, specifically for evidence of convergent validity.

References

Ainsworth, J.W. (2002). Why does it take a village? The mediation of neighborhood effects on educational achievement. *Social Forces, 81*(1), 117-152.

American Education Research Association, American Psychological Association, and the National Council on Measurement in Education. (2014). The Standards for Education and Psychological Testing. Washington, D.C.: AERA

Aneshensel, C.S. & Succoff, C.A. (1996). The neighborhood context of adolescent mental health. *Journal of Health and Social Behavior, 37*(4), 293-310.

Brisson, D. & Usher, C. L. (2007). Conceptualizing and measuring bonding social capital in low-income neighborhoods. *Journal of Social Service Research, 34*(1), 1-11, DOI:10.1300/J079v34n01_01

Bronfenbrenner, U. (1977). Toward an experimental ecology of human ecology. *American Psychologist, July,* 513-531.

Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental Psychology, 22*(6), 723-742.

Bronfenbrenner, U. (1993). The ecology of cognitive development: Research models and fugitive findings. In R.H. Wozniak & K. Fischer (Eds.). *Scientific Environments* (pp. 3 – 44). Hillsdale, NJ: Erlbaum.

Bronfenbrenner, U. & Morris, P.A. (2006). The bioecological model of human development. In R. M. Lerner (Ed), *Handbook of child psychology* (6th ed.). (pp. 793-828). Hoboken, NJ: Wiley.

Buckner, J. (1988). The development of an instrument to measure neighborhood cohesion. *American Journal of Neighborhood Psychology, 16*(6), 771 – 791.

Cagney, K.A., Glass, T.A., Skarupski, K.A., Barnes, L.L., Schwartz, B.S., & Mendes de Leon, C.F. (2009). Neighborhood-level cohesion and disorder: Measurement and validation in two older adult urban populations. *Journal of Gerontology: Social Sciences, 64B*(3), 415-423.

Chaffin, M., Kelleher, K., & Hollenberg, J. (1996). Onset of physical abuse and neglect: Psychiatric, substance abuse, and social risk factors from a prospective community data. *Child Abuse & Neglect, 20*(3), 191-203.

Chaskin, R.J. (1995). *Defining neighborhood: History, theory, and practice*. Chapin Hall Center for Children at the University of Chicago.

Chavis, D.M., Hogge, J.H., McMillan, D.W., & Wandersman, A. (1986). Sense of community through Brunswik's lens: A first look. *Journal of Community Psychology, 14*(1), 24-40.

Chavis, D.M. & Pretty, G.M.H. (1999). Sense of community: Advances in measurement and application. *Journal of Community Psychology, 27*(6), 635-642.

Cheadle, A., Wagner, E., Anderman, C., Walls, M., McBride, C., Bell, M.A., Catalano, R.F., Pettigrew, E. (1998). Measuring community mobilization in the Seattle Minority Youth Health Project. *Evaluation Review, 22*(6), 699-716.

Coulton, C.J., Crampton, D.S., Irwin, M., Spilsbury, J.C., & Korbin, J.E. (2007). How neighborhoods influence child maltreatment: A review of the literature and alternative pathways. *Child Abuse & Neglect, 31*,1117-142.

Coulton, C.J., Korbin, J.E., & Su, M. (1996). Measuring neighborhood context for young children in an urban area. *American Journal of Community Psychology, 24*(1), 5-32.

Coulton, C.J., Korbin, J.E., & Su, M. (1999). Neighborhoods and child maltreatment: A multi-level study. *Child Abuse & Neglect, 23*(11), 1019-1040.

Coulton, C.J., Korbin, J.E., Su, M., & Chow, J. (1995). Community level factors and child maltreatment rates. *Child Development, 66* (5), 1262-1276.

Curry, A., Latkin, C., & Davey-Rothwell, M. (2008). Pathways to depression: The impact of neighborhood violent crime on inner-city residents in Baltimore, Maryland, USA. *Social Science & Medicine, 67,* 23-30.

Drake, B. & Pandey, S. (1996). Understanding the relationship between neighborhood poverty and specific types of child maltreatment. *Child Abuse & Neglect, 20* (11), 1003-1018.

Echeverria, S., Diez-Roux, A.V., Shea, S., Borrell, L.N., & Jackson, S. (2008). Associations of neighborhood problems and neighborhood social cohesion with mental health and health behaviors: The Multi-Ethnic Study of Atherosclerosis. *Health & Place, 14,* 853-865.

Ewart, C.K. & Suchday, S. (2002). How urban poverty and violence affect health: Development and validation of a neighborhood stress index. *Health Psychology, 21*(3), 254-262.

Freisthler, B. (2004). A spatial analysis of social disorganization, alcohol access, and rates of child maltreatment in neighborhoods. *Children and Youth Services Review, 26,* 803-819.

Galster, G., Marcotte, D.E., Mandell, M., Wolman, H., & Augustine, N. (2007). The influence of neighborhood poverty during childhood on fertility, education, and earnings outcomes. *Housing Studies, 22*(5), 723-751.

Garbarino, J. & Kostelny, K. (1992). Child maltreatment as a community problem. *Child Abuse & Neglect, 16,* 455-464.

Garbarino, J. & Sherman, D. (1980). High-risk neighborhoods and high-risk families: The human ecology of child maltreatment. *Child Development, 51* (1). 188-198.

Gelman, A., Meng, X.L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica, 6*(4), 733-807.

Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement, 66*(6), 930-944.

Handal, P.J., Barling, P.W., & Morrissy, E. (1981). Development of perceived and preferred measures of physical and social characteristics of the residential environment and their relationship to satisfaction. *Journal of Community Psychology, 9,* 118-124.

Harding, D.J. (2003). Counterfactual models of neighborhood effects: The effect of neighborhood poverty on dropping out and teenage pregnancy. *American Journal of Sociology, 109*(3), 676-719.

Henry, D., Gorman-Smith, D., Schoeny, M., & Tolan, P. (2014). "Neighborhood Matters": Assessment of neighborhood social processes. *American Journal of Community Psychology, 54,* 187-204.

Hohl, B.C. (2013). *Working with the community to develop a new measure of neighborhood youth support.* (Unpublished doctoral dissertation). Temple University, Philadelphia, Pennsylvania.

Kling, J.R., Liebman, J.B., & Katz, L.F. (2006). Experimental analysis of neighborhood effects. *Econometrica, 75*(1), 83-119.

Kohen, D. E., Leventhal, R., Dahinten, V.S., & McIntosh, C.N. (2008). Neighborhood disadvantage: Pathways of effects for young children. *Child Development, 79* (1). 156-169.

Kutek, S.M., Turnbull, D., & Fairweather-Schmidt, A.K. (2011). Rural men's subjective well-being and the role of social support and sense of community: Evidence for the potential benefit of enhancing informal networks. *Australian Journal of Rural Health, 19,* 20-26.

Leventhal, T. & Brooks-Gunn, J. (2003). Moving to Opportunity: An experimental study of neighborhood effects on mental health. *American Journal of Public Health, 93*(9), 1576-1582.

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.

Merritt, D.H. (2009). Child abuse potential: Correlates with child maltreatment rates and structural measures of neighborhoods. *Children and Youth Services Review, 31,* 927-934.

Morrissy, E. & Handal, P.J. (1981). Characteristics of the Residential Environment Scale: Reliability and differential relationship to neighborhood satisfaction in divergent neighborhoods. *Journal of Community Psychology, 9,* 125-132.

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.

Omoto, A.M., & Packard, C.D. (2016). The power of connections: Psychological sense of community as a predictor of volunteerism. *The Journal of Social Psychology, 156*(3), 272 – 290.

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.

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.

Pfefferbaum, R.L., Pfefferbaum, B., Nitiema, P., Houston, J.B., & Van Horn, R.L. (2015). Assessing community resilience: An application of the expanded CART survey instrument with affiliated volunteer responders. *Community Resilience Assessment and Intervention, 29*(2), 181-199.

Pretty, G. M.H., Conroy, C., Dugay, J., Fowler, K., & Williams, D. (1996). Sense of community and its relevance to adolescents of all ages. *Journal of Community Psychology, 24*(4), 365-379.

Rachele, J.N., Wood, L., Nathan, A., Giskes, K., & Turrell, G. (2016). Neighbourhood disadvantage and smoking: Examining the role of neighbourhood-level psychosocial characteristics. *Health and Place, 40,* 98-105.

Riva, M., Gauvin, L., & Barnett, T.A. (1998). Toward the next generation of research into small area effects on health: A synthesis of multilevel investigations published since July 1998. *Journal of Epidemiology and Community Health, 61*(10), 853-861.

Romero, E., Richards, M.H., Harrison, P.R., Garbarino, J., & Mozley, M. (2015). The role of neighborhood in the development of aggression in urban African American youth: A multilevel analysis. *American Journal of Community Psychology, 56,* 156-169.

Roosa, M.S., Deng, S., Ryu, E., Burrell, G.L., Tein, J.Y., Jones, S. et al. (2005). Family and child characteristics linking neighborhood context and child externalizing behavior. *Journal of Marriage & Family, 67,* 515-529.

Ross, C.E. & Mirowsky, J. (1999). Disorder and decay: The concept and measurement of perceived neighborhood disorder. *Urban Affairs Review, 34*(3), 412-432.

Sampson, R.J. (2003). The neighborhood context of well-being. *Perspectives in Biology and Medicine, 46*(3), S53-S64.

Sampson, R.J. (2002). Assessing "neighborhood effects": Social processes and new directions in research. *Annual Review of Sociology, 28,* 443-478. Doi: 10.1146/annurev.soc.28.110601.141114

Sanbonmatsu, L., Kling, J.R., Duncan, G.J., & Brooks-Gunn, J. (2006). Neighborhoods and academic achievement: Results from the Moving to Opportunity Experiment. *The Journal of Human Resources, 41*(4), 649-691.

Schwartz, B.S., Glass, T.A., Bolla, K.I., Stewart, W.F., Glass, G., Rasmussen, M., Bressler, J., Shi, W., & Bandeen-Roche, K. (2004). Disparities in cognitive functioning by

race/ethnicity in the Baltimore Memory Study. *Environmental Health Perspectives, 112,* 314-320.

Shoffner, M.F. & Vacc, N.A. (2002). An analysis of the Community Safety Scale: A brief report. *Measurement and Evaluation in Counseling and Development, 35*,49-55.

Townley, G. & Kloos, B. (2009). Development of a measure of Sense of Community for individuals with serious mental illness residing in community settings. *Journal of Community Psychology, 37*(3), 362-380.

Wright, P.A. & Kloos, B. (2007). Housing environment and mental health outcomes: A levels of analysis perspective. Journal of Environmental Psychology, 27(1), 79-89.

Zaff, J., & Devlin, A.S. (1998). Sense of community in housing for the elderly. *Journal of Community Psychology, 26*(4), 381-398.

Zolotor, A.J., & Runyan, D.K. (2006). Social capital, family violence, and neglect. *Pediatrics, 117*(6), 1124-1131.

Chapter 2

Measurement invariance of the Sense of Community Index

Abstract

Neighborhood effects research indicates that neighborhood social characteristics play an important role in the functioning of residents (Curry, Latkin, & Davey-Rothwell, 2008; Echeverria et al., 2008; Kohen et al., 2008). Sense of community in particular has important effects on resident mental well-being (Fowler, Wareham-Fowler, & Barnes, 2013; Greenfield & Marks, 2010; Kitchen, Williams, & Chowhan, 2012; Li, Sun, He, & Chan, 2011; Prezza, Amici, Roberti, & Tedeschi, 2001). The Sense of Community Index (SCI) (Perkins et al., 1990) is the most commonly used instrument for measuring sense of community, particularly to compare rural, suburban, and urban neighborhoods (Obst, Smith, & Zinkiewicz, 2002; Kitchen, Williams, & Chowhan, 2012). However, evidence of measurement invariance of the SCI among these different types of neighborhoods has not been examined. Without evidence of measurement equivalence among these different neighborhood types, researchers will not know whether any differences identified between these neighborhoods are true differences or artifacts of measurement bias. This study examined the metric and scalar invariance of the SCI. Results indicated metric invariance and partial scalar invariance of the SCI. Recommendations were made regarding future use of the SCI in comparing neighborhoods of different area types.

Keywords: neighborhood effects, measurement invariance, Sense of Community Index, neighborhood measurement

Background

Findings from neighborhood effects research indicate that the social characteristics of a neighborhood play an important protective role in the relationship between economic characteristics and resident outcomes ranging from family functioning to health behaviors (Curry, Latkin, & Davey-Rothwell, 2008; Echeverria et al., 2008; Kohen et al., 2008). For example, Kohen and colleagues (2008) found that neighborhood cohesion mediated the relationship between neighborhood structural characteristics such as economic disadvantage and family functioning. Specifically, they found that economic disadvantage was significantly related to lower neighborhood cohesion, which in turn was significantly and negatively related to family functioning (Kohen et al., 2008). In addition, Echeverria and colleagues (2008) found that residents of less socially cohesive neighborhoods were more likely to have symptoms of depression, to smoke, and to not walk for exercise.

Neighborhood social characteristics are those features that are associated with how the residents of a neighborhood are socially connected to each other and their psychological or emotional experience of their neighborhood. Social cohesion, collective efficacy, and sense of community are examples of neighborhood social characteristics (*for a review see* Sampson, Morenoff, & Gannon-Rowley, 2002). In contrast, neighborhood structural characteristics have to do with economic and institutional qualities such as poverty, crime, employment, and education rates (*see* Coulton et al., 2007).

One of the more commonly studied neighborhood social characteristics is neighborhood sense of community. Sense of community is related, though distinct from, social cohesion and collective efficacy. In fact, the relationship is a hierarchical relationship where sense of community and collective efficacy are each dimensions of social cohesion (Buckner, 1988).

Neighborhood sense of community has been found to have protective effects against such negative outcomes as loneliness (Prezza, Amici, Roberti, & Tedeschi, 2001) and depression (Fowler, Wareham-Fowler, & Barnes, 2013; Kitchen, Williams, & Chowhan, 2012; Li, Sun, He, & Chan, 2011; O'Brien, Hassinger, & Dershem, 1994), and emerging research indicates that it may protect against the long term psychological effects of childhood physical and psychological abuse (Greenfield & Marks, 2010).

Sarason (1974) defined sense of community as "…[T]he sense that one was a part of a readily available, mutually supportive network of relationships upon which one could depend…" (pg. 1). Later, McMillan & Chavis (1986) proposed that sense of community consisted of four factors – membership, influence, integration and fulfillment of needs, and shared emotional connection. *Membership* is the feeling that one belongs to the group or is part of the community. *Influence* is the idea of mutual influence between the group and its members. *Integration and fulfillment of needs* represents the idea that being a member of the community is rewarding for the members. *Shared emotional connection* defines the emotional connection between group members based on shared history and identity.

The Sense of Community Index (SCI: Perkins et al., 1990) has been identified as the most widely used instrument for measuring sense of community (Chavis & Pretty, 1999). Originally developed by Chavis and colleagues (1986), the SCI was based on 44 items from the Neighborhood Participation Project survey that had been grouped into four categories based on McMillan and Chavis' (1986) sense of community framework. Perkins and colleagues (1990) later developed the 12-item short form of the SCI to measure sense of community based on the original SCI in their study of block association participation.

The SCI has been used to compare sense of community in rural, suburban, and urban neighborhoods (Obst, Smith, & Zinkiewicz, 2002; Kitchen, Williams, & Chowhan, 2012). Understanding the difference in sense of community levels between different types of neighborhoods is important because research indicates that sense of community may mediate the effect of neighborhood disorder on resident outcomes (Cantillon, Davidson, & Schweitzer, 2003). Low income urban neighborhoods have been found to have high levels of neighborhood disorder (Skogan, 1990) but are proposed to have lower levels of sense of community than other types of neighborhoods to act as a buffer between neighborhood risk and resident outcomes (Obst, Smith, & Zinkiewicz, 2002; Kitchen, Williams, & Chowhan, 2012). Interventions aimed at improving sense of community levels in at-risk urban neighborhoods may allow practitioners to improve resident outcomes despite other risk factors (Brodsky, O'Campo, & Aronson, 1999; Wandersman & Nation, 1998). However, without evidence of measurement invariance researchers will not know whether any differences in sense of community levels between neighborhoods are real, or an artifact of measurement invariance (Van de Schoot, Lugtig, & Hox, 2012).

Despite these important implications, measurement invariance, also known as measurement equivalence (Vanderberg & Lance, 2000) of the SCI with regard to neighborhood type has never been tested. Measurement equivalence or invariance means that two individuals with the same level of the construct of interest are expected to have the same observed score regardless of their group membership (Vandenberg & Lance, 2000).

In formal terms, let E be the expected score, τ_i be the true score of sense of community for person i, Y_i be the observed score on the SCI for person i, and P_1 and P_2 be two different populations of persons. For example, P1 could represent a population of urban residents and P2

could represent a population of rural residents. Then, measurement equivalence can be defined as,

$$E(Y_i|\tau_i, P_1) = E(Y_i|\tau_i, P_2) = E(Y_i|\tau_i).$$

That is, an individual's expected observed score on the SCI depends only on the true level of the person's sense of community, and not on the type of community in which he or she lives (i.e. rural, urban, suburban). Evidence of measurement invariance is important because without it researchers cannot be certain whether the differences found between groups are true differences or artifacts of measurement bias (Van de Schoot, Lugtig, & Hox, 2012). The lack of measurement equivalence would create a threat to the construct validity of the inferences based on scores from the instrument (Zumbo, 2007).

Within a confirmatory factor analytic framework there are three aspects of measurement equivalence: configural invariance, metric invariance, and scalar invariance. Configural invariance means the factor structure for the scale between the groups is the same. For example, the SCI was intended to have a four-factor structure based on the four dimensions of sense of community proposed by Chavis and McMillan (1986). Configural invariance would be indicated by that same four factor structure being evident across rural, suburban, and urban populations. Metric invariance means that the factor loadings of the items on the instrument are equivalent across each group. For example, the factor loading of item 1 of the SCI on the Needs Fulfillment latent construct would be the same across rural, suburban, and urban communities. Scalar invariance means the item intercepts are equal across all populations. The combination of these three facets of measurement equivalence means that participants who have the same true score on an item should also have the same observed score.

To date, there are very few studies examining the measurement invariance of the SCI. Furthermore, there is no examination of the SCI for measurement invariance across rural, suburban, and urban neighborhoods. Obst and White (2004) examined the factor structure of a 10-item version of the SCI which had been edited to be applicable to university campus communities. In this study, they found measurement invariance for that scale between neighborhoods, university campus, and student groups. However, since the instrument examined in this study had changed item wordings and excluded items from the scale, these results cannot be used as evidence for measurement invariance of the 12-item SCI. Coffman and BeLue (2009) examined the 12-item SCI for differential item functioning (DIF) between White and African American respondents using latent trait modeling. A two-parameter logistic model was fit separately to each group, then the item scores were tested for differences between groups using a chi-square analysis with the Bonferroni correction. None of the items exhibited DIF by race. While these results were promising and showed evidence of measurement invariance by race, they did not constitute evidence of measurement invariance across different neighborhood types.

Importance of measurement invariance

If neighborhood effects scholars are to understand the effect that differing levels of neighborhood sense of community have on residents, it is essential that researchers have confidence that the differences in levels of sense of community indicated by scores on the SCI are true differences and not artifacts of measurement bias. Without evidence of measurement invariance, SCI scores of neighborhoods from different areas cannot be meaningfully compared and any differences in mean scores could be due to measurement non-invariance instead of true differences (Brown, 2015; Messick, 1989; Vandenberg & Lance, 2000). Two studies are not sufficient to support a claim of measurement invariance of the SCI. As with any other form of

scientific inquiry, the study of measurement invariance requires a body of research to support the claims that the instrument functions the same across different groups, in this case types of neighborhoods. The purpose of this study is to examine the SCI for measurement invariance across rural, urban, and suburban neighborhoods.

Methodology

This study was a cross-sectional survey. An internet link to the instruments was shared through public posts on Facebook, posts on Facebook neighborhood group pages for neighborhoods across the United States, and linking advertisements on Facebook targeted toward adults in the United States. Facebook is an online social media and social networking site that allows users to connect with their friends, family members, neighbors, coworkers, or individuals who share similar interests. Users can share messages, links to websites, as well as digital images and videos either publicly or to specific groups. Public posts are not only visible to all Facebook users, but will also show up on searches conducted on online search engines such as Google. Users can also create and join groups with other users based on common interests or location such as workplace, neighborhood, or university.

Social media recruiting and online data collection

While the use of social media in research recruiting is a relatively new practice, there is emerging evidence that it is possible to collect fairly large sample sizes using it as a primary method of recruitment. For example, King, O'Rourke, and DeLongis (2014) compared recruiting using Facebook and traditional methods. They were able to recruit over 2,000 participants in 19 days, over 90% of whom reported accessing the study through Facebook (King et al., 2014). Moreover, the results indicated that using Facebook targeted advertising allowed for the recruitment of low-prevalence or underrepresented populations.

While online data collection continues to gain popularity, studies examining the effect of using online data collection versus offline data collection on sample characteristics is still sparse. Schillewaert and Meulemeester (2005) compared the response distributions of two online (pop-ups and online panels) and two offline (mail and telephone) data collection methods. The results indicated differences in responses for each methodology based on age, education, and professional situation. Specifically, online respondents tended to be younger, more highly educated, and have white collar jobs. However, more recent survey by the Pew Research Institute indicates that between 2000 and 2018, internet use by adults in the U.S. aged 65 and older increased from 14% to 66% (Pew Research Institute, 2018). Nonetheless, there are still differences in internet usage by income and education level. For example, as of January 2018 while 97% of adults making $75,000 and over reported using the internet, that internet usage for adults making less than $30,000 a year was at 81% (Pew Research Institute, 2018). In addition, while 97% of college graduates used the internet, only 65% of adults with less than a high school education did so. This being the case, there is the risk that a sample collected on line could be skewed toward higher income, more highly educated respondents.

Sampling

The sample was collected using snowball and convenience sampling. Neighborhood groups were identified through several searching protocols. This was done to acquire the greatest number of contacts and greatest level of visibility of the study possible, and to increase variability in respondent characteristics. First, “neighborhood group” was searched in Facebook’s search function. Next, neighborhood groups were identified through group suggestions under Facebook’s Group Discover page. Finally, neighborhood groups were identified through searching neighborhood groups by state alphabetically (i.e. Alabama, Alaska,

Arizona, and so on). Group administrators were contacted and asked if they would be willing to allow the survey to be posted on their page. The message included a description of the purpose of the research project as well as risks and benefits. A link was posted to the main page of the neighborhood group Facebook pages from which permission was obtained. Administrators and groups that declined permission were not contacted further. Group members accessed the informed consent and instrument materials through the posted link. All group rules and protocols, as well as all Facebook policies and regulations, were respected and followed. In total, 164 groups were contacted. Of those, 33 (20%) groups responded and gave permission, 4 groups responded and declined permission, and 127 did not respond at all.

To broaden the sampling base, targeted advertisements linking Facebook users to the study website were purchased at a cost-per-click rate for five days. “Cost per click” means that the advertiser is only charged for the advertisements on which users click, thus reducing costs for advertisements that are less effective. Facebook targeted linking advertisements allow links to external websites to be promoted in the advertisement bar on users’ news feed. The advertisements targeted Facebook users 18 years and older in the United States.

Instrumentation

The instruments for this study included a demographic questionnaire with items collecting information such as respondent age, race, and gender, and neighborhood information such as neighborhood area (rural, suburban, or urban), length of residency, and membership in neighborhood groups. Rural areas were defined as sparsely populated residential and agricultural areas outside of a city environment. Suburban areas were defined as residential or mixed-use residential/commercial areas outside, but within commuting distance of, the central city area.

Urban areas were defined as areas within a downtown or central city location with high population and building density.

Sense of Community. Sense of community was measured via the 12-item SCI (Perkins et al., 1990). The reported range of possible scores for this scale varies, as different authors have used different response scales. While Perkins and colleagues (1990) originally used a True/False response set, Obst and White (2004) used a 7-point Likert-type scale. For this study, the 7-point Likert type scale was used to allow for respondents to record degrees of agreement or disagreement. Possible scores for the 7-point response scale ranged from 12 to 84, with higher scores indicating higher levels of sense of community. The estimated total scale reliability coefficients vary. Perkins and colleagues (1990) calculated an estimated reliability coefficient of $\alpha = .80$, while Chipuer and Pretty (1999) found total score estimated reliability coefficients ranging from .64 to .69.

Peterson, Speer, and Hughey (2006) found evidence that the negatively worded items on the SCI adversely affected the evidence for factorial validity of the SCI. Other researchers have found evidence that including both positively and negatively worded items can lead to problems with factor structure (Pilotte & Gable, 1990; Schmitz & Baer, 2001), potentially creating a methods factor associated with the items' wording. Therefore, the negatively worded items in each subscale were replaced with a positively worded item. For example, item 2 "People in this neighborhood do not share the same values" was replaced with "People in this neighborhood share the same values". While having all positively worded items may raise concerns regarding response set bias, introducing negatively worded items can also impact the psychometric quality of the instrument (e.g. more measurement error, introducing scoring factors), cause problems with validity, and cause construct misspecification (*for a review, see* Dalal & Carter, 2015).

Given the evidence that the negatively worded items in the SCI may be adversely affecting the psychometric quality of the instrument, these items were replaced with positively worded items with the same content.

Data Collection

The data for this study was collected online using Qualtrics, an online data collection and analysis platform. Settings were selected to allow for the data to be collected completely anonymously, with no email addresses or IP addresses saved. The survey was available for five weeks, with reminders posted at weeks two and three. The items after those measuring demographic characteristics were randomized in Qualtrics to mitigate bias related to order effects (Nunnally, 1978). Initially, there was a poor response rate from urban and rural participants. Therefore, further neighborhood group requests were targeted to rural and urban neighborhood groups. These neighborhoods were identified through their location listing. For example, neighborhoods in the Bronx, NY were targeted as urban while neighborhoods in Union County, TN were targeted as rural. Approximately 35 additional rural and urban neighborhoods were messaged. Ten neighborhoods responded, all of which were rural. While rural participation increased after more targeted contacts, there continued to be a poor response rate from urban participants. Targeted messages were also sent to neighborhoods with larger minority populations to try to mitigate potential minority under representation.

Data Analysis

The data were screened and cleaned. Five cases with international data were deleted to mitigate validity issues due to differing interpretations of the concepts presented due to language barriers or cultural differences. A missing values analysis was then run. The results indicated that

less than 1% of data were missing on all variables. Missing data were handled using Full Information Maximum-Likelihood (FIML). Recent research suggests that FIML is a superior form of handling missing data in its efficiency and its ability to produce less biased point estimates (Allison, 2012; Olinsky, Chen, & Harlow, 2003; von Hippel, 2016).

Similarly worded items can introduce correlated error that is not based solely on the latent construct (Brown, 2015). Therefore, items from the SCI were examined prior to data analysis, as suggested by Brown (2015), for similar wording or common content to identify sources of potentially correlated error covariances in the measurement model. This was done in advance so that if it became necessary to refine the model using Modification Indices, decisions regarding which error terms to correlate would be conceptually sound (Brown, 2015). Ten of the twelve items were identified as having common content (*see Table 1).* For example, item 4 ("I can recognize most of the people who live in my neighborhood") and item 6 ("At least some of my neighbors know me") were identified as items with similar content. The common content of these items implies a similarity between items based on a construct other than sense of community.

Confirmatory factor analysis. To test for configural invariance, the single factor model was fit to the data from each type of area separately. To test for metric invariance, the factor loadings were fixed equal for the items across all three groups. A small root mean square error of approximation (RMSEA) and a statistically non-significant Chi-square change, $\Delta\chi^2$, would provide evidence of metric invariance. To test for metric and scalar invariance, factor loadings and intercepts were fixed equal for items across all three groups. Again, a small RMSEA and statistically non-significant $\Delta\chi^2$ statistic would suggest both metric and scalar invariance. The

estimated reliability coefficients of scores on the SCI for the different areas were calculated following the method described by Brown (2015).

Results

Sample Characteristics

A total of 787 complete responses were obtained. Of these, 245 (31%) were residents of rural neighborhoods, 419 (53%) were residents of suburban neighborhoods, and 123 (16%) were residents of urban neighborhoods. Approximately 90.5% of respondents were White/Caucasian, 2.2% were African American, 2.2% were Hispanic/Latino(a), 1.9% were East Asian/Asian American, and the remaining 3.1% included Middle Eastern, Pacific Islander, South Asian, and multi-racial respondents. The majority of respondents, 78.8%, were female and 20.9% were male. About 7.2% were between 18 – 25 years old, 32.9% between 26 – 35, 20.9% between 36 – 45, 17.3% between 46 – 55, and 21.7% were 56 or older. Most respondents (70.3%) were from the southeastern or south atlantic states, 11.5% were from the Midwest, 7.3% were from the North East, 6% were from the south central or south western states, and 4.9% were from the Pacific coast.

Reliability

The estimated reliability coefficient for the rural data was .77, the reliability coefficient for the suburban data was .75, and the reliability coefficient for the urban data was .77. While no study was found reporting estimated reliability coefficients for different neighborhood types, reliability coefficients for total scores and subscales on the SCI have been reported. Perkins et al. (1990) reported a total scores reliability coefficient of $\alpha = .80$ after using pairwise deletion of missing data. Long and Perkins (2003), using listwise deletion of missing data, found reliability

coefficients ranging from .69 to .75. Chipuer and Pretty (1999) calculated reliability coefficients ranging from .64 to .69. Therefore, the reliability coefficients found in this study were on the high end of previously reported ranges. Item means and skew and kurtosis statistics are reported in Table 2 (*see Appendix).*

Measurement Invariance

All analyses were conducted in SPSS AMOS version 23.0 (Arbuckle, 2014). First, the original four-factor model proposed by McMillan and Chavis (1986) was fitted to the total dataset with item loadings based on Peterson, Speer, and Hughey's (2006) factor analysis of the SCI. High correlations were estimated between the four factors, some of which surpassed unity (*see Table 3*), suggesting problematic fit for the four-factor model and suggesting that a one factor structure would be plausible. A one factor measurement model was then fit to the dataset to test for configural, metric, and scalar invariance.

The initial fit of the one factor model was poor, so the modification indices (MIs) for the standard residual covariances was checked. The MI output was compared to the list of SCI items identified as having similar wording or concepts. The error covariances between items identified as having common concepts were then correlated, and the single factor model was refit (*see Figure 1*). This model showed acceptable fit, with a standardized root mean residual (sRMR) of .03, a goodness of fit index (GFI) of .96, a normed fit index (NFI) of .96, a comparative fit index (CFI) of .97, and a RMSEA of .06 (95% CI = .05, .07, PCLOSE < .01).

To test for configural invariance, the single factor model with correlated error covariances was fitted to the rural data, the suburban data, and the urban data separately. The results for the rural one factor model with correlated error covariances yielded an overall chi-

square of $\chi^2 = 102.56$, $df = 46$, $p < .001$ with a sRMR of .04, a GFI of .94, a NFI of .92, a CFI of .96, and a RMSEA of .07. The results for the suburban one factor model with correlated error covariances yielded an overall chi-square of $\chi^2 = 105.37$, $df = 46$, $p < .001$ with a sRMR of .03, a GFI of .96, a NFI of .95, a CFI of .97, and a RMSEA of .06. The results for the urban one factor model with correlated covariances yielded an overall chi-square of $\chi^2 = 113.16$, $df = 46$, $p < .001$ with a sRMR of .06, a GFI of .86, an NFI of .86, a CFI of .109, and a RMSEA of .11. These findings were consistent with configural invariance.

A fixed factor loadings model with correlated error covariances was used to test for evidence of metric invariance. This model yielded a sRMR of .05, a RMSEA of .04 (.03, .04: PCLOSE = 1.00), and a CFI of .95, showing acceptable fit to the data. The chi-square change statistic was $\Delta\chi^2 = 26.68$, $df = 22$, $p = .22$. These findings were consistent with metric invariance.

Metric and scalar invariance were then tested using a fixed factor loadings and fixed intercepts model with correlated error covariances. This model showed acceptable fit with a sRMR of .05, a RMSEA of .04 (.04,.05: PCLOSE = .999), and a CFI of .94. However, the chi-square change was statistically significant ($\Delta\chi^2 = 103.16$, $df = 46$, $p < .00001$), indicating non-invariant intercepts and potential Differential Item Functioning (DIF). The model was then rerun, and potentially non-invariant intercepts were identified through the MIs. The model was then run again, and the intercepts for items 4,10, and 12 were allowed to vary between groups. Specifically, items 4 and 12 were allowed to vary for the rural group, item 4,10, and 12 were allowed to vary for the suburban group, and items 4, 10, and 12 were allowed to vary for the urban group.

The chi-square change statistic for this model was statistically significant ($\Delta\chi^2 = 52.15$, *df* = 6, $p < .000001$). These findings suggested DIF by area on items 4,10, and 12, which implied partial scalar invariance of the SCI. For clarity, the models for each area are presented first with only factor loadings and intercepts, and then with the correlated error covariances. The metric invariance models with bootstrap bias corrected 95% confidence intervals (CI) for item factor loadings and intercepts are shown in Figures 2, 3, and 4 (*see Appendix*). The bootstrap bias corrected 95% CIs for item error covariances are shown in Figures 5, 6, and 7 (*see Appendix*).

Discussion

This study examined the evidence for measurement invariance of the SCI. Specifically, the SCI was measured for evidence of configural invariance, metric invariance, and scalar invariance. While the SCI was initially intended to reflect the four-dimensional structure proposed by Chavis and McMillan (1986), the four-factor model fit the data poorly. This is not surprising, given prior studies' similar findings (Long & Perkins, 2003; Obst & White, 2004; Peterson, Speer, & Hughey, 2006). The one factor model with correlated error covariances showed an acceptable fit to the data across all three neighborhood types (rural, urban, and suburban), thus showing evidence of configural invariance for the SCI. The fixed factor loadings model showed acceptable fit to the data and yielded a statistically non-significant chi-square change statistic, thus providing evidence of metric invariance.

The fixed factor loadings and fixed intercepts model, however, while showing acceptable fit yielded a statistically significant chi-square change statistic, indicating the possible presence of non-invariant intercepts and possible DIF. Items 4, 10, and 12 were identified as potentially having non-invariant intercepts. Items 4 and 12 were then allowed to vary for the rural group, items 4, 10, and 12 were allowed to vary for the suburban group, and items 4, 10, and 12 were

allowed to vary for the urban group. This model yielded a statistically significant chi-square change statistic, indicating partial scalar invariance of the SCI.

Limitations

There are a few limitations to this study that should be considered. First, using self-select neighborhood area identification is a limitation because of the potential for inconsistent or incorrect area classifications. For example, an individual may indicate that they live in a "suburban" neighborhood because they live in a planned, named neighborhood development, when their neighborhood is actually located in a rural or urban area as defined by the item. This may have led to over- or underrepresentation of neighborhoods types which in turn could bias the results. Researchers could avoid this in future studies by quota sampling or purposive sampling of neighborhoods from areas they select themselves based on predetermined criteria or definitions of neighborhood types.

Another limitation to this study was the small urban sample size obtained. The small sample size may have been the source of problems with model fit for the urban data. A final limitation of this study is the under representation of minority populations in the data. This under representation may be an artifact of sampling method. Despite recruiting messages targeted to neighborhoods with high minority populations, these neighborhoods yielded very low response rates. As a result, the findings of this study may not be generalizable to urban, more heterogeneous, or high minority population neighborhoods. While emerging research has found measurement invariance of the SCI between White and African American respondents at the individual level (Coffman and BeLue, 2009), measurement invariance between other ethnicities has not been examined. In addition, no study has been conducted examining evidence of measurement invariance of the SCI between majority white neighborhoods and high minority

population neighborhoods. Future studies should be conducted utilizing random sampling or quota sampling to ensure appropriate representation of high minority population neighborhoods.

Conclusions and Implications

The SCI has been used to compare sense of community in rural, suburban, and urban neighborhoods (Kitchen, Williams, & Chowhan, 2012; Obst, Smith, & Zinkiewicz, 2002). In these studies, the assumption has been that the SCI functions the same way across all neighborhood types. The results of this study, however, challenge these assumptions. The findings indicated evidence for metric invariance and partial scalar invariance for the SCI. Specifically, the results indicated the presence of DIF by area on items 4,10, and 12. This means that the score an individual gets on these particular items depends not only on their level of sense of community, but on the type of neighborhood in which they reside. This in turn signifies that mean scores cannot be substantively compared across neighborhood types.

In terms of the continued use of the SCI, the most conservative implication of these findings is that, when comparing rural, urban, and suburban neighborhoods, items 4 ("I can recognize most of the people who live in my neighborhood"), 10 ("It is very important to me to live in this neighborhood"), and 12 ("I expect to live in this neighborhood a long time") should not be used. Problems with items 10 and 12 may in part stem from a lack of content validity of these items in measuring sense of community. While sense of community by definition refers to the social and emotional connection between residents, these two items more closely reflect the construct of neighborhood attachment, or attachment to place which is defined as an individual's emotional attachment to their neighborhood (Brown & Perkins, 1992). In addition, the wording of item 4 refers to recognition of neighbors. This is a more superficial level of interaction between residents than is indicated by the theoretical definition of sense of community, which is

characterized by feelings of emotional connectedness and belonging. Exclusion of these items will increase the validity of the inferences elicited by the scores on the SCI without substantially altering the SCI. This will allow the mean scores to be compared across different types of neighborhoods without bias.

It is possible that the construct of sense of community functions or is structured differently between different types of neighborhoods. This would mean that residents of urban, suburban, and rural neighborhoods each define or experience sense of community in different ways. Future research should test the measurement invariance of other sense of community instruments, such as the Brief Sense of Community Scale (BSCS: Peterson, Speer, & McMillan, 2008), among rural, urban, and suburban neighborhoods to see whether the problems identified in this study were limited to the SCI or can be extended to include the construct of sense of community itself.

Previous research on the SCI has indicated an unstable factor structure, with some studies yielding a one-factor structure and others changing the pattern of item loadings to produce a four-factor structure (Chipuer and Pretty, 1999; Long & Perkins, 2003; Obst & White, 2004). Peterson and colleagues (2008) proposed that this instability was due to the negatively worded items on the scale. However, this study replaced the negatively worded items with positively worded items with common content, but the four-factor model still did not provide a good fit to the data. This reflects more on the validity of the scale rather than the construct itself, as other instruments such as the BSCS (Peterson et al., 2008) successfully reflect the four-factor structure proposed by McMillan and Chavis (1986). Future studies should examine whether it is possible to revise the SCI in such a way that the items more accurately reflect the factor structure that it

was initially intended to mirror, or if instead the SCI should be discarded in favor of instruments that reflect the intended factor structure while exhibiting strong psychometric properties.

Future research should also include the examination of the SCI for measurement invariance by participant age and length of residence. The SCI has been used to study the impact of sense of community on various adolescent outcomes such as loneliness (Pretty, Andrewes, & Collett, 1994) and trauma related depression (Moscardino, Scrimin, Capello, & Altoe, 2010). However, in their study of the factor structure of the SCI between adult and adolescent respondents, Chipuer and Pretty (1999) found preliminary evidence that indicated an unstable factor structure (or possible configural noninvariance) between adult and adolescent respondents. In addition, length of residence is often examined as a predictor of SCI scores (Prezza et al., 2001; Wilson-Doenges, 2000), but to date there has been no examination of the SCI for measurement invariance by length of residence.

Future research should also include replication of the current findings using differing sampling and recruiting methods. While using social media as a recruitment tool for this study had the advantage of being efficient and cost effective, it had the disadvantage of not allowing for random sampling. Future replications of these findings may benefit from using traditional sampling methods or non-traditional sampling methods that are more conducive to obtaining a random sample of participants.

References

Brodsky, A., O'Campo, P., & Aronson, R.E. (1999). PSOC in community context: multilevel correlates of a measrure of psychological sense of community in low-income, urban neighborhoods. *Journal of Community Psychology, 27(*6), 659-679.

Brown, T. (2015). Confirmatory factor analysis for applied research (2nd ed.), New York: Guilford Press.

Brown, B. B. & Perkins, D. D. (1992). Disruptions in place attachment. In I. Altman & S. Low (Eds.), *Place attachment* (pp. 279 – 304). New York: Plenum.

Buckner, J. C. (1988). The development of an instrument to measure neighborhood cohesion. *American Journal of Community Psychology, 16*(6), 771-791.

Cantillon, D., Davidson, W. S., & Schweitzer, J.H. (2003). Measuring community social organization: Sense of community as a mediator in social disorganization theory. *Journal of Criminal Justice, 31,* 321 – 339.

Chavis, D.M. & Pretty, G.M.H. (1999). Sense of community: Advances in measurement and application. *Journal of Community Psychology, 27*(6), 635-642.

Chavis, D.M., Hogge, J.H., McMillan, D.W., & Wandersman, A. (1986). Sense of community through Brunswik's lens: A first look. *Journal of Community Psychology, 14*(1), 24-40.

Coffman, D.L. & BeLue, R. (2009). Disparities in sense of community: True race differences or differential item functioning? Journal of Community Psychology, 37(5), 547-558.

Curry, A., Latkin, C., & Davey-Rothwell, M. (2008). Pathways to depression: The impact of neighborhood violent crime on inner-city residents in Baltimore, Maryland, USA. *Social Science & Medicine, 67,* 23-30.

Dalal, D.K. & Carter, N.T. (2015). Negatively worded items negatively impact survey research. In C.E. Lance & R.J. Vandenberg (Eds.) *More statistical and methodological myths and urban legends* (pp. 112-132). New York, New York: Routledge.

Echeverria, S., Diez-Roux, A.V., Shea, S., Borrell, L.N., & Jackson, S. (2008). Associations of neighborhood problems and neighborhood social cohesion with mental health and health behaviors: The Multi-Ethnic Study of Atherosclerosis. *Health & Place, 14,* 853-865.

Fowler, K., Wareham-Fowler, S., & Barnes, C. (2013). Social context and depression severity and duration in Canadian men and women: exploring the influence of social support and sense of community belongingness. *Journal of Applied Social Psychology*, *43*(S1). E85–E96. Doi: 10.1111/jasp.12050

Greenfield, E. A., & Marks, N. F. (2010). Sense of community as a protective factor against long-term psychological effects of childhood violence. *Social Service Review*, *84*(1), 129-147.

King, D.B., O'Rourke, N., & DeLongis, A. (2014). Social media recruitment and online data collection: A beginner's guide and best practices for accessing low-prevalence and hard-to-reach populations. *Canadian Psychology/Psychologie Canadienne, 55*(4), 240-249.

Kitchen, P., Williams, A., & Chowhan, J. (2012). Sense of community belonging and health in Canada: A regional analysis. *Social Indicators Research*, *107*(1), 103-126.

Kohen, D. E., Leventhal, R., Dahinten, V.S., & McIntosh, C.N. (2008). Neighborhood disadvantage: Pathways of effects for young children. *Child Development, 79* (1). 156-169.

Li, Y., Sun, F., He, X., & Chan, K. S. (2011). Sense of community and depressive symptoms among older earthquake survivors following the 2008 earthquake in Chengdu China. *Journal of Community Psychology*, *39*(7), 776-785.

McMillan, D.W. & Chavis, D.W. (1986). Sense of Community: A definition and theory. Journal of Community Psychology, 14, 6-23.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.

Moscardino, U., Scrimin, S., Capello, F., & Altoe, G. (2010). Social support, sense of community, collectivistic values, and depressive symptoms in adolescent survivors of the 2004 Beslan terrorist attack. *Social Science & Medicine, 70,* 27 – 34.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

O'Brien, D. J., Hassinger, E. W., & Dershem, L. (1994). Community attachment and depression among residents in two rural Midwestern communities. *Rural Sociology*, *59*(2), 255-265.

Obst, P., Smith, S.G., & Zinkiewicz, L. (2002). An exploration of sense of community, part 3: Dimensions and predictors of psychological sense of community in geographical communities. Journal of Community Psychology, 30(1), 119-133.

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.

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.

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

Pew Research Center (2018). Internet/Broadband Fact Sheet. Retrieved from http://www.pewinternet.org/fact-sheet/internet-broadband/#

Pilotte, W.J. & Gable, R.K. (1990). The impact of positive and negative item stems on the validity of a computer anxiety scale. *Educational and Psychological Measurement, 50,* 603-610.

Pretty, G.M.H., Andrewes, L., & Collett, C. (1994). Exploring adolescents' sense of community and its relationship to loneliness. *Journal of Community Psychology, 22*(4), 346 – 358.

Prezza, M., Amici, M., Roberti, T.& Tedeschi, G. (2001). Sense of community referred to the whole town: Its relations with neighboring, loneliness, life satisfaction, and area of residents. Journal of Community Psychology, 29(1), 29-52.

Sarason, S.B. (1974). The psychological sense of community: Prospects for a community psychology. San Francisco: Jossey-Bass Publishers.

Schmitz, M.F. & Baer, J.C. (2001). The vicissitudes of measurement: A confirmatory factor analysis of the emotional autonomy scale. *Child Development, 72*(1), 207-219.

Skogan, W. G. (1990). Disorder and decline: crime and the spiral of decay in American cities. New York: The Free Press.

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. European Journal of Developmental Psychology, 9(4), 486-492, DOI: 10.1080/17405629.2012.686740

Vandenberg, R., & Lance, C. (2000). A Review and Synthesis of the Measurement Invariance Literature: suggestions, Practices, and Recommendations for Organizational Research. *Organizational research methods*, 3(1), 4 – 70.

Wandersman, A. & Nation, M. (1998). Urban neighborhoods and mental health: Psychological contributions to understanding toxicity, resilience, and interventions. *American Psychologist 54*(6), 647 – 656.

Wilson-Doenges, G. (2000). An exploration of sense of community and fear of crime in gated communities. *Environment and Behavior, 32*(5), 597 -611.

Zumbo, B. (2007). Three generations of DIF analysis: Where it has been, where it is now, and where it is going. *Language Assessment Quarterly, 4,* 223-233. dx.doi.org/10.1080/15434300701375832

Appendix

Tables.

Table 1. SCI items with similar wording or common content

Item	Similar content with
1	Items 5,10,12
2	Item 3
4	Item 6
5	Items 10, 12
10	Item 12

Table 2. SCI item skew and kurtosis statistics with standard error.

	Skew	Std. Error	Kurtosis	Std. Error
SCI1	-1.55	.09	2.09	.17
SCI2	-.59	.09	.16	.17
SCI3	-.58	.09	.07	.17
SCI4	-.31	.09	-1.14	.17
SCI5	-1.42	.09	1.58	.17
SCI6	-1.47	.09	2.34	.17
SCI7	-.54	.09	-.46	.17
SCI8	-.24	.09	-.89	.17
SCI9	-.69	.09	.07	.17
SCI10	-.23	.09	-.93	.17
SCI11	-1.26	.09	2.24	.17
SCI12	-.68	.09	-.79	.17

Table 3. Estimated correlations between factors of the four-factor model of the Sense of Community Index.

	NF	M	I
NF			
M	1.05		
I	.93	.92	
EC	1.44	1.36	1.06

Figures.

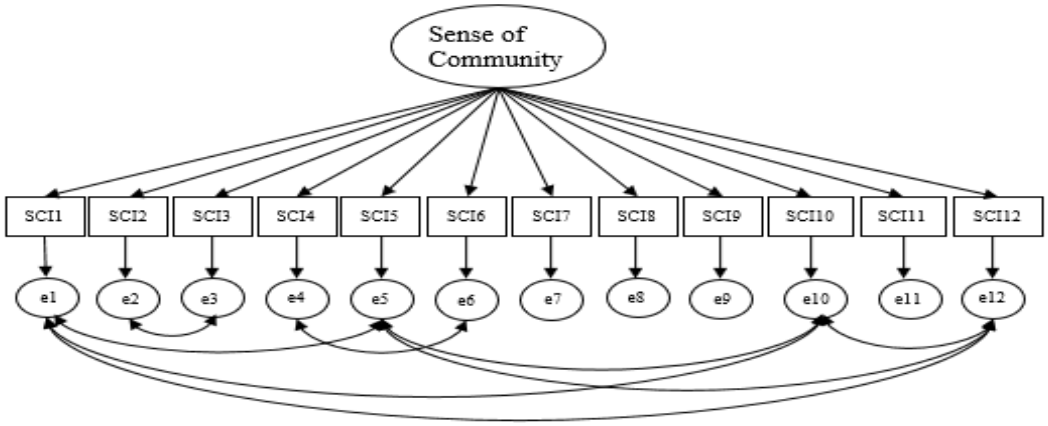

Figure 1. Single factor model of Sense of Community Index with correlated error covariances.

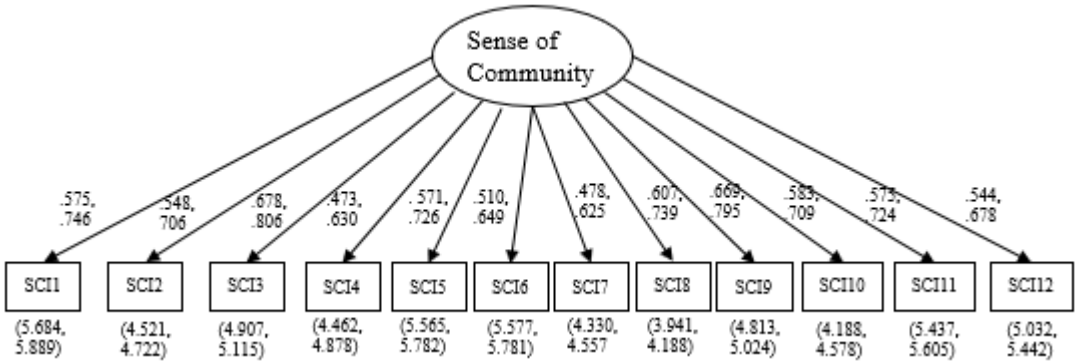

Figure 2. Bootstrap bias corrected 95% CIs for the standardized factor loadings and intercepts for the rural one factor model with correlated error variances.

Note: Numbers by single headed arrows indicate the CIs for the factor loadings. Numbers by boxes indicate the CIs for the intercepts

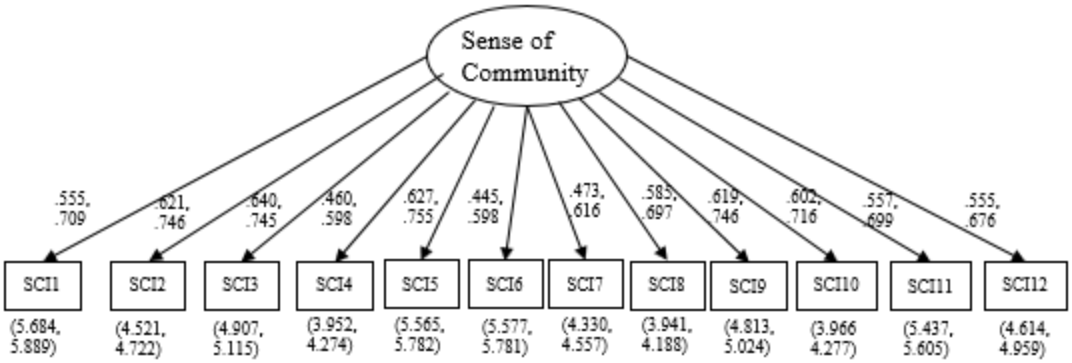

Figure 3. Bootstrap bias corrected 95% CIs for the standardized factor loadings and intercepts for the suburban one factor model with correlated error variances.

Note: Numbers by single headed arrows indicate the CIs for the factor loadings. Numbers by boxes indicate the CIs for the intercepts.

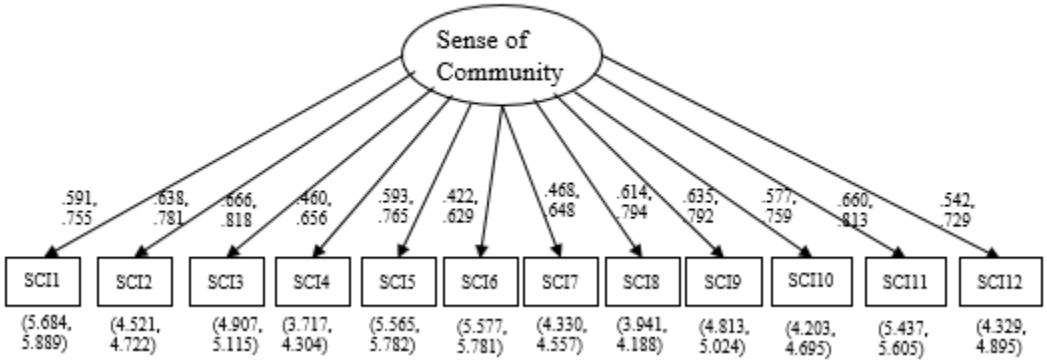

Figure 4. Bootstrap bias corrected 95% CIs for the standardized factor loadings and intercepts for the urban one factor model with correlated error variances.

Note: Numbers by single headed arrows indicate the CIs for the factor loadings. Numbers by boxes indicate the CIs for the intercepts

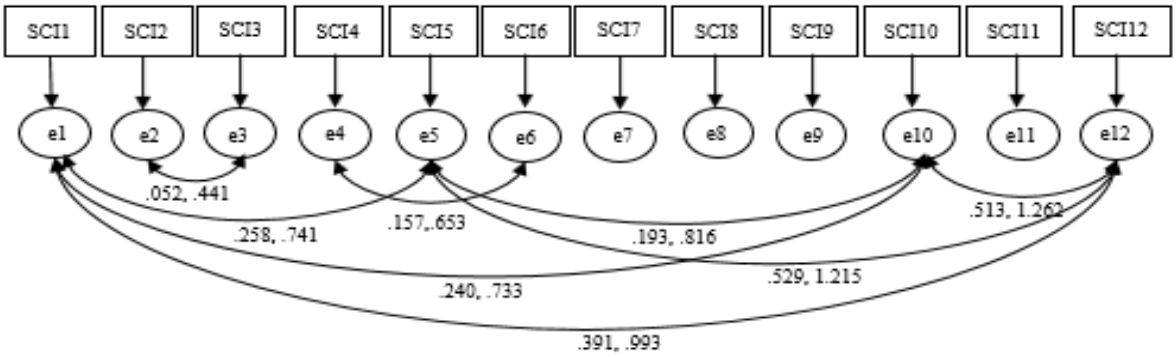

Figure 5. Bootstrap bias corrected 95% CIs for the rural model error covariances.

Note: Numbers below double headed arrows indicate CIs for the correlated error covariances.

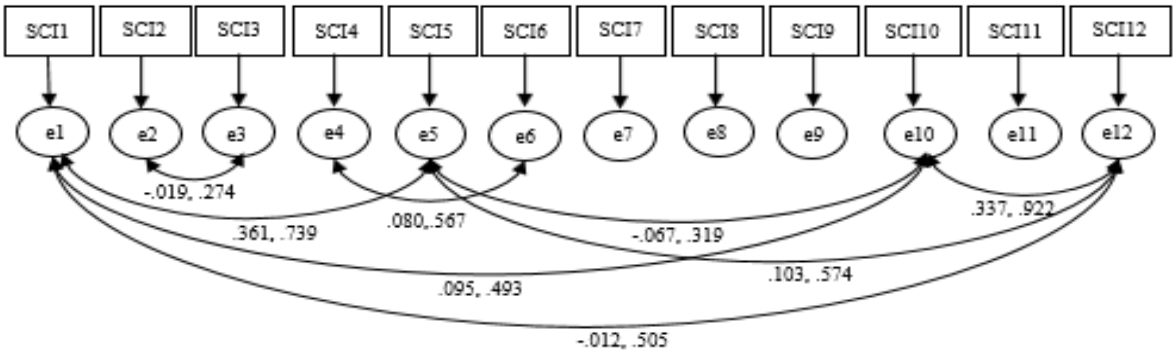

Figure 6. Bootstrap bias corrected 95% CIs for the suburban model error covariances.

Note: Numbers below double headed arrows indicate CIs for the correlated error covariances.

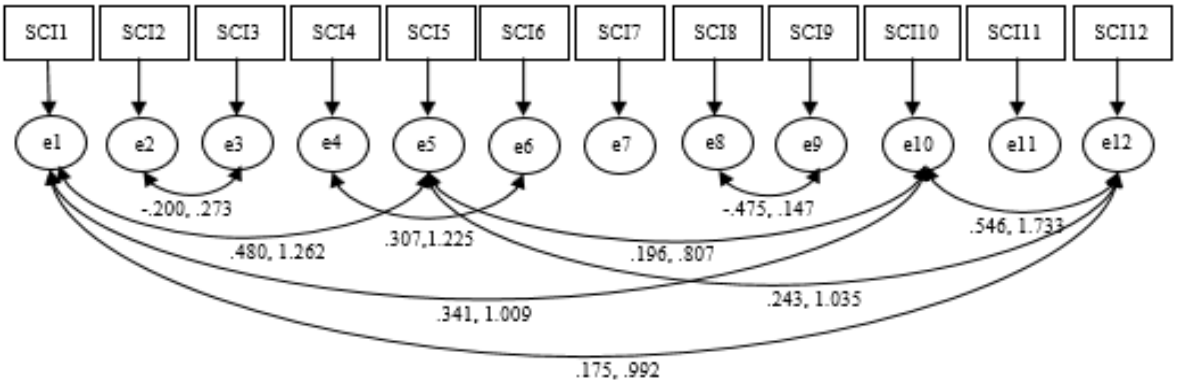

Figure 7. Bootstrap bias corrected 95% CIs for the urban model error covariances.

Note: Numbers below double headed arrows indicate CIs for the correlated error covariances.

Chapter 3

Convergent and divergent validity of the Sense of Community Index

Abstract

Sense of community is an important neighborhood social characteristic that has important effects on mental health and emotional wellbeing (Fowler, Wareham-Fowler, & Barnes, 2013; Kitchen, Williams, & Chowhan, 2012; Li, Sun, He, & Chan, 2011Prezza & Pacilli, 2007; Prezza et al., 2001; Wang et al., 2015). The Sense of Community Index (SCI: Perkins et al., 1990) is the most commonly used instrument to measure sense of community (Chavis & Pretty, 1999), however there continues to be a paucity of validity evidence to support such widespread use. This study examined the evidence of convergent and divergent validity for the SCI using confirmatory factor analysis. Results indicated that the SCI total scores loaded strongly and statistically significantly on the sense of community latent construct and did not cross-load onto the neighborhood disorder construct. Recommendations for future studies include further divergent validity studies with closely related but distinct constructs.

Keywords: neighborhood effects, construct validity, Sense of Community Index, neighborhood measurement

Background

The social characteristics of the environment in which one lives have an important influence on one's emotional functioning (Echeverria et al., 2008; Gary, Stark, & LaVeist, 2007; Latkin & Curry, 2003; Sampson, Morenoff, & Gannon-Rowley, 2002; Stockdale et al., 2007). For example, in their study of the impact of neighborhood disorder on mental health outcomes, Latkin and Curry (2003) found that resident perceptions of neighborhood characteristics predicted resident depression levels nine months later.

Neighborhood social characteristics are characteristics that concern how the residents of a neighborhood are connected to each other socially and their psychological or emotional experience of their neighborhood. Social cohesion, community attachment, and sense of community are examples of neighborhood social characteristics (*for a review see* Sampson, Morenoff, & Gannon-Rowley, 2002). Neighborhood effects research indicates that sense of community (SoC) has important impacts on individuals' emotional wellbeing and interaction with their community. For example, SoC has been shown to decrease loneliness (Prezza & Pacilli, 2007; Prezza et al., 2001), to support the psychological wellbeing of military wives (Wang et al., 2015), to protect against depression (Fowler, Wareham-Fowler, & Barnes, 2013; Kitchen, Williams, & Chowhan, 2012; Li, Sun, He, & Chan, 2011; O'Brien, Hassinger, & Dershem, 1994), and to increase life satisfaction (Prezza & Constantini, 1998) and community participation (Chavis & Wandersman, 1990; Cicognani et al., 2008).

The Sense of Community Index (SCI: Chavis et al., 1986) is the most widely used instrument for measuring SoC (Chavis & Pretty, 1999; Chipuer & Pretty, 1999). The original 44-item SCI was developed by Chavis and colleagues (1986) using data from the 1979 Neighborhood Participation Project. Perkins and colleagues (1990) later developed the 12-item

short form of the SCI to measure SoC in their study of block association participation. This study will examine the 12-item form SCI, as this is the form most commonly used (Chavis & Pretty, 1999). For example, it has been used to measure SoC in studies examining the association between SoC and participation in neighborhood substance abuse prevention activities (Peterson & Reid, 2003), adolescent well-being (Pretty, Andrewes, & Collett, 1994; Pretty, Conroy, Dugay, Fowler, & Williams, 1996), and has been adapted to examine SoC in the workplace (Pretty, McCarthy, & Catano, 1992). The SCI's wide use and the reliance on inferences based on its scores to understand and intervene in important neighborhood and resident problems makes the quality of its validity evidence crucial. Despite the widespread use of the SCI, there remains a paucity of validity studies focused on evidence for convergent and divergent validity of scores from the SCI.

Literature Review

There continues to be very little research regarding validity evidence for inferences from scores on the SCI. While the SCI has been examined for factorial validity in several studies (Long & Perkins, 2003; Obst & White, 2004; Peterson, Speer, & Hughey, 2006), the evidence for convergent validity is limited to the original development study of the 44-item scale by Chavis and colleagues (1986). In that study, total scores from the SCI were correlated with the score from one item intended to measure sense of community. The correlation between the SCI scores and the sense of community score was .52. No known studies exist examining the convergent validity of the 12-item short form SCI (Perkins et al., 1990).

In addition, only collecting evidence of convergent validity is insufficient because the association identified between instruments purported to be related could be due to construct irrelevant variance or common method variance (Messick, 1983). Evidence for divergent validity

must also be examined to determine that the intercorrelations of the instruments being examined are not due to their relationship to another construct. To date, no known study has been conducted to examine the evidence for divergent validity of inferences based on scores from the SCI. The paltry evidence for convergent validity coupled with the lack of divergent validity evidence raises significant questions regarding the construct validity of the SCI. The objective of this study is to examine evidence for convergent and divergent validity of the short form SCI (Perkins et al, 1990).

Methodology

This study was a cross-sectional survey utilizing online data collection. A link to the instruments was posted on the social media platform Facebook, and the instruments were administered through Qualtrics. Posts containing internet links to the instruments were posted publicly, on Facebook neighborhood social groups, and via targeted advertisements. Public posts are visible to all Facebook users and are searchable on online search engines. There is emerging evidence that the use of social media as a recruiting method can yield large sample sizes at a lower cost than traditional methods of recruitment. For example, in their study comparing using Facebook versus traditional recruitment methods, King, O'Rourke, and DeLongis (2014) found that of the 2,000 participants they were able to recruit for a study over 19 days, 90% reported that they accessed the study through Facebook.

Sampling

Convenience and snowball sampling were used as the sampling methods. Links to the study site were posted publicly on Facebook and on Facebook neighborhood group pages for five weeks. Reminders were then posted and two and three weeks. Neighborhood groups were identified first through a search using Facebook's search function. The search term used was

"neighborhood group". Then, additional groups were identified by browsing Facebook's Group Discover page. Finally, neighborhood groups were identified through an alphabetical search conducted by state (i.e. Alabama, Alaska, etc.). Group administrators were then contacted for permission to post the study link to their neighborhood groups in a message including information about the study and a copy of the informed consent information. Groups with administrators who gave consent were then contacted with a link and an informed consent post which was either posted by the researcher or by the group administrators. Group administrators who did not respond or who declined consent were not contacted again. Groups members who clicked on the link were taken first to the informed consent page, and then on to the instruments. All group and Facebook rules and protocols were observed.

In addition to public and group posts, targeted advertisements with links to the instruments were purchased for five days at a cost-per-click rate. Targeted linking advertisements allow links from external websites, such as the link to the study instruments, to be endorsed in the advertisement sidebar of Facebook users' feed. The advertisements were targeted at all Facebook users in the United States over the age of 18. "Cost per click" means that the purchaser is only charged for advertisements on which users click, which reduces the cost of advertising.

Instrumentation

The six instruments used for this study included the SCI (Perkins et al., 1990), the Sense of Community Scale (SCS: Nasar & Julian, 1995), the Brief Sense of Community Scale (BSCS: Peterson, Speer, & McMillan, 2008), the Perceptions of Neighborhood Conditions Scale (PNCS:Elo, Mykyta, Margolis, & Culhane, 2009), the Neighborhood Disorder Scale (NDS:Ross & Mirowsky, 1999), and the Perceptions of Neighborhood Disorder Scale (PNDS:Curry, Latkin,

& Davey-Rothwell, 2008). Previous research has indicated that sense of community may play a mediating role in the relationship between neighborhood structural characteristics and elements of neighborhood disorder (Lindblad, Manturuk, and Quercia, 2013; Warner & Rountree, 1997). Specifically, the findings indicate that higher levels of sense of community mitigate the effects of neighborhood structural risk factors on crime. These particular scales were chosen based on their psychometric quality. The characteristics of each of the instruments are discussed below.

Sense of Community. The 12-item short-form SCI (Perkins et al., 1990) was used in this study to measure sense of community. The reported range of possible scores for this scale varies, as different authors have used different response scales. While Perkins and colleagues (1990) originally used a True/False response set, Obst and White (2004) used a 7-point Likert-type scale. For this study, the 7-point Likert type scale was used to allow for respondents to record degrees of agreement or disagreement. Possible scores for the 7-point response scale ranged from 12 to 84. The total scale estimated reliability coefficient has been reported as high as $\alpha = .80$ (Perkins et al., 1990), with additional coefficient estimates ranging from .64 to .75 (Chipuer & Pretty, 1999; Long & Perkins, 2003).

Nasar and Julian's (1995) SCS is an 11-item scale based on Glynn's (1981) original sense of community instrument. The response format is a 5-point Likert scale, with responses ranging from strongly disagree to strongly agree. Higher scores indicate higher levels of SoC. The range of possible scores is from 11 to 55. Items 3, 7, and 8 of this scale are negatively worded items, which were reverse scored before the total score was calculated. The SCS was chosen because it was developed from Glynn's (1981) measure of psychological sense of community, which while having satisfactory psychometric properties, is too long for use in a battery of instruments. The SCS operates on four of the six factors from Glynn's (1981) original

instrument. Namely, supportive relationships in the community, similarity and relationship patterns of community residents, individual involvement in the community, and community security. The SCS has shown encouraging indications of construct validity and reliability. The estimated total scale reliability was α = .87 (Nasar & Julian, 1995).

Peterson, Speer, and McMillan (2008) developed the BSCS based on the four-factor theory of McMillan and Chavis (1986). The BSCS is an 8-item scale, with a 5-point Likert scale response set ranging from strongly disagree to strongly agree. Higher scores indicate higher levels of SoC. Possible scores ranged from 8 to 40. This scale included only positively worded items. The BSCS was chosen because it was developed using the same theoretical factor structure as the SCI and has shown promising evidence of both convergent and factorial validity (Peterson et al., 2008). The estimated total scale reliability coefficient was α = .92, and the subscale reliability coefficients ranged from .77 - .94 (Peterson et al., 2008).

Perceived Neighborhood Disorder. The PNCS (Elo, Mykyta, Margolis, & Culhane, 2009) measures neighborhood conditions with three domains: crime and safety, physical disorder, and social disorder. Responses are coded on a 10-point Likert scale from (1) rarely/not worried to (10) frequently/very worried. Higher scores indicate higher levels of neighborhood disorder. In order to preserve the clarity of the disorder construct, this study only used the 5-item physical disorder and 4-item social disorder scales, totaling a 9-item scale. Thus, the range of possible scores was from 9 to 90. The estimated reliability coefficient for the physical disorder subscale was α = .9 and the estimated reliability coefficient for the social disorder scale was α = .86. Neighborhood level estimated reliability coefficients for the subscales were both α = .65 (Elo et al., 2009).

The NDS (Ross & Mirowsky, 1999) is a 15-item scale measuring physical and social disorder. Responses were coded on a 4-point Likert scale from (1) strongly disagree to (5) strongly agree. In the development study for this scale, Ross and Mirowsky (1999) initially identified two factors of disorder and order. These two factors were also associated with positively worded items (order) and negatively worded items (disorder). To ascertain whether this factor structure was a result of bias, the authors fitted a model with three factors: one for disorder (which consisted of elements of social disorder and physical cues of disorder such as graffiti) one for decay (which consisted of physical deterioration of the neighborhood), and one factor for agreement bias. Once agreement was added to the model, all disorder items loaded on one disorder factor, with items measuring order loading negatively on the disorder factor. These items were reverse scored. Higher total scores represented higher levels of neighborhood disorder. Possible scores ranged from 15 to 75. The estimated total scale reliability coefficient was $\alpha = .92$ (Ross & Mirowsky, 1999).

The PNDS (Curry, Latkin, & Davey-Rothwell, 2008) is a seven-item scale based on Perkins, Meeks, & Taylor's (1992) Block Environmental Inventory. Responses are coded on a 3-point response scale labeled "not a problem" (0), "somewhat of a problem" (1), and "big problem" (2). Higher scores indicate higher levels of neighborhood disorder. Possible scores range from 0 to 14. The total scale estimated reliability coefficient was $\alpha = .99$ (Curry et al., 2008).

Data Collection

Participants were recruited through public and neighborhood group posts on the social media platform Facebook. Neighborhood group administrators were contacted and asked permission to post the link to the survey on their page. If permission was granted, a link was

posted on the neighborhood page with a statement regarding the purpose of the study and the voluntary nature of participation. The link directed participants to the informed consent page. Participants who consented were then directed to the full survey. A public post was also posted with the link to the survey which included informed consent. Reminders were posted one and two weeks after the initial post.

Data was collected through Qualtrics, an online data collection and analysis platform. Qualtrics allows the recording of participant response data without any identifying information, including location or IP address. In this way, participants can respond completely anonymously to the surveys. The surveys were available for four weeks.

Data Analysis

The data were screened and cleaned. Five cases were identified from participants outside of the United States. These cases were removed from the dataset in order to minimize validity issues caused by language or cultural differences. Results of a missing values analysis (MVA) indicated that less than 1% of data were missing on each variable. Full Information Maximum Likelihood (FIML) was used in AMOS to address missing data. Recent research indicates that FIML is an excellent method of addressing missing data, superior to multiple imputation in both small and large samples, because of its ability to produce less biased parameter estimates (Allison, 2012; Olinsky, Chen, & Harlow, 2003; von Hippel, 2016).

A confirmatory factor analysis (CFA) was performed to examine for convergent and divergent validity using AMOS by SPSS (Arbuckle, 2014). Negatively worded items on the SCS were reverse scored, then total scale scores were computed for all scales. A two-latent-construct model was constructed with the sense of community scales loading onto a sense of community

latent construct, and the neighborhood disorder scales loading on the neighborhood disorder latent construct. Evidence of convergent validity would be indicated by the scores on the SCI, as well as the scores on the other measures of sense of community, loading on the sense of community construct. Evidence of divergent validity would be indicated by the scores on the SCI not cross-loading onto the neighborhood disorder construct (Messick, 1989).

Results

Sample Characteristics

A final total of 788 completed cases were obtained. Of these respondents, 90.5% were White/Caucasian, 2.2% were African American, 2.2% were Hispanic/Latino, and 1.9% were East Asian/Asian American. The remaining respondents (3.1%) were Middle Eastern, Pacific Islander, South Asian, and multi-racial. Approximately 78.8% were female and 20.9% were male. The sample consisted mostly of individuals in their 30s and 40s, with 7.2% between 18 and 25 years old, 32.9% between 26 – 35, 20.9% between 36 – 45, 17.3% between 46 -55, and 21.7% aged 56 and older.

Correlations between Total Scores

For the initial analysis of convergent and divergent validity evidence, a correlation matrix was constructed using SPSS 25 (*see Table 4*: IBM, 2017). As expected, the SCI correlated highly and statistically significantly with the other measures of SoC. In addition, the correlations between the SCI with the PNCS and the PNDS were low, negative, and statistically significant. The correlations between the NDS and SoC scales were moderate (approximately .5), negative, and statistically significant. The correlations for the NDS were a bit higher than expected, indicating a stronger than expected relationships. If the scores from the SCI and other SoC scales

represent SoC, the correlations between scores on these scales should be high. Conversely, given current theory and previous research correlations between the scores from the SoC scales and the ND scales should be low and negative (Lindblad, Manturuk, and Quercia, 2013; Warner & Rountree, 1997). Overall, these results were consistent with the convergent and divergent validity of scores from the SCI.

Confirmatory Factor Analysis

A two-factor CFA model was fitted to the data to confirm the above evidence of convergent divergent validity (*see Figure 8*). Initial results showed poor fit of the model to the data ($\chi^2 = 216.25$ $df = 8$, $p < .001$). While the NFI was .94 and the CFI was .94, the RMSEA was .182 (.161, .203: PCLOSE < .001). The Modification Indices (MI) were then examined. The MIs indicated a serious fit problem with the NDS (MI = 129.42) and the sense of community measures. The standardized residual covariances were also examined. The NDS had high standardized residual covariances with the BSCS (-3.51), the SCS (-4.40), and the SCI (-5.46). Due to these results and the higher than expected correlation between the NDS and the SoC measures, the content validity of NDS items was examined to determine if any items were problematic. Three items - "I'm always having trouble with my neighbors", "In my neighborhood, people watch out for each other", and "I can trust people in my neighborhood" - were identified as potentially problematic items. Social disorder is represented by social features that indicate the presence of neighborhood disorder such as the neighborhood being unsafe, or inadequate police protection (Ross & Mirowsky, 1999). The three items identified seem to represent neighbor connections with each other, that is the sense of community construct, instead of neighborhood disorder. These results suggested fit would substantially improve if the scores on the NDS loaded on the Sense of Community latent construct.

A path was added allowing scores on the NDS to load on both the neighborhood disorder latent construct and the sense of community latent construct (*see Figure 9*). The addition of this path improved fit substantially with an overall chi-square of χ^2 =23.89, *df* = 7 *p* = .001. The NFI was .99, the CFI was .99, and the RMSEA was .06 (.03, .08: PCLOSE = .32). The standardized residual covariances were also reduced between the NDS and the BSCS (.96), the SCS (-.06), and the SCI (-.79). The SCI loaded statistically significantly on sense of community, and there were no cross loadings for the SCI (*see Figure 10*). These results confirm convergent and divergent validity for the SCI. In addition, the cross loading of ND on to the sense of community latent construct suggested that the scores on the ND represent of both sense of community and neighborhood disorder. Mean, skew, and kurtosis statistics are recorded on Table 5 (*See Appendix).*

Discussion

This study examined the evidence for convergent and divergent validity of inferences drawn from the scores on the SCI. Examination of the correlation matrix yielded high, statistically significant correlations between the SCI and other sense of community scales, and negative, small correlations with two of the three neighborhood disorder measures. These findings were confirmed by the CFA which showed the scores from the SCI and other SoC scales loaded on the sense of community latent construct, with no cross-loading onto the neighborhood disorder latent construct. Additional findings indicated that the NDS loaded on both the neighborhood disorder latent construct and the sense of community latent construct. These results suggest that certain items on the NDS may represent both sense of community and neighborhood disorder.

Limitations

There are some limitations which must be considered in the interpretation of the results of this study. First, the CFA model testing convergent and divergent validity only tested the SCI against one alternative construct, when there could be several rival constructs to explain the scores on the SCI (Messick, 1989). Future convergent and divergent validity studies should test the SCI against multiple constructs to rule out rival hypotheses for explanations of scores.

A second limitation to be considered is that all the instruments in this study are self-report methods. This could introduce common method variance, which could artificially inflate the association between instruments. Common method variance is measurement variance that is due to using the same method to measure a construct rather than to the construct the instrument was intended to measure (Messick, 1989). Using multiple methods, that is, conducting a multimethod multi-trait analysis, would allow for the parsing out of common method bias which would give a more accurate reflection of the association between instruments. Future studies of the convergent and divergent validity of the SCI should use multiple methods of measuring both SoC and the divergent constructs to address this limitation.

A final limitation is the underrepresentation of minority populations in the sample. Over 90% of the respondents in this study were White/Caucasian. Therefore, the findings of this study may not be generalizable to minority populations. Future research should expand the examination of convergent and divergent validity evidence to different ethnic groups to ensure that the findings of construct validity hold.

Conclusions and Implications

This study contributes to the literature by being the first rigorous examination of evidence for convergent and divergent validity of inferences drawn from scores from the SCI. The findings of this study support the use of the SCI by researchers and social work practitioners in measuring the construct of SoC. Neighborhood effects researchers and community organizers and practitioners who are concerned with evaluating SoC or testing interventions to improve SoC levels can use the SCI for these purposes.

The results of this study also strengthen the claims for interpreting prior results of studies examining relationships between SoC and resident outcomes. This means the results of studies using the SCI to examine the effect of SoC on outcomes such as depression, loneliness, or childhood trauma are bolstered by this evidence of construct validity of the SCI. Additionally, the results of this study offer support for the use of previous research using the SCI for informing policy decisions and neighborhood interventions.

Future studies should expand the evidence for convergent and divergent validity of the SCI with instruments that measure constructs that are theorized to be more closely related yet still distinct from sense of community such as neighborhood attachment or social cohesion. This will help to narrow the operational definition of SoC and help neighborhood researchers refine the distinctions between SoC and closely related constructs. Since no one study is generalizable to every neighborhood type, people group, or demographic group, future studies should attempt to replicate these findings using comparison groups among different ethnic, gender, or racial groups and among different neighborhood types.

Finally, future research should re-examine the factor structure of the NDS, including a factor for sense of community. As reported, three items from the NDS seem to measure sense of community or a similar construct rather than neighborhood disorder. Some of the problems with the factor loadings of the NDS in this study may result from these items. Researchers intending to use this instrument to measure neighborhood disorder should do so either with caution understanding the problems that are introduced by these three items or eliminate these items altogether.

References

Allison, P. (2012). Handling missing data by maximum likelihood. In *SAS global forum* (Vol. 2012). Haverford, PA, USA: Statistical Horizons.

Arbuckle, J.L. (2014). Amos (Version 23.0) [Computer Program]. Chicago: IBM SPS

Brown, T. (2015). Confirmatory factor analysis for applied research (2nd ed.), New York: Guilford Press.

Chavis, D.M., & Pretty, G.M. (1999). Sense of community: Advances in measurement and application. *Journal of Community Psychology, 27*(6), 635-742.

Chavis, D. M., & Wandersman, A. (1990). Sense of community in the urban environment: A catalyst for participation and community development. *American journal of community psychology*, *18*(1), 55-81.

Chavis, D. M., Hogge, J.H., McMillan, D.W., & Wandersman,A. (1986). Sense of community through Brunswick's lens: A first look. *Journal of Community Psychology, 14*(1), 24 – 40.

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.

Cicognani, E., Pirini, C., Keyes, C., Joshanloo, M., Rostami, R., & Nosratabadi, M. (2008). Social participation, sense of community and social well being: A study on American, Italian and Iranian university students. *Social Indicators Research*, *89*(1), 97-112.

Curry, A., Latkin, C., & Davey-Rothwell, M. (2008). Pathways to depression: The impact of neighborhood violent crime on inner-city residents in Baltimore, Maryland, USA. *Social Science & Medicine, 67*(1), 23-30.

Echeverria, S., Diez-Roux, A.V., Shea, S., Borrell, L.N., & Jackson, S. (2008). Associations of neighborhood problems and neighborhood social cohesion with mental health and health behaviors: The Multi-Ethnic Study of Atherosclerosis. *Health & Place, 14*,853-865, doi:10.1016/j.healthplace.2008.01.004

Elo, I.T., Mykyta, L., Margolis, R., & Culhane, J.F. (2009). Perceptions of neighborhood disorder: the role of individual and neighborhood characteristics. *Social Science Quarterly, 90*(5), 1298-1320. Doi: 2050/10.1111/j.1540-6237.2009.00657.x

Fowler, K.,Wareham-Fowler, S., & Barnes, C. (2013). Social context and depression severity and duration in Canadian men and women: Exploring the influence of social support and sense of community belongingness. *Journal of Applied Social Psychology, 43*(S1), E85 – E96. Doi: 10.1111/jasp.12050

Gary, T. L., Stark, S. A., & LaVeist, T. A. (2007). Neighborhood characteristics and mental health among African Americans and whites living in a racially integrated urban community. *Health & Place, 13*(2), 569-575. Doi.org/10.1016/j.healthplace.2006.06.001

Geis, K. J. & Ross, C. E. (1998). A new look at urban alienation: The effect of neighborhood disorder on perceived powerlessness. *Social Psychology Quarterly, 61*(3), 232-246. Doi: 10.2307/2787110

IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.

King, D. B., O'Rourke, N., & DeLongis A. (2014). Social media recruitment and online data collection: A beginner's guide and best practices for accessing low-prevalence and hard-to-reach populations. *Canadian Psychology/Psychologie Canadienne, 55*(4), 240 – 249. Dx.doi.org/10.1037.a0038087

Kitchen, P., Williams, A., & Chowhan, J. (2012). Sense of community belonging and health in Canada: A regional analysis. *Social Indicators Research, 107*(1), 103 – 126. Doi.org/10.1007/s11205-011-9830-9

Latkin, C. A. & Curry, A. D. (2003). Stressful neighborhoods and depression: A prospective study of the impact of neighborhood disorder. *Journal of Health and Social Behavior, 44*(1), 34 – 44. Doi: 10.2307/1519814

Li, Y., Sun, F., He, X., & Chan, K. S. (2011). Sense of community and depressive symptoms among older earthquake survivors following the 2008 earthquake in Chengdu China. *Journal of Community Psychology, 39*(7), 776-785. Doi.org:2050/10.1002/jcop.20469

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-org:2050/10.1002/jcop.10046

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.

Nasar, J. L. & Julian, D. A. (1995). The psychological sense of community in the neighborhood. *Journal of the American Planning Association, 61*(2), 178-184. Doi.org:2050/10.1080/01944369508975631.

O'Brien, D. J., Hassinger, E. W., & Dershem, L. (1994). Community attachment and depression among residents in two rural Midwestern communities. *Rural Sociology*, *59*(2), 255-265.

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.

Olinsky, A., Chen, S., & Harlow, L. (2003). The comparative efficacy of imputation methods for missing data in structural equation modeling. *European Journal of Operational Research, 151*(1), 53-79. Doi.org/10.1016/S0377-2217(02)00578-7

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.

Peterson, N. A. & Reid, R. J. (2003). Paths to psychological empowerment in an urban community: Sense of community and citizen participation in substance abuse prevention activities. *Journal of Community Psychology, 31*(1), 25-38. Doi: 10.1002/jcop.10034

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.20109Peterson, Speer, & McMillan, 2008

Pretty, G.M.H., Andrewes, L., & Collett, C. (1994). Exploring adolescents' sense of community and its relationship to loneliness. *Journal of Community Psychology*, *22*(4), 346 – 358.

Pretty, G. M. H., Conroy, C., Dugay, J., Fowler, K., & Williams, D. (1996). Sense of community and its relevance to adolescents of all ages. *Journal of Community Psychology, 24*(4), 365-379.

Pretty, G. M. H., McCarthy, M. E., & Catano, V. M. (1992). Psychological environments and burnout: Gender considerations within the corporation. *Journal of Organizational Behavior, 13*(7), 701-711.

Prezza, M., & Costantini, S. (1998). Sense of community and life satisfaction: Investigation in three different territorial contexts. *Journal of community & applied social psychology*, *8*(3), 181-194.

Prezza, M., & Pacilli, M. G. (2007). Current fear of crime, sense of community, and loneliness in Italian adolescents: The role of autonomous mobility and play during childhood. *Journal of Community Psychology, 35*(2), 151-170.

Prezza, M., Amici, M., Roberti, T., & Tedeschi, G. (2001). Sense of community referred to the whole town: Its relations with neighboring, loneliness, life satisfaction, and area of residence. *Journal of community psychology*, *29*(1), 29-52.

Ross, C. E. & Mirowsky, J. (1999). Disorder and decay: The concept and measurement of perceived neighborhood disorder. *Urban Affairs Review, 43*(3), 412-432.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing "neighborhood effects": Social processes and new directions in research. *Annual Review of Sociology, 28*,443-478. Doi: 10.1146/annurev.soc.28.110601.141114.

Stockdale, S. E., Wells, K. B., Tang, L., Belin, T. R., Zhang, L., & Sherbourne, C. D. (2007). The importance of social context: Neighborhood stressors, stress-buffering mechanisms, and alcohol, drug, and mental health disorders. *Social Science & Medicine, 65,* 1867 - 1881. Doi: 10.1016/j.sociscimed.2007.05.045.

Vandenberg, R., & Lance, C. (2000). A Review and Synthesis of the Measurement Invariance Literature: suggestions, Practices, and Recommendations for Organizational Research. *Organizational research methods*, 3(1), 4 – 70.

von Hippel, P.T. (2016). New confidence intervals and bias comparisons show that maximum likelihood can beat multiple imputation in small samples. *Structural Equation Modeling, 23*(3), 422 – 437.

Wang, Mei-Chuan, et al. "Finding resilience: The mediation effect of sense of community on the psychological well-being of military spouses." *Journal of Mental Health Counseling* 37.2 (2015): 164-174.

Appendix

Tables.

Table 4. Correlation matrix of scales.

	SCI	SCS	BSCS	PNCS	NDS	PNDS
SCI	1					
SCS	.77**	1				
BSCS	.82**	.78**	1			
PNCS	-.33**	-.28**	-.25**	1		
NDS	-.56**	-.50**	-.48**	.72**	1	
PNDS	-.30**	-.25**	-.22**	.78**	.73**	1

** Significant at the .01 level (2-tailed)

Table 5. Mean, skew, and kurtosis statistics for scale total scores.

	Mean	Std. dev.	Skew	Std. Error	Kurtosis	Std. Error
SCI	59.12	13.23	-.47	.09	-.23	.17
SCS	39.08	7.14	-.35	.09	.04	.17
BSCS	26.90	6.22	-.54	.09	.40	.17
PNCS	18.98	12.05	2.17	.09	5.89	.17
NDS	27.32	8.62	1.01	.09	1.54	.17
PNDS	8.71	2.23	1.80	.09	3.33	.17

Figures.

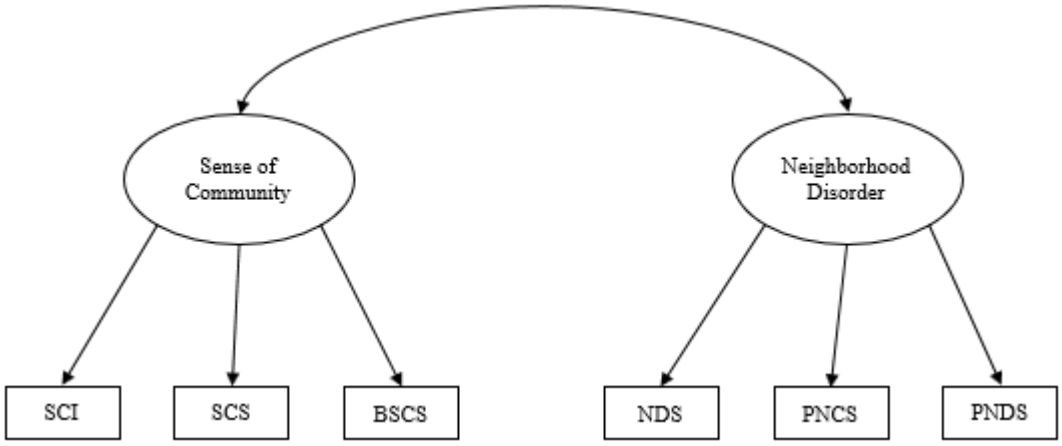

Figure 8. Simple Confirmatory Factor Analysis testing convergent and divergent validity of the SCI

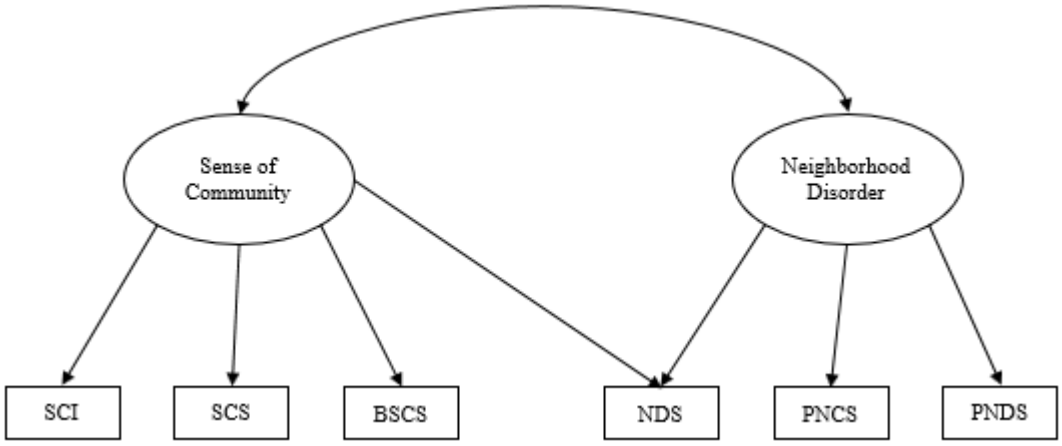

Figure 9. Simple Confirmatory Factor Analysis to test convergent and divergent validity of the SCI with added path between the sense of community latent construct and the NDS.

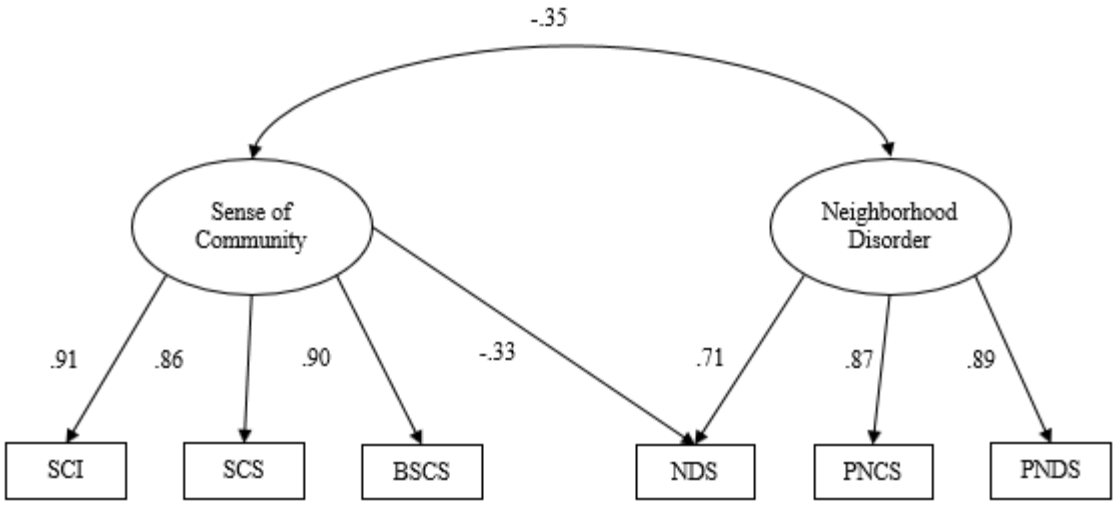

Figure 10. Standardized factor loadings of simple Confirmatory Factor Analysis to test convergent and divergent validity of the SCI.

Conclusion

This dissertation addressed the current state of the validation evidence for instruments measuring neighborhood social characteristics, and in particular the 12-item short form Sense of Community Index (SCI: Perkins et al., 1990). First, a critical literature review was conducted to evaluate the current state of the validation literature of instruments measuring neighborhood social characteristics. There were three research questions guiding this review: a) How many validation studies examining neighborhood social instruments exist? b) What is the quality of the validation studies? and c) What is the quality of the validity evidence presented for these instruments?

The results of the critical literature review indicated that there continues to be far too little validity evidence for these instruments, and too few validation studies. Fifty percent of the instruments included in this review were examined for only one type of validity evidence. Only two instruments were examined in more than one validity study. In order to trust the inferences made based on instrument scores, researchers must have evidence supporting the validity of those inferences. While the amount of validity evidence and validity studies in existence for each of these instruments is still too small, what evidence there is suggests that the instruments examined in the review have promising characteristics. Most of the instruments had moderate to high reliability coefficients, and the validity evidence presented in the studies was also promising.

The most promising instruments examined in this review were the SCI (Perkins et al., 1990), the Characteristics of the Residential Environment Scale (CRE: Handal, Barling, & Morrissy, 1981), and the Brief Sense of Community Scale (BSCS: Peterson, Speer, & McMillan, 2008). These instruments were identified as the most promising through examination of the

validity evidence presented by validation studies. The SCI (Perkins et al., 1990) was examined for evidence of content, convergent, and factorial validity. The CRE (Handal et al., 1981) was examined for evidence of convergent and criterion validity. The BSCS (Peterson et al., 2008) was examined for convergent and factorial validity.

Of these three, the SCI (Perkins et al., 1990) was the most studied instrument in the review. However, there remained certain validity issues for the SCI (Perkins et al., 1990) that should be addressed. For example, the evidence for convergent validity consisted of the correlation of total item scores with a one item scale designed to measure sense of community (Perkins et al., 1990). In addition, results of several studies investigating the factorial validity of the SCI (Perkins et al., 1990) indicated instability in the factor structure. These findings are important as the SCI is the most widely used instrument for measuring sense of community (Chavis & Pretty, 1999; Chipuer & Pretty, 1999). Given the instrument's wide use it is essential that the SCI (Perkins et al., 1990) have a strong foundation of validity evidence to support its use.

Recommendations from the critical literature review included the consideration of the theoretical underpinnings and intended use of the instruments being examined. The uses of the instruments included in this review require multiple types of validity evidence to be collected for each instrument. It was also recommended that validation of these instruments be viewed as a continuous process, requiring replication and broadening of the types of validation studies conducted. Finally, it was recommended that the SCI be examined for further, and stronger, evidence of various types of validity evidence, specifically for evidence of convergent validity.

The second and third papers aimed to fill the gaps in validity evidence for the SCI identified in the critical literature review. The second paper examined the measurement

invariance of the SCI using a confirmatory factor analysis. The SCI was examined for evidence of metric invariance and scalar invariance using a single factor model. Metric invariance was tested by using a fixed factor loadings model with correlated error covariances based on previously identified items with similar wording or common content. The model showed adequate fit to the data, demonstrating evidence for metric invariance. Scalar invariance was tested by using a fixed factor loadings and fixed intercepts model with correlated error covariances. While this model showed acceptable fit, the chi-square change statistic was statistically significant, which indicated possible differential item functioning (DIF). Modification indices suggested DIF by area on items 4, 10, and 12. Recommendations included dropping the two items which exhibited DIF when using the SCI to compare sense of community between rural, suburban, and urban neighborhoods.

The final paper examined the SCI for evidence of convergent and divergent validity. To examine for convergent and divergent validity, first a correlation matrix was constructed. The SCI correlated highly and statistically significantly with other measures of sense of community, and had low, negative, and statistically significant correlations with three measures of neighborhood disorder. These findings were confirmed with a two factor CFA model. Findings indicated that the SCI loaded strongly on the sense of community latent construct and had no cross loading on the neighborhood disorder latent construct. This indicates evidence of convergent and divergent validity for the SCI. However, the Neighborhood Disorder Scale (NDS: Ross & Mirowsky, 1999) loaded statistically significantly on both the neighborhood disorder latent construct and the sense of community latent construct. This indicates that the NDS may also measure sense of community. Recommendations included further validation of the SCI examining for divergent validity with constructs theorized to be related but distinct, such

as neighborhood attachment or social cohesion. Recommendations also included further examination of the factor structure and convergent/divergent validity of the NDS.

References

Chavis, D.M. & Pretty, G.M.H. (1999). Sense of community: Advances in measurement and application. *Journal of Community Psychology, 27*(6), 635-642.

Chipuer, H. M. & Pretty, G. M. H. (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.

Handal, P. J., Barling, P. W., & Morrissy, E. (1981). Development of perceived and preferred measures of physical and social characteristics of the residential environment and their relationship to satisfaction. *Journal of Community Psychology, 9*, 118-124.

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.

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.

Ross, C. E. & Mirowsky, J. (1999). Disorder and decay: The concept and measurement of perceived neighborhood disorder. *Urban Affairs Review, 43*(3), 412-432.

Vita

Erin Story completed a Bachelor of Psychology degree with a double major in Spanish Literature from the University of Tennessee, Knoxville in 2006 and a Master of Social Work degree from the University of Tennessee, Knoxville in 2009. Her practice experience includes community mental health and private practice therapy. She has provided mental health therapy services to adults, children, couples, groups, and families since 2009. She graduated with her Doctor of Philosophy degree in Social Work in August 2018. Her research interests include neighborhood effects research, neighborhood social measurement, and child maltreatment.